Math 221 Section 1 Spring 2025: lecture diary

Darij Grinberg

draft, June 6, 2025

(This is **NOT** a text or a set of notes. It is just an archive of what I write on my virtual blackboard in class. See `https://www.cip.ifi.lmu.de/~grinberg/t/24wd/24wd.pdf` for the actual notes.)

0. Preface

This is a course on **discrete mathematics**: the mathematics of finite, discrete objects, such as integers, finite sets. Integer sequences are also included even though they are infinite (since you only care about finite parts of them). We will not cover linear algebra or abstract algebra – while still discrete mathematics, these topics have their own classes.

The major topics we will cover in this course are

- **mathematical induction** and **recursion**;
- **elementary number theory** (divisibility, prime numbers, coprimality);
- basic **enumerative combinatorics** (counting and binomial coefficients), as well as the language of **maps** (aka **functions**).

We will not go very deep, nor will we be fully rigorous. Find more in the lecture notes; find even more in the references there.

Office hours (Korman 263):

- Wednesday 3–4PM;
- Thursday 11AM–noon.

1. Induction and recursion

1.1. The Tower of Hanoi

1.1.1. The puzzle

Let me start with a puzzle called the **Tower of Hanoi**.

You have 3 pegs (or rods). The first peg has n disks stacked on it. The disks have n different sizes, and are stacked in the order of their size, with the smallest one on top.

You can make a certain kind of moves ("**Hanoi moves**"): You can take the topmost disk from one peg and move it on top of another peg. However, this is only allowed if this disk is smaller than all other disks currently on the latter peg; i.e., you must never stack a larger disk atop a smaller disk.

Your **goal** is to move all n disks onto the third peg.

This game can be played online (e.g. `https://codepen.io/eliortabeka/pen/yOrrxG`, or `https://www.towersofhanoi.info/Play.aspx` for better graphics).

1.1.2. Some explorations

Let us analyze the case $n = 3$. Here is one strategy to win the game (i.e., achieve the goal) for $n = 3$:

1. Move the smallest disk from peg 1 to peg 3.
2. Move the middle disk from peg 1 to peg 2.
3. Move the smallest disk from peg 3 to peg 2.
4. Move the largest disk from peg 1 to peg 3.
5. Move the smallest disk from peg 2 to peg 1.
6. Move the middle disk from peg 2 to peg 3.
7. Move the smallest disk from peg 1 to peg 3.

So we can win in 7 moves for $n = 3$.

What about other values of n ? We ask ourselves:

> **Question 1.1.1.** **(a)** Can we always win the game?
> **(b)** If so, then what is the smallest # of moves we need?

Let us record the answers for small n's:

- For $n = 0$, we win in 0 moves (since all disks – i.e., all 0 of them – are on peg 3 already). This is the smallest #.
- For $n = 1$, we win in 1 move. This is the smallest #.
- For $n = 2$, we win in 3 moves. This is the smallest #.
- For $n = 3$, we win in 7 moves. Is this the smallest #?
- For $n = 4$, can we win? In how many moves?

Solving the problem by brute force gets exponentially harder as n increases. But maybe there is a pattern in our $n = 3$ strategy that we can detect and generalize to higher n ? We can summarize this strategy as follows:

1.–3. Move the two smaller disks from peg 1 onto peg 2 (using the $n = 2$ strategy, but with pegs 2 and 3 swapped).

4. Move the largest disk from peg 1 to peg 3.

5.–7. Move the two smaller disks from peg 2 onto peg 3 (using the $n = 2$ strategy, but with pegs 1 and 2 swapped).

So our $n = 3$ strategy consists of two (slightly modified) $n = 2$ strategies, with a single extra step (the moving of the largest disk) interspersed in between them. In particular, it requires $3 + 1 + 3 = 7$ steps.

Following this pattern, we now see a strategy for $n = 4$, winning the game in $7 + 1 + 7 = 15$ steps:

1.–7. Move the three smallest disks from peg 1 onto peg 2 (using the $n = 3$ strategy).

8. Move the largest disk from peg 1 to peg 3.

9.–15. Move the three smallest disks from peg 2 onto peg 3 (using the $n = 3$ strategy).

Thus, we don't just have a strategy for $n = 3$ and one for $n = 4$, but actually we have a "meta-strategy" that lets us win the game for n disks if we know how to win it for $n - 1$ disks. In a nutshell, it says "first move the $n - 1$ smaller disks onto peg 2; then move the largest disk onto peg 3; then move the $n - 1$ smaller disks onto peg 3". We will still call this a strategy.

1.1.3. The numbers m_n

Let us summarize what we gain from this.

Definition 1.1.2. For any integer $n \geq 0$, we let m_n denote the # of moves needed to win the Tower of Hanoi game with n disks. If the game cannot be won with n disks, then we set $m_n = \infty$ (not a number, just a symbol).

So our question (both parts **(a)** and **(b)**) comes down to computing m_n.
Here are some small values of m_n obtained using our strategy:

n	0	1	2	3	4	5	6	7	8
m_n	0	1	3	7	15	31	63	127	255

.

Note that our strategy yields $m_n = m_{n-1} + 1 + m_{n-1} = 2m_{n-1} + 1$ for each $n \geq 1$.

Right?

Not so fast! We have not actually computed m_n. We have only computed the number of steps that our strategy takes. Is this number of steps actually the minimal one, or can there be shortcuts? Who tells us that there isn't a faster strategy for (say) $n = 7$, which wins the game in (say) 109 moves?

So what we have really proved is the following:

Proposition 1.1.3. Let n be a positive integer. If m_{n-1} is an integer (i.e., not ∞), then $m_n \leq 2m_{n-1} + 1$.

Proof. Assume that m_{n-1} is an integer. Thus, we can win the game for $n-1$ disks in m_{n-1} moves. Let S be the strategy (= the sequence of moves) that does this. So the strategy S moves $n-1$ disks from peg 1 onto peg 3 in m_{n-1} moves.

Let S_{23} be the same strategy as S, but with the roles of pegs 2 and 3 swapped. Thus, S_{23} moves $n-1$ disks from peg 1 onto peg 2 in m_{n-1} moves.

Let S_{12} be the same strategy as S, but with the roles of pegs 1 and 2 swapped. Thus, S_{12} moves $n-1$ disks from peg 2 onto peg 3 in m_{n-1} moves.

Now, we proceed as follows to win the game with n disks:

A. We use strategy S_{23} to move the $n-1$ smaller disks from peg 1 onto peg 2. (This is allowed because the largest disk rests at the bottom of peg 1 and does not interfere with the movement of smaller disks.)

B. We move the largest disk from peg 1 onto peg 3. (This is allowed because all the other disks are on peg 2, so they don't hinder the movement.)

C. We use strategy S_{12} to move the $n-1$ smaller disks from peg 2 onto peg 3. (This is allowed because the largest disk is already on peg 3.)

This strategy wins the game for n disks in $m_{n-1} + 1 + m_{n-1} = 2m_{n-1} + 1$ moves. So $m_n \leq 2m_{n-1} + 1$. This proves the proposition. □

Now, we will prove that this inequality is actually an equality, so we do really have $m_n = 2m_{n-1} + 1$. In other words:

Proposition 1.1.4. Let n be a positive integer. If m_{n-1} is an integer (i.e., not ∞), then $m_n = 2m_{n-1} + 1$.

Proof. Again, assume that m_{n-1} is an integer.

We want to show that $m_n = 2m_{n-1} + 1$. It suffices to show that $m_n \geq 2m_{n-1} + 1$ (since $m_n \leq 2m_{n-1} + 1$ has already been proved above). In other words, it suffices to show that any winning strategy for n disks has at least $2m_{n-1} + 1$ many moves.

So let us consider a winning strategy T for n disks. Somewhere during this strategy T, the largest disk has to move (since it must eventually get from peg 1 to peg 3). Let us refer to these moves (the ones that move the largest disk) as the **special moves**. So we know that there is at least one special move.

Before the first special move, all the $n-1$ smaller disks must have moved from peg 1 to one of the other two pegs (otherwise, the largest disk would not be free or would not be able to find rest on another peg). Thus, before the first special move, we must have already won the ToH (Tower of Hanoi) game with $n-1$ disks. By definition of m_{n-1}, this requires at least m_{n-1} many moves. So we must have made at least m_{n-1} many moves before the first special move.

Now, consider what happens **after the last special move**. The last special move necessarily moves the largest disk to peg 3, but all the other $n-1$ disks have to still be moved on top of it. These other $n-1$ disks must be all on the same peg at the time of the last special move. Thus, we must still move these other $n-1$ disks from that peg to peg 3. This is again a ToH game with $n-1$ disks, and thus requires at least m_{n-1} moves to win. So we must make at least m_{n-1} many moves after the last special move.

Altogether, the strategy T thus has $\geq m_{n-1} + 1 + m_{n-1}$ moves, i.e., $\geq 2m_{n-1} + 1$ moves. Since T was completely arbitrary, this means that any winning strategy for n disks has $\geq 2m_{n-1} + 1$ moves. And that's what we wanted to show. □

Using the proposition we just proved, our table of values for m_n constructed above

n	0	1	2	3	4	5	6	7	8
m_n	0	1	3	7	15	31	63	127	255

is now justified. You can keep using the proposition to compute $m_9, m_{10}, m_{11}, \ldots$. Indeed, the equation

$$m_n = 2m_{n-1} + 1$$

is what is known as a **recursive formula** for the number m_n: a formula that lets you compute each m_n using the previous values $m_0, m_1, \ldots, m_{n-1}$.

Is this the best we can do? Can we find an **explicit formula** – i.e., a formula that gives m_n directly? It looks like

$$m_n \overset{?}{=} 2^n - 1.$$

Is there a way to see this without guessing? We can try:

$$\begin{aligned} m_n &= 2m_{n-1} + 1 \\ &= 2\left(2m_{n-2} + 1\right) + 1 \\ &= 4m_{n-2} + 2 + 1 \\ &= 4\left(2m_{n-3} + 1\right) + 2 + 1 \\ &= 8m_{n-3} + 4 + 2 + 1 \\ &= 8\left(2m_{n-4} + 1\right) + 4 + 2 + 1 \\ &= 16m_{n-4} + 8 + 4 + 2 + 1 \\ &= \cdots \\ &= 2^{n-1} + 2^{n-2} + \cdots + 4 + 2 + 1 \\ &= 1 + 2 + 4 + \cdots + 2^{n-2} + 2^{n-1}. \end{aligned}$$

Why is this the same as $2^n - 1$?

1.2. The Principle of Mathematical Induction

The easiest way to answer these questions is using a fundamental proof technique called **proof by induction**. This relies on the following principle:

> **Theorem 1.2.1** (Principle of Mathematical Induction, short PoMI)**.** Let b be an integer.
>
> Let $P(n)$ be a mathematical statement defined for each integer $n \geq b$.
>
> (For example, $P(n)$ can be "$n + 1 > n$" or "n is even" or "n is prime" or "there exists a prime number larger than n".)
>
> Assume the following:
>
> 1. The statement $P(b)$ holds (i.e., the statement $P(n)$ holds for $n = b$).
>
> 2. For each integer $n \geq b$, the implication $P(n) \Longrightarrow P(n+1)$ holds (i.e., if $P(n)$ holds, then $P(n+1)$ holds as well).
>
> Then, the statement $P(n)$ holds for every integer $n \geq b$.

Why is this principle true? Well, the assumptions yield that

$$\begin{aligned} &P(b) && \text{holds;} \\ P(b) &\Longrightarrow P(b+1) && \text{holds;} \\ P(b+1) &\Longrightarrow P(b+2) && \text{holds;} \\ P(b+2) &\Longrightarrow P(b+3) && \text{holds;} \\ &\ldots . \end{aligned}$$

By chaining together these implications, we see that $P(b)$ holds, $P(b+1)$ holds (since $P(b)$ and $P(b) \Longrightarrow P(b+1)$), and $P(b+2)$ holds (since $P(b+1)$ and $P(b+1) \Longrightarrow P(b+2)$), and $P(b+3)$ holds (for similar reasons and so on). Thus, by common sense, $P(n)$ must hold for each $n \geq b$ (because we eventually get to $P(n)$ if we follow this chain of reasoning). The PoMI is just a formalization of this common-sense argument.

Two classical metaphors for the PoMI:

- Think of an infinite daisy chain of lamps $P(b)$, $P(b+1)$, $P(b+2)$, ..., where each lamp $P(n)$ turns on the next one ($P(n+1)$) when it is turned on itself. The PoMI then says that turning on $P(b)$ will eventually turn on every lamp.
- Think of an infinite sequence of dominos standing in a row, close enough that if $P(n)$ falls, then it tips over $P(n+1)$. Then, the PoMI says that tipping over $P(b)$ will tip over all the dominos.

Now let us use the PoMI to prove our conjecture about m_n:

Theorem 1.2.2 (explicit answer to Tower of Hanoi)**.** For each integer $n \geq 0$, we let m_n be the # of moves needed to win the Tower of Hanoi game with n disks (or ∞ if the game cannot be won).

Then,

$$m_n = 2^n - 1 \qquad \text{for every integer } n \geq 0.$$

Proof. For each integer $n \geq 0$, we denote the statement "$m_n = 2^n - 1$" by $P(n)$. So our goal is to show that $P(n)$ holds for each integer $n \geq 0$.

According to the PoMI (applied to $b = 0$), it suffices to show that

1. the statement $P(0)$ holds;
2. for each integer $n \geq 0$, the implication $P(n) \Longrightarrow P(n+1)$ holds.

Proving these two claims will be our two goals; we call them Goal 1 and Goal 2.

Goal 1 is easy: It just requires us to prove $m_0 = 2^0 - 1$, which is clear because both sides are 0.

Now we attack Goal 2. Let $n \geq 0$ be an integer. We must prove the implication $P(n) \Longrightarrow P(n+1)$. So we assume that $P(n)$ holds, and we set out to prove that $P(n+1)$ holds.

Our assumption is that $P(n)$ holds, i.e., that $m_n = 2^n - 1$. In particular, m_n is an integer.

We need to show that $P(n+1)$ holds, i.e., that $m_{n+1} \overset{?}{=} 2^{n+1} - 1$.

The last proposition we showed says that $m_n = 2m_{n-1} + 1$ when $n \geq 1$. Applying this proposition to $n+1$ instead of n, we conclude that

$$\begin{aligned} m_{n+1} &= 2 \underbrace{m_n}_{=2^n-1} + 1 = 2 \cdot (2^n - 1) + 1 \\ &= \underbrace{2 \cdot 2^n}_{=2^{n+1}} - 2 + 1 = 2^{n+1} - 2 + 1 = 2^{n+1} - 1. \end{aligned}$$

In other words, $P(n+1)$ holds. Since we assumed $P(n)$ to prove this, we thus have proved $P(n) \Longrightarrow P(n+1)$. So Goal 2 is reached.

Thus, we can apply the PoMI, and we conclude that $P(n)$ holds for all $n \geq 0$. In other words, $m_n = 2^n - 1$ for all $n \geq 0$. This proves the theorem. □

1.3. Some more proofs by induction

A proof that uses the PoMI is called a **proof by induction** or **induction proof** or **inductive proof**. So our above proof of the theorem was a proof by induction. Let us see some more.

1.3.1. The sum of the first n positive integers

Theorem 1.3.1 ("Little Gauss formula"). For every integer $n \geq 0$, we have

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2}.$$

The LHS (= left hand side) here is understood as the sum of the first n positive integers. For $n = 0$, this is an empty sum (i.e., a sum with no addends), so its value is 0 by definition.

First proof of the theorem. We set

$$s_n := 1 + 2 + \cdots + n \qquad \text{for each } n \geq 0.$$

Thus, we must prove that $s_n \overset{?}{=} \frac{n(n+1)}{2}$ for each $n \geq 0$.

Let us denote this statement "$s_n = \frac{n(n+1)}{2}$" by $P(n)$. So we want to prove $P(n)$ for each $n \geq 0$. According to the PoMI, it suffices to show that

1. the statement $P(0)$ holds;
2. for each $n \geq 0$, the implication $P(n) \Longrightarrow P(n+1)$ holds.

Goal 1 is trivial: $P(0)$ says that $s_0 = \frac{0(0+1)}{2}$, which is true because both sides are 0.

For Goal 2, we fix an integer $n \geq 0$. We want to prove the implication $P(n) \Longrightarrow P(n+1)$. So we assume that $P(n)$ holds, and we set out to prove $P(n+1)$.

By assumption, $P(n)$ holds, i.e., we have

$$s_n = \frac{n(n+1)}{2}.$$

Our goal is to prove $P(n+1)$, that is, to prove that

$$s_{n+1} \overset{?}{=} \frac{(n+1)((n+1)+1)}{2}.$$

To do so, we observe that

$$\begin{aligned} s_{n+1} &= 1+2+\cdots+(n+1) = \underbrace{(1+2+\cdots+n)}_{=s_n} + (n+1) \\ &= \underbrace{s_n}_{=\frac{n(n+1)}{2}} + (n+1) = \frac{n(n+1)}{2} + (n+1) \\ &= (n+1)\left(\frac{n}{2}+1\right) = (n+1)\frac{n+2}{2} = \frac{(n+1)(n+2)}{2} \\ &= \frac{(n+1)((n+1)+1)}{2}. \end{aligned}$$

This is precisely $P(n+1)$. So we have proved $P(n) \Longrightarrow P(n+1)$. Thus, goal 2 is achieved.

Having achieved both goals, we now invoke the PoMI and conclude that $P(n)$ holds for all $n \geq 0$. Proof complete. □

This is not the only way to prove the theorem. Little Gauss did it without induction:

Second proof of the theorem. We have

$$\begin{aligned}
&2\cdot(1+2+\cdots+n)\\
&=(1+2+\cdots+n)+(1+2+\cdots+n)\\
&=(1+2+\cdots+n)+(n+(n-1)+\cdots+1)\\
&\qquad\qquad\qquad\text{(here, we reversed the second sum)}\\
&=(1+n)+(2+(n-1))+\cdots+(n+1)\\
&=\underbrace{(n+1)+(n+1)+\cdots+(n+1)}_{n\text{ times}}\\
&=n(n+1).
\end{aligned}$$

Dividing by 2 yields the claim. □

1.3.2. The sum of the squares of the first n positive integers

Theorem 1.3.2. For every integer $n\geq 0$, we have

$$1^2+2^2+\cdots+n^2=\frac{n(n+1)(2n+1)}{6}.$$

Proof. The following proof is almost a word-by-word copy of the first proof of the previous theorem; only the computations change.

We set

$$s_n:=1^2+2^2+\cdots+n^2,$$

and we define $P(n)$ to be the statement "$s_n=\dfrac{n(n+1)(2n+1)}{6}$". So we must prove $P(n)$ for all integers $n\geq 0$.

According to the PoMI, it suffices to show that

1. the statement $P(0)$ holds;
2. for each $n\geq 0$, the implication $P(n)\Longrightarrow P(n+1)$ holds.

Goal 1 is easy again, boiling down to $0=0$.

Now to Goal 2. We fix an integer $n\geq 0$; we assume that $P(n)$ holds; we set out to prove that $P(n+1)$ holds.

By assumption, $P(n)$ holds, i.e., we have

$$s_n=\frac{n(n+1)(2n+1)}{6}.$$

We want to prove that $P(n+1)$ holds, i.e., that we have

$$s_{n+1}\overset{?}{=}\frac{(n+1)((n+1)+1)(2(n+1)+1)}{6}.$$

To prove this, we observe that

$$\begin{aligned}
s_{n+1} &= 1^2 + 2^2 + \cdots + (n+1)^2 = \underbrace{\left(1^2 + 2^2 + \cdots + n^2\right)}_{=s_n} + (n+1)^2 \\
&= s_n + (n+1)^2 = \frac{n(n+1)(2n+1)}{6} + (n+1)^2 \\
&\qquad \left(\text{by our assumption } s_n = \frac{n(n+1)(2n+1)}{6}\right) \\
&= \frac{n(n+1)(2n+1) + 6(n+1)^2}{6} \\
&= \frac{(n+1)\cdot(n(2n+1) + 6(n+1))}{6} \\
&= \frac{(n+1)((n+1)+1)(2(n+1)+1)}{6}
\end{aligned}$$

(since an easy computation shows that $n(2n+1) + 6(n+1) = ((n+1)+1)(2(n+1)+1)$, as both sides equal $2n^2 + 7n + 6$). This shows that $P(n+1)$ holds, so we have proved the implication $P(n) \Longrightarrow P(n+1)$, and thus Goal 2 is achieved.

Having achieved both goals, we are done with the proof. □

Note that we were able to adapt the first proof of $1 + 2 + \cdots + n$ to $1^2 + 2^2 + \cdots + n^2$ with very little changes. However, the second proof could not be adapted at all. So induction is a rather flexible tool. On the downside, induction can help you prove a formula like $1 + 2 + \cdots + n = \frac{n(n+1)}{2}$, but it cannot help you discover such a formula. So if I ask you "is there an explicit formula for $1^3 + 2^3 + \cdots + n^3$", induction won't help you find it; it will only help you prove it once you've guessed it.

1.4. Notations for an induction proof

Here is some standard terminology that is commonly used in proofs by induction. Let's say you are proving a statement of the form $P(n)$ for every integer $n \geq b$ (where b is a fixed integer).

- The n is called the **induction variable**; you say that you **induct on** n. You don't have to call it n; it could just as well be called x and you are proving "$1 + 2 + \cdots + x = \frac{x(x+1)}{2}$ for every integer $x \geq 0$" by inducting on x.
- The proof of $P(b)$ (that is, Goal 1 in our above proofs) is called the **induction base** or the **base case**. In our above examples, this was always the proof of $P(0)$, but you can easily imagine b being something other than 0. (For example, if the goal is to prove "every integer $n \geq 4$ satisfies $2^n \geq n^2$", then $b = 4$, so that the base case is proving $2^4 \geq 4^2$.)

- The proof of "$P(n) \Longrightarrow P(n+1)$ for every $n \geq b$" (that is, Goal 2 in our above proofs) is called the **induction step**.

  In the induction step, the assumption that $P(n)$ holds is called the **induction hypothesis** (short **IH)** or the **induction assumption**, whereas the claim that $P(n+1)$ holds is called the **induction goal**. The induction step is complete when this goal is reached.

As an example, let us rewrite the above proof of the ToH formula $m_n = 2^n - 1$ using this language:

> **Theorem 1.4.1** (explicit answer to Tower of Hanoi)**.** For each integer $n \geq 0$, we let m_n be the # of moves needed to win the Tower of Hanoi game with n disks (or ∞ if the game cannot be won).
>
> Then,
> $$m_n = 2^n - 1 \qquad \text{for every integer } n \geq 0.$$

Proof. We induct on n.

Base case: The theorem holds for $n = 0$, since both m_0 and $2^0 - 1$ are 0.

Induction step: Let $n \geq 0$ be an integer. We assume that the theorem holds for n (this is what we called $P(n)$), and we set out to prove that the theorem holds for $n+1$ as well (this is what we called $P(n+1)$).

We have assumed that the theorem holds for n. In other words, $m_n = 2^n - 1$. In particular, m_n is an integer.

We need to show that the theorem holds for $n+1$, i.e., that $m_{n+1} \overset{?}{=} 2^{n+1} - 1$.

The last proposition we showed about the ToH game says that $m_n = 2m_{n-1} + 1$ when $n \geq 1$. Applying this proposition to $n+1$ instead of n, we conclude that

$$\begin{aligned} m_{n+1} &= 2 \underbrace{m_n}_{=2^n-1} + 1 = 2 \cdot (2^n - 1) + 1 \\ &= \underbrace{2 \cdot 2^n}_{=2^{n+1}} - 2 + 1 = 2^{n+1} - 2 + 1 = 2^{n+1} - 1. \end{aligned}$$

In other words, the theorem holds for $n+1$. Thus, the induction goal is reached, and the induction is complete. So the theorem is proved. □

1.5. The Fibonacci numbers

1.5.1. Definition

Our next applications of induction will be properties of the **Fibonacci sequence**. This is a sequence defined **recursively** – i.e., a given entry is not defined directly, but rather defined in terms of the previous entries. Specifically, it is defined as follows:

Definition 1.5.1. The **Fibonacci sequence** is the sequence $(f_0, f_1, f_2, \ldots)$ of nonnegative integers defined recursively by setting

$$f_0 = 0, \qquad f_1 = 1, \qquad \text{and}$$
$$f_n = f_{n-1} + f_{n-2} \qquad \text{for all } n \geq 2.$$

In other words, the Fibonacci sequence starts with the two entries 0 and 1, and then every next entry is the sum of the previous two entries.

The entries of this sequence are called the **Fibonacci numbers**. Here are some:

n	0	1	2	3	4	5	6	7	8	9	10	11	12	13
f_n	0	1	1	2	3	5	8	13	21	34	55	89	144	233

.

As we see, a recursive definition is a perfectly valid way of defining (e.g.) a sequence of numbers. It allows you to compute each entry of the sequence, as long as you do them in order. In a sense, the reason why this works is the same as why induction works: You can get to any integer $n \geq 0$ if you start with 0 and keep adding 1.

Note that a recursive definition can only depend on previous values, not on later values. Thus, for example, if we replaced $f_n = f_{n-1} + f_{n-2}$ by $f_n = f_{n+1} - f_{n-2}$, then we could not compute f_n, since (e.g.) computing f_2 would require computing f_3, which would require computing f_4, and so on.

1.5.2. The sum of the first n positive Fibonacci numbers

Let's prove our first property of Fibonacci numbers:

Theorem 1.5.2. For any integer $n \geq 0$, we have

$$f_1 + f_2 + \cdots + f_n = f_{n+2} - 1.$$

Proof. We induct on n.

Base case: For $n = 0$, the theorem claims that $f_1 + f_2 + \cdots + f_0 = f_{0+2} - 1$. Since $f_1 + f_2 + \cdots + f_0 = (\text{empty sum}) = 0$ and $f_{0+2} = f_2 = 1$, this boils down to $0 = 1 - 1$, which is true.

Induction step: Let $n \geq 0$ be an integer. Assume that the theorem holds for n. We must prove that the theorem holds for $n + 1$.

So we assumed that

$$f_1 + f_2 + \cdots + f_n = f_{n+2} - 1,$$

and we must prove that

$$f_1 + f_2 + \cdots + f_{n+1} \overset{?}{=} f_{(n+1)+2} - 1.$$

But

$$\begin{aligned} f_1 + f_2 + \cdots + f_{n+1} &= \underbrace{(f_1 + f_2 + \cdots + f_n)}_{\substack{=f_{n+2}-1 \\ \text{(by our IH = induction hypothesis)}}} + f_{n+1} \\ &= f_{n+2} - 1 + f_{n+1} = \underbrace{f_{n+2} + f_{n+1}}_{\substack{=f_{n+3} \\ \text{(by the recursive definition} \\ \text{of the Fibonacci sequence)}}} - 1 \\ &= f_{n+3} - 1 = f_{(n+1)+2} - 1. \end{aligned}$$

This is precisely what we wanted to prove – i.e., it says that the theorem holds for $n+1$. This completes the induction step. Thus the theorem is proved. □

1.6. Some more examples of induction

Theorem 1.6.1. For any integer $n \geq 0$, we have

$$2^0 + 2^1 + 2^2 + \cdots + 2^{n-1} = 2^n - 1.$$

Proof. We induct on n.

Base case: For $n = 0$, the theorem claims that $0 = 1 - 1$, which is true.

Induction step: Let $n \geq 0$ be an integer. Assume that the theorem holds for n, i.e., that we have

$$2^0 + 2^1 + 2^2 + \cdots + 2^{n-1} = 2^n - 1.$$

We must prove that the theorem also holds for $n+1$, i.e., that

$$2^0 + 2^1 + 2^2 + \cdots + 2^{(n+1)-1} \overset{?}{=} 2^{n+1} - 1.$$

However,

$$\begin{aligned} & 2^0 + 2^1 + 2^2 + \cdots + 2^{(n+1)-1} \\ &= 2^0 + 2^1 + 2^2 + \cdots + 2^n \\ &= \underbrace{\left(2^0 + 2^1 + 2^2 + \cdots + 2^{n-1}\right)}_{\substack{=2^n-1 \\ \text{(by the IH)}}} + 2^n \\ &= (2^n - 1) + 2^n = \underbrace{2 \cdot 2^n}_{=2^{n+1}} - 1 = 2^{n+1} - 1. \end{aligned}$$

This is precisely the goal, so the induction is complete. □

The theorem we just proved can be generalized:

Theorem 1.6.2. Let x and y be any two numbers. Then, for any integer $n \geq 0$, we have

$$(x-y)\left(x^{n-1}+x^{n-2}y+x^{n-3}y^2+\cdots+x^2y^{n-3}+xy^{n-2}+y^{n-1}\right)=x^n-y^n.$$

Here, the big sum in the parentheses is the sum of all products x^iy^j where i and j are nonnegative integers with $i+j=n-1$. (This is an empty sum when $n=0$.)

Before we prove this, let's give some examples:

- For $n=2$, the theorem says that
$$(x-y)(x+y)=x^2-y^2.$$
- For $n=3$, the theorem says that
$$(x-y)\left(x^2+xy+y^2\right)=x^3-y^3.$$
- For $n=4$, the theorem says that
$$(x-y)\left(x^3+x^2y+xy^2+y^3\right)=x^4-y^4.$$
- For $x=2$ and $y=1$, the theorem says that
$$(2-1)\left(2^{n-1}+2^{n-2}1+2^{n-3}1^2+\cdots+2^21^{n-3}+2\cdot 1^{n-2}+1^{n-1}\right)=2^n-1^n.$$
That is,
$$2^{n-1}+2^{n-2}+2^{n-3}+\cdots+2^2+2+1=2^n-1.$$
This is precisely the previous theorem. So our new theorem generalizes the previous theorem.

Let us now prove it:

Proof of the theorem. We induct on n.

Base case: For $n=0$, the claim

$$(x-y)\left(x^{n-1}+x^{n-2}y+x^{n-3}y^2+\cdots+x^2y^{n-3}+xy^{n-2}+y^{n-1}\right)=x^n-y^n$$

is true, since the LHS is 0 (since the second factor is an empty sum), while the RHS is $x^0-y^0=1-1=0$.

Induction step: Let $n\geq 0$ be an integer. Assume that the theorem holds for n. That is, assume that

$$(x-y)\left(x^{n-1}+x^{n-2}y+x^{n-3}y^2+\cdots+x^2y^{n-3}+xy^{n-2}+y^{n-1}\right)=x^n-y^n.$$

We must prove that the theorem also holds for $n+1$. In other words, we must prove that

$$(x-y)\left(x^n + x^{n-1}y + x^{n-2}y^2 + \cdots + x^3y^{n-3} + x^2y^{n-2} + xy^{n-1} + y^n\right) = x^{n+1} - y^{n+1}.$$

We begin by extracting the y^n addend from the long sum:

$$\begin{aligned}
&(x-y)\left(x^n + x^{n-1}y + x^{n-2}y^2 + \cdots + x^3y^{n-3} + x^2y^{n-2} + xy^{n-1} + y^n\right)\\
&= (x-y)\left(x^n + x^{n-1}y + x^{n-2}y^2 + \cdots + x^3y^{n-3} + x^2y^{n-2} + xy^{n-1}\right)\\
&\qquad + (x-y)\,y^n\\
&= (x-y)\,x\left(x^{n-1} + x^{n-2}y + x^{n-3}y^2 + \cdots + x^2y^{n-3} + xy^{n-2} + y^{n-1}\right)\\
&\qquad + (x-y)\,y^n\\
&= x\underbrace{(x-y)\left(x^{n-1} + x^{n-2}y + x^{n-3}y^2 + \cdots + x^2y^{n-3} + xy^{n-2} + y^{n-1}\right)}_{\substack{=x^n-y^n\\ \text{(by the IH)}}}\\
&\qquad + (x-y)\,y^n\\
&= x\left(x^n - y^n\right) + (x-y)\,y^n = x^{n+1} - xy^n + xy^n - y^{n+1} = x^{n+1} - y^{n+1}.
\end{aligned}$$

This is precisely the induction goal. So the induction is complete. □

Another useful particular case of this theorem is the case $y = 1$ (but x can be anything):

> **Corollary 1.6.3.** Let x be a number. Let $n \geq 0$ be an integer. Then,
>
> $$(x-1)\left(x^0 + x^1 + \cdots + x^{n-1}\right) = x^n - 1.$$
>
> Thus, if $x \neq 1$, we can divide this equality by $x-1$, and obtain
>
> $$x^0 + x^1 + \cdots + x^{n-1} = \frac{x^n - 1}{x - 1}.$$
>
> This is the classical formula for geometric sums.

1.7. How not to use induction

Induction proofs can be slippery:

Theorem 1.7.1 (Fake theorem)**.** In any set of $n \geq 1$ horses, all the horses are the same color.

Proof. We induct on n.

Base case: This is clearly true for $n = 1$, since a single horse always has the same color as itself.

Induction step: Let $n \geq 1$ be an integer. We assume that the theorem holds for n, i.e., that any n horses are the same color.

We must prove that it also holds for $n+1$, i.e., that any $n+1$ horses are the same color.

So let $H_1, H_2, \ldots, H_{n+1}$ be $n+1$ horses.

By our IH, the first n horses $H_1, H_2, \ldots, H_n$ are the same color.

Again by our IH, the last n horses $H_2, H_3, \ldots, H_{n+1}$ are the same color.

Now consider the first horse H_1 and the last horse H_{n+1}. They both have the same color as the middle horses $H_2, H_3, \ldots, H_n$ (by the previous two paragraphs). Thus, all the $n+1$ horses $H_1, H_2, \ldots, H_{n+1}$ have the same color, right? □

When a claim is as obviously wrong as this one, there is an easy way to find the mistake in the proof: You just look at some example where you know the claim is wrong, and you try to let the proof convince you that it is true, and you look out for wrong conclusions. The first wrong conclusion you see is where the error lies.

The fake theorem is already false for $n = 2$, so the induction step (for $n = 1$) must be at fault. In this induction step, we claim that H_1 and $H_{n+1} = H_2$ have the same color as the "middle horses" $H_2, H_3, \ldots, H_1$. But there are no middle horses! So we are comparing H_1 and H_2 to something that does not exist, and so the argument does not work.

Note that a single broken link in the induction chain ($P(1)$ did not imply $P(2)$) brought down the entire chain!

1.8. More on the Fibonacci numbers

Recall the Fibonacci numbers:

n	0	1	2	3	4	5	6	7	8	9	10	11	12	13
f_n	0	1	1	2	3	5	8	13	21	34	55	89	144	233

.

They are defined by $f_0 = 0$ and $f_1 = 1$ and $f_n = f_{n-1} + f_{n-2}$ for all $n \geq 0$. Let us now prove some more properties of these numbers.

1.8.1. The addition theorem

> **Theorem 1.8.1** (addition theorem for the Fibonacci numbers)**.** We have
>
> $$f_{n+m+1} = f_n f_m + f_{n+1} f_{m+1} \qquad \text{for all } n, m \geq 0.$$

Proof. We induct on n. To this purpose, for every integer $n \geq 0$, we define the statement $P(n)$ to say

"for all integers $m \geq 0$, we have $f_{n+m+1} = f_n f_m + f_{n+1} f_{m+1}$".

(Don't forget the "for all integers $m \geq 0$" part! The statement $P(n)$ is not a single equality $f_{n+m+1} = f_n f_m + f_{n+1} f_{m+1}$, but rather combines infinitely many such equalities, one for each m. If we fixed a value of m and defined $P(n)$ to only say $f_{n+m+1} = f_n f_m + f_{n+1} f_{m+1}$ for this specific m, then our proof below would not work.)

We shall now prove the statement $P(n)$ for all $n \geq 0$ by induction on n.

Base case: We must prove $P(0)$. In other words, we must prove that

"for all integers $m \geq 0$, we have $f_{0+m+1} = f_0 f_m + f_{0+1} f_{m+1}$".

This is easy: For all integers $m \geq 0$, we have $\underbrace{f_0}_{=0} f_m + \underbrace{f_{0+1}}_{=f_1=1} f_{m+1} = f_{m+1} = f_{0+m+1}$.

Induction step: Let $n \geq 0$ be an integer. We assume that $P(n)$ holds. We must show that $P(n+1)$ holds.

Our IH says that $P(n)$ holds, i.e., that

"for all integers $m \geq 0$, we have $f_{n+m+1} = f_n f_m + f_{n+1} f_{m+1}$".

Our goal is to show that $P(n+1)$ holds, i.e., that

"for all integers $m \geq 0$, we have $f_{(n+1)+m+1} = f_{n+1} f_m + f_{(n+1)+1} f_{m+1}$".

To prove this, we let $m \geq 0$ be an integer. Then,

$$\begin{aligned} f_{n+1} f_m + f_{(n+1)+1} f_{m+1} &= f_{n+1} f_m + \underbrace{f_{n+2}}_{=f_{n+1}+f_n} f_{m+1} \\ &= f_{n+1} f_m + (f_{n+1} + f_n) f_{m+1} \\ &= f_{n+1} f_m + f_{n+1} f_{m+1} + f_n f_{m+1} \\ &= f_{n+1} \underbrace{(f_m + f_{m+1})}_{=f_{m+2}} + f_n f_{m+1} \\ &= f_{n+1} f_{m+2} + f_n f_{m+1} = f_n f_{m+1} + f_{n+1} f_{m+2}. \end{aligned}$$

Now, recall that the IH says that

"for all integers $m \geq 0$, we have $f_{n+m+1} = f_n f_m + f_{n+1} f_{m+1}$".

The m in this statement is a bound variable, i.e., it has nothing to do with the m that we have fixed. Thus, we are free to apply our IH not to the current m, but to any other m as well. In particular, we can apply it to $m+1$ instead of m. Thus, we get

$$f_{n+(m+1)+1} = f_n f_{m+1} + f_{n+1} f_{(m+1)+1}$$

In other words,

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

Comparing this with the previous computation $f_{n+1} f_m + f_{(n+1)+1} f_{m+1} = f_n f_{m+1} + f_{n+1} f_{m+2}$, we obtain

$$f_{n+1} f_m + f_{(n+1)+1} f_{m+1} = f_{n+m+2} = f_{(n+1)+m+1}.$$

So we have shown that $f_{(n+1)+m+1} = f_{n+1} f_m + f_{(n+1)+1} f_{m+1}$ for all integers $m \geq 0$. This is precisely our induction goal $P(n+1)$. So the induction is complete. □

More exercises along these lines on HW#1.

1.8.2. Divisibility of Fibonacci numbers

Our next theorem involves divisibility of integers. We will study this in more detail in a later chapter (elementary number theory), but for now let us recall the definition:

Definition 1.8.2. Let a and b be two integers. We say that a **divides** b (and we write $a \mid b$) if there exists an integer c such that $b = ac$. Equivalently, we can say that b **is divisible by** a in this case.

For example, $2 \mid 4$ and $3 \mid 12$ and $10 \mid 30$ and $0 \mid 0$ and $5 \mid 0$, but $2 \nmid 3$ and $0 \nmid 1$.

Now we can state a divisibility property of Fibonacci numbers:

Theorem 1.8.3. If $a, b \geq 0$ are two integers that satisfy $a \mid b$, then $f_a \mid f_b$.

Proof. It is reasonable to try induction. However, inducting on a does not lead us anywhere: Going from a to $a+1$, we find no way to use the IH, since the condition $a+1 \mid b$ has nothing to do with $a \mid$anything.

Inducting on b does not work either: $a \mid b$ does not imply $a \mid b+1$ or vice versa; in fact, quite the opposite.

What can we do? Give up on induction?

One thing we haven't tried so far is to introduce a new variable and then induct on that new variable. Let us do this.

We observe that two integers $a, b \geq 0$ satisfy $a \mid b$ if and only if there exists an integer $c \geq 0$ such that $b = ac$. So we can restate our theorem as follows:

Restated theorem: "For any integers $a, c \geq 0$, we have $f_a \mid f_{ac}$".

Now, we shall prove this restated theorem by induction on c. In other words, for each $c \geq 0$, we shall prove the statement

$$P(c) := (\text{"for any integer } a \geq 0 \text{, we have } f_a \mid f_{ac}\text{"}).$$

Base case: We must prove $P(0)$. In other words, we must prove that

$$\text{"for any integer } a \geq 0 \text{, we have } f_a \mid f_{a \cdot 0}\text{"}.$$

But this is easy, since $f_{a \cdot 0} = f_0 = 0$ is divisible by everything.

Induction step: Let $c \geq 0$ be an integer. We assume that $P(c)$ holds, i.e., that

$$\text{for any integer } a \geq 0 \text{, we have } f_a \mid f_{ac}.$$

We must prove that $P(c+1)$ holds, i.e., that

$$\text{for any integer } a \geq 0 \text{, we have } f_a \mid f_{a(c+1)}.$$

Let $a \geq 0$ be any integer. Then, the IH yields $f_a \mid f_{ac}$. In other words, $f_{ac} = f_a p$ for some integer p. Now,

$$\begin{aligned} f_{a(c+1)} &= f_{ac+a} = f_{ac+(a-1)+1} \\ &= \underbrace{f_{ac}}_{=f_a p} f_{a-1} + f_{ac+1} \underbrace{f_{(a-1)+1}}_{=f_a} \qquad \left(\begin{array}{c}\text{by the addition theorem} \\ f_{n+m+1} = f_n f_m + f_{n+1} f_{m+1}\end{array}\right) \\ &= f_a p f_{a-1} + f_{ac+1} f_a = f_a \underbrace{\left(p f_{a-1} + f_{ac+1}\right)}_{\text{an integer}}. \end{aligned}$$

This shows that $f_a \mid f_{a(c+1)}$, thus completing our induction step (i.e., proving $P(c+1)$). This completes the proof, right?

Yeah, up to one little gap: We used the addition theorem $f_{n+m+1} = f_n f_m + f_{n+1} f_{m+1}$ for $n = ac$ and $m = a-1$. For this to work, we need ac and $a-1$ to be ≥ 0, since the addition theorem requires $n, m \geq 0$. Of course, $ac \geq 0$, but $a-1$ might fail to be ≥ 0. The latter happens exactly when $a = 0$. So our argument does not work for $a = 0$.

The easiest way to fix this is to just manually prove our theorem for $a = 0$. This is very easy: If $a = 0$, then both a and ac are 0, so that $f_a \mid f_{ac}$ boils down to $f_0 \mid f_0$, which is obvious.

Let us rewrite our induction step to make this correction:

Induction step: Let $c \geq 0$ be an integer. We assume that $P(c)$ holds, i.e., that

$$\text{for any integer } a \geq 0 \text{, we have } f_a \mid f_{ac}.$$

We must prove that $P(c+1)$ holds, i.e., that

$$\text{for any integer } a \geq 0 \text{, we have } f_a \mid f_{a(c+1)}.$$

Let $a \geq 0$ be any integer. Then, the IH yields $f_a \mid f_{ac}$. In other words, $f_{ac} = f_a p$ for some integer p. Now, we are in one of the following two cases:

Case 1: We have $a = 0$.

Case 2: We have $a \neq 0$.

In Case 1, we have $a = 0$, and thus our goal (to prove $f_a \mid f_{a(c+1)}$) follows immediately from both a and $a(c+1)$ being 0.

Now consider Case 2. In this case, $a \neq 0$, so that $a \geq 1$ and thus $a - 1 \geq 0$. Now,

$$\begin{aligned} f_{a(c+1)} &= f_{ac+a} = f_{ac+(a-1)+1} \\ &= \underbrace{f_{ac}}_{=f_a p} f_{a-1} + f_{ac+1} \underbrace{f_{(a-1)+1}}_{=f_a} && \left(\begin{array}{c}\text{by the addition theorem} \\ f_{n+m+1} = f_n f_m + f_{n+1} f_{m+1} \\ \text{for } n = ac \text{ and } m = a-1\end{array}\right) \\ &= f_a p f_{a-1} + f_{ac+1} f_a = f_a \underbrace{(p f_{a-1} + f_{ac+1})}_{\text{an integer}}. \end{aligned}$$

This shows that $f_a \mid f_{a(c+1)}$ in Case 2 as well.

Thus, $f_a \mid f_{a(c+1)}$ is proved in both Cases 1 and 2, hence always holds. This completes our induction step (i.e., proving $P(c+1)$) and thus the whole proof. □

1.8.3. Binet's formula

Is there an explicit formula for f_n, that is, a formula that does not rely on the previous Fibonacci numbers?

Theorem 1.8.4 (Binet's formula)**.** Let

$$\varphi = \frac{1+\sqrt{5}}{2} \approx 1.618\ldots \qquad \text{and} \qquad \psi = \frac{1-\sqrt{5}}{2} \approx -0.618\ldots.$$

Then,

$$f_n = \frac{\varphi^n - \psi^n}{\sqrt{5}} \qquad \text{for every integer } n \geq 0.$$

Some remarks:

- The number φ is the so-called **golden ratio**, and is famous for many properties, including the fact that $\varphi^2 = \varphi + 1$, or equivalently $\varphi = 1 + \dfrac{1}{\varphi}$. The number ψ is its so-called conjugate and also satisfies $\psi^2 = \psi + 1$.

- The numbers f_n are integers, but Binet's formula expresses them in terms of the irrationals φ and ψ.
- As n grows large, ψ^n approaches 0, whereas φ^n grows exponentially. Thus, f_n also grows exponentially, with growth rate $\varphi \approx 1.618\ldots$.

Two questions:

1. How do we prove Binet's formula?
2. How could we find Binet's formula if we didn't already know it?

I will only answer question 1 in this course. For question 2, see an advanced linear algebra class (eigenvalues and diagonalization) or a class on algebraic combinatorics (generating functions).

First, let us try to prove Binet's formula by induction on n:

Attempted proof of Binet's formula. We induct on n:

Base case: For $n = 0$, we have $f_n = f_0 = 0$ and $\dfrac{\varphi^0 - \psi^0}{\sqrt{5}} = \dfrac{1-1}{\sqrt{5}} = 0$. So Binet's formula holds for $n = 0$.

Induction step: Let $n \geq 0$ be an integer.

Assume (as the IH) that Binet's formula holds for n; that is, assume that $f_n = \dfrac{\varphi^n - \psi^n}{\sqrt{5}}$.

We must prove that Binet's formula holds for $n+1$. In other words, we must prove that $f_{n+1} = \dfrac{\varphi^{n+1} - \psi^{n+1}}{\sqrt{5}}$.

We have

$$\begin{aligned} f_{n+1} &= f_n + f_{n-1} && \left(\begin{array}{c}\text{assuming that } n \geq 1; \\ \text{the } n = 0 \text{ case can be done by hand}\end{array}\right) \\ &= \frac{\varphi^n - \psi^n}{\sqrt{5}} + f_{n-1} && (\text{by the IH}). \end{aligned}$$

What now? The IH says nothing about f_{n-1}. We are stuck. □

Let us see how to fix this by introducing a more advanced version of induction.

1.9. Strong induction

1.9.1. Reminder on regular induction

Recall the (original) principle of induction:

Theorem 1.9.1 (Principle of Mathematical Induction, short PoMI)**.** Let b be an integer.

Let $P(n)$ be a mathematical statement defined for each integer $n \geq b$.

(For example, $P(n)$ can be "$n+1>n$" or "n is even" or "n is prime" or "there exists a prime number larger than n".)

Assume the following:

1. **Base case:** The statement $P(b)$ holds (i.e., the statement $P(n)$ holds for $n=b$).

2. **Induction step:** For each integer $n \geq b$, the implication $P(n) \Longrightarrow P(n+1)$ holds (i.e., if $P(n)$ holds, then $P(n+1)$ holds as well).

Then, the statement $P(n)$ holds for every integer $n \geq b$.

We can restate this principle slightly by renaming the n in the induction step as $n-1$ (so that instead of $P(n) \Longrightarrow P(n+1)$ we are now proving $P(n-1) \Longrightarrow P(n)$):

Theorem 1.9.2 (Principle of Mathematical Induction, short PoMI)**.** Let b be an integer.

Let $P(n)$ be a mathematical statement defined for each integer $n \geq b$.

(For example, $P(n)$ can be "$n+1>n$" or "n is even" or "n is prime" or "there exists a prime number larger than n".)

Assume the following:

1. **Base case:** The statement $P(b)$ holds (i.e., the statement $P(n)$ holds for $n=b$).

2. **Induction step:** For each integer $n > b$, the implication $P(n-1) \Longrightarrow P(n)$ holds (i.e., if $P(n-1)$ holds, then $P(n)$ holds as well).

Then, the statement $P(n)$ holds for every integer $n \geq b$.

The idea behind this principle (in either form) is that the base case gives us $P(b)$ whereas the induction step gives us the implications

$$\begin{aligned} P(b) &\Longrightarrow P(b+1), \\ P(b+1) &\Longrightarrow P(b+2), \\ P(b+2) &\Longrightarrow P(b+3), \\ &\ldots. \end{aligned}$$

In the domino metaphor, the base case tips over the first domino, and the induction step ensures that each domino falls from the impact of the previous domino's falling.

1.9.2. Strong induction

Now assume that the $b+2$-domino (i.e., $P(b+2)$) falls not from the impact of the previous domino $P(b+1)$, but rather from the combined force of $P(b)$ and $P(b+1)$. In other words, instead of the implication $P(b+1) \Longrightarrow P(b+2)$, imagine you have the implication

$$(P(b) \text{ AND } P(b+1)) \Longrightarrow P(b+2).$$

This would still suffice, because $P(b)$ and $P(b+1)$ have already fallen. Likewise, we could just as well replace the implication $P(b+2) \Longrightarrow P(b+3)$ by the weaker implication

$$(P(b) \text{ AND } P(b+1) \text{ AND } P(b+2)) \Longrightarrow P(b+3).$$

More generally, for each $n > b$, instead of proving the implication $P(n-1) \Longrightarrow P(n)$, it will suffice to prove the weaker implication

$$\underbrace{(P(b) \text{ AND } P(b+1) \text{ AND } P(b+2) \text{ AND } \cdots \text{ AND } P(n-1))}_{\text{i.e., the statement } P(k) \text{ holds for each } k \in \{b, b+1, \ldots, n-1\}} \Longrightarrow P(n)$$

(so that the domino $P(n)$ is tipped over by the combined force of all the preceding dominos, not just its left neighbor).

This induction principle is called **strong induction**. Explicitly, it says the following:

> **Theorem 1.9.3** (Principle of Strong Induction)**.** Let b be an integer.
> Let $P(n)$ be a mathematical statement defined for each integer $n \geq b$.
> Assume the following:
>
> 1. **Base case:** The statement $P(b)$ holds.
> 2. **Induction step:** For each integer $n > b$, the implication
>
> $$(P(b) \text{ AND } P(b+1) \text{ AND } P(b+2) \text{ AND } \cdots \text{ AND } P(n-1)) \Longrightarrow P(n)$$
>
> holds.
>
> Then, the statement $P(n)$ holds for every integer $n \geq b$.

Proofs using this principle are called **proofs by strong induction**. They differ from regular induction proofs in that in their induction steps, you can use not only $P(n-1)$ but also $P(n-2)$ and $P(n-3)$ and so on all the way down to

$P(b)$. Thus, you are proving

$$\begin{aligned} &P(b), \\ P(b) &\Longrightarrow P(b+1), \\ (P(b) \text{ AND } P(b+1)) &\Longrightarrow P(b+2), \\ (P(b) \text{ AND } P(b+1) \text{ AND } P(b+2)) &\Longrightarrow P(b+3), \\ &\ldots, \end{aligned}$$

and concluding that all $P(n)$'s hold.

1.9.3. Example: Proof of Binet's formula

Let us now prove Binet's formula by strong induction.

Proof of Binet's formula. We strongly induct on n (i.e., we use strong induction):

Base case: For $n = 0$, we have $f_n = f_0 = 0$ and $\dfrac{\varphi^0 - \psi^0}{\sqrt{5}} = \dfrac{1-1}{\sqrt{5}} = 0$. So Binet's formula holds for $n = 0$.

Induction step: Let $n > 0$ be an integer.

Assume (as the IH) that Binet's formula holds for 0, for 1, for 2, ..., for $n-1$. In other words, assume that $f_k = \dfrac{\varphi^k - \psi^k}{\sqrt{5}}$ for each $k \in \{0, 1, \ldots, n-1\}$.

We must prove that Binet's formula holds for n. In other words, we must prove that $f_n = \dfrac{\varphi^n - \psi^n}{\sqrt{5}}$.

We have

$$\begin{aligned}
f_n &= f_{n-1} + f_{n-2} \\
&= \frac{\varphi^{n-1} - \psi^{n-1}}{\sqrt{5}} + \frac{\varphi^{n-2} - \psi^{n-2}}{\sqrt{5}} \\
&\qquad \left(\begin{array}{c} \text{because the IH tells us that Binet's formula} \\ \text{holds for } n-1 \text{ and for } n-2 \end{array} \right) \\
&= \frac{1}{\sqrt{5}} \left(\varphi^{n-1} - \psi^{n-1} + \varphi^{n-2} - \psi^{n-2} \right) \\
&= \frac{1}{\sqrt{5}} \left(\underbrace{\left(\varphi^{n-1} + \varphi^{n-2} \right)}_{\substack{= \varphi^{n-2}\varphi + \varphi^{n-2} \\ = \varphi^{n-2}(\varphi+1) \\ = \varphi^{n-2}\varphi^2 \\ = \varphi^n}} - \underbrace{\left(\psi^{n-1} + \psi^{n-2} \right)}_{\substack{= \psi^{n-2}(\psi+1) \\ = \psi^{n-2}\psi^2 \\ = \psi^n}} \right) \\
&= \frac{1}{\sqrt{5}} \left(\varphi^n - \psi^n \right) = \frac{\varphi^n - \psi^n}{\sqrt{5}}.
\end{aligned}$$

In other words, Binet's formula holds for n. Induction complete, right?

Almost. There is a gap here very similar to the gap we had in our proof of $f_a \mid f_b$: We applied the IH to $n-1$ and $n-2$ without checking that $n-1$ and $n-2$ are ≥ 0 (which is a condition for Binet's formula).

While $n > 0$ ensures that $n - 1 \geq 0$, there is no guarantee that $n - 2 \geq 0$ unless we rule out the case $n = 1$. So we need an extra argument for the case $n = 1$.

In this case, we can verify Binet's formula simply by comparing both sides: $f_1 = 1$ and

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

so $f_1 = \dfrac{\varphi^1 - \psi^1}{\sqrt{5}}$, and thus Binet's formula holds for $n = 1$. □

Let us summarize: We have used strong induction on n to prove Binet's formula. We needed it because the stronger hypothesis in a strong induction gives us the "memory" to include the $(n-2)$-case when proving the n-case.

Note that we had to handle the cases $n = 0$ and $n = 1$ by hand. The $n = 0$ case was the base case. The $n = 1$ case was an extra case inside the induction step. We can think of this $n = 1$ case as an "extra base case", even though formally it was part of the induction step.

1.9.4. Baseless strong induction

You can actually reformulate the principle of strong induction in a form that does not have a de-jure base case at all:

> **Theorem 1.9.4** (Principle of Strong Induction, baseless version)**.** Let b be an integer.
>
> Let $P(n)$ be a mathematical statement defined for each integer $n \geq b$.
>
> Assume the following:
>
> 1. **Induction step:** For each integer $n \geq b$, the implication
>
> $$(P(b) \text{ AND } P(b+1) \text{ AND } P(b+2) \text{ AND } \cdots \text{ AND } P(n-1)) \Longrightarrow P(n)$$
>
> holds.
>
> Then, the statement $P(n)$ holds for every integer $n \geq b$.

Why does this work? Why were we able to get rid of the base case? Because the induction step now allows n to be b, in which case the implication

$$(P(b) \text{ AND } P(b+1) \text{ AND } P(b+2) \text{ AND } \cdots \text{ AND } P(n-1)) \Longrightarrow P(n)$$

is simply saying that

$$(\text{true}) \Longrightarrow P(b)$$

(since a conjunction of 0 statements is just “true”), and this is the same as just saying $P(b)$ unconditionally. So we did not really lose the base case; we just packed it inside the induction step.

In practice, this is not all that useful, because even with the $n = b$ case inside the induction step, you will often have to prove it by a separate argument, so it becomes a “de-facto base case” just like $n = 1$ was in our proof of Binet’s formula.

1.9.5. Example: Prime factorization exist

Another example of a strong induction proof comes from elementary number theory. We recall a basic definition:

> **Definition 1.9.5.** A **prime** (or **prime number**) means an integer $p > 1$ whose only positive divisors are 1 and p.

So the primes (in increasing order) are

$$2, \ 3, \ 5, \ 7, \ 11, \ 13, \ 17, \ 19, \ 23, \ 29, \ \ldots.$$

There are infinitely many primes, as you will get to prove on a HW.

Now we claim the following:

Theorem 1.9.6. Every positive integer is a product of finitely many primes.

This includes 1, because an empty product is defined to be 1. Some more interesting examples:

$$\begin{aligned} 2023 &= 7 \cdot 17 \cdot 17; \\ 2024 &= 2 \cdot 2 \cdot 2 \cdot 11 \cdot 23; \\ 2025 &= 5 \cdot 5 \cdot 3 \cdot 3 \cdot 3 \cdot 3. \end{aligned}$$

Also, $2 = 2$ is a product of a single prime (2 itself).

How do we prove the theorem in general?

Proof. We must prove the statement

$$P(n) = (\text{“}n \text{ is a product of finitely many primes”})$$

for each integer $n \geq 1$.

We shall prove this by strong induction on n (original version, not the baseless one).

Base case: $P(1)$ is true, as we already saw (1 is an empty product).

Induction step: Let $n > 1$. We must prove the implication

$$(P(1) \text{ AND } P(2) \text{ AND } P(3) \text{ AND } \cdots \text{ AND } P(n-1)) \Longrightarrow P(n).$$

So we assume that $P(1)$ AND $P(2)$ AND $P(3)$ AND $\cdots$ AND $P(n-1)$ holds. We must prove that $P(n)$ holds. In other words, we must prove that n is a product of finitely many primes.

We are in one of the following two cases:

Case 1: The only positive divisors of n are 1 and n.

Case 2: There is a positive divisor d of n that is neither 1 nor n.

Consider Case 1 first. In this case, n is itself a prime (by the definition of a prime), and thus the product of finitely many primes (just itself). So $P(n)$ holds in Case 1.

Now consider Case 2. In this case, there is a positive divisor d of n that is neither 1 nor n. Consider this d. Then, $1 < d < n$ (why? $1 < d$ is easy; $d < n$ is because every positive divisor of n is at most n; we will actually prove this in more detail later). Moreover, $\frac{n}{d}$ is an integer (since d is a divisor of n), and also satisfies $1 < \frac{n}{d} < n$ (since $1 < d < n$).

So both d and $\frac{n}{d}$ are integers among $1, 2, \ldots, n-1$ (since they are > 1 and $< n$). As a consequence, $P(d)$ and $P\left(\frac{n}{d}\right)$ hold by the IH. In other words, both d and $\frac{n}{d}$ are products of primes:

$$d = p_1 p_2 \cdots p_k \qquad \text{and} \qquad \frac{n}{d} = q_1 q_2 \cdots q_\ell$$

for some primes $p_1, p_2, \ldots, p_k$ and $q_1, q_2, \ldots, q_\ell$. Now, multiplying these equalities together, we find

$$\begin{aligned} n &= d \cdot \frac{n}{d} = (p_1 p_2 \cdots p_k) \cdot (q_1 q_2 \cdots q_\ell) \\ &= p_1 p_2 \cdots p_k q_1 q_2 \cdots q_\ell. \end{aligned}$$

Thus, n is a product of finitely many primes. This proves $P(n)$ in Case 2.

So now $P(n)$ is proved in both cases, and we are done. □

The above proof contains the elementary recursive algorithm for finding prime factorizations (i.e., for factoring a positive integer n into a product of primes): Look for positive divisors of n other than 1 and n. If no such exist, then n itself is prime. If you find such a divisor d, then solve the problem for d and $\frac{n}{d}$ and multiply.

1.9.6. Example: Paying with 3-cent and 5-cent coins

> **Exercise 1.9.1.** Assume that you have an infinite supply of 3-cent coins and of 5-cent coins. What denominations can you pay with it?

Here is a table of small denominations:

0 cents	yes	
1 cent	no	
2 cents	no	
3 cents	yes	
4	no	
5	yes	
6	yes: $2 \cdot 3$	
7	no	.
8	yes: $3 + 5$	
9	yes: $3 \cdot 3$	
10	yes: $2 \cdot 5$	
11	yes: $2 \cdot 3 + 5$	
12	yes: $4 \cdot 3$	
13	yes: $3 + 2 \cdot 5$	
14	yes: $5 + 3 \cdot 3$	

There is no visible regular pattern in the first few answers, but it appears that any denomination ≥ 8 cents can be paid. Why?

We can notice that if a denomination k (that is, k cents) can be paid, then so can $k + 3$ (just add a 3-cent coin). Thus, the "yes" for $k = 8$ implies a "yes" for $11, 14, 17, \ldots$. Likewise, the "yes" for $k = 3$ implies a "yes" for $6, 9, 12, 15, \ldots$, and the "yes" for $k = 10$ implies a "yes" for $13, 16, 19, \ldots$. Together, these three sequences ("arithmetic progressions") cover all the integers ≥ 8. So our guess was right: We can pay every denomination ≥ 8.

Let us formalize this argument as an induction proof.

We define $\mathbb{N}$ to be the set of all nonnegative integers:

$$\mathbb{N} = \{0, 1, 2, \ldots\}.$$

Proposition 1.9.7. For any integer $n \geq 8$, we can pay n cents with 3-cent and 5-cent coins. In other words, any integer $n \geq 8$ can be written as $n = 3a + 5b$ with $a, b \in \mathbb{N}$.

Proof. We use strong induction on n:

Base case: For $n = 8$, the claim is true, since $8 = 3 \cdot 1 + 5 \cdot 1$.

Induction step: Fix an integer $n > 8$. Assume that the proposition is already proved for all integers $8, 9, \ldots, n-1$. We must prove that it also holds for n.

We are in one of the following three cases (since $n > 8$):

Case 1: We have $n = 9$.

Case 2: We have $n = 10$.

Case 3: We have $n \geq 11$.

In Case 1, we are done, since $n = 9 = 3 \cdot 3 + 5 \cdot 0$.

In Case 2, we are done, since $n = 10 = 3 \cdot 0 + 5 \cdot 2$.

Consider Case 3. In this case, $n \geq 11$, so that $n - 3 \geq 8$, and thus $n - 3$ is one of the integers $8, 9, \ldots, n-1$. Thus, we can apply the induction hypothesis to $n - 3$. We conclude that $n - 3 = 3c + 5d$ for some $c, d \in \mathbb{N}$. Hence,

$$n = 3 + 3c + 5d = 3(c+1) + 5d.$$

Thus, $n = 3a + 5b$ for $a = c + 1$ and $b = d$. This proves the claim for our n, so the induction is complete. □

Note that the above proof had one "de-jure base case" (the case $n = 8$) and two "de-facto base cases" ($n = 9$ and $n = 10$). We could just as well have used the baseless form of strong induction, which would have made all three into "de-facto base cases". This would be a bit more uniform.

2. Sums and products

In this chapter, we will study finite sums $a_1 + a_2 + \cdots + a_n$ and finite products $a_1 a_2 \cdots a_n$. We will introduce a shorter (and also clearer) notation for them, and we will study their properties and many examples. Note that you have seen some examples of finite sums already:

$$\begin{aligned}
1+2+\cdots+n &= \frac{n\left(n+1\right)}{2};\\
1^2+2^2+\cdots+n^2 &= \frac{n\left(n+1\right)\left(2n+1\right)}{6};\\
f_1+f_2+\cdots+f_n &= f_{n+2}-1;\\
q^0+q^1+\cdots+q^{n-1} &= \frac{q^n-1}{q-1} \qquad \left(\text{for } q\neq 1\right);\\
x^{n-1}+x^{n-2}y+x^{n-3}y^2+\cdots+x^2y^{n-3}+xy^{n-2}+y^{n-1} &= \frac{x^n-y^n}{x-y} \qquad \left(\text{for } x\neq y\right).
\end{aligned}$$

The last of these equalities shows that this "naive" way of writing finite sums eventually gets tiresome. When sums get more complicated, it gets harder and harder to understand what the "$\cdots$" means; it is an exercise in continuing a pattern, and can be ambiguous and annoying. So we need a better notation.

This notation comes in form of the "sigma notation" $\sum$ for sums and the "pi notation" $\prod$ for products. For instance, the sum

$$x^{n-1}+x^{n-2}y+x^{n-3}y^2+\cdots+x^2y^{n-3}+xy^{n-2}+y^{n-1}$$

becomes

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

In a programming language (fake Python), this would be

$$\texttt{sum(}x^{n-1-k}y^k \texttt{ for } k\in\left\{0,1,\ldots,n-1\right\}\texttt{)}$$

or

$$\texttt{sum(}x^{n-1-k}y^k \texttt{ for } k \texttt{ from } 0 \texttt{ to } n-1\texttt{).}$$

2.1. Finite sums

Definition 2.1.1. Let u and v be two integers. Let $a_u, a_{u+1}, \ldots, a_v$ be some numbers. Then,

$$\sum_{k=u}^{v} a_k$$

is defined to be the sum

$$a_u + a_{u+1} + \cdots + a_v$$

(in more detail: $a_u + a_{u+1} + a_{u+2} + \cdots + a_{v-1} + a_v$). It is called the **sum of the numbers a_k where k ranges from u to v**. When $v < u$, this sum is called **empty** and defined to be 0. When $v = u$, this sum contains only one addend, which is a_u, and thus equals a_u.

For example:

$$\begin{aligned}
\sum_{k=5}^{10} k &= 5+6+7+8+9+10 = 45;\\
\sum_{k=5}^{10} k^k &= 5^5+6^6+7^7+8^8+9^9+10^{10};\\
\sum_{k=5}^{10} k^{k+1} &= 5^6+6^7+7^8+8^9+9^{10}+10^{11};\\
\sum_{k=5}^{5} k &= 5;\\
\sum_{k=5}^{4} k &= 0 \qquad \text{(an empty sum)};\\
\sum_{k=5}^{3} k &= 0 \qquad \text{(an empty sum)};\\
\sum_{k=0}^{n-1} q^k &= q^0+q^1+q^2+\cdots+q^{n-1} = \frac{q^n-1}{q-1} \qquad \text{when } q \neq 1\\
&\qquad (\text{but } = n \text{ when } q=1);
\end{aligned}$$

$$\begin{aligned}
\sum_{k=0}^{n-1} x^k y^{n-1-k} &= x^0y^{n-1}+x^1y^{n-2}+x^2y^{n-3}+\cdots+x^{n-1}y^0\\
&= y^{n-1}+xy^{n-2}+x^2y^{n-3}+\cdots+x^{n-1}\\
&= x^{n-1}+x^{n-2}y+x^{n-3}y^2+\cdots+y^{n-1}.
\end{aligned}$$

So one of the theorems we proved above says that

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

for any numbers x and y and any $n \in \mathbb{N}$.

The variable k is not set in stone; you can replace it by any other variable that is not otherwise occupied. For instance,

$$\sum_{k=u}^{v} a_k = \sum_{i=u}^{v} a_i = \sum_{\ddot{a}=u}^{v} a_{\ddot{a}} = \sum_{\spadesuit=u}^{v} a_{\spadesuit}.$$

(Just don't write $\sum_{u=u}^{v} a_u$.)

Here are some more examples: For any $n \in \mathbb{N}$, we have

$$\sum_{k=1}^{n} k = 1 + 2 + \cdots + n = \frac{n(n+1)}{2} \qquad \text{(by Little Gauss)};$$
$$\sum_{k=1}^{n} k^2 = 1^2 + 2^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6};$$
$$\sum_{k=1}^{n} 1 = \underbrace{1 + 1 + \cdots + 1}_{n \text{ times}} = n \cdot 1 = n;$$
$$\begin{aligned} \sum_{k=1}^{n} (2k-1) &= (2 \cdot 1 - 1) + (2 \cdot 2 - 1) + (2 \cdot 3 - 1) + \cdots + (2n-1) \\ &= 1 + 3 + 5 + \cdots + (2n-1) \\ &= (\text{the sum of the first } n \text{ odd positive integers}). \end{aligned}$$

We have not computed this last sum, but let's do this now. I will use two "laws of summation":

- We have
$$\sum_{k=u}^{v} (a_k - b_k) = \sum_{k=u}^{v} a_k - \sum_{k=u}^{v} b_k$$
for any integers u, v and any numbers a_k, b_k. Indeed, if we rewrite this without summation signs, it takes the form
$$\begin{aligned} &(a_u - b_u) + (a_{u+1} - b_{u+1}) + \cdots + (a_v - b_v) \\ &= (a_u + a_{u+1} + \cdots + a_v) - (b_u + b_{u+1} + \cdots + b_v). \end{aligned}$$
This is rather clear; a formal proof would use induction on v.

- We have
$$\sum_{k=u}^{v} \lambda a_k = \lambda \sum_{k=u}^{v} a_k$$
for any integers u, v and any numbers λ, a_k. Indeed, this is just saying
$$\lambda a_u + \lambda a_{u+1} + \cdots + \lambda a_v = \lambda (a_u + a_{u+1} + \cdots + a_v).$$

Rules like these are dime a dozen, and you should find them all obvious (even though rigorous proofs would require some work – mostly induction).

Let us now compute our sum:

$$\begin{aligned}\sum_{k=1}^{n}(2k-1) &= \underbrace{\sum_{k=1}^{n} 2k}_{=2\sum_{k=1}^{n} k} - \underbrace{\sum_{k=1}^{n} 1}_{=n} \\ &= 2 \underbrace{\sum_{k=1}^{n} k}_{=\frac{n(n+1)}{2}} - n = 2 \cdot \frac{n(n+1)}{2} - n \\ &= n(n+1) - n = n^2 + n - n = n^2.\end{aligned}$$

As another illustration of how to use our notation, we can rewrite Gauss's proof of the Little Gauss formula

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$$

using sum notation. We need three new rules for this:

- We have
$$\sum_{k=u}^{v}(a_k + b_k) = \sum_{k=u}^{v} a_k + \sum_{k=u}^{v} b_k$$
for any integers u, v and any numbers a_k, b_k.

- We have
$$\sum_{k=u}^{v} a_k = \sum_{k=u}^{v} a_{u+v-k}$$
for any integers u, v and any numbers a_k. This is just saying that
$$a_u + a_{u+1} + \cdots + a_{v-1} + a_v = a_v + a_{v-1} + \cdots + a_{u+1} + a_u,$$
that is, a sum does not change if you reverse the order of its addends.

- For any integers $u \leq v$ and any number λ, we have
$$\sum_{k=u}^{v} \lambda = (v-u+1)\lambda.$$
(This is just saying that a sum of $v-u+1$ many equal addends λ is $(v-u+1)\lambda$. Keep in mind that $v-u+1$ – not $v-u$ – is the number of addends here!)

Now, Gauss's proof of his formula takes the following shape:

$$\begin{aligned}
&2 \cdot \sum_{k=1}^{n} k \\
&= \sum_{k=1}^{n} k + \sum_{k=1}^{n} k \\
&= \sum_{k=1}^{n} k + \sum_{k=1}^{n} (n+1-k) \\
&\qquad\qquad \text{(here, we reversed the second sum)} \\
&= \sum_{k=1}^{n} \underbrace{(k + (n+1-k))}_{=n+1} \\
&= \sum_{k=1}^{n} (n+1) \\
&= n(n+1).
\end{aligned}$$

We have found closed-form expressions for several sums. Not every sum has a closed-form expression. For example, there are no closed forms for $\sum\limits_{k=1}^{n} \dfrac{1}{k}$ and $\sum\limits_{k=1}^{n} k^k$.

Some more terminology:

The notation $\sum\limits_{k=u}^{v} a_k$ is called **sigma notation** or **finite sum notation** (\sum in LaTeX). The symbol $\sum$ itself is called the **summation sign**. The numbers u and v are called the **lower limit** and the **upper limit** of the summation. The variable k is called the **summation index** or the **running index**, and is said to **range** (or **run**) from u to v. The numbers a_k are called the **addends** of the finite sum.

There are many similarities between finite sums $\sum\limits_{k=u}^{v} a_k$ and integrals $\int_u^v f(x)\,dx$. But don't take this analogy too far – for example, $\sum\limits_{k=u}^{u} a_k = a_u$ but $\int_u^u f(x)\,dx = 0$.

Two more rules for finite sums are worth stating:

- The "splitting-off rule": For any integers $u \leq v$ and any numbers $a_u, a_{u+1}, \ldots, a_v$, we have
$$\sum_{k=u}^{v} a_k = \sum_{k=u}^{v-1} a_k + a_v = a_u + \sum_{k=u+1}^{v} a_k.$$

This rule allows us to split off the first or the last addend from a finite sum. This is useful in proofs by induction.

- More generally, any finite sum can be split at any point:

$$\sum_{k=u}^{v} a_k = \sum_{k=u}^{w} a_k + \sum_{k=w+1}^{v} a_k$$

for any integers $u-1 \leq w \leq v$ and any a_k. This is just saying that

$$\begin{aligned} & a_u + a_{u+1} + \cdots + a_v \\ & = (a_u + a_{u+1} + \cdots + a_w) + (a_{w+1} + a_{w+2} + \cdots + a_v) . \end{aligned}$$

Finite sum notation exists in several forms. The $\sum\limits_{k=u}^{v} a_k$ form is the most common. But there are other variants. For example,

$$\sum_{k \in \{1,2,\ldots,n\} \text{ is even}} k = 2+4+6+\cdots+\underbrace{(\text{the last even integer } \leq n)}_{=2\left\lfloor \frac{n}{2} \right\rfloor} .$$

We won't use these variants much, but they might come useful in a homework problem.

2.2. Finite products

Finite products are analogous to finite sums, just using multiplication instead of addition:

> **Definition 2.2.1.** Let u and v be two integers. Let $a_u, a_{u+1}, \ldots, a_v$ be some numbers. Then,
>
> $$\prod_{k=u}^{v} a_k$$
>
> is defined to be the product
>
> $$a_u a_{u+1} \cdots a_v$$
>
> (in more detail: $a_u a_{u+1} a_{u+2} \cdots a_{v-1} a_v$). It is called the **product of the numbers a_k where k ranges from u to v**. When $v < u$, this product is called **empty** and defined to be 1.

For example:

$$\begin{aligned}
&\prod_{k=5}^{10} k = 5 \cdot 6 \cdot 7 \cdot 8 \cdot 9 \cdot 10; \\
&\prod_{k=5}^{10} k^k = 5^5 \cdot 6^6 \cdot 7^7 \cdot 8^8 \cdot 9^9 \cdot 10^{10}; \\
&\prod_{k=5}^{5} \frac{1}{k} = \frac{1}{5}; \\
&\prod_{k=6}^{5} \frac{1}{k} = 1 \qquad \text{(an empty product)}; \\
&\prod_{k=1}^{n} a = \underbrace{aa\cdots a}_{n \text{ times}} = a^n; \\
&\prod_{k=1}^{n} a^k = a^1 a^2 \cdots a^n \\
&\qquad = a^{1+2+\cdots+n} \qquad \left(\begin{array}{c}\text{by one of the laws of exponents:} \\ a^{i_1}a^{i_2}\cdots a^{i_n} = a^{i_1+i_2+\cdots+i_n}\end{array}\right) \\
&\qquad = a^{n(n+1)/2}.
\end{aligned}$$

The notation $\prod\limits_{k=u}^{v} a_k$ is called **finite product notation** (\prod in LaTeX). The symbol $\prod$ itself is called the **product sign**. The numbers u and v are called the **lower limit** and the **upper limit** of the product. The variable k is called the **product index** or the **running index**, and is said to **range** (or **run**) from u to v. The numbers a_k are called the **factors** of the finite sum.

Almost all rules for finite sums have analogues for finite products. For instance:

- The "splitting-off rule": For any integers $u \le v$ and any numbers $a_u, a_{u+1}, \ldots, a_v$, we have
$$\prod_{k=u}^{v} a_k = \left(\prod_{k=u}^{v-1} a_k\right) a_v = a_u \prod_{k=u+1}^{v} a_k.$$
This rule allows us to split off the first or the last factor from a finite product. This is useful in proofs by induction.

- We have
$$\prod_{k=u}^{v} (a_k/b_k) = \left(\prod_{k=u}^{v} a_k\right) / \left(\prod_{k=u}^{v} b_k\right)$$
for any integers u, v and any numbers a_k, b_k, as long as the numbers b_k are nonzero.

2.3. Factorials

Recall that $\mathbb{N} = \{0, 1, 2, \ldots\}$.

> **Definition 2.3.1.** For any $n \in \mathbb{N}$, we define the positive integer $n!$ (called the **factorial** of n, and often pronounced "n **factorial**") by
>
> $$n! = \prod_{k=1}^{n} k = 1 \cdot 2 \cdot \cdots \cdot n.$$
>
> This is the product of the first n positive integers.

> **Proposition 2.3.2** (recursion of the factorials)**.** For any positive integer n, we have
>
> $$n! = (n-1)! \cdot n.$$

Proof. Let n be a positive integer. Then,

$$n! = 1 \cdot 2 \cdot \cdots \cdot n = \underbrace{(1 \cdot 2 \cdot \cdots \cdot (n-1))}_{=(n-1)!} \cdot n = (n-1)! \cdot n.$$

□

2.4. Binomial coefficients: Definition

We shall now define one of the most important families of numbers in mathematics:

> **Definition 2.4.1.** Let n and k be two numbers. Then, we define a number $\dbinom{n}{k}$ as follows:
>
> - If $k \in \mathbb{N}$, then we set
>
> $$\binom{n}{k} := \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}.$$
>
>   (The numerator here is the product of k factors, where the first factor is n and each further factor is 1 smaller than the previous. You can also write this product as $\prod\limits_{i=0}^{k-1} (n-i)$.)
>
> - If $k \notin \mathbb{N}$, then we set
>
> $$\binom{n}{k} := 0.$$

The number $\dbinom{n}{k}$ is called *"n* **choose** *k"*, and is known as the **binomial coefficient** of n and k. Do not mistake the notation $\dbinom{n}{k}$ for a vector $\begin{pmatrix} n \\ k \end{pmatrix}$.

Example 2.4.2. For any number n, we have

$$\begin{aligned}
\binom{n}{3} &= \frac{n\,(n-1)\,(n-2)}{3!} = \frac{n\,(n-1)\,(n-2)}{6};\\
\binom{n}{2} &= \frac{n\,(n-1)}{2!} = \frac{n\,(n-1)}{2};\\
\binom{n}{1} &= \frac{n}{1!} = n;\\
\binom{n}{0} &= \frac{(\text{empty product})}{0!} = \frac{1}{1} = 1;\\
\binom{n}{2.5} &= 0 \qquad (\text{since } 2.5 \notin \mathbb{N})\,;\\
\binom{n}{-1} &= 0 \qquad (\text{since } -1 \notin \mathbb{N})\,.
\end{aligned}$$

For any $k \in \mathbb{N}$, we have

$$\binom{0}{k} = \frac{0\,(0-1)\,(0-2)\cdots(0-k+1)}{k!} = \begin{cases} 1, & \text{if } k = 0; \\ 0, & \text{if } k \neq 0 \end{cases}$$

(since the product $0\,(0-1)\,(0-2)\cdots(0-k+1)$ always equals 0 unless $k = 0$, in which case it is empty) and

$$\binom{-1}{k} = \frac{(-1)\,(-2)\cdots(-k)}{k!} = (-1)^k \underbrace{\frac{1\cdot 2\cdot \cdots \cdot k}{k!}}_{=1} = (-1)^k\,.$$

Let us tabulate the values of $\dbinom{n}{k}$ for nonnegative integers n and k:

	$k=0$	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$
$n=0$	1	0	0	0	0	0	0
$n=1$	1	1	0	0	0	0	0
$n=2$	1	2	1	0	0	0	0
$n=3$	1	3	3	1	0	0	0
$n=4$	1	4	6	4	1	0	0
$n=5$	1	5	10	10	5	1	0
$n=6$	1	6	15	20	15	6	1

.

What patterns do we see on this table?

- We seem to have $\dbinom{n}{n} = 1$ for each $n \in \mathbb{N}$. This is indeed easy.
- As a matrix, the table is lower-triangular – i.e., everything above the main diagonal is 0. In other words, $\dbinom{n}{k} = 0$ whenever $n \in \mathbb{N}$ and $k > n$.
- Symmetry: $\dbinom{n}{k} = \dbinom{n}{n-k}$ whenever $n \in \mathbb{N}$.
- Binomial formula:
$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} \qquad \text{for any numbers } a, b.$$
For example,
$$(a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4.$$
- Pascal's recursion: $\dbinom{n}{k} = \dbinom{n-1}{k-1} + \dbinom{n-1}{k}$ for each n, k. In other words, each entry is the sum of the entry above it with the entry left-above it.
- Each binomial coefficient $\dbinom{n}{k}$ (with $n \in \mathbb{N}$) is an integer.

Let us prove these patterns. We begin with the "sea of zeroes":

Proposition 2.4.3. Let $n \in \mathbb{N}$ and $k > n$. Then, $\dbinom{n}{k} = 0$.

Proof. If $k \notin \mathbb{N}$, then this is obvious from the definition. So let's assume that $k \in \mathbb{N}$. Then,
$$\binom{n}{k} = \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}.$$
But the product in the numerator has $n - n = 0$ as one of its factors (since $n < k$), and thus is 0. So $\dbinom{n}{k} = 0$.

For instance, $\dbinom{3}{5} = \dfrac{3 \cdot 2 \cdot 1 \cdot 0 \cdot (-1)}{5!} = 0$. □

Remark 2.4.4. The requirement $n \in \mathbb{N}$ is important! In fact,

$$\binom{1.5}{3} = \frac{1.5 \cdot 0.5 \cdot (-0.5)}{3!} \neq 0.$$

In light of this proposition, we can reformat our table by removing all the 0's above the diagonal, and making the table more symmetric: (picture of Pascal's triangle drawn on the whiteboard; you can find a bigger version in the notes.)

Let us now prove the more serious properties.

2.5. Binomial coefficients: Properties

2.5.1. Pascal's identity

Theorem 2.5.1 (Pascal's identity, aka the recurrence of the binomial coefficients)**.** For any numbers n and k, we have

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

For instance, $\binom{7}{3} = \binom{6}{2} + \binom{6}{3}$, which is saying that $35 = 15 + 20$. But the theorem also holds for negative or non-integer n and k.

Proof. Let n and k be two numbers. We are in one of the following three cases:

Case 1: The number k is a positive integer.

Case 2: We have $k = 0$.

Case 3: None of the above.

We consider Case 1 first. In this case, k is a positive integer, so that both k and $k-1$ belong to $\mathbb{N}$. Thus, by the definition of BCs (= binomial coefficients),

$$\begin{aligned}
\binom{n}{k} &= \frac{n\left(n-1\right)\left(n-2\right)\cdots\left(n-k+1\right)}{k!};\\
\binom{n-1}{k-1} &= \frac{\left(n-1\right)\left(n-2\right)\left(n-3\right)\cdots\left(\left(n-1\right)-\left(k-1\right)+1\right)}{\left(k-1\right)!}\\
&= \frac{\left(n-1\right)\left(n-2\right)\left(n-3\right)\cdots\left(n-k+1\right)}{\left(k-1\right)!};\\
\binom{n-1}{k} &= \frac{\left(n-1\right)\left(n-2\right)\left(n-3\right)\cdots\left(\left(n-1\right)-k+1\right)}{k!}\\
&= \frac{\left(n-1\right)\left(n-2\right)\left(n-3\right)\cdots\left(n-k\right)}{k!}.
\end{aligned}$$

If we set $a := (n-1)(n-2)(n-3)\cdots(n-k+1)$, then we can rewrite these as

$$\binom{n}{k} = \frac{na}{k!}; \qquad \binom{n-1}{k-1} = \frac{a}{(k-1)!}; \qquad \binom{n-1}{k} = \frac{a(n-k)}{k!},$$

so our goal is to show that

$$\frac{na}{k!} = \frac{a}{(k-1)!} + \frac{a(n-k)}{k!}.$$

Rewriting $\dfrac{a}{(k-1)!}$ as $\dfrac{ak}{k!}$ (because $k! = (k-1)! \cdot k$), this takes the form

$$\frac{na}{k!} = \frac{ak}{k!} + \frac{a(n-k)}{k!},$$

boiling down to $na = ak + a(n-k)$, which is obvious.

Now consider Case 2. In this case, $k = 0$. So our goal

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

rewrites as

$$\binom{n}{0} = \binom{n-1}{-1} + \binom{n-1}{0}, \qquad \text{that is,}$$
$$1 = 0 + 1, \qquad \text{which is true.}$$

Finally, consider Case 3. In this case, $k \notin \mathbb{N}$, whence $k-1 \notin \mathbb{N}$. So our goal

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

rewrites as

$$0 = 0 + 0, \qquad \text{which is true.}$$

So we are done in all cases. □

Pascal's triangle is highly useful for proving facts about BCs (binomial coefficients) by induction on n.

2.5.2. The factorial formula

The binomial coefficients $\dbinom{n}{k}$ are defined for all numbers n and k. However, when n and k are nonnegative integers with $k \leq n$ (that is, when $n \in \mathbb{N}$ and $k \in \{0, 1, \ldots, n\}$; this is exactly the case that you see in Pascal's triangle), there is a particularly simple formula for $\dbinom{n}{k}$:

Theorem 2.5.2 (factorial formula)**.** Let $n \in \mathbb{N}$ and $k \in \{0, 1, \ldots, n\}$. Then,

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

Proof. The definition of $\binom{n}{k}$ yields

$$\binom{n}{k} = \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}.$$

Multiplying both sides by $k!$, we obtain

$$\begin{aligned}
k! \cdot \binom{n}{k} &= n(n-1)(n-2)\cdots(n-k+1) \\
&= (n-k+1)(n-k+2)\cdots n \\
&= \frac{1 \cdot 2 \cdot \cdots \cdot n}{1 \cdot 2 \cdot \cdots \cdot (n-k)} \\
&\qquad \left(\begin{array}{c}\text{since } n-k+1,\ n-k+2,\ \ldots,\ n \\ \text{are precisely the factors of } 1 \cdot 2 \cdot \cdots \cdot n \\ \text{that are missing from } 1 \cdot 2 \cdot \cdots \cdot (n-k)\end{array}\right) \\
&= \frac{n!}{(n-k)!}.
\end{aligned}$$

Dividing this by $k!$, we find

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

□

Remark 2.5.3. The factorial formula has its assumptions $n \in \mathbb{N}$ and $k \in \{0, 1, \ldots, n\}$; it does not hold (or even make sense) without them. So you can use it to compute $\binom{10}{4}$, but not to compute $\binom{2.5}{3}$ or $\binom{-1}{3}$.

2.5.3. The symmetry of binomial coefficients

Theorem 2.5.4 (symmetry of Pascal's triangle)**.** Let $n \in \mathbb{N}$, and let k be any number. Then,

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

Proof. We are in one of the following cases:

Case 1: We have $k \in \{0, 1, \ldots, n\}$.

Case 2: We have $k \in \mathbb{Z}$ and $k > n$.

Case 3: We have $k \in \mathbb{Z}$ and $k < 0$.

Case 4: We have $k \notin \mathbb{Z}$.

In Case 1, the factorial formula can be applied to k, yielding

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

but the factorial formula can also be applied to $n - k$ (because $k \in \{0, 1, \ldots, n\}$ entails $n - k \in \{0, 1, \ldots, n\}$), yielding

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

which is visibly the same number. So we have proved the formula $\binom{n}{k} = \binom{n}{n-k}$ in Case 1.

In the other three cases, the formula boils down to $0 = 0$, albeit for slightly different reasons in each case. (See the notes for details.) □

Next time: binomial formula.

Remark 2.5.5. The symmetry theorem does not hold for negative or non-integer n. For instance, if n is negative and k is nonnegative, then $\binom{n}{k} \neq 0$ but $\binom{n}{n-k} = 0$.

Corollary 2.5.6. For each $n \in \mathbb{N}$, we have $\binom{n}{n} = 1$.

Proof. For each $n \in \mathbb{N}$, the symmetry of Pascal's triangle yields $\binom{n}{n} = \binom{n}{n-n} = \binom{n}{0} = 1$. □

Again, the corollary fails for negative n: For instance, $\binom{-1}{-1} = 0$.

2.5.4. Pascal's triangle consists of integers

Perhaps surprisingly given their definition as fractions, the binomial coefficients $\binom{n}{k}$ are integers, at least when $n \in \mathbb{N}$:

Theorem 2.5.7. For any $n \in \mathbb{N}$ and any number k, we have $\dbinom{n}{k} \in \mathbb{N}$.

Proof. Idea: You can compute all the $\dbinom{n}{k}$ by Pascal's recursion starting with the $\dbinom{0}{i}$ (which are all 0 or 1). Pascal's recursion only involves addition, so you never have a chance to get a non-integer or a negative integer. So all the $\dbinom{n}{k}$ are $\in \mathbb{N}$.

Formally: We induct on n.

Base case: The theorem holds for $n = 0$, since any number k satisfies

$$\binom{0}{k} = \begin{cases} 1, & \text{if } k = 0; \\ 0, & \text{if } k \neq 0 \end{cases} \in \mathbb{N}.$$

Induction step: We will step from $n-1$ to n. So we fix a positive integer n, and we assume (as IH) that the theorem holds for $n-1$ instead of n. In other words, we assume that

$$\binom{n-1}{k} \in \mathbb{N} \qquad \text{for all numbers } k.$$

Our goal is to prove that the theorem holds for n as well. In other words, we must prove that

$$\binom{n}{k} \in \mathbb{N} \qquad \text{for all numbers } k.$$

But this is easy: Pascal's recursion yields

$$\binom{n}{k} = \underbrace{\binom{n-1}{k-1}}_{\substack{\in\mathbb{N} \\ \text{(by the IH,} \\ \text{applied to } k-1 \\ \text{instead of } k\text{)}}} + \underbrace{\binom{n-1}{k}}_{\substack{\in\mathbb{N} \\ \text{(by the IH)}}} \in \mathbb{N}$$

for all numbers k. So the induction step is complete, and the theorem is proved. □

The theorem we just proved cries out for a better explanation: Does $\dbinom{n}{k}$ count something? Certainly, in that case, $\dbinom{n}{k} \in \mathbb{N}$ would be clear.

And indeed, $\dbinom{n}{k}$ counts something:

Theorem 2.5.8 (combinatorial interpretation of BCs)**.** Let $n \in \mathbb{N}$, and let k be any number. Let A be any n-element set (i.e., a set that has exactly n distinct elements). Then,

$$\binom{n}{k} \text{ is the number of } k\text{-element subsets of } A.$$

This theorem will be proved in the last chapter, as we learn more about finite sets and their sizes. Note that the k-element subsets of A are also known as **combinations without replacement**. This explains why $\binom{n}{k}$ is called "n choose k": It is the number of ways to choose k out of n given elements, provided that the order does not matter and the chosen elements must be distinct.

Note that the theorem says nothing about $\binom{n}{k}$ when n is negative or non-integer. In fact, in this cases, you cannot find any n-element set A, so you cannot apply the theorem.

2.5.5. Upper negation

Theorem 2.5.9 (upper negation formula)**.** For any number n and $k \in \mathbb{Z}$, we have

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

Proof. If $k \notin \mathbb{N}$, then this is clear because both sides are 0.

Thus, consider only the case when $k \in \mathbb{N}$. In this case,

$$\begin{aligned}\binom{-n}{k} &= \frac{(-n)(-n-1)(-n-2)\cdots(-n-k+1)}{k!} \\ &= (-1)^k \frac{n(n+1)(n+2)\cdots(n+k-1)}{k!}\end{aligned}$$

and

$$\begin{aligned}\binom{n+k-1}{k} &= \frac{(n+k-1)(n+k-2)\cdots((n+k-1)-k+1)}{k!} \\ &= \frac{(n+k-1)(n+k-2)\cdots n}{k!} \\ &= \frac{n(n+1)(n+2)\cdots(n+k-1)}{k!}.\end{aligned}$$

Comparing these equalities, the claim becomes clear. □

Corollary 2.5.10. For any $n \in \mathbb{Z}$ and any number k, we have $\binom{n}{k} \in \mathbb{Z}$.

Proof. If $n \geq 0$, this has already been proved in a theorem.

If $k \notin \mathbb{N}$, then this is clear since $\binom{n}{k} = 0$.

In the remaining case, upper negation can be used to reduce this to the nonnegative case. See a reference in the notes for the details. □

For example, $\binom{-3}{5} \in \mathbb{Z}$. This has no direct combinatorial interpretation, since there is no (-3)-element set.

2.5.6. Finding Fibonacci numbers in Pascal's triangle

Theorem 2.5.11. For any $n \in \mathbb{N}$, the Fibonacci number f_{n+1} is

$$\begin{aligned} f_{n+1} &= \binom{n-0}{0} + \binom{n-1}{1} + \binom{n-2}{2} + \cdots + \binom{n-n}{n} \\ &= \sum_{k=0}^{n} \binom{n-k}{k}. \end{aligned}$$

For instance, for $n = 5$, this is saying

$$\begin{aligned} f_6 &= \binom{5}{0} + \binom{4}{1} + \binom{3}{2} + \binom{2}{3} + \binom{1}{4} + \binom{0}{5} \\ &= 1 + 4 + 3 + 0 + 0 + 0 = 8. \end{aligned}$$

We will prove this in the last chapter.

2.6. The binomial formula

The **binomial formula** is one of the most important properties of BCs (and in fact gave them their name):

Theorem 2.6.1 (binomial formula, aka binomial theorem)**.** Let a and b be any numbers, and let $n \in \mathbb{N}$. Then,

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

Restating this without summation signs:

$$(a+b)^n = \binom{n}{0} a^0 b^n + \binom{n}{1} a^1 b^{n-1} + \cdots + \binom{n}{n} a^n b^0.$$

Equivalently:

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

Proof. The three formulas are clearly equivalent: The second just restates the first, whereas the third follows from the first upon swapping a and b (since $b + a = a + b$). So it suffices to show the first formula.

We do this by induction on n:

Base case: For $n = 0$, the binomial formula says that

$$\underbrace{(a+b)^0}_{=1} = \underbrace{\sum_{k=0}^{0} \binom{0}{k} a^k b^{0-k}}_{=\binom{0}{0} a^0 b^{0-0} = 1\cdot 1\cdot 1 = 1} .$$

Induction step: Let $n \in \mathbb{N}$. We assume (as the IH) that the binomial formula holds for n. In other words, we assume that

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

We must show that the binomial formula holds for $n+1$. In other words, we must show that

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

Indeed, we have

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)^n \cdot (a+b) \\
&= \left(\sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} \right) \cdot (a+b) \qquad \text{(by the IH)} \\
&= \left(\sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} \right) \cdot a + \left(\sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} \right) \cdot b \\
&= \sum_{k=0}^{n} \binom{n}{k} \underbrace{a^k b^{n-k} a}_{=a^{k+1} b^{n-k}} + \sum_{k=0}^{n} \binom{n}{k} a^k \underbrace{b^{n-k} b}_{\substack{=b^{n-k+1} \\ =b^{n+1-k}}} \\
&\qquad \left(\text{by the formula } \left(\sum_{s=u}^{v} a_s \right) c = \sum_{s=u}^{v} a_s c \right) \\
&= \sum_{k=0}^{n} \binom{n}{k} a^{k+1} b^{n-k} + \sum_{k=0}^{n} \binom{n}{k} a^k b^{n+1-k}.
\end{aligned}$$

On the other hand, Pascal's recursion yields

$$\begin{aligned}\binom{n+1}{k} &= \binom{n+1-1}{k-1} + \binom{n+1-1}{k} \\ &= \binom{n}{k-1} + \binom{n}{k} \qquad \text{for each } k.\end{aligned}$$

Thus,

$$\begin{aligned}
&\sum_{k=0}^{n+1}\binom{n+1}{k}a^kb^{n+1-k} \\
&= \sum_{k=0}^{n+1}\left(\binom{n}{k-1}+\binom{n}{k}\right)a^kb^{n+1-k} \\
&= \sum_{k=0}^{n+1}\left(\binom{n}{k-1}a^kb^{n+1-k}+\binom{n}{k}a^kb^{n+1-k}\right) \\
&= \underbrace{\sum_{k=0}^{n+1}\binom{n}{k-1}a^kb^{n+1-k}}_{\substack{=\binom{n}{0-1}a^0b^{n+1-0}+\sum_{k=1}^{n+1}\binom{n}{k-1}a^kb^{n+1-k} \\ =\sum_{k=1}^{n+1}\binom{n}{k-1}a^kb^{n+1-k} \\ \text{(since }\binom{n}{0-1}=0\text{)}}} + \underbrace{\sum_{k=0}^{n+1}\binom{n}{k}a^kb^{n+1-k}}_{\substack{=\sum_{k=0}^{n}\binom{n}{k}a^kb^{n+1-k}+\binom{n}{n+1}a^{n+1}b^{n+1-(n+1)} \\ =\sum_{k=0}^{n}\binom{n}{k}a^kb^{n+1-k} \\ \text{(since }\binom{n}{n+1}=0\text{)}}} \\
&= \sum_{k=1}^{n+1}\binom{n}{k-1}a^kb^{n+1-k}+\sum_{k=0}^{n}\binom{n}{k}a^kb^{n+1-k}
\end{aligned}$$

For comparison,

$$LHS = \sum_{k=0}^{n}\binom{n}{k}a^{k+1}b^{n-k}+\sum_{k=0}^{n}\binom{n}{k}a^kb^{n+1-k}.$$

So both *LHS* and *RHS* are written as sums of two sums each, and the right sums agree. It remains to show that the left sums also agree, i.e., that

$$\sum_{k=0}^{n}\binom{n}{k}a^{k+1}b^{n-k} = \sum_{k=1}^{n+1}\binom{n}{k-1}a^kb^{n+1-k}.$$

But this equality holds because both sums contain the same addends: If you write them out without summation signs, they both become

$$\binom{n}{0}a^1b^n+\binom{n}{1}a^2b^{n-1}+\binom{n}{2}a^3b^{n-2}+\cdots+\binom{n}{n}a^{n+1}b^0.$$

This argument can be made rigorous using an important summation rule called **substitution**. In its simplest form, this rule says that

$$\sum_{k=u}^{v} c_k = \sum_{k=u+\delta}^{v+\delta} c_{k-\delta}$$

for any integers u, v, δ and any numbers $c_u, c_{u+1}, \ldots, c_v$. Intuitively this is obvious (both sides are $c_u + c_{u+1} + \cdots + c_v$).

When we use this rule to rewrite a sum of the form $\sum\limits_{k=u}^{v} c_k$ in the form $\sum\limits_{k=u+\delta}^{v+\delta} c_{k-\delta}$, we say that we are **substituting** $k - \delta$ for k in the sum. For instance, taking $u = 4$ and $v = 7$ and $c_k = k^k$ and $\delta = -3$, then we get

$$\sum_{k=4}^{7} k^k = \sum_{k=1}^{4} (k+3)^{k+3}.$$

Now, substituting $k - 1$ for k in the sum $\sum\limits_{k=0}^{n} \binom{n}{k} a^{k+1} b^{n-k}$, we obtain

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

This is the equality we wanted to show.

So the induction step is complete, and the binomial formula is proved. □

3. Elementary number theory

Number theory is commonly understood as the study of integers, and particularly of those properties of integers that do not work or make much sense for more general types of numbers. Divisibility is one such property; prime numbers are another. We will only cover the basics of the field; more advanced texts are references in the notes. (See also abstract algebra classes like Math 331/2 for a more general viewpoint.)

3.1. Divisibility

3.1.1. Definition

Definition 3.1.1. Let a and b be two integers.

We write $a \mid b$ (and we say that "a **divides** b", or "b is **divisible** by a", or "b is a **multiple** of a", or "a is a **divisor** of b") if there exists an integer c such that $b = ac$.

We write $a \nmid b$ if we don't have $a \mid b$.

Example 3.1.2. **(a)** We have $4 \mid 12$, because $12 = 4 \cdot 3$.

(b) We have $4 \nmid 11$, since there is no integer c such that $11 = 4c$.

(c) We have $1 \mid b$ for every integer b, since $b = 1b$.

(d) We have $a \mid a$ for every integer a, since $a = a \cdot 1$. In particular, $0 \mid 0$, against certain authors.

(e) We have $a \mid 0$ for every integer a, since $0 = a \cdot 0$.

(f) An integer b satisfies $0 \mid b$ if and only if $b = 0$.

In particular:

Definition 3.1.3. **(a)** An integer n is said to be **even** if $2 \mid n$.

(b) An integer n is said to be **odd** if $2 \nmid n$.

You likely already know some properties of even and odd numbers:

1. A sum of two even numbers is even.
2. A sum of an even and an odd number is odd.
3. A sum of two odd numbers is even.

These claims are not actually obvious. Well, claim 1 is easy, claim 2 is still pretty easy, but what about claim 3? With the tools at hand right now, this is far from obvious. So let us understand divisibility better, so that all three claims become straightforward.

3.1.2. Basic properties

Proposition 3.1.4. Let a and b be two integers. Then:

(a) We have $a \mid b$ if and only if $|a| \mid |b|$ (to read as "$|a|$ divides $|b|$").

(b) If $a \mid b$ and $b \neq 0$, then $|a| \leq |b|$.

(c) If $a \mid b$ and $b \mid a$, then $|a| = |b|$.

(d) Assume that $a \neq 0$. Then, $a \mid b$ if and only if $\frac{b}{a} \in \mathbb{Z}$.

Proof. **(a)** Part **(a)** just says that the divisibility $a \mid b$ does not depend on the signs of a and b; in other words, it says that we can replace the numbers a and b by their absolute values without changing the truth/falsity of $a \mid b$.

Clearly, in order to prove this, it suffices to show the equivalences

$$(a \mid b) \iff (a \mid -b) \iff (-a \mid b),$$

because then you are allowed to replace any of a and b by its negative and thus also by its absolute value (since $|x|$ is always either x or $-x$).

But both equivalences are easy:

Proof of $(a \mid b) \implies (a \mid -b)$: Assume that $a \mid b$. Thus, $b = ac$ for some integer c. Using this c, we then have $-b = -ac = a(-c)$, so that $a \mid -b$.

Proof of $(a \mid -b) \implies (a \mid b)$: Assume that $a \mid -b$. Thus, $-b = ac$ for some integer c. Using this c, we then have $b = -ac = a(-c)$, so that $a \mid b$.

Proof of $(a \mid -b) \implies (-a \mid b)$: Assume that $a \mid -b$. Thus, $-b = ac$ for some integer c. Using this c, we then have $b = -ac = (-a)c$, so that $-a \mid b$.

Proof of $(-a \mid b) \implies (a \mid -b)$: Exercise.

Thus, part **(a)** is proved.

(b) Assume that $a \mid b$ and $b \neq 0$. We must prove that $|a| \leq |b|$.

Let $x = |a|$ and $y = |b|$. Then, from $a \mid b$, we obtain $x \mid y$ by part **(a)**. Moreover, $y = |b| > 0$ because $b \neq 0$. And of course, $x \geq 0$.

From $x \mid y$, we obtain $y = xz$ for some integer z. Consider this z.

If z was ≤ 0, then we would have $y = \underbrace{x}_{\geq 0}\underbrace{z}_{\leq 0} \leq 0$, but this would contradict $y > 0$. So $z > 0$. Thus, $z \geq 1$ (since z is an integer). Multiplying this inequality with the nonnegative x, we obtain $xz \geq x$. But $y = xz \geq x$, meaning that $x \leq y$, meaning that $|a| \leq |b|$. This proves part **(b)**.

(c) Assume that $a \mid b$ and $b \mid a$. We must prove that $|a| = |b|$.

If $a = 0$, then this is easy (since $a \mid b$ and $a = 0$ imply $b = 0$). If $b = 0$, then this is equally easy. Thus, from now on, we assume that neither a nor b is 0. Thus, part **(b)** yields $|a| \leq |b|$. But we can also apply part **(b)** with the roles of a and b interchanged, and obtain $|b| \leq |a|$. Combining these two inequalities, we obtain $|a| = |b|$.

(d) *Proof of $(a \mid b) \implies \left(\frac{b}{a} \in \mathbb{Z}\right)$:* Assume that $a \mid b$. Thus, $b = ac$ for some integer c. This integer c must therefore be $\frac{b}{a}$ (just by solving $b = ac$ for c). So $\frac{b}{a}$ is an integer, i.e., we have $\frac{b}{a} \in \mathbb{Z}$.

Proof of $\left(\frac{b}{a} \in \mathbb{Z}\right) \implies (a \mid b)$: If $\frac{b}{a} \in \mathbb{Z}$, then $b = a \cdot \underbrace{\frac{b}{a}}_{\in \mathbb{Z}}$, so that $a \mid b$. □

Theorem 3.1.5 (rules for divisibility). **(a)** We have $a \mid a$ for each $a \in \mathbb{Z}$. (This is called **reflexivity of divisibility**.)

(b) If $a, b, c \in \mathbb{Z}$ such that $a \mid b$ and $b \mid c$, then $a \mid c$. (This is called **transitivity of divisibility**.)

(c) If $a_1, a_2, b_1, b_2 \in \mathbb{Z}$ satisfy $a_1 \mid b_1$ and $a_2 \mid b_2$, then $a_1 a_2 \mid b_1 b_2$. (This is called **multiplying two divisibilities**.)

(d) If $d, a, b \in \mathbb{Z}$ satisfy $d \mid a$ and $d \mid b$, then $d \mid a + b$. (In other words, a sum of two multiples of d is again a multiple of d.)

Proof. **(a)** Let $a \in \mathbb{Z}$. Then, $a = a \cdot 1$, so that $a \mid a$. This proves part **(a)**.

(b) Let $a, b, c \in \mathbb{Z}$. Assume that $a \mid b$ and $b \mid c$. We must show $a \mid c$.

From $a \mid b$, we see that there is an integer x such that $b = ax$.

From $b \mid c$, we see that there is an integer y such that $c = by$.

Consider these x and y. We have

$$c = \underbrace{b}_{=ax} y = axy.$$

Thus, there exists an integer z such that $c = az$ (namely, $z = xy$). In other words, $a \mid c$. This proves part **(b)**.

(c) Let $a_1, a_2, b_1, b_2 \in \mathbb{Z}$ be such that $a_1 \mid b_1$ and $a_2 \mid b_2$. We must show $a_1 a_2 \mid b_1 b_2$.

As in part **(b)**, unravel the definitions: $a_1 \mid b_1$ means that $b_1 = a_1 x_1$ for some integer x_1, whereas $a_2 \mid b_2$ means that $b_2 = a_2 x_2$ for some integer x_2. Then,

$$b_1 b_2 = (a_1 x_1)(a_2 x_2) = (a_1 a_2) \underbrace{(x_1 x_2)}_{\text{an integer}},$$

and this shows that $a_1 a_2 \mid b_1 b_2$.

(d) Let $d, a, b \in \mathbb{Z}$ be such that $d \mid a$ and $d \mid b$. We must show $d \mid a + b$.

As in part **(b)**, unravel the definitions: $d \mid a$ means that $a = dx$ for some integer x, whereas $d \mid b$ means that $b = dy$ for some integer y. Then,

$$a + b = dx + dy = d \underbrace{(x + y)}_{\text{an integer}},$$

whence $d \mid a + b$. □

Part **(b)** of the theorem tells us that divisibilities can be chained together: If $a \mid b$ and $b \mid c$, then $a \mid c$. Therefore you will often see a statement of the form "$a \mid b$ and $b \mid c$" shortened to "$a \mid b \mid c$", just like for chains of inequalities. More generally, the statement

$$a_1 \mid a_2 \mid \cdots \mid a_k$$

means that each of the numbers $a_1, a_2, \ldots, a_k$ divides the next (i.e., that $a_1 \mid a_2$ and so on). By induction on k, this is easily seen to entail that $a_1 \mid a_k$.

3.1.3. Divisibility criteria

How can you spot divisibilities between specific numbers? For small values of a, there are several known **divisibility criteria/rules**. Here are some:

Theorem 3.1.6. Let $b \in \mathbb{N}$. Write b in decimal notation.

(a) We have $2 \mid b$ if and only if the last digit of b is 0 or 2 or 4 or 6 or 8.

(b) We have $5 \mid b$ if and only if the last digit of b is 0 or 5.

(c) We have $10 \mid b$ if and only if the last digit of b is 0.

(d) We have $3 \mid b$ if and only if the sum of all digits of b is divisible by 3.

(e) We have $9 \mid b$ if and only if the sum of all digits of b is divisible by 9.

Example 3.1.7. Let $b = 64216$. Then, $2 \mid b$ and $5 \nmid b$ and $10 \nmid b$. Is $3 \mid b$? The sum of digits of b is $6 + 4 + 2 + 1 + 6 = 19$, which is not divisible by 3. Hence, $3 \nmid b$. Likewise, $9 \nmid b$. (In fact, $9 \nmid b$ follows from $3 \nmid b$, because if we had $9 \mid b$, then we would have $3 \mid 9 \mid b$.)

Some parts of the above theorem are easy to prove: Part **(c)** just follows from the fact that multiplying a number by 10 simply inserts a 0 digit at its end (in decimal). Parts **(a)** and **(b)** are somewhat trickier. Parts **(d)** and **(e)** are harder. To give them simple proofs, we will now introduce another relations between integers, known as **congruence modulo** n.

3.2. Congruence modulo n

3.2.1. Definition

Definition 3.2.1. Let $n, a, b \in \mathbb{Z}$. We say that a is **congruent to** b **modulo** n if and only if $n \mid a - b$.

We will use the notation "$a \equiv b \bmod n$" for "a is congruent to b modulo n".

We will use the notation "$a \not\equiv b \bmod n$" for "a is not congruent to b modulo n".

Example 3.2.2. **(a)** Is $3 \equiv 7 \bmod 2$? This would mean that $2 \mid 3 - 7$, which is true ($3 - 7 = -4 = 2\,(-2)$). So yes, $3 \equiv 7 \bmod 2$.

(b) Is $3 \equiv 6 \bmod 2$? No, since $2 \nmid 3 - 6$.

(c) We have $a \equiv b \bmod 1$ for any integers a and b. This is because $1 \mid a - b$.

(d) We have $a \equiv b \bmod 0$ for two integers a and b if and only if $a = b$.

(e) For any two integers a and b, we have $a + b \equiv a - b \bmod 2$, since $(a + b) - (a - b) = 2b$ is divisible by 2.

The word "modulo" originates with Gauss and should be read as something like "with respect to". You can translate the statement "a is congruent to b modulo n" as "a equals b up to a multiple of n". In fact,

$$a \equiv b \bmod n \qquad \text{if and only if} \qquad a = b + nc \text{ for some } c \in \mathbb{Z}.$$

As we will soon see, congruence modulo 2 is essentially parity: Two integers a and b are congruent to each other modulo 2 if and only if they are either both even or both odd.

3.2.2. Basic properties

Proposition 3.2.3. Let $n, a \in \mathbb{Z}$. Then, $a \equiv 0 \mod n$ if and only if $n \mid a$.

Proof. We have $(a \equiv 0 \mod n) \Longleftrightarrow (n \mid a - 0) \Longleftrightarrow (n \mid a)$. □

Proposition 3.2.4. Let $n \in \mathbb{Z}$. Then:

(a) We have $a \equiv a \mod n$ for every $a \in \mathbb{Z}$. (This is called **reflexivity of congruence**.)

(b) If $a, b \in \mathbb{Z}$ satisfy $a \equiv b \mod n$, then $b \equiv a \mod n$. (This is called **symmetry of congruence**.)

(c) If $a, b, c \in \mathbb{Z}$ satisfy $a \equiv b \mod n$ and $b \equiv c \mod n$, then $a \equiv c \mod n$. (This is called **transitivity of congruence**.)

(d) If $a_1, a_2, b_1, b_2 \in \mathbb{Z}$ satisfy

$$a_1 \equiv b_1 \mod n \qquad \text{and} \qquad a_2 \equiv b_2 \mod n,$$

then

$$\begin{aligned} a_1 + a_2 &\equiv b_1 + b_2 \mod n; \\ a_1 - a_2 &\equiv b_1 - b_2 \mod n; \\ a_1 a_2 &\equiv b_1 b_2 \mod n. \end{aligned}$$

In other words, two congruences modulo n can be added, subtracted or multiplied.

(e) Let $m \in \mathbb{Z}$ be such that $m \mid n$. Then, if $a, b \in \mathbb{Z}$ satisfy $a \equiv b \mod n$, then $a \equiv b \mod m$. (In other words, you can always replace the n in a valid congruence $a \equiv b \mod n$ by a divisor of this n.)

Proof. **(a)** Let $a \in \mathbb{Z}$. Then, $n \mid a - a$ since $a - a = 0 = n \cdot 0$. So $a \equiv a \mod n$.

(b) Let $a, b \in \mathbb{Z}$ satisfy $a \equiv b \mod n$. Why is $b \equiv a \mod n$?

From $a \equiv b \mod n$, we obtain $n \mid a - b \mid b - a$ (since $b - a = (a - b)(-1)$) and therefore $b \equiv a \mod n$.

(c) Assume that $a, b, c \in \mathbb{Z}$ satisfy $a \equiv b \mod n$ and $b \equiv c \mod n$. Why is $a \equiv c \mod n$?

From $a \equiv b \mod n$, we have $n \mid a - b$.

From $b \equiv c \mod n$, we have $n \mid b - c$.

So $a - b$ and $b - c$ are two multiples of n. Thus, their sum is also a multiple of n. But this sum is $(a - b) + (b - c) = a - c$. So $n \mid a - c$. In other words, $a \equiv c \mod n$.

(d) Let $a_1, a_2, b_1, b_2 \in \mathbb{Z}$ satisfy

$$a_1 \equiv b_1 \mod n \qquad \text{and} \qquad a_2 \equiv b_2 \mod n.$$

We must prove that

$$\begin{aligned} a_1 + a_2 &\equiv b_1 + b_2 \mod n; \\ a_1 - a_2 &\equiv b_1 - b_2 \mod n; \\ a_1 a_2 &\equiv b_1 b_2 \mod n. \end{aligned}$$

From $a_1 \equiv b_1 \mod n$, we obtain $n \mid a_1 - b_1$. In other words, $a_1 - b_1 = nc_1$ for some integer c_1. Thus, $a_1 = b_1 + nc_1$.

Likewise, $a_2 = b_2 + nc_2$ for some integer c_2.

Thus,

$$\begin{aligned} a_1 + a_2 &= (b_1 + nc_1) + (b_2 + nc_2) \\ &= (b_1 + b_2) + n(c_1 + c_2). \end{aligned}$$

This differs from $b_1 + b_2$ by a multiple of n (namely, $n(c_1 + c_2)$). In other words, $(a_1 + a_2) - (b_1 + b_2)$ is a multiple of n. In other words, $a_1 + a_2 \equiv b_1 + b_2 \mod n$.

Likewise,

$$\begin{aligned} a_1 - a_2 &= (b_1 + nc_1) - (b_2 + nc_2) \\ &= (b_1 - b_2) + n(c_1 - c_2). \end{aligned}$$

Thus, $a_1 - a_2 \equiv b_1 - b_2 \mod n$.

Finally,

$$\begin{aligned} a_1 a_2 &= (b_1 + nc_1)(b_2 + nc_2) \\ &= b_1 b_2 + b_1 nc_2 + b_2 nc_1 + nc_1 nc_2 \\ &= b_1 b_2 + n(b_1 c_2 + b_2 c_1 + nc_1 c_2). \end{aligned}$$

This differs from $b_1 b_2$ by a multiple of n. In other words, $a_1 a_2 - b_1 b_2$ is a multiple of n. Thus, $a_1 a_2 \equiv b_1 b_2 \mod n$.

(e) Let $a, b \in \mathbb{Z}$ satisfy $a \equiv b \mod n$. Then, we must prove that $a \equiv b \mod m$.

But this is easy: $m \mid n \mid a - b$ (since $a \equiv b \mod n$). So $a \equiv b \mod m$. □

Part **(b)** of the proposition we just proved says that congruences can be turned around: If $a \equiv b \mod n$, then $b \equiv a \mod n$. (This is very different from divisibilities: $a \mid b$ and $b \mid a$ rarely hold together.)

Part **(c)** says that congruences modulo the same n can be chained together: If $a \equiv b \mod n$ and $b \equiv c \mod n$, then $a \equiv c \mod n$. Thus, instead of writing "$a \equiv b \mod n$ and $b \equiv c \mod n$", we will often just say "$a \equiv b \equiv c \mod n$". More generally, the statement

$$a_1 \equiv a_2 \equiv \cdots \equiv a_k \mod n$$

shall mean that each of the numbers $a_1, a_2, \ldots, a_k$ is congruent to the next modulo n. By induction on k, it is easy to see that such a chain of congruences always entails that **any two** of the a_i's are congruent modulo n.

Part **(d)** says that congruences modulo n (for a fixed integer n) can be added, subtracted and multiplied. However,

- they cannot be divided by each other: $2 \equiv 0 \bmod 2$ and $2 \equiv 2 \bmod 2$ but $\dfrac{2}{2} \not\equiv \dfrac{0}{2} \bmod 2$.
- they cannot be taken to each other's power: $2 \equiv 2 \bmod 2$ and $2 \equiv 0 \bmod 2$ but $2^2 \not\equiv 2^0 \bmod 2$.

However, we can take a congruence to a k-th power for a fixed $k \in \mathbb{N}$:

Exercise 3.2.1. Let $n, a, b \in \mathbb{Z}$ be such that $a \equiv b \bmod n$. Let $k \in \mathbb{N}$. Then, $a^k \equiv b^k \bmod n$.

3.2.3. Proving the divisibility criteria

Now let us prove the divisibility criterion for 9 ("We have $9 \mid b$ if and only if the sum of all digits of b is divisible by 9"). We restate this as follows:

Proposition 3.2.5. Let $m \in \mathbb{N}$. Let s be the sum of digits of m written in decimal (e.g., if $m = 320$, then $s = 3 + 2 + 0 = 5$).

Then, $9 \mid m$ if and only if $9 \mid s$.

Proof. We will prove that $m \equiv s \bmod 9$.

Indeed, let the integer m have decimal representation $m_d m_{d-1} \cdots m_0$ (where m_d is the leading digit). Thus,

$$m = m_d \cdot 10^d + m_{d-1} \cdot 10^{d-1} + \cdots + m_0 \cdot 10^0 \qquad \text{and}$$
$$s = m_d + m_{d-1} + \cdots + m_0.$$

However, $10 \equiv 1 \bmod 9$ (since $9 \mid 10 - 1$). Therefore, by the exercise we just gave, for each $k \in \mathbb{N}$ we have $10^k \equiv 1^k \bmod 9$. Multiplying this congruence by the obvious congruence $m_k \equiv m_k \bmod 9$, we obtain

$$m_k \cdot 10^k \equiv m_k \cdot 1^k \bmod 9 \qquad \text{for every } k \in \{0, 1, \ldots, d\}.$$

In other words,

$$m_k \cdot 10^k \equiv m_k \bmod 9 \qquad \text{for every } k \in \{0, 1, \ldots, d\}.$$

In other words,

$$\begin{aligned} m_d \cdot 10^d &\equiv m_d \mod 9; \\ m_{d-1} \cdot 10^{d-1} &\equiv m_{d-1} \mod 9; \\ m_{d-2} \cdot 10^{d-2} &\equiv m_{d-2} \mod 9; \\ &\ldots; \\ m_0 \cdot 10^0 &\equiv m_0 \mod 9. \end{aligned}$$

Adding these congruences all together, we obtain

$$\begin{aligned} & m_d \cdot 10^d + m_{d-1} \cdot 10^{d-1} + \cdots + m_0 \cdot 10^0 \\ & \equiv m_d + m_{d-1} + \cdots + m_0 \mod 9. \end{aligned}$$

In other words,

$$m \equiv s \mod 9.$$

Now, if $9 \mid m$, then $m \equiv 0 \mod 9$, so $0 \equiv m \equiv s \mod 9$, thus $s \equiv 0 \mod 9$, which means that $9 \mid s$.

Conversely, if $9 \mid s$, then $s \equiv 0 \mod 9$, whence $m \equiv s \equiv 0 \mod 9$, therefore $9 \mid m$.

Combining these, we see that $9 \mid m$ if and only if $9 \mid s$. This proves the proposition. □

The same proof can be used for the divisibility rule for 3, since 3 also satisfies $3 \mid 10 - 1$.

The divisibility rules for 2 and 5 can be proved similarly but are also pretty easy by themselves.

There are also divisibility rules for 11 (exercise) and 7 (less easy to use).

3.3. Division with remainder

3.3.1. The theorem

What comes next is the most fundamental theorem of number theory:

> **Theorem 3.3.1** (division-with-remainder theorem). Let n be an integer. Let d be a positive integer. Then, there exists a **unique** pair (q, r) of integers
>
> $$q \in \mathbb{Z} \qquad \text{and} \qquad r \in \{0, 1, \ldots, d-1\}$$
>
> such that
>
> $$n = qd + r.$$

Before we prove this, let us introduce notations:

Definition 3.3.2. Let n be an integer. Let d be a positive integer. Let (q, r) be the unique pair from the previous theorem. Then:

- The number q is called the **quotient** of the division of n by d, and is denoted by $n//d$.
- The number r is called the **remainder** of the division of n by d, and is denoted by $n\%d$.
- The pair (q, r) is called the **quo-rem pair** of n and d.

So the theorem is claiming that there is a unique quo-rem pair (q, r) for n and d. Before we have proved it, we must take care to always say "**a** quotient" and "**a** remainder". The definite articles must first be "unlocked" by proving the theorem.

Example 3.3.3. We have $8//5 = 1$ and $8\%5 = 3$, since

$$1 \in \mathbb{Z} \qquad \text{and} \qquad 3 \in \{0, 1, \ldots, 4\}$$

such that

$$8 = 1 \cdot 5 + 3.$$

Example 3.3.4. We have $19//5 = 3$ and $19\%5 = 4$.

Example 3.3.5. We have $(-7)//5 = -2$ and $(-7)\%5 = 3$.

3.3.2. The proof

How do we prove the theorem? Let us first show uniqueness:

Proof of the uniqueness part. Fix an integer n and a positive integer d. We must show that there is **at most one** quo-rem pair (q, r) of n and d. In other words, we must show that there are no two distinct quo-rem pairs of n and d.

We shall prove this by contradiction. So we assume that (q_1, r_1) and (q_2, r_2) are two distinct quo-rem pairs of n and d. We want to find a contradiction.

Since (q_1, r_1) is a quo-rem pair of n and d, we have

$$q_1 \in \mathbb{Z} \qquad \text{and} \qquad r_1 \in \{0, 1, \ldots, d-1\} \qquad \text{and} \qquad n = q_1 d + r_1.$$

Since (q_2, r_2) is a quo-rem pair of n and d, we have

$$q_2 \in \mathbb{Z} \qquad \text{and} \qquad r_2 \in \{0, 1, \ldots, d-1\} \qquad \text{and} \qquad n = q_2 d + r_2.$$

Subtracting the equalities $n = q_1 d + r_1$ and $n = q_2 d + r_2$ from each other, we obtain

$$\begin{aligned} 0 &= (q_1 d + r_1) - (q_2 d + r_2) \\ &= (r_1 - r_2) - (q_2 - q_1)\, d. \end{aligned}$$

In other words,

$$r_1 - r_2 = (q_2 - q_1)\, d.$$

But $r_1 - r_2$ is a difference of two elements of $\{0, 1, \ldots, d-1\}$, hence has absolute value $< d$. Meanwhile, $(q_2 - q_1)\, d$ is a multiple of d, hence has absolute value $\geq d$ unless it is just 0. Since these two numbers are equal, they must both be 0. So $r_1 - r_2 = 0$ and $(q_2 - q_1)\, d = 0$. This easily yields $r_1 = r_2$ and $q_1 = q_2$, so that $(q_1, r_1) = (q_2, r_2)$, which contradicts the assumption that the two pairs (q_1, r_1) and (q_2, r_2) are distinct. This completes the proof of uniqueness. (See the notes for this in more detail.)

So we now know that the quo-rem pair of n and d is unique, but we still need to prove that it exists.

We will do this by strong induction on n. However, there is a problem: n is just an integer, not a nonnegative integer, not even an integer ≥ -9000. The induction principle does not give us a way to prove a statement for all integers with no lower bound.

We will surpass this obstacle in a straightforward way: We will use induction to prove our claim for $n \geq 0$; then we will derive the general case from this.

Here is the $n \geq 0$ case:

Lemma 3.3.6. Let $n \in \mathbb{N}$, and let d be a positive integer. Then, there exists a quo-rem pair of n and d.

Proof of the lemma. Fix d. We apply strong induction on n:

Induction step: Let $n \in \mathbb{N}$. Assume (as the IH) that the lemma holds for all nonnegative integers smaller than n instead of n. In other words, for every nonnegative integer $k < n$, there exists a quo-rem pair of k and d. Now we must prove that the lemma also holds for n, that is, that there exists a quo-rem pair of n and d.

If $n < d$, then such a pair can be easily constructed: it is $(0, n)$. Indeed, this satisfies the requirement for a quo-rem pair since $n = 0d + n$ and $n \in \{0, 1, \ldots, d-1\}$.

Otherwise, we have $n \geq d$, so that $n - d \in \mathbb{N}$. Thus, we can apply the IH to $n - d$ instead of n (since $n - d < n$). We conclude that there exists a quo-rem pair of $n - d$ and d; let's call this pair (q, r). Thus,

$$n - d = qd + r \qquad \text{and} \qquad q \in \mathbb{Z} \qquad \text{and} \qquad r \in \{0, 1, \ldots, d-1\}.$$

Now,

$$n = \underbrace{n - d}_{=qd+r} + d = qd + r + d = (q+1)\, d + r.$$

So $(q+1, r)$ is a quo-rem pair of n and d. This shows that there is a quo-rem pair of n and d. Thus, the induction step is complete, and the lemma is proved. □

We still have to prove the existence of a quo-rem pair of n and d. And we have to prove this in general, not just for $n \in \mathbb{N}$. So let us do this now.

We use a trick: We derive the negative case from the positive one by a "long jump".

Namely, we assume that $n < 0$ (since the case $n \geq 0$ is already handled by the lemma). Then, the product $\underbrace{(1-d)}_{\substack{\leq 0 \\ (\text{since } d \geq 1)}} \underbrace{n}_{<0}$ is nonnegative. Hence, we can apply the lemma to $(1-d)\,n$ instead of n. We conclude that there is a quo-rem pair (q, r) for $(1-d)\,n$ and n. Thus,

$$(1-d)\,n = qd + r.$$

In other words,

$$n - dn = qd + r,$$

so that

$$n = dn + qd + r = (n+q)\,d + r.$$

This shows that $(n+q, r)$ is a quo-rem pair of n and d. Thus, the quo-rem pair exists, even when n is negative. This completes the proof. □

3.3.3. An application: even and odd integers

Recall that an integer n is **even** or **odd** if it is divisible or not divisible by 2. Now we shall prove the following:

Proposition 3.3.7. Let n be an integer.

(a) The integer n is even if and only if there exists some $k \in \mathbb{Z}$ such that $n = 2k$.

(b) The integer n is odd if and only if there exists some $k \in \mathbb{Z}$ such that $n = 2k+1$.

Proof. **(a)** This is a direct consequence of the definition of divisibility.

(b) The claim is an "if and only if" statement: it combines the two implications

$$(n \text{ is odd}) \Longrightarrow (\text{there exists some } k \in \mathbb{N} \text{ such that } n = 2k+1)$$

and

$$(n \text{ is odd}) \Longleftarrow (\text{there exists some } k \in \mathbb{N} \text{ such that } n = 2k+1)\,.$$

Let us prove these two implications (I will just call them the $\Longrightarrow$ and $\Longleftarrow$ directions) separately:

$\Longrightarrow$: Assume that n is odd. Consider the quo-rem pair (q, r) of n and 2 (we know that it exists, by the preceding theorem). Thus, $n = q \cdot 2 + r$ and $q \in \mathbb{Z}$ and $r \in \{0, 1, \ldots, 2-1\} = \{0, 1\}$. So r is either 0 or 1.

If r was 0, then $n = q \cdot 2 + \underbrace{r}_{=0} = q \cdot 2 = 2q$, which would render n even, not odd. So r cannot be 0. Thus, r must be 1. So $n = \underbrace{q \cdot 2}_{=2q} + \underbrace{r}_{=1} = 2q + 1$. So there exists some $k \in \mathbb{N}$ such that $n = 2k + 1$, namely $k = q$. This proves the $\Longrightarrow$ direction.

$\Longleftarrow$: Assume that there exists some $k \in \mathbb{N}$ such that $n = 2k + 1$. Consider this k. We must prove that n is odd, i.e., not even.

Assume the contrary. Thus, n is even. In other words, $n = 2c$ for some integer c. Consider this c.

So n is both $2k + 1$ and $2c = 2c + 0$. This shows that both pairs $(k, 1)$ and $(c, 0)$ are quo-rem pairs of n and 2. But the preceding theorem says that a quo-rem pair is unique. So these two quo-rem pairs $(k, 1)$ and $(c, 0)$ must be equal. But they differ at least in their second entries, so contradiction! This proves the $\Longleftarrow$ direction. □

Corollary 3.3.8. **(a)** The sum of any two even integers is even.

(b) The sum of any even integer with any odd integer is odd.

(c) The sum of any two odd integers is even.

Proof. **(c)** Let a and b be two odd integers. We must prove that $a + b$ is even.

By the previous proposition, we can write a as $a = 2k + 1$ for some integer k. Likewise, we can write b as $b = 2\ell + 1$ for some integer ℓ. Using these k and ℓ, we find

$$a + b = (2k + 1) + (2\ell + 1) = 2k + 2\ell + 2 = 2(k + \ell + 1),$$

which is clearly even. So part **(c)** is proved.

(a), **(b)** Analogous or even easier. □

This corollary is specific to the number 2. It is not true that the sum of any two integers that are not divisible by 3 must be divisible by 3.

3.3.4. Basic properties of quotients and remainders

Proposition 3.3.9. Let $n \in \mathbb{Z}$, and let d be a positive integer. Then:

(a) We have $n\%d \in \{0, 1, \ldots, d-1\}$ and $n\%d \equiv n \mod d$. (In other words, the remainder $n\%d$ is an integer in $\{0, 1, \ldots, d-1\}$ that is congruent to n modulo d.)

(b) We have $d \mid n$ if and only if $n\%d = 0$.

(c) If $c \in \{0, 1, \ldots, d-1\}$ satisfies $c \equiv n \mod d$, then $c = n\%d$. (In other words, the remainder $n\%d$ is the only integer in $\{0, 1, \ldots, d-1\}$ that is congruent to n modulo d.)

(d) We have $n = (n//d)\,d + (n\%d)$.
(e) If $n \in \mathbb{N}$, then $n//d \in \mathbb{N}$.

Proof. We set $q := n//d$ and $r := n\%d$, so that (q,r) is a quo-rem pair of n and d, and we have $n = qd + r$ and $q \in \mathbb{Z}$ and $r \in \{0, 1, \ldots, d-1\}$. This immediately yields part **(d)** and the first half of **(a)**. Now to the rest:

(a) We have $n\%d = r = n - qd$ (since $n = qd + r$). Hence, $n\%d - n = -qd$ is divisible by d. In other words, $n\%d \equiv n \bmod d$.

(c) Let $c \in \{0, 1, \ldots, d-1\}$ satisfy $c \equiv n \bmod d$. We must prove that $c = n\%d$.

From $c \equiv n \bmod d$, we obtain $d \mid c - n$, which means that $c - n = dx$ for some integer x. Consider this x. Then, solving $c - n = dx$ for n, we find $n = -dx + c = (-x)\,d + c$. This shows that $(-x, c)$ is a quo-rem pair of n and d. But (q,r) is also a quo-rem pair of n and d. Since the quo-rem pair is unique, we thus conclude that $(-x, c) = (q, r)$. In other words, $-x = q$ and $c = r$. In particular, $c = r = n\%d$.

(b) We must prove that $d \mid n$ if and only if $r = 0$.

$\Longrightarrow$: Assume that $d \mid n$. We must prove that $r = 0$. In other words, we must prove that $0 = r$.

We have $0 \in \{0, 1, \ldots, d-1\}$ and $0 \equiv n \bmod d$ (since $d \mid n$ and thus $n \equiv 0 \bmod d$). Thus, part **(c)** of this proposition (applied to $c = 0$) yields $0 = n\%d = r$, just as desired.

$\Longleftarrow$: Assume that $r = 0$. Then, $n = qd + \underbrace{r}_{=0} = qd = dq$, so that $d \mid n$.

So part **(b)** is proved.

(e) See the notes. Alternatively, use the proof of the lemma above. □

Corollary 3.3.10. Let $n \in \mathbb{Z}$. Then:
(a) The integer n is even if and only if $n\%2 = 0$.
(b) The integer n is odd if and only if $n\%2 = 1$.

Proof. **(a)** Easy consequence of part **(b)** of the above proposition.
(b) Contrapositive of part **(a)**, since $n\%2$ is either 0 or 1. □

Quotients and remainders are closely connected to the so-called floor function:

Definition 3.3.11. The **integer part** (aka **floor**) of a real number x is defined to be the largest integer that is $\leq x$. It is denoted by $\lfloor x \rfloor$.

For example,

$$\lfloor 3.8 \rfloor = 3, \qquad \lfloor 4.2 \rfloor = 4, \qquad \lfloor 5 \rfloor = 5, \qquad \left\lfloor \sqrt{2} \right\rfloor = 1,$$
$$\left\lfloor \sqrt{5} \right\rfloor = 2, \qquad \lfloor \pi \rfloor = 3, \qquad \lfloor 0.5 \rfloor = 0, \qquad \lfloor -1.2 \rfloor = -2.$$

Now there is a direct connection to quotients and remainders:

Proposition 3.3.12 (quotient=floor)**.** Let $n \in \mathbb{Z}$, and let d be a positive integer. Then,

$$n//d = \left\lfloor \frac{n}{d} \right\rfloor \qquad \text{and} \qquad n\%d = n - d \cdot \left\lfloor \frac{n}{d} \right\rfloor .$$

Proof. See the notes. □

3.3.5. Base-b representation of nonnegative integers

Division with remainder is the main ingredient in the fact that every integer can be uniquely expressed in decimal notation, or, more generally, in base-b notation for any given integer $b > 1$.

For instance,

$$\begin{aligned} 3401 &= 3 \cdot 1000 + 4 \cdot 100 + 0 \cdot 10 + 1 \cdot 1 \\ &= 3 \cdot 10^3 + 4 \cdot 10^2 + 0 \cdot 10^1 + 1 \cdot 10^0. \end{aligned}$$

This can be done for any nonnegative integer n instead of 3401. This can also be done with any fixed integer $b > 1$ instead of 10, except that the coefficients ("generalized digits") will then be integers between 0 and $b - 1$. This is called the "base-b representation" of n.

For instance, let us find the base-4 representation of 3401: This will be a representation of the form

$$3401 = r_6 4^6 + r_5 4^5 + r_4 4^4 + r_3 4^3 + r_2 4^2 + r_1 4^1 + r_0 4^0,$$

where each r_i is a "base-4 digit" (i.e., an element of $\{0, 1, 2, 3\}$). We are tacitly assuming that these 7 digits $r_0, r_1, \ldots, r_6$ are enough; if not, we can extend the sum to allow higher powers of 4.

How do we find these base-4 digits $r_0, r_1, \ldots, r_6$?

We start by identifying r_0. We rewrite the equation as follows:

$$\begin{aligned} 3401 &= r_6 4^6 + r_5 4^5 + r_4 4^4 + r_3 4^3 + r_2 4^2 + r_1 4^1 + r_0 4^0 \\ &= \left(r_6 4^5 + r_5 4^4 + r_4 4^3 + r_3 4^2 + r_2 4^1 + r_1 4^0 \right) \cdot 4 + r_0. \end{aligned}$$

This equation (combined with $r_0 \in \{0, 1, 2, 3\}$) shows that

$$\begin{aligned} r_6 4^5 + r_5 4^4 + r_4 4^3 + r_3 4^2 + r_2 4^1 + r_1 4^0 &= 3401//4 = 850 \qquad \text{and} \\ r_0 &= 3401\%4 = 1. \end{aligned}$$

So r_0 has been found!

Let us next identify r_1. We have

$$\begin{aligned} 850 &= r_6 4^5 + r_5 4^4 + r_4 4^3 + r_3 4^2 + r_2 4^1 + r_1 4^0 \\ &= \left(r_6 4^4 + r_5 4^3 + r_4 4^2 + r_3 4^1 + r_2 4^0 \right) \cdot 4 + r_1, \end{aligned}$$

so that

$$\begin{aligned} r_6 4^4 + r_5 4^3 + r_4 4^2 + r_3 4^1 + r_2 4^0 &= 850//4 = 212 \qquad \text{and} \\ r_1 &= 850\%4 = 2. \end{aligned}$$

So r_1 has been found.

Let us next identify r_2. We have

$$\begin{aligned} 212 &= r_6 4^4 + r_5 4^3 + r_4 4^2 + r_3 4^1 + r_2 4^0 \\ &= \left(r_6 4^3 + r_5 4^2 + r_4 4^1 + r_3 4^0 \right) \cdot 4 + r_2, \end{aligned}$$

so that

$$\begin{aligned} r_6 4^3 + r_5 4^2 + r_4 4^1 + r_3 4^0 &= 212//4 = 53 \qquad \text{and} \\ r_2 &= 212\%4 = 0. \end{aligned}$$

So r_2 has been found.

Likewise, we find $r_3, r_4, \ldots$:

$$\begin{aligned} 53 &= r_6 4^3 + r_5 4^2 + r_4 4^1 + r_3 4^0 \\ &= \left(r_6 4^2 + r_5 4^1 + r_4 4^0 \right) + r_3, \end{aligned}$$

so

$$\begin{aligned} r_6 4^2 + r_5 4^1 + r_4 4^0 &= 53//4 = 13 \qquad \text{and} \\ r_3 &= 53\%4 = 1. \end{aligned}$$

Next,

$$13 = r_6 4^2 + r_5 4^1 + r_4 4^0 = \left(r_6 4^1 + r_5 4^0 \right) + r_4,$$

so

$$\begin{aligned} r_6 4^1 + r_5 4^0 &= 13//4 = 3 \qquad \text{and} \\ r_4 &= 13\%4 = 1. \end{aligned}$$

Next,

$$3 = r_6 4^1 + r_5 4^0 = \left(r_6 4^0 \right) \cdot 4 + r_5,$$

so

$$\begin{aligned} r_6 4^0 &= 3//4 = 0 \qquad \text{and} \\ r_5 &= 3\%4 = 3. \end{aligned}$$

Of course, this gives $r_6 = 0$, and we are done. So altogether, the base-4 representation of 3401 is

$$3401 = \underbrace{r_6}_{=0} 4^6 + \underbrace{r_5}_{=3} 4^5 + \underbrace{r_4}_{=1} 4^4 + \underbrace{r_3}_{=1} 4^3 + \underbrace{r_2}_{=0} 4^2 + \underbrace{r_1}_{=2} 4^1 + \underbrace{r_0}_{=1} 4^0.$$

In analogy to the decimal system, we can state this as "the number 3401 written in base-4 is 0311021". Commonly you omit the leading zeroes, so this becomes 311021.

The method we just used can be used for any integer $b > 1$ instead of 4: To find the "base-b digits" of a nonnegative integer n, we first divide n by b with remainder, then divide the resulting quotient again by b with remainder, then divide the resulting quotient again by b with remainder, and so on, until we are eventually left with the quotient 0. The remainders obtained in this process will be the base-b digits of n (from right to left).

Let us summarize this as a theorem:

Theorem 3.3.13. Let $b > 1$ be an integer. Let $n \in \mathbb{N}$. Then:

(a) We can write n in the form

$$n = r_k \cdot b^k + r_{k-1} \cdot b^{k-1} + \cdots + r_1 \cdot b^1 + r_0 \cdot b^0$$

with

$$k \in \mathbb{N} \qquad \text{and} \qquad r_0, r_1, \ldots, r_k \in \{0, 1, \ldots, b-1\}.$$

(b) If $n < b^{k+1}$ for some $k \in \mathbb{N}$, then we can write n in the form

$$n = r_k \cdot b^k + r_{k-1} \cdot b^{k-1} + \cdots + r_1 \cdot b^1 + r_0 \cdot b^0$$

with

$$r_0, r_1, \ldots, r_k \in \{0, 1, \ldots, b-1\}.$$

(c) These $r_0, r_1, \ldots, r_k$ are unique (when k is given). Moreover, they can be explicitly computed by the formula

$$r_i = \left(n//b^i\right) \% b \qquad \text{for each } i \in \{0, 1, \ldots, k\}.$$

In particular,

$$\begin{aligned} r_0 &= n \% b, \\ r_1 &= (n//b) \% b, \\ r_2 &= \left(n//b^2\right) \% b, \\ &\ldots. \end{aligned}$$

Proof. See the notes. Essentially, parts **(a)** and **(b)** have already been explained in the above example. As for part **(c)**, we have given a different way of computing r_i, namely,

$$r_i = \underbrace{\left(\left(\left(\left(n//b\right)//b\right)//b\right)\cdots//b\right)}_{\text{take the quotient } i \text{ times}} \%b,$$

so we need to do some work to ensure that the new formula $r_i = \left(n//b^i\right)\%b$ also works (and gives the same result). □

3.3.6. Congruence in terms of remainders

Proposition 3.3.14. Let d be a positive integer. Let a and b be two integers. Then, $a \equiv b \bmod d$ if and only if $a\%d = b\%d$.

Proof. See notes. □

Corollary 3.3.15. Let a and b be two integers. Then, $a \equiv b \bmod 2$ holds if and only if a and b are either both even or both odd.

Proof. See notes. □

3.3.7. The birthday lemma

If you have lived for exactly n days, then you are $n//365$ years and $n\%365$ days old (assuming for simplicity that there are no leapyears). On any "normal" day, the latter number $n\%365$ increases by 1 while the former number $n//365$ stays unchanged. But on a birthday, the former number $n//365$ increases by 1 while the latter number $n\%365$ is reset to 0. This intuitively clear observation is worth stating as a proposition, as it is not special to 365:

Proposition 3.3.16 (birthday lemma)**.** Let $n \in \mathbb{Z}$, and let d be a positive integer. Then:

(a) If $d \mid n$, then

$$n//d = \left(\left(n-1\right)//d\right)+1 \qquad \text{and}$$
$$n\%d = 0 \qquad \text{and} \qquad \left(n-1\right)\%d = d-1.$$

(b) If $d \nmid n$, then

$$n//d = \left(n-1\right)//d \qquad \text{and}$$
$$n\%d = \left(\left(n-1\right)\%d\right)+1.$$

Proof. See notes. □

Proposition 3.3.17. Let $n \in \mathbb{Z}$, and let d be a positive integer. Then:

(a) If $d \mid n$, then
$$\left\lfloor \frac{n}{d} \right\rfloor = \left\lfloor \frac{n-1}{d} \right\rfloor + 1.$$

(b) If $d \nmid n$, then
$$\left\lfloor \frac{n}{d} \right\rfloor = \left\lfloor \frac{n-1}{d} \right\rfloor .$$

Proof. This is just a restatement of the quotient parts of the above proposition. □

3.4. Greatest common divisors

3.4.1. Definition

Definition 3.4.1. Let a and b be two integers.

(a) The **common divisors** of a and b are the integers that divide a and simultaneously divide b.

(b) The **greatest common divisor** of a and b is the largest among the common divisors of a and b, unless $a = b = 0$; in the latter case, we define it to be 0 instead.

We denote the greatest common divisor of a and b by $\gcd(a, b)$, and we even call it the **gcd** of a and b.

For instance:

- What is $\gcd(4, 6)$? It must be 2, since the common divisors of 4 and 6 are 1 and 2 (and their negatives).
- What is $\gcd(0, 5)$? It is 5.
- What is $\gcd(15, 24)$? It is 3.
- What is $\gcd(0, 0)$? It is 0 by definition.

Remark 3.4.2. Let $a, b \in \mathbb{Z}$. Why is $\gcd(a, b)$ well-defined?

For $a = b = 0$, it is because we just defined it to be 0. In all other cases, we argue that there are only finitely many common divisors of a and b, but there is at least one of them. The latter follows from the fact that 1 is always a common divisor of a and b. The former follows from the fact that at least one of a and b is nonzero and thus has only finitely many divisors (in fact, if $a \neq 0$, then each divisor of a has absolute value $\leq |a|$).

3.4.2. Basic properties

First, we collect a bunch of properties of gcds:

Proposition 3.4.3. **(a)** We have $\gcd(a,b) \in \mathbb{N}$ for any $a,b \in \mathbb{Z}$.

(b) We have $\gcd(a,0) = \gcd(0,a) = |a|$ for any $a \in \mathbb{Z}$.

(c) We have $\gcd(a,b) = \gcd(b,a)$ for all $a,b \in \mathbb{Z}$.

(d) If $a,b,c \in \mathbb{Z}$ satisfy $b \equiv c \mod a$, then $\gcd(a,b) = \gcd(a,c)$.

(e) We have $\gcd(a,b) = \gcd(a,\ ua+b)$ for any $a,b,u \in \mathbb{Z}$.

(f) We have $\gcd(a,b) = \gcd(a,\ b\%a)$ for any positive integer a and any $b \in \mathbb{Z}$.

(g) We have $\gcd(a,b) \mid a$ and $\gcd(a,b) \mid b$ for any $a,b \in \mathbb{Z}$.

(h) We have $\gcd(-a,b) = \gcd(a,b)$ and $\gcd(a,-b) = \gcd(a,b)$ for any $a,b \in \mathbb{Z}$.

(i) If $a,b \in \mathbb{Z}$ satisfy $a \mid b$, then $\gcd(a,b) = |a|$.

Proof. Many of these are easy; all proofs can be found in the notes. We will focus on the not completely easy ones – that is, **(d)**, **(e)** and **(f)**:

(d) Let $a,b,c \in \mathbb{Z}$ satisfy $b \equiv c \mod a$. We must show that $\gcd(a,b) = \gcd(a,c)$.

If $a = 0$, then this is clear, since $b \equiv c \mod a$ implies $b = c$ in this case. So we assume that $a \neq 0$ from now on. Then, the greatest common divisors are literally the largest of the common divisors.

So we must prove that the largest of the common divisors of a and b is also the largest of the common divisors of a and c. We will show something even better: The set of common divisors of a and b is the set of common divisors of a and c. For this purpose, we need to prove the following two claims:

> *Claim 1:* Each common divisor of a and b is a common divisor of a and c.
>
> *Claim 2:* Each common divisor of a and c is a common divisor of a and b.

Proof of Claim 1. Let d be a common divisor of a and b. We must prove that d is also a common divisor of a and c. Clearly, it suffices to show that $d \mid c$.

But $d \mid a \mid b - c$ (since $b \equiv c \mod a$). In other words, $b \equiv c \mod d$. So $c \equiv b \equiv 0 \mod d$ (since $d \mid b$). In other words, $d \mid c$. This proves Claim 1. □

Proof of Claim 2. Same argument, with the roles of b and c swapped. □

So we have proved part **(d)**.

(e) Let $a, b, u \in \mathbb{Z}$. We must show that $\gcd(a, b) = \gcd(a,\ ua + b)$.

We have $b \equiv ua + b \mod a$ since $b - (ua + b) = -ua = a(-u)$ is divisible by a. So part **(e)** follows from part **(d)**.

(f) Let a be a positive integer, and let $b \in \mathbb{Z}$. We must prove that $\gcd(a, b) = \gcd(a,\ b\%a)$.

We have $b\%a \equiv b \mod a$ (by properties of remainders). So this also follows from part **(d)**. □

Corollary 3.4.4 (Euclidean recursion for the gcd)**.** Let $a, b \in \mathbb{Z}$ with $b > 0$. Then

$$\gcd(a, b) = \gcd(b,\ a\%b).$$

Proof. Part **(f)** of the above proposition (applied to b and a instead of a and b) says that $\gcd(b, a) = \gcd(b,\ a\%b)$. But $\gcd(b, a) = \gcd(a, b)$. Comparing, we see $\gcd(a, b) = \gcd(b,\ a\%b)$. □

3.4.3. The Euclidean algorithm

Applying this corollary repeatedly, we can compute gcds rather quickly: for example,

$$\begin{aligned}
\gcd(93,\ 18) &= \gcd(18,\ 93\%18) && \text{(by the corollary)} \\
&= \gcd(18,\ 3) \\
&= \gcd(3,\ 18\%3) && \text{(by the corollary)} \\
&= \gcd(3,\ 0) \\
&= |3| = 3
\end{aligned}$$

and

$$\begin{aligned}
\gcd(1145,\ 739) &= \gcd(739,\ 1145\%739) \\
&= \gcd(739,\ 406) \\
&= \gcd(406,\ 739\%406) \\
&= \gcd(406,\ 333) \\
&= \gcd(333,\ 406\%333) \\
&= \gcd(333,\ 73) \\
&= \gcd(73,\ 333\%73) \\
&= \gcd(73,\ 41) \\
&= \gcd(41,\ 73\%41) \\
&= \gcd(41,\ 32) \\
&= \gcd(32,\ 41\%32) \\
&= \gcd(32,\ 9) \\
&= \gcd(9,\ 32\%9) \\
&= \gcd(9,\ 5) \\
&= \gcd(5,\ 9\%5) \\
&= \gcd(5,\ 4) \\
&= \gcd(4,\ 5\%4) \\
&= \gcd(4,\ 1) \\
&= \gcd(1,\ 4\%1) \\
&= \gcd(1,\ 0) = |1| = 1.
\end{aligned}$$

These two computations are instances of a general algorithm for computing $\gcd(a, b)$ for any two numbers $a \in \mathbb{Z}$ and $b \in \mathbb{N}$. This algorithm proceeds as follows:

- If $b = 0$, then the gcd is $|a|$.
- If $b > 0$, then we replace a and b by b and $a\%b$, and recurse (i.e., we apply the method again to b and $a\%b$ instead of a and b).

In Python:

```
def gcd(a, b):  # for b nonnegative
    if b == 0:
        return abs(a)
    return gcd(b, a%b)
```

This algorithm is called the **Euclidean algorithm**. We can tweak it to allow negative b as follows:

```
def gcd(a, b):  # for b arbitrary
    if b < 0:
        return gcd(a, -b)
    if b == 0:
        return abs(a)
    return gcd(b, a%b)
```

Let us convince ourselves that the algorithm really terminates (i.e., does not get stuck in an endless loop):

Proposition 3.4.5. Let $a \in \mathbb{Z}$ and $b \in \mathbb{N}$. Then, the Euclidean algorithm terminates after at most b steps. (A "step" here means a single replacement of (a,b) by $(b,\ a\%b)$.)

Proof. In each step of the Euclidean algorithm, b is replaced by $a\%b$. But $a\%b \leq b-1$ because remainders upon division by b belong to $\{0,1,\ldots,b-1\}$. So the second entry of the pair decreases by at least 1 in each step. This cannot go on for longer than b steps, without decreasing b below 0. So the algorithm cannot run for longer than b steps. □

This proposition overestimates the running time of the algorithm. In truth, the Euclidean algorithm terminates after at most $\log_2(ab)+2$ steps. The proof is outlined in the notes. The main idea is that in each step (except for the first and the last step), the **product** of the two entries decreases by a factor of at least 2: we have $b\,(a\%b) \leq \dfrac{ab}{2}$ whenever $a > b$.

3.4.4. Bezout's theorem and the extended Euclidean algorithm

The Euclidean algorithm can be adapted so that it doesn't just compute $\gcd(a,b)$, but also expresses $\gcd(a,b)$ as an "integer linear combination" of a and b – that is, as a multiple of a plus a multiple of b. This lets us prove the following theorem:

Theorem 3.4.6 (Bezout's theorem for integers)**.** Let a and b be two integers. Then, there exist two integers x and y such that

$$\gcd(a,b) = xa + yb.$$

We will soon prove this. But first, let us get an intuition for this: For $a = 4$ and $b = 7$, this is saying that $\gcd(4,7) = x \cdot 4 + y \cdot 7$ for some integers x and y. Can you find such x and y ? One answer is $(x,y) = (2,-1)$, since $2 \cdot 4 + (-1) \cdot 7 = 1 = \gcd(4,7)$. Another answer is $(x,y) = (-5,3)$, and there are infinitely many more.

Note that this is a variant of the coin problem from the first chapter: Can we pay 1 cent with 4-cent coins and 7-cent coins? But now we are allowing change. So Bezout's theorem solves the coin problem with change, not just for

4-cent and 7-cent coins, but for any pair of denominations. It says that if a and b are two integers, then $\gcd(a,b)$ cents can be paid using a-cent coins and b-cent coins if change is available. What denominations can be paid **without** change is a much trickier question, and we will come back to it.

> **Definition 3.4.7.** Let a and b be two integers. Then, a **Bezout pair** for (a,b) means a pair (x,y) of two integers satisfying $\gcd(a,b) = xa + yb$.

So Bezout's theorem says that for any two integers a and b, there exists a Bezout pair for (a,b).

How will we prove it? We will explain how a Bezout pair for (a,b) can be constructed from a Bezout pair for $(b,\ a\%b)$. Since $a\%b < b$, this is a reduction of our problem to a lesser/lower/simpler case, so it furnishes a recursive algorithm. (Of course, will also need to deal with the case $b = 0$ first, and with the case $b < 0$ at the end.)

So, how do we find a Bezout pair (x,y) for (a,b), given a Bezout pair (u,v) for $(b,\ a\%b)$? We have

$$ub + v\,(a\%b) = \gcd(b,\ a\%b) = \gcd(a,b)$$

(by the corollary in the previous subsection). Writing q for $a//b$, then $a\%b = a - qb$, so we can rewrite the left hand side $ub + v\,(a\%b)$ as $ub + v\,(a - qb) = ub + va - vqb = va + (u - vq)\,b$. So we obtain

$$va + (u - vq)\,b = ub + v\,(a\%b) = \gcd(a,b)\,.$$

In other words, the pair $(v,\ u - vq)$ is a Bezout pair for (a,b).

Next, let us deal with the case $b = 0$. We need to find a Bezout pair for $(a,0)$, that is, a pair (x,y) of integers such that $xa + y0 = \gcd(a,0)$. Since $\gcd(a,0) = |a|$, this can be achieved by choosing x to be either 1 or -1 depending on whether $a \geq 0$ or $a < 0$. The y does not matter. So one answer is $(\operatorname{sign} a,\ 0)$, where $\operatorname{sign} a = \begin{cases} 1, & \text{if } a \geq 0; \\ -1, & \text{if } a < 0. \end{cases}$

This observation allows us to construct an algorithm for finding Bezout pairs, as long as b is nonnegative:

```
def sign(a):
    if a >= 0:
        return 1
    return -1
def bezout_pair(a, b):  # for b nonnegative
    if b == 0:
        return (sign(a), 0)
    (u, v) = bezout_pair(b, a%b)
    q = a//b
    return (v, u - v*q)
```

This is called the **extended Euclidean algorithm**. It can easily be adapted to negative b (you just need to multiply the second entry of the pair by -1 if b is negative).

This algorithm proves Bezout's theorem. (Formally speaking, the proof is an induction proof, and is explained as such in the notes.)

3.4.5. The universal property of the gcd

Bezout's theorem is one of the workhorses of number theory. In particular, it lies beneath the proof of the **universal property of the gcd**:

Theorem 3.4.8 (Universal property of the gcd)**.** Let $a, b, m \in \mathbb{Z}$. Then, we have the equivalence

$$(m \mid a \text{ and } m \mid b) \iff (m \mid \gcd(a, b)).$$

In other words, m is a common divisor of a and b if and only if $m \mid \gcd(a, b)$. In other words, the common divisors of a and b are the divisors of $\gcd(a, b)$. So $\gcd(a, b)$ is not the greatest among the common divisors of a and b, but is actually divisible by them all.

Proof of the theorem. $\Longleftarrow$: If $m \mid \gcd(a, b)$, then $m \mid \gcd(a, b) \mid a$ and $m \mid \gcd(a, b) \mid b$.

$\Longrightarrow$: Assume that $m \mid a$ and $m \mid b$. We must prove that $m \mid \gcd(a, b)$.

Bezout's theorem lets us write $\gcd(a, b)$ as $xa + yb$ for some integers x and y. So we just need to show that $m \mid xa + yb$. But this is clear, since both a and b and therefore both xa and yb are multiples of m. □

3.4.6. Factoring out a common factor from a gcd

Theorem 3.4.9. Let $s, a, b \in \mathbb{Z}$. Then,

$$\gcd(sa, sb) = |s| \cdot \gcd(a, b).$$

Proof. Let $g = \gcd(a, b)$ and $h = \gcd(sa, sb)$. So we must prove that $h = |s| \cdot g$.

- We have $sg \mid sa$ because $g \mid a$. Similarly, $sg \mid sb$. So sg divides both sa and sb. Hence, by the universal property of the gcd, $sg \mid \gcd(sa, sb) = h$. Because divisibility is sign-agnostic, this entails $|s| \cdot g \mid h$.

- Let us now prove that $h \mid |s| \cdot g$.

  Bezout's theorem shows that $\gcd(a,b) = xa + yb$ for two integers x and y. Consider them. So $g = \gcd(a,b) = xa + yb$. Therefore,

  $$\begin{aligned} sg = s(xa + yb) &= x \underbrace{sa}_{=hu \text{ for some } u} + y \underbrace{sb}_{=hv \text{ for some } v} \\ &= xhu + yhv = h \underbrace{(xu + yv)}_{\text{an integer}}. \end{aligned}$$

  Thus, $h \mid sg$. Therefore, $h \mid |s| \cdot g$, since signs don't matter.

So now we have shown that the two integers $|s| \cdot g$ and h divide each other mutually. Thus, they are equal up to sign. But they are both nonnegative, so they are just equal, qed. □

3.5. Coprime integers

3.5.1. Definition and examples

Definition 3.5.1. Two integers a and b are said to be **coprime** (or **relatively prime**) if $\gcd(a,b) = 1$.

This is a symmetric relation: If a and b are coprime, then b and a are coprime.

Example 3.5.2. **(a)** An integer n is coprime to 2 if and only if n is odd. Indeed, $\gcd(n,2)$ is always a divisor of 2 and is nonnegative, so it must be 1 or 2. If n is even, then $\gcd(n,2) = 2$. If n is odd, then $\gcd(n,2) = 1$ (since $2 \nmid n$).

(b) An integer n is coprime to 3 if and only if n is not divisible by 3. (Same reason as for 2.)

(c) An integer n is coprime to 4 if and only if n is odd. Indeed, $\gcd(n,4)$ is always a divisor of 4 and is nonnegative, so it must be 1, 2 or 4. If n is even, then this gcd is at least 2, because 2 is a common divisor of n and 4. If n is odd, then this gcd must be 1, since neither 2 nor 4 divides n.

(d) An integer n is coprime to 5 if and only if n is not divisible by 5. (Same reason as for 2 and 3.)

(e) An integer n is coprime to 6 if and only if n is odd and not divisible by 3. Indeed, if n is even, then 2 is a common divisor of n and 6 and thus $\gcd(n,6) \geq 2$. If n is divisible by 3, then 3 is a common divisor of n and 6 and thus $\gcd(n,6) \geq 3$. But if n is neither even nor divisible by 3, then $\gcd(n,6)$ can be neither 2 nor 3 nor 6 and so must be 1. (The positive divisors of 6 are $1, 2, 3, 6$.)

(f) An integer n is always coprime to 1, since the only nonnegative divisor of 1 is 1.

3.5.2. Three theorems about coprimality

Theorem 3.5.3 (coprime divisors theorem)**.** Let $a, b, c \in \mathbb{Z}$ satisfy $a \mid c$ and $b \mid c$. Assume that a and b are coprime. Then, $ab \mid c$.

In words: A product of two coprime divisors of c is again a divisor of c.

Proof. We know that ab is a common divisor of ca and cb (in fact, $ab \mid ca$ because $b \mid c$, whereas $ab \mid cb$ because $a \mid c$). By the universal property of the gcd, this entails

$$ab \mid \gcd(ca, cb) = |c| \cdot \underbrace{\gcd(a, b)}_{\substack{=1 \\ \text{(since } a \text{ and } b \text{ are coprime)}}} = |c| \mid c.$$

□

For example:

- $3 \mid 24$ and $8 \mid 24$, and the divisors 3 and 8 are coprime. So the theorem yields $3 \cdot 8 \mid 24$.
- $6 \mid 12$ and $3 \mid 12$ but $6 \cdot 3 \nmid 12$, which is OK because 6 and 3 are not coprime.

Theorem 3.5.4 (coprime removal theorem)**.** Let $a, b, c \in \mathbb{Z}$ be such that $a \mid bc$. Assume that a is coprime to b. Then, $a \mid c$.

Proof. We have $a \mid bc$ and $b \mid bc$. So $ab \mid bc$ by the preceding theorem. Now, we cancel b to obtain $a \mid c$.

(Careful: Cancelling b only works if $b \neq 0$. But the $b = 0$ case is easy: If a is coprime to 0, then $a = \pm 1$, since $\gcd(a, 0) = |a|$.)

In the notes I give a different proof. □

For example:

- We have $6 \mid 7 \cdot 12$, but 6 is coprime to 7. So the theorem yields $6 \mid 12$.

Theorem 3.5.5 (coprime product theorem)**.** Let $a, b, c \in \mathbb{Z}$. Assume that each of the numbers a and b is coprime to c. Then, ab is coprime to c.

Proof. Let $g = \gcd(ab, c)$. We must prove that $g = 1$.

We have $g \mid ab$ and $g \mid c \mid ac$. Thus, g is a common divisor of ab and ac, so the universal property of the gcd yields

$$g \mid \gcd(ab, ac) = |a| \cdot \underbrace{\gcd(b, c)}_{=1} = |a| \mid a.$$

Combining this with $g \mid c$, we obtain that g is a common divisor of a and c. Hence, by the universal property of the gcd, we see that $g \mid \gcd(a, c) = 1$. Thus, $g = 1$ (because g is nonnegative), qed. □

3.5.3. Reducing a fraction

Theorem 3.5.6. Let a and b be two integers that are not both 0. Let $g = \gcd(a, b)$. Then, the integers $\frac{a}{g}$ and $\frac{b}{g}$ are coprime.

Proof. We have $g > 0$ and $g \mid a$ and $g \mid b$. Hence, the quotients $\frac{a}{g}$ and $\frac{b}{g}$ are integers.

Now, recall that $\gcd(sa, sb) = |s| \cdot \gcd(a, b)$. As a consequence,

$$\gcd\left(g \cdot \frac{a}{g}, g \cdot \frac{b}{g}\right) = |g| \cdot \gcd\left(\frac{a}{g}, \frac{b}{g}\right).$$

Thus,

$$\gcd\left(\frac{a}{g}, \frac{b}{g}\right) = \gcd\left(\underbrace{g \cdot \frac{a}{g}}_{=a}, \underbrace{g \cdot \frac{b}{g}}_{=b}\right) / \underbrace{|g|}_{=g} = \gcd(a, b) / g = 1$$

(since $g = \gcd(a, b)$). This means that $\frac{a}{g}$ and $\frac{b}{g}$ are coprime. □

The importance of this theorem lies in the consequence that if you have a fraction $\frac{a}{b}$ of two integers, then cancelling $\gcd(a, b)$ from both a and b yields a **reduced fraction** (i.e., a fraction whose numerator and denominator are coprime).

3.6. Prime numbers

3.6.1. Definition

Definition 3.6.1. An integer $n > 1$ is said to be **prime** (or **a prime**) if the only positive divisors of n are 1 and n.

The first few primes (= prime numbers) are

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, \ldots.$$

There are infinitely many of them (see HW4).

3.6.2. The friend-or-foe lemma

The first property of primes that we will show is a crucial result I call the **friend-or-foe lemma**:

Lemma 3.6.2 (friend-or-foe lemma)**.** Let p be a prime. Let $n \in \mathbb{Z}$. Then, n is either divisible by p or coprime to p, but not both.

Proof. The number p is prime, so its only positive divisors are 1 and p. Thus, $\gcd(n,p)$ must be either 1 or p (since $\gcd(n,p)$ is a positive divisor of n and p).

If it is 1, then n is coprime to p (by the definition of "coprime").

If it is p, then n is divisible by p (since $p = \gcd(n,p) \mid n$).

So it is always one or the other. It cannot be both, because if n is divisible by p, then p is a common divisor of n and p, and thus $\gcd(n,p) \geq n > 1$. □

3.6.3. Binomial coefficients and primes

A look at Pascal's triangle might reveal a pattern: If p is a prime, then all the numbers in the p-th row of the triangle – except for the two 1's at both ends – are divisible by p. This is indeed true:

Theorem 3.6.3. Let p be a prime. Let $k \in \{1, 2, \ldots, p-1\}$. Then, $p \mid \dbinom{p}{k}$.

Proof. HW3 Exercise 5 **(a)** says that

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

Thus, $p \mid k\dbinom{p}{k}$. If we can show that p is coprime to k, then we can use the coprime removal theorem to remove the k from this divisibility, and conclude the desired result $p \mid \dbinom{p}{k}$.

So why is p coprime to k ? Well, the friend-or-foe lemma yields that k is either divisible by p or coprime to p. But k cannot be divisible by p, since $k \in \{1, 2, \ldots, p-1\}$. Therefore, k must be coprime to p. In other words, p is coprime to k. And we are done. □

3.6.4. Fermat's little theorem

It is easy to see that every integer a satisfies $a^2 \equiv a \bmod 2$. Indeed, the difference $a^2 - a = (a-1)\,a$ is divisible by 2 because at least one of the factors $a-1$ and a is divisible by 2.

Likewise, every integer a satisfies $a^3 \equiv a \bmod 3$. Indeed, the difference $a^3 - a = (a-1)\,a\,(a+1)$ is divisible by 3.

This pattern does not persist for 4: Indeed, $a^4 \equiv a \bmod 4$ fails for $a = 2$.

But the pattern reemerges for 5: Every integer a satisfies $a^5 \equiv a \bmod 5$. This is harder to show than the previous claims about 2 and 3, but there is still a straightforward computation that proves it (check all possible values of $a\%5$).

The pattern disappears for 6 again ($a^6 \not\equiv a \bmod 6$ for $a = 2$), but it reappears for 7.

It turns out that this is no coincidence:

> **Theorem 3.6.4** (Fermat's Little Theorem)**.** Let p be a prime. Let $a \in \mathbb{Z}$. Then,
>
> $$a^p \equiv a \bmod p.$$

Proof. We induct on a. This will only cover the case $a \geq 0$, so we will have to handle the case $a < 0$ by a separate argument later.

Base case: For $a = 0$, we have $a^p \equiv a \bmod p$ because both sides are 0.

Induction step: Let $a \in \mathbb{N}$. Assume (as IH) that $a^p \equiv a \bmod p$. We must prove that $(a+1)^p \equiv a + 1 \bmod p$.

The binomial formula yields

$$\begin{aligned}
(a+1)^p &= \sum_{k=0}^{p} \binom{p}{k} a^k \underbrace{1^{p-k}}_{=1} = \sum_{k=0}^{p} \binom{p}{k} a^k \\
&= \underbrace{\binom{p}{0}}_{=1} \underbrace{a^0}_{=1} + \sum_{k=1}^{p-1} \underbrace{\binom{p}{k}}_{\substack{\text{is divisible by } p \\ \text{(by the preceding} \\ \text{theorem)}}} a^k + \underbrace{\binom{p}{p}}_{=1} a^p \\
&= 1 + \underbrace{(\text{some multiple of } p)}_{\equiv 0 \bmod p} + \underbrace{a^p}_{\equiv a \bmod p} \\
&\equiv 1 + a = a + 1 \bmod p.
\end{aligned}$$

So the induction step is complete.

Thus, Fermat's Little Theorem is proved in the case when $a \geq 0$. It remains to handle the case $a < 0$.

We reduce this case to the case $a \geq 0$ by replacing the negative integer a by a "nonnegative proxy" $b = a\%p$. Indeed, $b = a\%p \equiv a \bmod p$, so $b^p \equiv a^p \bmod p$. But because $b \geq 0$, the already-handled nonnegative case of Fermat's Little Theorem shows that $b^p \equiv b \bmod p$. Thus, $a^p \equiv b^p \equiv b \equiv a \bmod p$.

Now the theorem is proved for all a. □

> **Remark 3.6.5.** Note that Fermat's Little Theorem holds not only when p is prime, but also when p is 1 or 561 or 1105 or 1729 or 2465 or various other such "lucky" numbers (they are called Carmichael numbers).

3.6.5. Prime divisor separation theorem

You can think of the primes as "inseparable" positive integers: They cannot be written as products of two smaller positive integers.

One useful consequence of this "inseparability" is that if a prime p divides a product ab of two integers, then it must divide one of the two factors a and b, because it cannot be "separated" into a part that divides a and a part that divides b. This is merely a heuristic argument, but the result is true:

Theorem 3.6.6 (prime divisor separation theorem)**.** Let p be a prime. Let $a, b \in \mathbb{Z}$ be such that $p \mid ab$. Then, $p \mid a$ or $p \mid b$.

Proof. We shall prove the following equivalent statement: "If $p \nmid a$, then $p \mid b$".

Assume that $p \nmid a$. So we must prove $p \mid b$.

The friend-or-foe lemma tells us that a is either divisible by p or coprime to p. Since $p \nmid a$, we thus conclude that a is coprime to p. Thus, by the coprime removal theorem, we can remove the a from $p \mid ab$, obtaining $p \mid b$, as desired. □

The theorem would fail if p was a non-prime like 4. For instance, $4 \mid 6 \cdot 2$ but $4 \nmid 6$ and $4 \nmid 2$.

We can extend the theorem to several factors:

Corollary 3.6.7. Let p be a prime. Let $a_1, a_2, \ldots, a_k \in \mathbb{Z}$ such that $p \mid a_1 a_2 \cdots a_k$. Then, $p \mid a_i$ for some $i \in \{1, 2, \ldots, k\}$.

Proof. Induct on k. □

3.6.6. p-valuations: definition

Lemma 3.6.8. Let p be a prime. Let n be a nonzero integer. Then, there exists a largest $m \in \mathbb{N}$ such that $p^m \mid n$.

For example, 120 is divisible by 2^3 but not by any higher power of 2.

Proof. For high enough m, we have $p^m > |n|$, which precludes $p^m \mid n$ from happening.

Of course, $p^0 \mid n$ since $p^0 = 1$. So the set of all $m \in \mathbb{N}$ such that $p^m \mid n$ is nonempty but finite. Thus, it has a maximum element. □

Definition 3.6.9. Let p be a prime.

(a) Let n be a nonzero integer. Then, $v_p(n)$ shall denote the largest $m \in \mathbb{N}$ such that $p^m \mid n$. This number $v_p(n)$ is called the ***p*-valuation** or the ***p*-adic valuation** of n. In other words, $v_p(n)$ is the number of times you can divide n by p before you get a nonzero remainder.

(b) In order to have $v_p(n)$ defined not just for nonzero n but also for $n = 0$ (thus for all integers n), we also define $v_p(0)$ to be ∞. This symbol ∞ is not an actual number, but it can be made to behave like a number at least in some ways. In particular, we can involve it in addition and comparison, using the rules

$$\begin{aligned} k+\infty=\infty+k &= \infty \qquad \text{for each } k \in \mathbb{Z}; \\ \infty+\infty &= \infty; \\ k<\infty \text{ and } \infty>k & \qquad \text{for each } k \in \mathbb{Z}. \end{aligned}$$

So ∞ is like "a mythical number larger than any actual number". We cannot subtract ∞.

Examples:

$$\begin{aligned} v_3(99) &= 2 \qquad \left(\text{since } 3^2 \mid 99 \text{ but } 3^3 \nmid 99\right); \\ v_3(98) &= 0 \qquad \left(\text{since } 3^0 \mid 98 \text{ but } 3^1 \nmid 98\right); \\ v_3(96) &= 1 \qquad \left(\text{since } 3^1 \mid 96 \text{ but } 3^2 \nmid 96\right); \\ v_3(0) &= \infty. \end{aligned}$$

Note that the definition above can be generalized to all $p>1$ (not just primes). But the primeness of p causes some properties to hold that would otherwise be false.

3.7. p-valuations: basic properties

Lemma 3.7.1. Let p be a prime. Let $i \in \mathbb{N}$ and $n \in \mathbb{Z}$. Then, $p^i \mid n$ if and only if $v_p(n) \geq i$.

Proof. If $n=0$, then this is clear (since $p^i \mid 0$ always holds, and $v_p(0)=\infty \geq i$ always holds as well).

So it remains to deal with the case $n \neq 0$. In this case, $v_p(n)$ is defined as the largest $m \in \mathbb{N}$ such that $p^m \mid n$. Thus, if $i \leq v_p(n)$, then $p^i \mid p^{v_p(n)} \mid n$. Conversely, if $i > v_p(n)$, then $p^i \nmid n$, since $v_p(n)$ was the **largest** $m \in \mathbb{N}$ such that $p^m \mid n$. Thus, we conclude that $i \leq v_p(n)$ if and only if $p^i \mid n$. This is just the claim of the lemma. □

Recall some standard notations: For any two numbers x and y, we let $\min\{x,y\}$ denote the smaller of these two numbers, and we let $\max\{x,y\}$ denote the larger of them. More generally, if S is a set of numbers, then $\min S$ means the smallest element of S (if it exists) and $\max S$ means the largest element of S (if it exists).

Theorem 3.7.2 (properties of p-valuations)**.** Let p be a prime. Then:

(a) We have $v_p(ab) = v_p(a) + v_p(b)$ for any $a, b \in \mathbb{Z}$.

(b) We have $v_p(a+b) \geq \min\{v_p(a), v_p(b)\}$ for any $a, b \in \mathbb{Z}$.

(c) We have $v_p(1) = 0$.

(d) We have $v_p(p) = 1$.

(e) For any prime $q \neq p$, we have $v_p(q) = 0$.

Proof. Parts **(c)**–**(e)** are easy. Part **(b)** is also pretty easy: If $k = \min\{v_p(a), v_p(b)\}$, then both a and b are multiples of p^k, whence $a+b$ is also a multiple of p^k, but this means that $v_p(a+b) \geq k$.

Remains to prove **(a)**. Fix $a, b \in \mathbb{Z}$. We must prove that $v_p(ab) = v_p(a) + v_p(b)$.

If a or b is 0, then this is obvious from the fact that $\infty = \infty + k = k + \infty$.

So let us assume that neither a nor b is zero (whence $ab \neq 0$ as well). Let

$$n = v_p(a) \qquad \text{and} \qquad m = v_p(b).$$

We must show that $v_p(ab) = n + m$.

Since $n = v_p(a)$, we have $p^n \mid a$, so that $a = p^n x$ for some integer x. Consider this x. If we have $p \mid x$, then we would have $p^{n+1} \mid a$ (since $p \mid x$ would yield $p^n p \mid p^n x = a$, thus $p^{n+1} \mid a$), which would contradict the fact that $n = v_p(a)$. So x is not divisible by p. By the friend-or-foe lemma, this entails that x is coprime to p.

So we have written a as $a = p^n x$ where x is an integer coprime to p.

Likewise, we can write b as $b = p^m y$ where y is an integer coprime to p.

Multiplying these two equalities, we find

$$ab = (p^n x)(p^m y) = p^{n+m}(xy).$$

This immediately shows that $p^{n+m} \mid ab$. Moreover, if we had $p^{n+m+1} \mid ab$, then we would have $p^{n+m+1} \mid ab = p^{n+m}(xy)$ and thus (by cancelling p^{n+m}) we would get $p \mid xy$; but this would contradict the fact that xy is not divisible by p (which is because x and y are coprime to p, so that their product xy is also coprime to p (by the coprime product theorem), hence $p \nmid xy$ (by friend-or-foe)). So we cannot have $p^{n+m+1} \mid ab$.

Therefore, $v_p(ab) = n + m$, completing our proof. □

Corollary 3.7.3. Let p be a prime. Then, for any k integers $a_1, a_2, \ldots, a_k$, we have

$$v_p(a_1 a_2 \cdots a_k) = v_p(a_1) + v_p(a_2) + \cdots + v_p(a_k).$$

Proof. Induct on k. The base case ($k = 0$) uses $v_p(1) = 0$. The induction step uses part **(a)** of the theorem. □

3.7.1. Back to Hanoi

Let us write down the 2-valuations of the positive integers:

n	1	2	3	4	5	6	7	8	9	10
$v_2(n)$	0	1	0	2	0	1	0	3	0	1

.

This sequence (going on forever) is called the **ruler sequence**. It appears in rather unexpected places:

> **Proposition 3.7.4.** Let $n \in \mathbb{N}$. Recall the strategy we used for solving the Tower of Hanoi puzzle with n disks. Let $k \in \{1, 2, \ldots, 2^n - 1\}$. Then, the k-th move of the strategy is moving the $(v_2(k) + 1)$-st smallest disk.

Proof. See the notes. □

> **Remark 3.7.5.** A "Tower of Hanoi" backup scheme is a backup scheme for data using several backup drives. Every odd day, you back up to the first drive. Every even day that is not divisible by 4, you back up to the second drive. Every day divisible by 4 but not by 8, you back up to the third drive. And so on. (On the k-th day, you back up to the $(v_2(k) + 1)$-st drive.)

3.7.2. The p-valuation of $n!$

> **Theorem 3.7.6** (de Polignac's formula)**.** Let p be a prime. Let $n \in \mathbb{N}$. Then,
>
> $$\begin{aligned} v_p(n!) &= \left\lfloor \frac{n}{p^1} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \left\lfloor \frac{n}{p^3} \right\rfloor + \cdots \\ &= \left(n//p^1\right) + \left(n//p^2\right) + \left(n//p^3\right) + \cdots . \end{aligned}$$

The proof can be found in the notes (§3.6.11).

As an example, for $p = 2$ and $n = 13$, the formula is saying that

$$\begin{aligned} v_2(13!) &= \left\lfloor \frac{13}{2^1} \right\rfloor + \left\lfloor \frac{13}{2^2} \right\rfloor + \left\lfloor \frac{13}{2^3} \right\rfloor + \cdots \\ &= \lfloor 6.5 \rfloor + \lfloor 3.25 \rfloor + \lfloor 1.625 \rfloor + \lfloor 0.8125 \rfloor + \lfloor 0.40625 \rfloor + \cdots \\ &= 6 + 3 + 1 + \underbrace{0 + 0 + 0 + \cdots}_{\text{just zeroes here}} \\ &= 6 + 3 + 1 = 10. \end{aligned}$$

And indeed, $13! = 2^{10} 3^5 5^2 7 \cdot 11 \cdot 13$.

3.7.3. Prime factorization

We are now ready to prove one of the most important properties of primes: the fact that every positive integer can be uniquely decomposed into a product of some primes. For instance,

$$200 = 2 \cdot 2 \cdot 2 \cdot 5 \cdot 5.$$

The word "uniquely" means "uniquely up to order"; i.e., it is saying that any two ways of decomposing a given positive integer n into a product of primes differ only in the order of their factors.

Let us introduce a name for this:

> **Definition 3.7.7.** Let n be a positive integer. A **prime factorization** of n means a finite list $(p_1, p_2, \ldots, p_k)$ of primes such that
>
> $$n = p_1 p_2 \cdots p_k.$$

Thus, $(2, 2, 2, 5, 5)$ is a prime factorization 200; $(2, 5, 2, 5, 2)$ is another. There are more, but each of them contains the number 2 three times and the number 5 twice (and no other numbers).

Let us state this as a general claim:

> **Theorem 3.7.8** (Fundamental Theorem of Arithmetic)**.** Let n be a positive integer. Then:
>
> **(a)** There exists a prime factorization of n.
>
> **(b)** This prime factorization is unique up to reordering its entries. In other words, if $(p_1, p_2, \ldots, p_k)$ and $(q_1, q_2, \ldots, q_\ell)$ are two prime factorizations of n, then $(p_1, p_2, \ldots, p_k)$ can be obtained from $(q_1, q_2, \ldots, q_\ell)$ by reordering the entries.
>
> **(c)** Let $(p_1, p_2, \ldots, p_k)$ be a prime factorization of n. Let p be any prime. Then, the number of times that p appears in the list $(p_1, p_2, \ldots, p_k)$ (in other words, the number of $i \in \{1, 2, \ldots, k\}$ such that $p_i = p$) is $v_p(n)$.

Proof. **(a)** This was shown as an example of a strong induction proof.

(c) By the definition of a prime factorization, we have $n = p_1 p_2 \cdots p_k$. Hence,

$$\begin{aligned} v_p(n) &= v_p(p_1 p_2 \cdots p_k) \\ &= v_p(p_1) + v_p(p_2) + \cdots + v_p(p_k) \end{aligned}$$

by the last corollary we proved. But each addend $v_p(p_i)$ is 0 or 1; namely, it is 1 if $p_i = p$ and 0 otherwise. So the sum on the RHS is a sum of k numbers, each of which is a 0 or a 1; therefore it equals the number of 1's among its addends.

But this number is the number of $i \in \{1, 2, \ldots, k\}$ that satisfy $p_i = p$. So we conclude that

$$\begin{aligned} v_p(n) &= (\text{number of } i \in \{1, 2, \ldots, k\} \text{ that satisfy } p_i = p) \\ &= (\text{number of times } p \text{ appears in } (p_1, p_2, \ldots, p_k)). \end{aligned}$$

This proves **(c)**.

(b) This is an easy consequence of **(c)**. In fact, let $(p_1, p_2, \ldots, p_k)$ be $(q_1, q_2, \ldots, q_\ell)$ are two prime factorizations of n. Then, each prime p appears equally often in $(p_1, p_2, \ldots, p_k)$ as it appears in $(q_1, q_2, \ldots, q_\ell)$, since part **(c)** tells us that it appears $v_p(n)$ times in each prime factorization of n. So the two prime factorizations contain the same primes the same number of times, and can only differ in the order in which these primes appear, e.g., $(2, 2, 2, 5, 5)$ vs. $(2, 5, 2, 5, 2)$. □

Note that actually finding a prime factorization of n is not very easy computationally; there are no really fast algorithms like the Euclidean algorithm for gcds.

More examples:

$$\begin{aligned} 2025 &= 5 \cdot 405 = 5 \cdot 5 \cdot 81 = 5 \cdot 5 \cdot 3 \cdot 3 \cdot 3 \cdot 3; \\ 2024 &= 2 \cdot 2 \cdot 2 \cdot 11 \cdot 23. \end{aligned}$$

3.8. Least common multiples

We previously studied greatest common divisors (gcds). Now we shall talk about least common multiples (lcms). These are a kind of counterpart to gcds.

Definition 3.8.1. Let a and b be two integers.

(a) The **common multiples** of a and b are the integers that are divisible by both a and b.

(b) The **least common multiple** (aka the **lowest common multiple**) of a and b is defined as follows:

- If a and b are nonzero, then it is the smallest positive common multiple of a and b.
- Otherwise, it is 0.

It is denoted by $\operatorname{lcm}(a, b)$.

For example,

$$\begin{aligned} &\operatorname{lcm}(3,4) = 12, \qquad \operatorname{lcm}(6,4) = 12, \qquad \operatorname{lcm}(6,8) = 24, \\ &\operatorname{lcm}(2,4) = 4, \qquad \operatorname{lcm}(0,5) = 0, \qquad \operatorname{lcm}(-2,3) = 6. \end{aligned}$$

Note that the lcm of two positive integers is a well-known concept: When you bring two fractions (of integers) to their lowest common denominator, this lowest common denominator is actually the lcm of their individual denominators.

Here are some facts about lcms:

Theorem 3.8.2. Let a and b be two integers. Then:

(a) The lcm of a and b exists.

(b) We have $\operatorname{lcm}(a,b) \in \mathbb{N}$.

(c) We have $\operatorname{lcm}(a,b) = \operatorname{lcm}(b,a)$.

(d) We have $a \mid \operatorname{lcm}(a,b)$ and $b \mid \operatorname{lcm}(a,b)$.

(e) We have $\operatorname{lcm}(-a,b) = \operatorname{lcm}(a,b)$ and $\operatorname{lcm}(a,-b) = \operatorname{lcm}(a,b)$.

Proof. All of these follow easily from the definitions. For part **(a)**, observe that two nonzero integers a and b have at least one positive common multiple, for instance $|ab|$. □

The universal property of the gcd has a counterpart for the lcm:

Theorem 3.8.3 (universal property of the lcm)**.** Let $a, b, m \in \mathbb{Z}$. Then, we have the equivalence

$$(a \mid m \text{ and } b \mid m) \iff (\operatorname{lcm}(a,b) \mid m).$$

Proof. $\Longleftarrow$: If $\operatorname{lcm}(a,b) \mid m$, then $a \mid \operatorname{lcm}(a,b) \mid m$ and $b \mid \operatorname{lcm}(a,b) \mid m$.

$\Longrightarrow$: Assume that $a \mid m$ and $b \mid m$. We must prove that $\operatorname{lcm}(a,b) \mid m$.

If one of a and b is 0, then this is easy. Hence, we only need to consider the case when a and b are nonzero.

In this case, let $\ell = \operatorname{lcm}(a,b)$. Recall that ℓ is defined as the smallest positive common multiple of a and b. Hence, ℓ is a positive integer and is a multiple of both a and b. Let $q = m // \ell$ and $r = m\%\ell$. Then,

$$q \in \mathbb{Z} \qquad \text{and} \qquad r \in \{0, 1, \ldots, \ell - 1\} \qquad \text{and} \qquad m = q\ell + r.$$

Thus, $r < \ell$.

Now, I claim that r is a multiple of a. Indeed, from $m = q\ell + r$, we obtain $r = \underbrace{m}_{\text{a multiple of } a} - q \underbrace{\ell}_{\text{a multiple of } a}$, which shows that r is a multiple of a. Similarly, r is a multiple of b. Hence, r is a common multiple of a and b. Since $r < \ell$, this entails that r cannot be positive, since ℓ is the smallest positive common multiple of a and b.

So $r \in \{0, 1, \ldots, \ell - 1\}$ but r cannot be positive. Thus, $r = 0$. Consequently, $m\%\ell = r = 0$, and thus $\ell \mid m$. In other words, $\operatorname{lcm}(a,b) \mid m$. □

Theorem 3.8.4. Let a and b be two integers. Then,

$$\gcd(a,b) \cdot \operatorname{lcm}(a,b) = |ab| .$$

Proof. See the notes. (Mostly an application of universal properties.) □

Both gcd and lcm have easily computable p-valuations:

Theorem 3.8.5. Let p be a prime. Let a and b be two integers. Then,

$$v_p(\gcd(a,b)) = \min\{v_p(a),\ v_p(b)\} \qquad \text{and}$$
$$v_p(\operatorname{lcm}(a,b)) = \max\{v_p(a),\ v_p(b)\} .$$

Proof. Not hard using universal properties. □

For example, knowing that $18 = 2 \cdot 3^2$ and $12 = 2^2 \cdot 3$, we obtain

$$\gcd(18,12) = 2 \cdot 3 = 6 \qquad \text{and}$$
$$\operatorname{lcm}(18,12) = 2^2 \cdot 3^2 = 36.$$

Gcds and lcms can also be defined for multiple numbers (not just for two numbers). Their properties are mostly analogous to the two-number case, with some exceptions ($\gcd(a,b) \cdot \operatorname{lcm}(a,b) = |ab|$ generalizes not to $\gcd(a,b,c) \cdot \operatorname{lcm}(a,b,c) = |abc|$ but rather to $\gcd(a,b,c) \cdot \operatorname{lcm}(bc,ca,ab) = |abc|$).

3.9. Sylvester's $xa + yb$ theorem (or the Chicken McNugget theorem)

We now return to the problem of paying with a-cent coins and b-cent coins.

For this entire section, we let a and b be two positive integers.

Definition 3.9.1.

Definition 3.9.2. (a) A **$\mathbb{Z}$-linear combination** (short: **$\mathbb{Z}$-LC**) of a and b means a number of the form

$$xa + yb \qquad \text{for some } x, y \in \mathbb{Z}.$$

In other words, it means a number of cents that you can pay with a-cent coins and b-cent coins if you can get change.

(b) A **$\mathbb{N}$-linear combination** (short: **$\mathbb{N}$-LC**) of a and b means a number of the form

$$xa + yb \qquad \text{for some } x, y \in \mathbb{N}.$$

In other words, it means a number of cents that you can pay with a-cent coins and b-cent coins without getting change.

For example, a proposition we proved long ago says that any integer $n \geq 8$ is an $\mathbb{N}$-LC of 3 and 5. Moreover, the $\mathbb{N}$-LCs of 3 and 5 are

$$0,\ 3,\ 5,\ 6,\ \underbrace{8,\ 9,\ 10,\ 11,\ 12,\ \ldots}_{\text{all integers } n\geq 8}.$$

What do the $\mathbb{N}$-LCs of a and b look like in general?

First, let us describe the $\mathbb{Z}$-LCs of a and b:

Proposition 3.9.3. The $\mathbb{Z}$-LCs of a and b are just the multiples of $\gcd(a,b)$.

Proof. Any $\mathbb{Z}$-LC of a and b is a multiple of $\gcd(a,b)$, since it has the form $xa+yb$, and you can factor out $\gcd(a,b)$ from both a and b.

Conversely, any multiple of $\gcd(a,b)$ is a $\mathbb{Z}$-LC of a and b, because Bezout's theorem allows us to write $\gcd(a,b)$ as such a $\mathbb{Z}$-LC, and then we can get any multiple as well ($\gcd(a,b) = xa+yb \Longrightarrow 5\gcd(a,b) = 5xa+5yb$). □

Now what about the $\mathbb{N}$-LCs? Obviously, each $\mathbb{N}$-LC of a and b is a $\mathbb{Z}$-LC of a and b, so it is a multiple of $\gcd(a,b)$. But not every multiple of $\gcd(a,b)$ is an $\mathbb{N}$-LC. The question is: which ones are?

We can replace the numbers a and b by $\dfrac{a}{\gcd(a,b)}$ and $\dfrac{b}{\gcd(a,b)}$. Thus, we can WLOG assume that a and b are coprime. So we need to understand the $\mathbb{N}$-LCs of two coprime numbers a and b. For example, the $\mathbb{N}$-LCs of 3 and 5 are

$$0,\ 3,\ 5,\ 6,\ \underbrace{8,\ 9,\ 10,\ 11,\ 12,\ \ldots}_{\text{all integers } n\geq 8}.$$

For a more complicated example, the $\mathbb{N}$-LCs of 5 and 9 are

$$0,\ 5,\ 9,\ 10,\ 14,\ 15,\ 18,\ 19,\ 20,\ 23,\ 24,\ 25,\ 27,\ 28,\ 29,\ 30,\ \underbrace{32,\ 33,\ 34,\ \ldots}_{\text{all integers } \geq 32}.$$

There is no obvious pattern here, but there are some things you might notice: All integers $n \geq 32$ are $\mathbb{N}$-LCs of 5 and 9. Of the remaining 32 nonnegative integers $0,1,\ldots,31$, exactly half (so 16) are $\mathbb{N}$-LCs of 5 and 9. This generalizes:

Theorem 3.9.4 (Sylvester's two-coin theorem, or Chicken McNugget theorem)**.** Assume that the two positive integers a and b are coprime. Then:

(a) Every integer $n > ab-a-b$ is an $\mathbb{N}$-LC of a and b.

(b) The number $ab-a-b$ is **not** an $\mathbb{N}$-LC of a and b.

(c) Among the first $(a-1)(b-1)$ nonnegative integers $0,1,\ldots,ab-a-b$, exactly half are $\mathbb{N}$-LCs of a and b.

(d) Let $n \in \mathbb{Z}$. Then, exactly one of the two numbers n and $ab-a-b-n$ is an $\mathbb{N}$-LC of a and b.

Proof. See the notes (§3.8). Part **(b)** is the best place to start; then part **(d)**; then parts **(c)** and **(a)**. □

4. An informal introduction to enumeration

Enumeration is a fancy word for counting – i.e., answering questions of the form "how many ... are there?". Here are some examples:

- How many ways are there to choose 3 odd integers between 0 and 20, if the order matters (i.e., we count the choice $1, 3, 5$ as different from the choice $5, 1, 3$)? (The answer is 1000.)

- How many ways are there to choose 3 odd integers between 0 and 20, if the order does not matter? (The answer is 220.)

- How many ways are there to choose 3 distinct odd integers between 0 and 20, if the order matters? (The answer is 720.)

- How many ways are there to choose 3 distinct odd integers between 0 and 20, if the order does not matter? (The answer is 120.)

- How many prime factorizations does 200 have, where we count different orderings as distinct? (The answer is 10.)

- How many ways are there to tile a 2×15-rectangle with dominos (i.e., rectangles of size 1×2 or 2×1)? (The answer is 987. A Fibonacci number! Why?)

- How many addends do you get when you expand the product $(a+b)(c+d+e)(f+g)$? (The answer is 12.)

- How many different monomials do you get when you expand the product $(a-b)\left(a^2+ab+b^2\right)$? (Trap problem. The answer is 2, since $(a-b)\left(a^2+ab+b^2\right) = a^3-b^3$.)

- How many positive divisors does 12 have? (The answer is 6, namely $1, 2, 3, 4, 6, 12$.)

We will solve a few of these problems informally in this chapter, and then (in Chapter 6) make the underlying concepts rigorous and solve more such problems.

4.1. A refresher on sets

You have seen some set theory in prerequisite classes, but let me remind and maybe also elaborate on it.

Formally, the notion of a set is fundamental and cannot be defined.

Informally, a **set** is a collection of objects that knows which objects it contains and which objects it doesn't.

That is, if S is a set and p is any object, then S can either contain p (in which case we write $p \in S$) or not contain p (for this we write $p \notin S$). There is no such thing as "containing p twice".

The objects that a set S contains are called the **elements** of S; they are said to **belong to** S or **lie in** S or **be contained in** S.

A set can be finite or infinite (i.e., contain finitely many or infinitely many objects). It can be empty (i.e., contain nothing) or nonempty.

For example, consider the set of all odd integers. This set contains all odd integers and nothing else; it is infinite. Generally, "the set of X" means the set that contains X and nothing else.

When a set is finite, it can be written down by listing all its elements. For example, the set of all odd integers between 0 and 10 can be written as

$$\{1,3,5,7,9\}.$$

The braces { and } are there to signal that we mean the set of all the elements, not the single elements themselves. These braces are called "set braces".

Some more examples of finite sets are

$$\begin{aligned}
&\{1,2,3,4,5\}, \\
&\{1,2\}, \\
&\{1\}, \\
&\{\} \qquad (\text{the empty set, also denoted } \varnothing), \\
&\{1,2,\ldots,1000\}.
\end{aligned}$$

Some infinite sets can also be written in this form:

$$\begin{aligned}
&\{1,2,3,\ldots\} \qquad (\text{this is the set of all positive integers}), \\
&\{0,1,2,\ldots\} \qquad (\text{this is the set of all nonnegative integers}), \\
&\{4,5,6,\ldots\} \qquad (\text{this is the set of all integers } \geq 4), \\
&\{-1,-2,-3,\ldots\} \qquad (\text{this is the set of all nonnegative integers}), \\
&\{\ldots,-2,-1,0,1,2,\ldots\} \qquad (\text{this is the set of all integers}).
\end{aligned}$$

Some others cannot. For example, how would you list all real numbers? or even all rational numbers?

Another way to describe a set is just by putting a description of its element in set braces. For example,

$$\{\text{all integers}\}, \qquad \{\text{all integers between 3 and 9 inclusive}\}, \\ \{\text{all real numbers}\}.$$

Often you want to define a set that contains all objects of a certain type that satisfy a certain condition. For instance, say you want the set of all integers x that satisfy $x^2 < 13$. There is a notation for this:

$$\left\{x \text{ is an integer} \mid x^2 < 13\right\}.$$

The vertical bar | here should be read as "such that"; it is not a divisibility sign or an absolute value bracket. You can also use a colon (:) instead. The part before the vertical bar says what type of objects you are considering (in our case, the integers x); the part after this bar imposes a condition (or several) on these objects (in our case, $x^2 < 13$). What you get is the set of all objects of the former type that satisfy the latter condition. For instance,

$$\begin{aligned}
&\left\{x \text{ is an integer } \mid x^2 < 13\right\} \\
&= \{\text{all integers whose square is smaller than } 13\} \\
&= \{-3, -2, -1, 0, 1, 2, 3\}.
\end{aligned}$$

Some sets have standard names:

$$\begin{aligned}
\mathbb{Z} &= \{\text{all integers}\} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}; \\
\mathbb{N} &= \{\text{all nonnegative integers}\} = \{0, 1, 2, \ldots\} \\
&\qquad\qquad (\text{some authors use this for } \{1, 2, 3, \ldots\} \text{ instead}); \\
\mathbb{Q} &= \{\text{all rational numbers}\}; \\
\mathbb{R} &= \{\text{all real numbers}\}; \\
\mathbb{C} &= \{\text{all complex numbers}\}; \\
\varnothing &= \{\} = (\text{the empty set}).
\end{aligned}$$

Using these notations, we can rewrite

$$\left\{x \text{ is an integer } \mid x^2 < 13\right\} \text{ as } \left\{x \in \mathbb{Z} \mid x^2 < 13\right\}.$$

Yet another way to define a set is when you let a variable range over a given set and collect a certain derived quantity. For instance,

$$\left\{x^2 + 2 \mid x \in \{1, 3, 5, 7, 9\}\right\}$$

means the set whose elements are the numbers $x^2 + 2$ for all $x \in \{1, 3, 5, 7, 9\}$. Thus,

$$\begin{aligned}
&\left\{x^2 + 2 \mid x \in \{1, 3, 5, 7, 9\}\right\} \\
&= \left\{1^2 + 2,\ 3^2 + 2,\ 5^2 + 2,\ 7^2 + 2,\ 9^2 + 2\right\} \\
&= \{3,\ 11,\ 27,\ 51,\ 83\}.
\end{aligned}$$

In general, if S is a given set, then the notation

$$\{\text{an expression} \mid x \in S\}$$

stands for the set whose elements are the values of the given expression for all $x \in S$.

Some more examples of this:

$$\begin{aligned}\left\{\frac{x+1}{x} \ \mid\ x \in \{1,2,3,4,5\}\right\} &= \left\{\frac{1+1}{1},\ \frac{2+1}{2},\ \frac{3+1}{3},\ \frac{4+1}{4},\ \frac{5+1}{5}\right\} \\ &= \left\{2,\ \frac{3}{2},\ \frac{4}{3},\ \frac{5}{4},\ \frac{6}{5}\right\}\end{aligned}$$

and

$$\begin{aligned}\left\{x^2\%5 \ \mid\ x \in \mathbb{N}\right\} &= \left\{0^2\%5,\ 1^2\%5,\ 2^2\%5,\ 3^2\%5,\ \ldots\right\} \\ &= \{0,\ 1,\ 4,\ 4,\ 1,\ 0,\ 1,\ 4,\ 4,\ 1,\ 0,\ 1,\ 4,\ 4,\ \ldots\} \\ &= \{0,1,4\}\end{aligned}$$

(since $(x+5)^2\%5 = x^2\%5 = (-x)^2\%5$ for all integers x).

A set cannot contain an element more than once. It also does not come with an ordering of its elements. Thus,

$$\{1,2\} = \{2,1\} = \{1,1,2\} = \{1,2,1\} = \{1,2,1,2,1\},$$

since each of these five sets contains 1 and 2 and nothing else.

Note that sets can contain any mathematical objects, not just numbers. In particular, they can contain other sets. Make sure you understand what the sets

$$\{1,2,3\}, \qquad \{\{1,2,3\}\}, \qquad \{\{1,2\},\{3\}\}, \qquad \{\{1\},\{2\},\{3\}\}$$

are and why they are different.

Sets can be compared and combined in several ways:

Definition 4.1.1. Let A and B be two sets.

(a) We say that A is a **subset** of B (and we write $A \subseteq B$) if every element of A is an element of B.

(b) We say that A is a **superset** of B (and we write $A \supseteq B$) if every element of B is an element of A. This is tantamount to saying $B \subseteq A$.

(c) We say that $A = B$ if the sets A and B contain the same elements. This is tantamount to saying that both $A \subseteq B$ and $A \supseteq B$ hold.

(d) We define the **union** of A and B to be the set

$$\begin{aligned}A \cup B :&= \{\text{all elements that are contained in } A \text{ or } B\} \\ &= \{x \ \mid\ x \in A \text{ or } x \in B\}.\end{aligned}$$

(As usual in mathematics, the "or" is non-exclusive. So this includes also the elements that are contained in both A and B.)

(e) We define the **intersection** of A and B to be the set

$$\begin{aligned} A \cap B &:= \{\text{all elements that are contained in both } A \text{ and } B\} \\ &= \{x \mid x \in A \text{ and } x \in B\}. \end{aligned}$$

(f) We define the **set difference** of A and B to be the set

$$\begin{aligned} A \setminus B &:= \{\text{all elements of } A \text{ that do not lie in } B\} \\ &= \{x \mid x \in A \text{ and } x \notin B\} = \{x \in A \mid x \notin B\}. \end{aligned}$$

Some authors call this $A - B$.

(g) We say that A and B are **disjoint** if $A \cap B = \varnothing$ (that is, A and B have no element in common).

Beware: "disjoint" $\neq$ "distinct".

For example,

$$\begin{aligned} \{1,3,5\} &\subseteq \{1,2,3,4,5\}; \\ \{1,2,3,4,5\} &\supseteq \{1,3,5\}; \\ \text{we don't have } \{5,6,7\} &\subseteq \{1,2,3,4,5\}; \\ \{1,2,3\} &= \{3,2,1\}; \\ \{1,3,5\} \cup \{3,6\} &= \{1,3,5,3,6\} = \{1,3,5,6\}; \\ \{1,3,5\} \cap \{3,6\} &= \{3\}; \\ \{1,3,5\} \setminus \{3,6\} &= \{1,5\}; \\ \{3,6\} \setminus \{1,3,5\} &= \{6\}; \\ \{1,2,4\} \cap \{3,5\} &= \varnothing, \end{aligned}$$

so that the sets $\{1,2,4\}$ and $\{3,5\}$ are disjoint.

We can likewise define unions and intersections of multiple sets:

$$\begin{aligned} A_1 \cup A_2 \cup \cdots \cup A_k &= \{x \mid x \in A_i \text{ for some } i \in \{1,2,\ldots,k\}\}; \\ A_1 \cap A_2 \cap \cdots \cap A_k &= \{x \mid x \in A_i \text{ for all } i \in \{1,2,\ldots,k\}\}. \end{aligned}$$

(The latter requires $k > 0$.)

Several sets $A_1, A_2, \ldots, A_k$ are called **disjoint** if any two of them are disjoint (i.e., if $A_i \cap A_j = \varnothing$ for each $i \neq j$). For instance, $\{1,2\}$, $\{2,3\}$, $\{3,1\}$ are not disjoint, even though the total intersection $\{1,2\} \cap \{2,3\} \cap \{3,1\} = \varnothing$.

4.2. Counting, informally

Now let us see how the elements of a set can be counted. A formal definition of "counting" will be given later, so we are still doing not-quite-rigorous work.

Recall that the set of all odd integers between 0 and 10 has 5 elements $(1, 3, 5, 7, 9)$. More generally:

> **Proposition 4.2.1.** Let $n \in \mathbb{N}$. Then, there are exactly $(n+1)\ //2 = \left\lfloor \dfrac{n+1}{2} \right\rfloor$ odd integers between 0 and n (inclusive).

Informal proof. The equality $(n+1)\ //2 = \left\lfloor \dfrac{n+1}{2} \right\rfloor$ follows from the quotient=floor proposition. It remains to show that there are exactly $\left\lfloor \dfrac{n+1}{2} \right\rfloor$ many odd integers between 0 and n. ("Between" always means "between inclusively".)

We prove this by induction on n:

Base case: For $n = 0$, the claim is true $(0 = 0)$.

Induction step: Let n be a positive integer. Assume (as the IH) that the claim is true for $n-1$. That is, assume that there are exactly $\left\lfloor \dfrac{n}{2} \right\rfloor$ odd integers between 0 and $n-1$. We must now show that the claim is also true for n, i.e., that there are exactly $\left\lfloor \dfrac{n+1}{2} \right\rfloor$ odd integers between 0 and n.

IOW (= In other words): The IH says that

$$(\text{\# of odd integers between 0 and } n-1) = \left\lfloor \frac{n}{2} \right\rfloor,$$

and our goal is to prove that

$$(\text{\# of odd integers between 0 and } n) = \left\lfloor \frac{n+1}{2} \right\rfloor.$$

We are in one of the following cases:

Case 1: The number n is even.

Case 2: The number n is odd.

Consider Case 1. In this case, n is even, i.e., not odd. Thus, the odd integers between 0 and n are just the odd integers between 0 and $n-1$. Hence,

$$\begin{aligned}&(\text{\# of odd integers between 0 and } n)\\&= (\text{\# of odd integers between 0 and } n-1) = \left\lfloor \frac{n}{2} \right\rfloor.\end{aligned}$$

On the other hand,

$$\left\lfloor \frac{n+1}{2} \right\rfloor = \left\lfloor \frac{n}{2} \right\rfloor \qquad (\text{by the birthday lemma}).$$

Comparing these two equalities, we conclude that

$$(\text{\# of odd integers between 0 and } n) = \left\lfloor \frac{n+1}{2} \right\rfloor,$$

as desired.

Now to Case 2. In this case, n is odd. Thus, the odd integers between 0 and n are just the odd integers between 0 and $n-1$ along with the new odd integer n. Therefore,

$$\begin{aligned}
&(\text{\# of odd integers between 0 and } n) \\
&= (\text{\# of odd integers between 0 and } n-1) + 1 \\
&= \left\lfloor \frac{n}{2} \right\rfloor + 1 \qquad (\text{by the IH}).
\end{aligned}$$

On the other hand,

$$\left\lfloor \frac{n+1}{2} \right\rfloor = \left\lfloor \frac{n}{2} \right\rfloor + 1 \qquad (\text{by the birthday lemma, since } 2 \mid n+1).$$

Comparing these two equalities, we again obtain

$$(\text{\# of odd integers between 0 and } n) = \left\lfloor \frac{n+1}{2} \right\rfloor,$$

as desired.

So the goal is achieved in both Cases 1 and 2, and the proof is complete. □

Incidentally, it is worth stating the formula for the number of all integers (not just odd ones) in a given interval:

> **Proposition 4.2.2.** Let $a, b \in \mathbb{Z}$ be such that $a \leq b+1$.
>
> Then, there are exactly $b-a+1$ numbers in the set
>
> $$\{a,\ a+1,\ a+2,\ \ldots,\ b\} := \{i \in \mathbb{Z} \ \mid\ a \leq i \leq b\}.$$

Proof. Induct on b. □

Likewise, if $a \leq b-1$, then there are $b-a-1$ numbers in the set

$$\{a+1,\ a+2,\ \ldots,\ b-1\} := \{i \in \mathbb{Z} \ \mid\ a < i < b\}.$$

Note: In Python, `range(a, b)` means $\{a, a+1, \ldots, b-1\}$ (to be fully precise, it means the list $(a, a+1, \ldots, b-1)$); this contains exactly $b-a$ numbers.

> **Convention 4.2.3.** The symbol “#” means “number”.

4.3. Counting subsets

4.3.1. Counting them all

How many subsets does the set $\{1, 2, 3\}$ have? Let us list them:

$$\begin{array}{cccc} \varnothing, & \{1\}, & \{2\}, & \{3\}, \\ \{1,3\}, & \{2,3\}, & \{1,2\}, & \{1,2,3\}. \end{array}$$

So there are 8 of them. (Note that every set A satisfies $\varnothing \subseteq A$ and $A \subseteq A$.)

Likewise:

- The set $\{1, 2\}$ has 4 subsets: $\varnothing$, $\{1\}$, $\{2\}$, $\{1, 2\}$.
- The set $\{1\}$ has 2 subsets: $\varnothing$ and $\{1\}$.
- The set $\{\}$ has 1 subset: itself.
- The subset $\{1, 2, 3, 4\}$ has 16 subsets.

The pattern here is hard to miss:

Theorem 4.3.1. Let $n \in \mathbb{N}$. Then,

$$(\text{\# of subsets of } \{1, 2, \ldots, n\}) = 2^n.$$

Informal proof. We induct on n.

Base case ($n = 0$)*:* This is saying that $(\text{\# of subsets of } \{1, 2, \ldots, 0\}) = 2^0$. But this is clear, since the set $\{1, 2, \ldots, 0\}$ is empty and thus has only 1 subset (itself).

Induction step: We proceed from $n - 1$ to n. So let n be a positive integer. We assume (as IH) that the theorem holds for $n - 1$ instead of n. Our goal is to prove that it holds for n.

So our IH says that

$$(\text{\# of subsets of } \{1, 2, \ldots, n-1\}) = 2^{n-1}.$$

Our goal is to prove that

$$(\text{\# of subsets of } \{1, 2, \ldots, n\}) = 2^n.$$

We define

- a **red set** to be a subset of $\{1, 2, \ldots, n\}$ that contains n;
- a **green set** to be a subsets of $\{1, 2, \ldots, n\}$ that does not contain n.

Each subset of $\{1, 2, \ldots, n\}$ is either red or green but not both. Hence,

$$(\text{\# of subsets of } \{1, 2, \ldots, n\}) = (\text{\# of red sets}) + (\text{\# of green sets}).$$

(This is an instance of what is known as the **sum rule for two sets**: If some objects are classified into two types, then these objects can be counted by counting the objects of each type and adding the results. We will formalize this soon.)

Now let us count the red sets and the green sets separately.

- The green sets are easy: They are just the subsets of $\{1, 2, \ldots, n-1\}$. So

  $$(\text{\# of green sets}) = (\text{\# of subsets of } \{1, 2, \ldots, n-1\}) = 2^{n-1}$$

  (by the IH).

- The problem of counting the red sets can be reduced to counting the green sets: Indeed, the red sets are just the green sets with the element n inserted into them. To be more precise: Each green set can be turned into a red set by inserting n. Conversely, each red set can be turned into a green set by removing n. These two operations are mutually inverse, so they set up a one-to-one correspondence between the green sets and the red sets. Thus

  $$(\text{\# of red sets}) = (\text{\# of green sets}) = 2^{n-1}.$$

Combining what we have shown,

$$\begin{aligned}(\text{\# of subsets of } \{1, 2, \ldots, n\}) &= \underbrace{(\text{\# of red sets})}_{=2^{n-1}} + \underbrace{(\text{\# of green sets})}_{=2^{n-1}} \\ &= 2^{n-1} + 2^{n-1} = 2 \cdot 2^{n-1} = 2^n,\end{aligned}$$

as desired. So the induction step is complete, and the theorem proved. □

More generally:

Theorem 4.3.2. Let $n \in \mathbb{N}$. Let S be any n-element set. Then,

$$(\text{\# of subsets of } S) = 2^n.$$

Informal proof. This follows from the previous theorem, since we can rename the n elements of S as $1, 2, \ldots, n$. □

4.3.2. Counting the subsets of a given size

Let us now refine our counting question: We want to count not all subsets of $\{1,2,\ldots,n\}$, but only those of a given size k. Here, the **size** (or **cardinality**) of a set means the # of its elements, i.e., how many distinct elements it has. (For example, the size of $\{1,4,1,2\}$ is 3.) A set of size k is also known as a k**-element set**.

For instance, $\{1,2,3,4\}$ is a 4-element set. How many 2-element subsets does it have? It has six:

$$\{1,2\}, \qquad \{1,3\}, \qquad \{1,4\}, \qquad \{2,3\}, \qquad \{2,4\}, \qquad \{3,4\}.$$

More generally, the answer to the question "how many k-element subsets does a given n-element set have" turns out to be the binomial coefficient $\dbinom{n}{k}$:

Theorem 4.3.3 (combinatorial interpretation of BCs)**.** Let $n \in \mathbb{N}$, and let k be any number. Let S be an n-element set. Then,

$$(\text{\# of } k\text{-element subsets of } S) = \dbinom{n}{k}.$$

Informal proof. We induct on n (without fixing k). That is, we use induction on n to prove the statement

$$P(n) := \left(\begin{array}{c} \text{"for any number } k \text{ and any } n\text{-element set } S, \\ \text{we have } (\text{\# of } k\text{-element subsets of } S) = \dbinom{n}{k}\text{"} \end{array} \right)$$

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

Base case: This is the case $n = 0$. Thus, an n-element set S is just an empty set. Hence, it has one 0-element subset, and no other subsets. So we find

$$(\text{\# of } k\text{-element subsets of } S) = \begin{cases} 1, & \text{if } k = 0; \\ 0, & \text{if } k \neq 0. \end{cases}$$

But

$$\dbinom{n}{k} = \dbinom{0}{k} = \begin{cases} 1, & \text{if } k = 0; \\ 0, & \text{if } k \neq 0. \end{cases}$$

Comparing these two equalities, we get $P(0)$.

Induction step: Let n be a positive integer. Assume (as IH) that $P(n-1)$ holds. We must prove that $P(n)$ holds.

So let us consider any number k and any n-element set S. We must prove that

$$(\text{\# of } k\text{-element subsets of } S) = \dbinom{n}{k}.$$

We rename the n elements of S as $1, 2, \ldots, n$, so we must prove that

$$(\text{\# of } k\text{-element subsets of } \{1, 2, \ldots, n\}) = \binom{n}{k}.$$

To prove this, we define

- a **red set** to be a k-element subset of $\{1, 2, \ldots, n\}$ that contains n;
- a **green set** to be a k-element subset of $\{1, 2, \ldots, n\}$ that does not contain n.

Just as in the previous induction proof, we have

$$\begin{aligned} &(\text{\# of } k\text{-element subsets of } \{1, 2, \ldots, n\}) \\ &= (\text{\# of red sets}) + (\text{\# of green sets}). \end{aligned}$$

Now let us count them:

- The green sets are just the k-element subsets of $\{1, 2, \ldots, n-1\}$. So

  $$\begin{aligned} (\text{\# of green sets}) &= (\text{\# of } k\text{-element subsets of } \{1, 2, \ldots, n-1\}) \\ &= \binom{n-1}{k} \end{aligned}$$

  (by the IH, since $\{1, 2, \ldots, n-1\}$ is an $(n-1)$-element set).

- Now for the red sets.

  If T is a red set, then $T \setminus \{n\}$ is a $(k-1)$-element subset of $\{1, 2, \ldots, n-1\}$. In other words, $T \setminus \{n\}$ is a blue set, where we define a **blue set** to mean a $(k-1)$-element subset of $\{1, 2, \ldots, n-1\}$.

  Conversely, if U is a blue set, then $U \cup \{n\}$ is a red set.

  This sets up a one-to-one correspondence between the red sets and the blue sets. Thus,

  $$\begin{aligned} (\text{\# of red sets}) &= (\text{\# of blue sets}) \\ &= (\text{\# of } (k-1)\text{-element subsets of } \{1, 2, \ldots, n-1\}) \\ &= \binom{n-1}{k-1} \end{aligned}$$

  (by the IH, applied to $\{1, 2, \ldots, n-1\}$ instead of S and $k-1$ instead of k).

Altogether,

$$\begin{aligned}
&(\text{\# of } k\text{-element subsets of } \{1,2,\ldots,n\}) \\
&= \underbrace{(\text{\# of red sets})}_{=\binom{n-1}{k-1}} + \underbrace{(\text{\# of green sets})}_{=\binom{n-1}{k}} \\
&= \binom{n-1}{k-1} + \binom{n-1}{k} = \binom{n}{k} \qquad (\text{by Pascal's recursion}).
\end{aligned}$$

This completes the induction step, so the theorem is proved. □

The above proof actually constructs a recursive algorithm for listing all the k-element subsets of $\{1,2,\ldots,n\}$.

We will see $\binom{n}{k}$ come out as an answer to several other counting questions.

4.4. Tuples (aka lists)

4.4.1. Definition and disambiguation

We shall now come to counting lists. First of all, what is a finite list? Here is a somewhat awkward definition:

Definition 4.4.1. A **finite list** (aka **tuple**) is a list consisting of finitely many objects. The objects appear in this list in a specified order, and they don't have to be distinct.

A finite list is delimited using parentheses: i.e., the list that contains the objects $a_1, a_2, \ldots, a_n$ in this order is written $(a_1, a_2, \ldots, a_n)$.

"Specified order" means that the list has a well-defined first entry, a well-defined second entry, and so on. Thus, two lists $(a_1, a_2, \ldots, a_n)$ and $(b_1, b_2, \ldots, b_m)$ are considered equal if and only if

- we have $n = m$, and
- we have $a_i = b_i$ for each $i \in \{1,2,\ldots,n\}$.

For example,

$$\begin{aligned}
(1,2) &\neq (2,1); \\
(1,2) &\neq (1,2,2); \\
(1,1,2) &\neq (1,2,2).
\end{aligned}$$

Definition 4.4.2. **(a)** The **length** of a list $(a_1, a_2, \ldots, a_n)$ is defined to be n.

(b) A list of length 2 is called a **pair** (or **ordered pair**).

(c) A list of length 3 is called a **triple**.

(d) A list of length 4 is called a **quadruple**.

(e) A list of length n is called an **n-tuple**.

For instance, $(1,3,2,2)$ is a list of length 4, that is, a 4-tuple (aka quadruple). For another example, $(5,8)$ is a pair, i.e., a 2-tuple.

Lists of length 1 contain just a single entry. For instance, (3) is a list containing only the entry 3.

4.4.2. Counting pairs

Now let us count some pairs:

- How many pairs (a,b) are there with $a,b \in \{1,2,3\}$? There are nine:

  $$\begin{array}{ccc} (1,1), & (1,2), & (1,3), \\ (2,1), & (2,2), & (2,3), \\ (3,1), & (3,2), & (3,3). \end{array}$$

  I have deliberately arranged them in a square pattern – a table with 3 rows and 3 columns, where the row determines the first entry of each pair, and the column determines the second entry. Thus, there is a total of $3 \cdot 3 = 9$ such pairs.

- How many pairs (a,b) are there with $a,b \in \{1,2,3\}$ and $a < b$? There are three:

  $$(1,2), \qquad (1,3), \qquad (2,3).$$

- How many pairs (a,b) are there with $a,b \in \{1,2,3\}$ and $a = b$? There are three:

  $$(1,1), \qquad (2,2), \qquad (3,3).$$

- How many pairs (a,b) are there with $a,b \in \{1,2,3\}$ and $a > b$? There are three:

  $$(2,1), \qquad (3,1), \qquad (3,2).$$

Let us generalize this:

Proposition 4.4.3. Let $n \in \mathbb{N}$. Then:

(a) The # of pairs (a,b) with $a,b \in \{1,2,\ldots,n\}$ is n^2.

(b) The # of pairs (a,b) with $a,b \in \{1,2,\ldots,n\}$ and $a < b$ is $1 + 2 + \cdots + (n-1)$.

(c) The # of pairs (a,b) with $a,b \in \{1,2,\ldots,n\}$ and $a = b$ is n.

(d) The # of pairs (a,b) with $a,b \in \{1,2,\ldots,n\}$ and $a > b$ is $1 + 2 + \cdots + (n-1)$.

Informal proof. **(a)** These pairs can be arranged in a table with n rows and n columns, where the rows determine the first entry and the columns determine the second entry. Here is how this table looks like:

$$\begin{array}{cccc} (1,1)\,, & (1,2)\,, & \ldots, & (1,n)\,, \\ (2,1)\,, & (2,2)\,, & \ldots, & (2,n)\,, \\ \vdots & \vdots & \cdots & \vdots \\ (n,1)\,, & (n,2)\,, & \ldots, & (n,n)\,. \end{array}$$

So there are $n \cdot n = n^2$ such pairs.

(b) In the table we have just shown, a pair (a,b) satisfies $a < b$ if and only if it is placed above the main diagonal. Thus, the # of such pairs is the # of cells above the main diagonal. But this # is

$$0 + 1 + 2 + \cdots + (n-1)\,,$$

because there are 0 such cells in the first column, 1 such cell in the second, 2 in the third, and so on. Hence,

$$\begin{aligned} &(\text{\# of pairs } (a,b) \text{ with } a,b \in \{1,2,\ldots,n\} \text{ and } a < b) \\ &= 0 + 1 + 2 + \cdots + (n-1) \\ &= 1 + 2 + \cdots + (n-1)\,. \end{aligned}$$

(c) A pair (a,b) with $a = b$ is just a pair of the form (a,a). So the number of such pairs is the number of possible a's; but this number is n.

(d) This follows from part **(b)**, since the pairs (a,b) satisfying $a < b$ are in one-to-one correspondence with the pairs (a,b) satisfying $a > b$ (just swap the two entries). □

The proposition we just proved has a nice consequence: For any $n \in \mathbb{N}$, we have

$$\begin{aligned} n^2 &= (\text{\# of pairs } (a,b) \text{ with } a,b \in \{1,2,\ldots,n\}) \qquad (\text{by part } \mathbf{(a)}) \\ &= (\text{\# of pairs } (a,b) \text{ with } a,b \in \{1,2,\ldots,n\} \text{ and } a < b) \\ &\qquad + (\text{\# of pairs } (a,b) \text{ with } a,b \in \{1,2,\ldots,n\} \text{ and } a = b) \\ &\qquad + (\text{\# of pairs } (a,b) \text{ with } a,b \in \{1,2,\ldots,n\} \text{ and } a > b) \\ &= \underbrace{(1 + 2 + \cdots + (n-1)) + n}_{=1+2+\cdots+n} + \underbrace{(1 + 2 + \cdots + (n-1))}_{=(1+2+\cdots+n)-n} \\ &= (1 + 2 + \cdots + n) + (1 + 2 + \cdots + n) - n \\ &= 2 \cdot (1 + 2 + \cdots + n) - n. \end{aligned}$$

Solving this for $1 + 2 + \cdots + n$, we find

$$1 + 2 + \cdots + n = \frac{n^2 + n}{2} = \frac{n\,(n+1)}{2}.$$

So we have recovered the Little Gauss formula.

Exercise 4.4.1. How many pairs (a, b) are there with $a \in \{1, 2, 3\}$ and $b \in \{1, 2, 3, 4, 5\}$?

Solution. There are 15 of them, since you can arrange them in a table with 3 rows and 5 columns (just as above). □

More generally, by the same reasoning, we find the following:

Theorem 4.4.4. Let $n, m \in \mathbb{N}$. Let A be an n-element set. Let B be an m-element set. Then,

$$(\text{\# of pairs } (a, b) \text{ with } a \in A \text{ and } b \in B) = nm.$$

Likewise for triples:

Theorem 4.4.5. Let $n, m, p \in \mathbb{N}$. Let A be an n-element set. Let B be a m-element set. Let C be a p-element set. Then,

$$(\text{\# of triples } (a, b, c) \text{ with } a \in A \text{ and } b \in B \text{ and } c \in C) = nmp.$$

Informal proof. You can think of these triples as occupying the cells of a 3-dimensional table, but this kind of visualization gets slippery as the dimensions increase.

So here is a better approach: Re-encode each triple (a, b, c) as a pair $((a, b), c)$ (a pair whose first entry is itself a pair). This is a pair whose first entry comes from the set U of all pairs (a, b) with $a \in A$ and $b \in B$, whereas its second entry comes from C. Note that U is an nm-element set, since

$$\begin{aligned}(\text{\# of elements of } U) &= (\text{\# of pairs } (a, b) \text{ with } a \in A \text{ and } b \in B) \\ &= nm \qquad (\text{by the previous theorem}).\end{aligned}$$

Now, we have re-encoded each triple (a, b, c) as a pair $((a, b), c)$ with $(a, b) \in U$ and $c \in C$. Thus,

$$\begin{aligned}&(\text{\# of triples } (a, b, c) \text{ with } a \in A \text{ and } b \in B \text{ and } c \in C) \\ &= (\text{\# of pairs } ((a, b), c) \text{ with } (a, b) \in U \text{ and } c \in C) \\ &= (\text{\# of pairs } (u, c) \text{ with } u \in U \text{ and } c \in C) \\ &= (nm)\, p \qquad \left(\begin{array}{c}\text{by the previous theorem, since } U \text{ is} \\ \text{an } nm\text{-element set and } C \text{ is a } p\text{-element set}\end{array}\right) \\ &= nmp.\end{aligned}$$

□

4.4.3. Cartesian products

There is a general notation for sets of pairs:

Definition 4.4.6. Let A and B be two sets.

The set of all pairs (a,b) with $a \in A$ and $b \in B$ is denoted by $A \times B$, and is called the **Cartesian product** (or just **product**) of the sets A and B.

For instance, $\{1,2\} \times \{7,8,9\}$ is the set of all pairs (a,b) with $a \in \{1,2\}$ and $b \in \{7,8,9\}$. Explicitly, the elements of this set are the pairs

$$\begin{array}{ccc} (1,7), & (1,8), & (1,9), \\ (2,7), & (2,8), & (2,9). \end{array}$$

Likewise, the set $\{1,2\} \times \{2,3\}$ consists of the pairs

$$\begin{array}{cc} (1,2), & (1,3), \\ (2,2), & (2,3). \end{array}$$

A similar notation exists for sets of triples, of quadruples or of k-tuples in general:

Definition 4.4.7. Let $A_1, A_2, \ldots, A_k$ be k sets.

The set of all k-tuples $(a_1, a_2, \ldots, a_k)$ with $a_1 \in A_1$ and $a_2 \in A_2$ and $\cdots$ and $a_k \in A_k$ is denoted by

$$A_1 \times A_2 \times \cdots \times A_k,$$

and is called the **Cartesian product** (or just **product**) of the sets $A_1, A_2, \ldots, A_k$.

For example, the set $\{1,2\} \times \{5\} \times \{2,7,6\}$ consists of the triples

$$\begin{array}{ccc} (1,5,2), & (1,5,7), & (1,5,6), \\ (2,5,2), & (2,5,7), & (2,5,6). \end{array}$$

In total, there are $3 \cdot 1 \cdot 2$ such triples, by the last theorem we proved.

Using the notations we just introduced, we can restate the last two theorems as follows:

Theorem 4.4.8 (product rule for two sets)**.** Let $n, m \in \mathbb{N}$. Let A be an n-element set. Let B be an m-element set. Then, $A \times B$ is an nm-element set.

Theorem 4.4.9 (product rule for three sets)**.** Let $n, m, p \in \mathbb{N}$. Let A be an n-element set. Let B be a m-element set. Let C be a p-element set. Then, $A \times B \times C$ is an nmp-element set.

More generally:

Theorem 4.4.10 (product rule for k sets)**.** Let $A_1, A_2, \ldots, A_k$ be k sets. If each A_i is an n_i-element set, then $A_1 \times A_2 \times \cdots \times A_k$ is an $n_1 n_2 \cdots n_k$-element set.

You can prove this by induction on k, using the product rule for two sets and the same "re-encode a tuple as a nested pair" trick that we used to prove the case of three sets.

4.4.4. Counting strictly increasing tuples (informally)

Above, we have shown that for any given $n \in \mathbb{N}$, the # of pairs (a, b) with $a, b \in \{1, 2, \ldots, n\}$ and $a < b$ is

$$1 + 2 + \cdots + (n-1) = \frac{(n-1)\,n}{2} = \binom{n}{2}.$$

What is the # of triples (a, b, c) of elements of $\{1, 2, \ldots, n\}$ such that $a < b < c$?

Such a triple (a, b, c) always determines a 3-element subset $\{a, b, c\}$ of $\{1, 2, \ldots, n\}$ (and yes, it is really a 3-element set, since $a < b < c$ ensures that a, b, c are distinct). Conversely, any 3-element subset of $\{1, 2, \ldots, n\}$ becomes a triple (a, b, c) satisfying $a < b < c$ if we list its elements in increasing order. Hence, the triples (a, b, c) of elements of $\{1, 2, \ldots, n\}$ satisfying $a < b < c$ are just the 3-element subsets of $\{1, 2, \ldots, n\}$ in disguise. Hence,

$$\begin{aligned}
&(\text{\# of triples } (a,b,c) \text{ with } a,b,c \in \{1,2,\ldots,n\} \text{ and } a<b<c) \\
&= (\text{\# of 3-element subsets of } \{1,2,\ldots,n\}) \\
&= \binom{n}{3}
\end{aligned}$$

(by the combinatorial interpretation of BCs).

More generally, for any $k \in \mathbb{N}$, we have

$$\begin{aligned}
&(\text{\# of } k\text{-tuples } (a_1,a_2,\ldots,a_k) \text{ of elements of } \{1,2,\ldots,n\} \text{ satisfying } a_1<a_2<\cdots<a_k) \\
&= \binom{n}{k}
\end{aligned}$$

(by the same logic).

In contrast, if you drop the $a_1 < a_2 < \cdots < a_k$ condition, you have

$$\begin{aligned}
&(\text{\# of } k\text{-tuples } (a_1,a_2,\ldots,a_k) \text{ of elements of } \{1,2,\ldots,n\}) \\
&= \underbrace{nn\cdots n}_{k \text{ times}} \qquad (\text{by the product rule for } k \text{ sets}) \\
&= n^k.
\end{aligned}$$

Other counting problems don't have answers this simple. For instance,

$$\begin{aligned} & (\text{\# of } k\text{-tuples } (a_1, a_2, \ldots, a_k) \text{ of elements of } \{1, 2, \ldots, n\} \\ & \qquad \text{whose largest entry is } a_k) \\ & = 1^{k-1} + 2^{k-1} + 3^{k-1} + \cdots + n^{k-1}, \end{aligned}$$

but there is no way to express this without a "$\cdots$" or a "$\sum$" sign. For each specific k, we can simplify this:

$$\begin{aligned} 1^0 + 2^0 + \cdots + n^0 &= 1 + 1 + \cdots + 1 = n; \\ 1^1 + 2^1 + \cdots + n^1 &= 1 + 2 + \cdots + n = \frac{n\left(n+1\right)}{2}; \\ 1^2 + 2^2 + \cdots + n^2 &= \frac{n\left(n+1\right)\left(2n+1\right)}{6}; \\ 1^3 + 2^3 + \cdots + n^3 &= \frac{n^2\left(n+1\right)^2}{4}; \\ 1^4 + 2^4 + \cdots + n^4 &= \frac{n\left(n+1\right)\left(2n+1\right)\left(3n+3n^2-1\right)}{30}; \\ & \ldots \end{aligned}$$

Such a closed-form expression for $1^m + 2^m + \cdots + n^m$ exists for any specific value of m (references in the notes), but it gets messier the larger m gets.

In the next two chapters, we will learn what it means for a set to have n elements, and we will state precisely the rules that we have used in our above arguments. First we will need to get familiar with the notion of **maps** (= **functions**).

5. Maps (aka functions)

5.1. Functions, informally

One of the main notions is mathematics is that of a **function**, aka **map**, aka **mapping**, aka **transformation**.

Intuitively, a function is a "black box" that takes inputs and transforms them into outputs. For example, the "$f\left(t\right) = t^2$" function takes a real number t and outputs its square t^2.

You can thus think of a function as a rule for producing an output from an input. This leads us to the following **provisional** definition of a function:

> **Definition 5.1.1** (Informal definition of a function)**.** Let X and Y be two sets. A **function** from X to Y is (provisionally) a rule that transforms each element of X into some element of Y.

> If this function is called f, then the result of applying it to a given $x \in X$ (that is, the output it produces when x is given as input) will be called $f(x)$ (sometimes fx).

This "definition" kicks the can down the road: It defines "function" in terms of "rule", but what is a rule? But it gives the right intuition for what a function is for. Here are some comments that should clarify this definition:

- A function has to "work" for each element of X. It cannot decline to operate on some elements! Thus, "take the reciprocal" is **not** a function from $\mathbb{R}$ to $\mathbb{R}$, since it does not operate on 0. However, "take the reciprocal" is a function from $\mathbb{R} \setminus \{0\}$ to $\mathbb{R}$.

- A function must not be ambiguous. Each input must produce exactly one output. So "take your number to a random power" is not a function.

- We write "$f : X \to Y$" for "f is a function from X to Y".

- Instead of saying "$f(x) = y$", we can say "f transforms x into y" or "f sends x to y" or "f maps x to y" or "f takes the value y at x" or "y is the value of f at x" or "y is the image of x under f" or "applying f to x yields y" or "f takes x to y" or "$f : x \mapsto y$". All these statements are synonyms.

  For instance, if f is the "take the square" function from $\mathbb{R}$ to $\mathbb{R}$, then $f(2) = 2^2 = 4$, so that f transforms 2 into 4, or sends 2 to 4, or takes the value 4 at 2, or $f : 2 \mapsto 4$.

  Do not confuse the $\to$ and $\mapsto$ arrows. (LaTeX: \to vs. \mapsto)

- So the **value** of a function f at an input x is the output $f(x)$.

- The notation

  $$X \to Y, \qquad x \mapsto (\text{some expression involving } x)$$

  (where X and Y are two sets) means "the function from X to Y that sends each element x of X to the expression to the right of the $\mapsto$ symbol".

  For example,

  $$\mathbb{R} \to \mathbb{R}, \qquad x \mapsto x^2$$

  is the "take the square" function (sending each element x of $\mathbb{R}$ to x^2). For another example,

  $$\mathbb{R} \to \mathbb{R}, \qquad x \mapsto \frac{x}{\sin x + 15}$$

is the function that takes the sine of the input, then adds 15, then divides the input by the result. (Note that this is well-defined, since $\sin x + 15$ is always positive.)

For yet another example,

$$\begin{aligned} \mathbb{R} &\to \mathbb{R}, \\ x &\mapsto 2 \end{aligned}$$

is the function that sends each input x to 2; this is an example of a constant function.

For yet another example,

$$\begin{aligned} \mathbb{Z} &\to \mathbb{Q}, \\ x &\mapsto 2^x \end{aligned}$$

is a function (sending each integer x to 2^x). Here are some of its values:

x	-2	-1	0	1	2
2^x	$\frac{1}{4}$	$\frac{1}{2}$	1	2	4

A more complicated example is the function

$$\begin{aligned} \mathbb{Z} &\to \mathbb{Q}, \\ x &\mapsto \begin{cases} \dfrac{1}{x-1}, & \text{if } x \neq 1; \\ 5, & \text{if } x = 1. \end{cases} \end{aligned}$$

- The notation

  $$\begin{aligned} f : X &\to Y, \\ x &\mapsto (\text{some expression involving } x) \end{aligned}$$

  means that we take the function from X to Y that sends each $x \in X$ to the expression on the right of the $\mapsto$ symbol, and we call this function f.

  (Or, if a function named f already exists, then this notation means that this f **is** the function from X to Y that sends each $x \in X$ to the expression on the right of the $\mapsto$ symbol.)

  For instance, if I write

  $$\begin{aligned} f : \mathbb{R} &\to \mathbb{R}, \\ x &\mapsto x^2 + 1, \end{aligned}$$

  then f is the function from $\mathbb{R}$ to $\mathbb{R}$ that sends each $x \in \mathbb{R}$ to $x^2 + 1$.

- If the set X is finite, then a function $f : X \to Y$ can be specified by listing all its values. For instance, I can define a function $h : \{1, 5, 7\} \to \mathbb{N}$ by setting

  $$\begin{aligned} h(1) &= 10, \\ h(5) &= 35, \\ h(7) &= 10. \end{aligned}$$

  These values have been chosen at whim. A function does not have to behave in any "natural" way.

- If f is a function from X to Y, then the sets X and Y are part of the function. Thus,

  $$\begin{aligned} g_1 : \mathbb{Z} &\to \mathbb{Q}, \\ x &\mapsto 2^x \end{aligned}$$

  and

  $$\begin{aligned} g_2 : \mathbb{N} &\to \mathbb{Q}, \\ x &\mapsto 2^x \end{aligned}$$

  and

  $$\begin{aligned} g_3 : \mathbb{N} &\to \mathbb{N}, \\ x &\mapsto 2^x \end{aligned}$$

  are three distinct functions! We distinguish between them so that we can speak of the "domain" and the "target" of a function. Namely, the **domain** of a function $f : X \to Y$ is defined to be the set X, whereas the **target** of f is defined to be the set Y. So g_1 has domain $\mathbb{Z}$ while g_2 and g_3 have domain $\mathbb{N}$. Meanwhile, g_1 and g_2 have target $\mathbb{Q}$, whereas g_3 has target $\mathbb{N}$.

- When are two functions equal? In programming, functions are often understood to be algorithms, and two algorithms can be different even if they compute the same thing. In mathematics, however, only the domain, the target and the output values matter; the way they are computed does not. Two algorithms that (always) compute the same thing count for only one function.

  Thus equality of functions can be defined as follows:

  Two functions $f_1 : X_1 \to Y_1$ and $f_2 : X_2 \to Y_2$ are said to be **equal** if and only if

  $$X_1 = X_2 \qquad \text{and} \qquad Y_1 = Y_2 \qquad \text{and}$$
  $$f_1(x) = f_2(x) \qquad \text{for each } x \in X_1.$$

An example of two equal functions is

$$\begin{aligned} f_1 : \mathbb{R} &\to \mathbb{R}, \\ x &\mapsto x^2 \end{aligned}$$

and

$$\begin{aligned} f_2 : \mathbb{R} &\to \mathbb{R}, \\ x &\mapsto |x|^2 , \end{aligned}$$

since each $x \in \mathbb{R}$ satisfies $x^2 = |x|^2$.

At this point, we still don't have a rigorous definition of a function. Our next goal is to fix this. We will first define the more general concept of a **relation**, and then characterize functions as relations with a certain property.

5.2. Relations

Relations (to be specific: binary relations) are another concept that you have already seen on examples:

- The relation $\subseteq$ is a relation between two sets. For example, we have $\{1,3\} \subseteq \{1,2,3,4\}$ but not $\{1,5\} \subseteq \{1,2,3,4\}$.
- The order relations $\leq$ and $<$ and $>$ and $\geq$ are relations between integers or rational numbers or real numbers. For example, $1 \leq 5$ but $1 \not\leq -1$.
- The containment relation $\in$ is a relation between an object and a set. For instance, $3 \in \mathbb{N}$ but $-7 \notin \mathbb{N}$.
- The equality relation $=$ is a relation between objects.
- The divisibility relation $\mid$ is a relation between two integers.
- The relation "coprime" is a relation between two integers.
- In plane geometry, relations abound: "parallel", "perpendicular", "congruent", "similar", "directly similar", "lies on", "tangent", ...
- For any given integer n, the relation "congruent modulo n" is a relation between two integers. Let me call it $\overset{n}{\equiv}$, so that $a \overset{n}{\equiv} b$ just means $a \equiv b \mod n$. For example, $2 \overset{3}{\equiv} 8$ but $2 \overset{3}{\not\equiv} 7$.

What do these relations all have in common? They can be applied to pairs of objects. Applying a relation to a pair of objects gives a statement that can be true or false. For example, applying the relation "coprime" to $(5, 8)$ yields the statement "5 is coprime to 8", which is true. Applying it to $(5, 10)$ yields the statement "5 is coprime to 10", which is false.

A general relation R relates elements of a set X with elements of a set Y. For any pair $(x, y) \in X \times Y$ (that is, for any pair consisting of an element x of X and an element y of Y), we can apply the relation R to the pair (x, y), obtaining a statement "$x\ R\ y$" which is either true or false. To describe the relation R, we need to know which pairs (x, y) do satisfy $x\ R\ y$ and which pairs don't. In other words, we need to know the **set** of all pairs $(x, y) \in X \times Y$ that satisfy $x\ R\ y$. For a rigorous definition of a relation, we simply take the relation R to **be** this set of pairs. In other words:

Definition 5.2.1. Let X and Y be two sets. A **relation** from X to Y is a subset of $X \times Y$ (that is, a set of pairs (x, y) with $x \in X$ and $y \in Y$).

If R is a relation from X to Y, and if $(x, y) \in X \times Y$ is any pair, then

- we write $x\ R\ y$ if $(x, y) \in R$;
- we write $x \not R\ y$ if $(x, y) \notin R$.

All the relations we have seen so far can be recast in terms of this definition:

- The divisibility relation $\mid$ is a subset of $\mathbb{Z} \times \mathbb{Z}$, namely the subset

$$\begin{aligned}&\{(x, y) \in \mathbb{Z} \times \mathbb{Z} \mid x \text{ divides } y\} \\ &= \{(x, y) \in \mathbb{Z} \times \mathbb{Z} \mid \text{there exists some } z \in \mathbb{Z} \text{ such that } y = xz\} \\ &= \{(x, xz) \mid x, z \in \mathbb{Z}\}.\end{aligned}$$

  For instance, the pairs $(2, 4)$ and $(3, 9)$ and $(4, 16)$ and $(6, 12)$ belong to this subset, while the pairs $(3, 7)$ and $(3, 5)$ and $(0, 5)$ do not.

- The coprimality relation ("coprime to") is a subset of $\mathbb{Z} \times \mathbb{Z}$, namely the subset

$$\begin{aligned}&\{(x, y) \in \mathbb{Z} \times \mathbb{Z} \mid x \text{ is coprime to } y\} \\ &= \{(x, y) \in \mathbb{Z} \times \mathbb{Z} \mid \gcd(x, y) = 1\}.\end{aligned}$$

  It contains, for instance, $(5, 8)$ and $(4, 9)$ but not $(4, 10)$.

- For any $n \in \mathbb{Z}$, the "congruent modulo n" relation $\overset{n}{\equiv}$ is a subset of $\mathbb{Z} \times \mathbb{Z}$, namely the subset

$$\begin{aligned}&\{(x, y) \in \mathbb{Z} \times \mathbb{Z} \mid x \equiv y \bmod n\} \\ &= \{(x,\ x + nz) \mid x, z \in \mathbb{Z}\}.\end{aligned}$$

- If A is any set, then the **equality relation** on A is the subset E_A of $A \times A$ given by

  $$\begin{aligned} E_A &= \{(x,y) \in A \times A \mid x = y\} \\ &= \{(x,x) \mid x \in A\}. \end{aligned}$$

  Two elements x and y of A satisfy $x \ E_A \ y$ if and only if they are equal.

- We can literally take any subset of $X \times Y$ (where X and Y are two sets) and it will be a relation from X to Y; it does not have to be defined by any "meaningful" rule. For example, here is a relation from $\{1,2,3\}$ to $\{5,6,7\}$:

  $$\{(1,6), (1,7), (3,5)\}.$$

  Equivalently, it can be specified by the table

	5	6	7
1	no	yes	yes
2	no	no	no
3	yes	no	no

  If we call this relation R, then we have $1 \ R \ 6$ and $1 \ R \ 7$ and $3 \ R \ 5$ but $1 \not R \ 5$ etc.

A good way to visualize a relation R from a set X to a set Y (at least when X and Y are finite) is by drawing the sets X and Y as blobs, drawing their elements as nodes within the blob, and drawing an arrow from the x-node to the y-node for each pair $(x,y) \in R$.

5.3. Functions, formally

We can now give a rigorous definition of functions:

> **Definition 5.3.1.** Let X and Y be two sets. A **function** from X to Y means a relation R from X to Y that has the following property:
>
> - **Output uniqueness:** For each $x \in X$, there exists **exactly one** $y \in Y$ such that $x \ R \ y$. (In terms of blobs: From each x-node, exactly one arrow emanates.)
>
> If R is a function from X to Y, and if x is an element of X, then the unique element $y \in Y$ such that $x \ R \ y$ is called $R(x)$.

In our last example, the relation

$$\{(1,6), (1,7), (3,5)\}$$

is not a function from $\{1,2,3\}$ to $\{5,6,7\}$, for two different reasons:

- Output uniqueness fails for $x = 1$, since $1\ R\ y$ holds both for $y = 6$ and for $y = 7$.
- Output uniqueness fails for $x = 2$, since $2\ R\ y$ holds for no $y \in \{5, 6, 7\}$.

We can fix this by replacing the pair $(1, 7)$ by an arbitrary pair of the form $(2, y)$ for some $y \in \{5, 6, 7\}$.

Now we have two definitions of a function: the provisional way ("rule for transforming elements of X into elements of Y") and the formal way ("binary relation that satisfies output uniqueness"). These two definitions are equivalent:

- If f is a function in the provisional sense, then the corresponding formal function is the binary relation
$$\{(x, f(x)) \mid x \in X\}.$$
- Conversely, if R is a function in the formal sense (i.e., a binary relation satisfying output uniqueness), then the corresponding provisional function is the "rule" that sends each $x \in X$ to the unique $y \in Y$ that satisfies $x\ R\ y$.

We shall thus think of the two concepts of a function as interchangeable.

Remark 5.3.2. Strictly speaking, a binary relation from X to Y is not just a subset of $X \times Y$, but really the triple (X, Y, R) where R is a subset of $X \times Y$. This way, the binary relation "remembers" what X and Y are. That is, X and Y are part of the data of a binary relation.

5.4. Some more examples of functions

Here are some more examples and non-examples of functions.

Example 5.4.1. Consider the function

$$f_0 : \{1, 2, 3, 4\} \to \{1, 2, 3, 4\}$$

that sends $1, 2, 3, 4$ to $3, 2, 3, 3$, respectively. As a rigorous/formal function, it is the binary relation R that satisfies

$$1\ R\ 3, \qquad 2\ R\ 2, \qquad 3\ R\ 3, \qquad 4\ R\ 3,$$

and nothing else. In other words, it is the relation

$$\{(1, 3),\ (2, 2),\ (3, 3),\ (4, 3)\}.$$

Example 5.4.2. What about the function

$$\begin{aligned} f_1 : \{1,2,3,4\} &\to \{1,2,3\}, \\ n &\mapsto n \qquad ? \end{aligned}$$

Such a function f_1 does not exist, since it would send 4 to 4, but $4 \notin \{1,2,3\}$.

Lesson learned: Make sure that the expression to the right of the "$\mapsto$" arrow always is an actual element of the target (in our case, $\{1,2,3\}$).

Example 5.4.3. Consider the function

$$\begin{aligned} f_2 : \{1,2,3,\ldots\} &\to \{1,2,3,\ldots\}, \\ n &\mapsto (\text{the number of positive divisors of } n). \end{aligned}$$

As a relation, it is

$$\{(1,1), (2,2), (3,2), (4,3), (5,2), (6,4), (7,2), (8,4), (9,3), \ldots\}.$$

Thus, $f_2(1) = 1$ and $f_2(2) = 2$ and $f_2(3) = 2$ and so on.

Example 5.4.4. What about the function

$$\begin{aligned} \widetilde{f_2} : \mathbb{Z} &\to \{1,2,3,\ldots\}, \\ n &\mapsto (\text{the number of positive divisors of } n) \qquad ? \end{aligned}$$

There is no such function, because $\widetilde{f_2}(0)$ would be undefined (0 has infinitely many positive divisors).

Example 5.4.5. What about the function

$$\begin{aligned} f_3 : \{1,2,3,\ldots\} &\to \{1,2,3,\ldots\}, \\ n &\mapsto (\text{the smallest prime divisor of } n) \qquad ? \end{aligned}$$

There is no such function, since $f_3(1)$ is not defined (1 has no prime divisors, thus no smallest prime divisor).

However, there is a relation "y is the smallest prime divisor of x" from $\{1,2,3,\ldots\}$ to $\{1,2,3,\ldots\}$, but this relation fails output uniqueness for $x = 1$, and thus is not a function.

We can "fix" f_3 by removing 1 from its domain. Thus, we end up with an actual function

$$\begin{aligned} \widetilde{f_3} : \{2,3,4,\ldots\} &\to \{1,2,3,\ldots\}, \\ n &\mapsto (\text{the smallest prime divisor of } n). \end{aligned}$$

Example 5.4.6. What about the function

$$\begin{aligned} f_4 : \mathbb{Q} &\to \mathbb{Z}, \\ \frac{a}{b} &\mapsto a \qquad \text{(for } a, b \in \mathbb{Z} \text{ with } b \neq 0\text{)} \qquad ? \end{aligned}$$

In words: This is a function that takes a rational number as input, writes it as a ratio of two integers, and outputs the numerator. Is there such a function?

Again, the answer is **no**. For instance, it would satisfy both

$$\begin{aligned} f_4(0.5) &= f_4\left(\frac{1}{2}\right) = 1 \qquad \text{and} \\ f_4(0.5) &= f_4\left(\frac{3}{6}\right) = 3, \end{aligned}$$

which clearly cannot hold at the same time. In other words, it fails output uniqueness. The underlying issue is that a rational number can be written as a fraction in several ways, which lead to different numerators.

5.5. Well-definedness

The issues we have seen in the last few examples (supposed functions failing to exist either because their outputs make no sense, or because their outputs don't lie in Y, or because their outputs are ambiguous) are known as **well-definedness** issues. Often, mathematicians say "a function is well-defined" when they mean that this function exists (i.e., its definition does not suffer from these issues).

For example, we just saw that the function

$$\begin{aligned} f_4 : \mathbb{Q} &\to \mathbb{Z}, \\ \frac{a}{b} &\mapsto a \qquad \text{(for } a, b \in \mathbb{Z} \text{ with } b \neq 0\text{)} \end{aligned}$$

is not well-defined (i.e., does not exist). But it can be shown that the function

$$\begin{aligned} f_5 : \mathbb{Q} &\to \mathbb{Q}, \\ \frac{a}{b} &\mapsto \frac{a^2}{b^2} \qquad \text{(for } a, b \in \mathbb{Z} \text{ with } b \neq 0\text{)} \end{aligned}$$

is well-defined. Likewise, the function

$$\begin{aligned} f_1 : \{1, 2, 3, 4\} &\to \{1, 2, 3\}, \\ n &\mapsto n \end{aligned}$$

is not well-defined (since $f_1(4)$ would be $4 \notin \{1,2,3\}$), but the function

$$\begin{aligned} f_6 : \{1,2,3,4\} &\to \{1,2,3\}, \\ n &\mapsto 1 + (n\%3) \end{aligned}$$

is well-defined (its outputs values are $2,3,1,2$),

In general, when you are given a function described in terms of "what it does to an input", you can check that it exists (i.e., is well-defined) in the following way ("checklist"):

1. Show that the output makes sense. (For example, "the number of prime divisors of n" makes no sense when $n = 0$ unless you allow ∞. Our example $\widetilde{f_2}$ failed here.)
2. Show that the output is in the target. (For example, f_1 failed at this step.)
3. Show that the output is unambiguous (i.e., different ways of writing the input give the same output). (For example, f_5 failed at this step.)

5.6. The identity function

> **Definition 5.6.1.** For any set A, there is an **identity function** $\operatorname{id}_A : A \to A$. This is the function that sends each element $a \in A$ to a itself. In other words, it is precisely the relation
>
> $$\begin{aligned} E_A &= \{(x,y) \in A \times A \mid x = y\} \\ &= \{(x,x) \mid x \in A\}. \end{aligned}$$
>
> seen among our examples of relations.

5.7. More examples; multivariate functions

As we said before, a function can be described either by a rule, or by a list of values, or as a relation. For instance, the "take the square" function on real numbers is the function

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R}, \\ x &\mapsto x^2. \end{aligned}$$

As a relation, it is the set

$$\left\{ \left(x, x^2\right) \mid x \in \mathbb{R} \right\}.$$

When the domain of a function f is a Cartesian product of several sets (i.e., its inputs are tuples), f is called a **multivariate** function. For instance, the function

$$\begin{aligned} f : \mathbb{Z} \times \mathbb{Z} &\to \mathbb{Z}, \\ (a,b) &\mapsto a+b \end{aligned}$$

(which sends each pair (a,b) of two integers to their sum $a+b$) is a multivariate function, known as the addition of integers. Its input is a pair of two integers, i.e., it really has two inputs (a and b). As a relation, it is the subset

$$\begin{aligned} &\{((a,b), a+b) \ \mid \ a,b \in \mathbb{Z}\} \\ &= \{((a,b), c) \ \mid \ a,b,c \in \mathbb{Z} \text{ such that } c = a+b\} \end{aligned}$$

of $(\mathbb{Z} \times \mathbb{Z}) \times \mathbb{Z}$. Other multivariate functions are

$$\begin{aligned} \mathbb{Z} \times \mathbb{Z} &\to \mathbb{Z}, \\ (a,b) &\mapsto a-b \end{aligned}$$

and

$$\begin{aligned} \mathbb{Z} \times \mathbb{Z} &\to \mathbb{Z}, \\ (a,b) &\mapsto ab. \end{aligned}$$

We can try to define a multivariate function

$$\begin{aligned} \mathbb{Z} \times \mathbb{Z} &\to \mathbb{Z}, \\ (a,b) &\mapsto a/b, \end{aligned}$$

but this fails for two reasons: a/b makes no sense when $b = 0$, and can fail to be in the target $\mathbb{Z}$ even when $b \neq 0$. We can fix this, e.g., by defining the multivariate function

$$\begin{aligned} \mathbb{Z} \times (\mathbb{Z} \setminus \{0\}) &\to \mathbb{Q}, \\ (a,b) &\mapsto a/b \end{aligned}$$

instead.

When f is a multivariate function whose inputs are k-tuples, we commonly use the shorthand notation $f(a_1, a_2, \ldots, a_k)$ for its values $f((a_1, a_2, \ldots, a_k))$.

5.8. Composition of functions

5.8.1. Definition

The most important ways to transform functions into other functions is **composition**:

Definition 5.8.1. Let X, Y and Z be three sets. Let $f : Y \to Z$ and $g : X \to Y$ be two functions. Then,

$$f \circ g \qquad \text{(pronounced "}f\text{ composed with }g\text{" or "}f\text{ after }g\text{")}$$

denotes the function

$$\begin{aligned} X &\to Z, \\ x &\mapsto f\left(g\left(x\right)\right). \end{aligned}$$

In other words, $f \circ g$ is the function that first applies g and then applies f. It is called the **composition** of f with g.

In terms of relations, if we view f and g as two relations F and G, then $f \circ g$ is the relation

$$\left\{(x, z) \ \mid \ \text{there exists } y \in Y \text{ such that } x \ G \ y \text{ and } y \ F \ z\right\}.$$

Example 5.8.2. Consider the two functions

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R}, \\ x &\mapsto x^3 \end{aligned}$$

and

$$\begin{aligned} g : \mathbb{R} &\to \mathbb{R}, \\ x &\mapsto \frac{1}{x^2+7}. \end{aligned}$$

Then, for any real $x \in \mathbb{R}$, we have

$$\left(f \circ g\right)\left(x\right) = f\left(g\left(x\right)\right) = f\left(\frac{1}{x^2+7}\right) = \left(\frac{1}{x^2+7}\right)^3$$

and

$$\left(g \circ f\right)\left(x\right) = g\left(f\left(x\right)\right) = g\left(x^3\right) = \frac{1}{\left(x^3\right)^2+7} = \frac{1}{x^6+7}.$$

Note that these results are different. Thus, $f \circ g \neq g \circ f$ in general!

Example 5.8.3. Consider the two functions $f : \{1, 2, 3\} \to \{1, 2, 3, 4\}$ and $g : \{1, 2, 3, 4\} \to \{1, 2, 3\}$ given by the following tables of values:

i	1	2	3
$f\left(i\right)$	1	3	2

i	1	2	3	4
$g\left(i\right)$	2	1	3	2

.

The composition $f \circ g : \{1,2,3,4\} \to \{1,2,3,4\}$ then has

$$(f \circ g)(1) = f(g(1)) = g(2) = 3.$$

We can obtain the blobs-and-arrows picture for $f \circ g$ by taking the blobs-and-arrows pictures for f and g (with the target-blob of g reused as the domain-blob of f) and removing the middle blob.

5.8.2. Basic properties

As we saw already, the compositions $f \circ g$ and $g \circ f$ are not usually the same (even if they are both defined, which is not always the case). In other words, composition of functions is not commutative. However, a few other nice properties do hold:

Theorem 5.8.4 (associativity of composition)**.** Let X, Y, Z, W be four sets. Let $f : Z \to W$, $g : Y \to Z$ and $h : X \to Y$ be three functions. Then,

$$(f \circ g) \circ h = f \circ (g \circ h).$$

Proof. Both $(f \circ g) \circ h$ and $f \circ (g \circ h)$ have domain X and target W. Moreover, for each $x \in X$, we have

$$\begin{aligned} ((f \circ g) \circ h)(x) &= (f \circ g)(h(x)) && \text{(by the definition of composition)} \\ &= f(g(h(x))) && \text{(by the definition of composition)} \end{aligned}$$

and

$$\begin{aligned} (f \circ (g \circ h))(x) &= f\left(\underbrace{(g \circ h)(x)}_{=g(h(x))}\right) && \text{(by the definition of composition)} \\ &= f(g(h(x))). \end{aligned}$$

Comparing these, we get

$$((f \circ g) \circ h)(x) = (f \circ (g \circ h))(x).$$

Since this holds for each $x \in X$, we conclude that $(f \circ g) \circ h = f \circ (g \circ h)$ (since two functions with the same domain and the same target are equal if they give the same outputs for the same inputs).

□

This theorem is intuitively quite obvious: It is just saying that when you do three things (applying h, applying g and applying f) in succession, you can group them as you wish: either "first do h followed by g, then do f" or "first do h, then do g followed by f".

Thanks to this theorem, we can write compositions of several functions without parentheses: e.g., we can just write $f \circ g \circ h$.

The following property of composition of functions is even easier. Recall that id_A means the identity map on a given set A (so $\operatorname{id}_A(a) = a$ for each $a \in A$).

Theorem 5.8.5. Let $f : X \to Y$ be a function. Then,

$$f \circ \operatorname{id}_X = \operatorname{id}_Y \circ f = f.$$

Proof. For each $x \in X$, we have

$$(f \circ \operatorname{id}_X)(x) = f\left(\underbrace{\operatorname{id}_X(x)}_{=x}\right) = f(x).$$

Thus, $f \circ \operatorname{id}_X = f$ (since both $f \circ \operatorname{id}_X$ and f are functions from X to Y). Similarly, $\operatorname{id}_Y \circ f = f$. □

5.9. Jectivities (injectivity, surjectivity and bijectivity)

Now we introduce some important properties of functions, which have to do with how often they attain certain values. I call them "jectivity properties", as they are called in-, sur- and bijectivity.

Definition 5.9.1. Let $f : X \to Y$ be a function. Then:

(a) We say that f is **injective** (aka **one-to-one**, aka an **injection**) if

for each $y \in Y$, there exists **at most one** $x \in X$ such that $f(x) = y$.

In other words: We say that f is injective if there are no two distinct elements x_1, x_2 with $f(x_1) = f(x_2)$.

In other words: if each $y \in Y$ is taken as a value at most once by f.

(b) We say that f is **surjective** (aka **onto**, aka a **surjection**) if

for each $y \in Y$, there exists **at least one** $x \in X$ such that $f(x) = y$.

In other words: We say that f is surjective if every element of Y is an output value of f.

(c) We say that f is **bijective** (aka a **one-to-one correspondence**, aka a **bijection**) if

for each $y \in Y$, there exists **exactly one** $x \in X$ such that $f(x) = y$.

Thus, f is bijective if and only if f is both injective and surjective.

In terms of blobs and arrows:

- f is injective if and only if each element of Y is hit at most once;
- f is surjective if and only if each element of Y is hit at least once;
- f is bijective if and only if each element of Y is hit exactly once.

Some examples:

- The function
$$\begin{aligned} f : \mathbb{N} &\to \mathbb{N}, \\ k &\mapsto k^2 \end{aligned}$$
is injective (since different nonnegative integers have different squares) but not surjective (for example, 3 is not an output value, since no nonnegative integer has square 3). So it is not bijective.

- Let $S = \{0, 1, 4, 9, 16, \ldots\}$ be the set of all perfect squares (= squares of integers). Then, the function
$$\begin{aligned} g : \mathbb{N} &\to S, \\ k &\mapsto k^2 \end{aligned}$$
is injective (for the same reasons as f) and surjective (since each perfect square is the square of a nonnegative integer). So it is bijective.

  Take note: $f \neq g$ because the targets are different. Other than that, f and g are indistinguishable. So the choice of target matters for surjectivity.

- Let $S = \{0, 1, 4, 9, 16, \ldots\}$ be the set of all perfect squares again. Consider the function
$$\begin{aligned} g_{\mathbb{Z}} : \mathbb{Z} &\to S, \\ k &\mapsto k^2, \end{aligned}$$
which differs from g only in its choice of domain. This function $g_{\mathbb{Z}}$ is still surjective, but it is not injective, since $g(-2) = g(2)$. Of course, this entails that $g_{\mathbb{Z}}$ is not bijective.

- The function
$$\begin{aligned} h : \mathbb{N} &\to \mathbb{N}, \\ k &\mapsto k//2 = \left\lfloor \frac{k}{2} \right\rfloor \end{aligned}$$
is surjective (since each $y \in \mathbb{N}$ can be written as $h(2y)$) but not injective (since $h(0) = h(1)$).

- Let $E = \{0, 2, 4, 6, \ldots\}$ be the set of all even nonnegative integers. The function
$$\begin{aligned} h_{\text{even}} : E &\to \mathbb{N}, \\ k &\mapsto k//2 = \left\lfloor \frac{k}{2} \right\rfloor \end{aligned}$$
is bijective (= both injective and surjective).

- Let $O = \{1, 3, 5, 7, \ldots\}$ be the set of all odd nonnegative integers. The function
$$\begin{aligned} h_{\text{odd}} : O &\to \mathbb{N}, \\ k &\mapsto k//2 = \left\lfloor \frac{k}{2} \right\rfloor \end{aligned}$$
is bijective, too.

- The function
$$\begin{aligned} f : \mathbb{Z} \times \mathbb{Z} &\to \mathbb{Z}, \\ (a, b) &\mapsto a + b \end{aligned}$$
(that is, addition of integers) is not injective (for example, $f(1,3) = f(2,2)$ or $f(0,1) = f(1,0)$). But it is surjective (since each integer $y \in \mathbb{Z}$ can be written as $f(y, 0)$).

If your function $f : X \to Y$ is represented by a table of values (each column having an $x \in X$ on top and its image $f(x)$ on the bottom), then

- f is injective if and only if the bottom row has no two equal entries;
- f is surjective if and only if the bottom row contains each element of Y at least once;
- f is bijective if and only if the bottom row contains each element of Y exactly once.

5.10. Inverses

5.10.1. Definition and examples

Bijective maps have a special power: they can be **inverted**. Here is what this means:

Definition 5.10.1. Let $f : X \to Y$ be a function. An **inverse** of f means a function $g : Y \to X$ such that

$$f \circ g = \operatorname{id}_Y \qquad \text{and} \qquad g \circ f = \operatorname{id}_X .$$

In other words, an **inverse** of f means a function $g : Y \to X$ such that

$$\begin{aligned} f(g(y)) &= y && \text{for each } y \in Y, && \text{and} \\ g(f(x)) &= x && \text{for each } x \in X. \end{aligned}$$

Roughly speaking, an inverse of f thus means a map that undoes f and is also undone by f.

Not every function f has an inverse. We will soon see which ones do and which ones don't.

5.10.2. Invertibility is bijectivity by another name

Theorem 5.10.2. Let $f : X \to Y$ be a map. Then, f has an inverse if and only if f is bijective.

Proof. This appears in the notes (Theorem 5.10.2). Essentially, the proof of $\Longrightarrow$ is an easy verification, whereas the proof of the $\Longleftarrow$ arrow is made by explicitly constructing an inverse of f: namely, the map

$$\begin{aligned} g : Y &\to X, \\ y &\mapsto (\text{the unique } x \in X \text{ such that } f(x) = y) \end{aligned}$$

(and the uniqueness here comes from the bijectivity of f). □

This is a fundamental result used all over mathematics.

5.10.3. Uniqueness of the inverse

Theorem 5.10.3. Let $f : X \to Y$ be a map. Then, f has at most one inverse.

Proof. What does "at most one inverse" mean? It means that f has no two distinct inverses. In other words, it means that any two inverses of f are identical.

So let us prove this. Let g_1 and g_2 be two inverses of f. We must show that $g_1 = g_2$.

Since g_1 is an inverse of f, we have $g_1 \circ f = \operatorname{id}$ and $f \circ g_1 = \operatorname{id}$ (noting that these two ids are id_X and id_Y, respectively, but we abbreviate them both as id).

Since g_2 is an inverse of f, we have $g_2 \circ f = \operatorname{id}$ and $f \circ g_2 = \operatorname{id}$.

Now, the "triple composition" $g_1 \circ f \circ g_2$ (which makes sense because composition is associative) can be simplified in two ways:

$$g_1 \circ \underbrace{f \circ g_2}_{=\mathrm{id}} = g_1 \circ \mathrm{id} = g_1$$

and

$$\underbrace{g_1 \circ f}_{=\mathrm{id}} \circ g_2 = \mathrm{id} \circ g_2 = g_2.$$

Comparing these two equalities, we conclude $g_1 = g_2$, as desired. □

Definition 5.10.4. Let $f : X \to Y$ be a map that has an inverse. Then, this inverse is unique (by the previous theorem), and will be called f^{-1}.

So

$$f^{-1} \circ f = \mathrm{id}_X \qquad \text{and} \qquad f \circ f^{-1} = \mathrm{id}_Y .$$

Equivalently,

$$\begin{aligned} f^{-1}(f(x)) &= x && \text{for all } x \in X; \\ f\left(f^{-1}(y)\right) &= y && \text{for all } y \in Y. \end{aligned}$$

5.10.4. Examples

Here are some examples of maps that have or have not inverses:

- Let f be the map $\{1,2,3\} \to \{7,8,9\}$, $k \mapsto k+6$. This map f has an inverse, namely the map $\{7,8,9\} \to \{1,2,3\}$, $k \mapsto k-6$.
- Let f be the map $\{1,2,3\} \to \{7,8,9\}$, $k \mapsto 10-k$. This map f has an inverse, namely the map $\{7,8,9\} \to \{1,2,3\}$, $k \mapsto 10-k$. This is because
  $$10-(10-k) = k \qquad \text{for each } k.$$
  More generally,
  $$a-(a-k) = k \qquad \text{for each } k.$$
- Let $\mathbb{R}_{\geq 0} = \{\text{all nonnegative real numbers}\}$. Then, the function
  $$f : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}, \quad x \mapsto x^2$$
  has an inverse. This inverse is the function
  $$f^{-1} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}, \quad x \mapsto \sqrt{x}.$$

- In contrast, the function

  $$\begin{aligned} f : \mathbb{R} &\to \mathbb{R}, \\ x &\mapsto x^2 \end{aligned}$$

  has no inverse, since it is not bijective. It is neither injective nor surjective.

- The function

  $$\begin{aligned} f : \mathbb{R} &\to \mathbb{R}, \\ x &\mapsto x^3 \end{aligned}$$

  has an inverse, namely

  $$\begin{aligned} f^{-1} : \mathbb{R}_{\geq 0} &\to \mathbb{R}_{\geq 0}, \\ x &\mapsto \sqrt[3]{x}. \end{aligned}$$

- Let $E = \{0, 2, 4, 6, 8, \ldots\}$. The function

  $$\begin{aligned} f : E &\to \mathbb{N}, \\ k &\mapsto k/2 \end{aligned}$$

  has an inverse, namely,

  $$\begin{aligned} f^{-1} : E &\to \mathbb{N}, \\ k &\mapsto 2k. \end{aligned}$$

5.10.5. Inverses of inverses and compositions

Let us prove some general facts about inverses:

Proposition 5.10.5. Let X be any set. Then, the identity map $\operatorname{id}_X : X \to X$ is bijective, and its inverse is itself.

Proof. $\operatorname{id}_X \circ \operatorname{id}_X = \operatorname{id}_X$. □

Theorem 5.10.6. Let $f : X \to Y$ be a map that has an inverse $f^{-1} : Y \to X$. Then, f^{-1} also has an inverse, namely f.

Proof. Since f^{-1} is an inverse of f, we have $f \circ f^{-1} = \operatorname{id}$ and $f^{-1} \circ f = \operatorname{id}$. But the same two equalities are saying that f is an inverse of f^{-1}. □

Theorem 5.10.7 (Socks-and-shoes formula)**.** Let $g : X \to Y$ and $f : Y \to Z$ be two bijective maps. Then, the composition $f \circ g : X \to Z$ is bijective as well, and its inverse is

$$(f \circ g)^{-1} = g^{-1} \circ f^{-1}.$$

Proof. We need to show that $g^{-1} \circ f^{-1}$ is an inverse of $f \circ g$. This is obvious from blobs and arrows, but can also be checked rigorously:

$$\begin{aligned}\left(g^{-1} \circ f^{-1}\right) \circ (f \circ g) &= g^{-1} \circ \underbrace{f^{-1} \circ f}_{=\mathrm{id}} \circ g = g^{-1} \circ \underbrace{\mathrm{id} \circ g}_{=g} \\ &= g^{-1} \circ g = \mathrm{id}\end{aligned}$$

and similarly $(f \circ g) \circ \left(g^{-1} \circ f^{-1}\right) = \mathrm{id}$. □

Note that $(f \circ g)^{-1} \neq f^{-1} \circ g^{-1}$ in general. In fact, $f^{-1} \circ g^{-1}$ might not even exist.

There is a section in the notes (§5.11) which contains several solved exercises.

5.11. Isomorphic sets

> **Definition 5.11.1.** Let X and Y be two sets. We say that these two sets X and Y are **isomorphic as sets** (or, for short, **isomorphic**, or **in bijection**, or **in one-to-one correspondence**, or **equinumerous**) if there exists a bijective map from X to Y.

This relation "isomorphic" is symmetric (i.e., if X and Y are isomorphic, then Y and X are isomorphic), because any bijection has an inverse which is also a bijection.

Some examples:

- The sets $\{1,2\}$ and $\{1,2,3\}$ are **not** isomorphic. In fact, there is no surjective map $f : \{1,2\} \to \{1,2,3\}$, thus no bijective map either.

- The sets $\{1,2,3\}$ and $\{7,8,9\}$ are isomorphic. In fact,

$$\begin{aligned}\{1,2,3\} &\to \{7,8,9\}, \\ k &\mapsto k+6\end{aligned}$$

  is a bijective map.

- The sets $\{1,2,3\}$ and $\{3,8,9\}$ are isomorphic. For example, this is because the map $f : \{1,2,3\} \to \{3,8,9\}$ with values

x	1	2	3
$f(x)$	3	8	9

  is bijective.

- The sets $\{1,2,3\}$ and $\{1,3,5\}$ are isomorphic. In fact, the map
$$\begin{aligned}\{1,2,3\} &\to \{1,3,5\},\\ k &\mapsto 2k-1\end{aligned}$$
is bijective.

- The sets $\mathbb{N}$ and $E := \{\text{all even nonnegative integers}\}$ are isomorphic, since the map
$$\begin{aligned}\mathbb{N} &\to E,\\ n &\mapsto 2n\end{aligned}$$
is bijective.

- The sets $\mathbb{N}$ and $O := \{\text{all odd nonnegative integers}\}$ are isomorphic, since the map
$$\begin{aligned}\mathbb{N} &\to O,\\ n &\mapsto 2n+1\end{aligned}$$
is a bijection.

- The sets $\mathbb{N}$ and $\mathbb{Z}$ are isomorphic, since there is a bijection from $\mathbb{N}$ to $\mathbb{Z}$ that sends
$$\begin{aligned}&0,1,2,3,4,5,6,7,8,\ldots && \text{to}\\ &0,-1,1,-2,2,-3,3,-4,4,-5,5,\ldots, && \text{respectively.}\end{aligned}$$
Explicitly, this map f can be defined by the following formula (note: not the same as in the notes)
$$\begin{aligned}f(n) &= \begin{cases} n/2, & \text{if } n \text{ is even;}\\ -(n+1)/2, & \text{if } n \text{ is odd}\end{cases}\\ &= (-1)^n \lceil n/2 \rceil .\end{aligned}$$

- The sets $\mathbb{N}$ and $\mathbb{Q} = \left\{ \frac{a}{b} \mid a,b \in \mathbb{Z} \text{ and } b \neq 0 \right\}$ are isomorphic, since there is a bijection from $\mathbb{N}$ to $\mathbb{Q}$ that sends
$$0,1,2,3,4,5,6,7,8,\ldots \qquad \text{to}$$
$$\underbrace{\frac{-1}{1},\frac{0}{1},\frac{1}{1}}_{\substack{\text{all reduced fractions}\\ \text{whose num and denom}\\ \text{are } \leq 1 \text{ in absolute value}\\ \text{(ordered from min to max)}}},\quad \underbrace{\frac{-2}{2},\frac{-1}{2},\frac{1}{2},\frac{2}{1}}_{\substack{\text{all reduced fractions}\\ \text{whose num and denom}\\ \text{are } \leq 2 \text{ but not } \leq 1 \text{ in abs val}\\ \text{(ordered from min to max)}}},\quad \underbrace{\frac{-3}{1},\frac{-3}{2},\frac{-2}{3},\frac{-1}{3},\frac{1}{3},\frac{2}{3},\frac{3}{2},\frac{3}{1}}_{\substack{\text{all reduced fractions}\\ \text{whose num and denom}\\ \text{are } \leq 3 \text{ but not } \leq 2\\ \text{(ordered from min to max)}}},\ldots \qquad \text{respectivel}$$
(To be fully precise, we must only allow **fully reduced fractions**, i.e., fractions $\frac{a}{b}$ with $a \in \mathbb{Z}$ and $b \in \{1,2,3,\ldots\}$ and $\gcd(a,b) = 1$.)

- The sets $\mathbb{N}$ and $\mathbb{N} \times \mathbb{N}$ are isomorphic, since there is a bijection f from $\mathbb{N}$ to $\mathbb{N} \times \mathbb{N}$ that sends

$$\begin{array}{c} 0,1,2,3,4,5,6,7,8,9,\ldots \qquad \text{to} \\ \underbrace{(0,0)}_{\substack{\text{the pairs} \\ \text{with sum 0}}}, \quad \underbrace{(1,0),\ (0,1)}_{\substack{\text{the pairs with sum 1} \\ \text{(in order of increasing} \\ \text{y-coordinate/second entry)}}}, \quad \underbrace{(2,0),\ (1,1),\ (0,2)}_{\substack{\text{the pairs with sum 2} \\ \text{(in order of increasing} \\ \text{y-coordinate/second entry)}}}, \ \ldots. \end{array}$$

  Actually, there is an explicit formula for the inverse $f^{-1} : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ of this bijection:

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

  This is the so-called Cantor pairing function.

- The sets $\mathbb{N}$ and $\mathbb{R}$ are **not** isomorphic; the set $\mathbb{R}$ has "a lot more" elements than $\mathbb{N}$. This is called "uncountability of $\mathbb{R}$", and a proof can be find in any text on set theory.

6. Enumeration revisited

6.1. Counting, formally

6.1.1. Definition

As you have surely noticed, isomorphic sets have the same number of elements (i.e., size). We will now use this to **define** the size of a set!

First some notations:

Definition 6.1.1. **(a)** If $n \in \mathbb{N}$, then $[n]$ means the set $\{1,2,\ldots,n\}$.

For instance, $[5] = \{1,2,3,4,5\}$ and $[1] = \{1\}$ and $[0] = \varnothing$.

(b) If $a,b \in \mathbb{Z}$, then $[a,b]$ means the set

$$\{a, a+1, a+2, \ldots, b\} = \{x \in \mathbb{Z} \mid a \leq x \leq b\}.$$

This is called an **integer interval**. If $a > b$, then this is the empty set.

Now we can define the size of a finite set:

Definition 6.1.2. Let $n \in \mathbb{N}$. A set S is said to have **size** n if S is isomorphic to $[n]$ (that is, if there is a bijection from S to $[n]$).

Examples:

- The set {cat, dog, rat} has size 3, since the map

$$\begin{aligned} \{\text{cat, dog, rat}\} &\to [3], \\ \text{cat} &\mapsto 1, \\ \text{dog} &\mapsto 2, \\ \text{rat} &\mapsto 3. \end{aligned}$$

- The set $[4,7] = \{4,5,6,7\}$ has size 4, since the map

$$\begin{aligned} [4,7] &\to [4], \\ x &\mapsto x-3 \end{aligned}$$

  is a bijection.

- The set $\mathbb{N}$ is infinite, so there is no bijection from $\mathbb{N}$ to $[n]$ for any $n \in \mathbb{N}$. So $\mathbb{N}$ does not have size n for any $n \in \mathbb{N}$.

Here is another equivalent definition of size:

Definition 6.1.3. We define the notion of a "set of size n" recursively as follows:

(a) A set S is said to have **size** 0 if and only if it is empty.

(b) Let n be a positive integer. A set S is said to have **size** n if and only if there exists an $s \in S$ such that $S \setminus \{s\}$ has size $n-1$.

In other words, a set S has size n (for $n > 0$) if and only if we can remove a single element from it and obtain a set of size $n-1$.

The following is not obvious, but can be proved:

Theorem 6.1.4. (a) The above two definitions of size are equivalent.

(b) The size of a finite set is determined uniquely – i.e., a set cannot have two different sizes.

Now we introduce some notations for sizes of sets:

Definition 6.1.5. (a) An **n-element set** (for some $n \in \mathbb{N}$) means a set of size n.

(b) A set is said to be **finite** if it has size n for some $n \in \mathbb{N}$.

(c) If S is a finite set, then $|S|$ shall denote the size of S.

(LaTeX: | or \left| ... \right|. Or use \abs{...} if you are using my macros.)

(d) This size $|S|$ is also called the **cardinality** of S, or the **number** of elements of S.

For example,

$$|\{\text{cat, dog, rat}\}| = 3 \qquad \text{and} \qquad |[4,7]| = 4.$$

The number of odd integers between 4 and 10 is

$$|\{\text{odd integers between 4 and 10}\}| = |\{5,7,9\}| = 3.$$

6.1.2. Rules for sizes of finite sets

Let us now formally state some of the basic counting rules we have been using so far:

Theorem 6.1.6 (Bijection principle)**.** Let A and B be two finite sets. Then, $|A| = |B|$ if and only if there exists a bijection from A to B.

Theorem 6.1.7. For each $n \in \mathbb{N}$, we have $|[n]| = n$.

Theorem 6.1.8. Let S be a set. Then:
(a) We have $|S| = 0$ if and only if $S = \varnothing$.
(b) We have $|S| = 1$ if and only if $S = \{s\}$ for a single element s.
(c) We have $|S| = 2$ if and only if $S = \{s, t\}$ for two distinct elements s, t.

Theorem 6.1.9. Let S be a finite set. Let t be an object such that $t \notin S$. Then,

$$|S \cup \{t\}| = |S| + 1.$$

Theorem 6.1.10 (Sum rule for two sets)**.** Let A and B be two disjoint finite sets (i.e., we have $A \cap B = \varnothing$). Then, $A \cup B$ is again finite, and

$$|A \cup B| = |A| + |B|.$$

Theorem 6.1.11 (Sum rule for k sets)**.** Let $A_1, A_2, \ldots, A_k$ be k disjoint finite sets (i.e., we have $A_i \cap A_j = \varnothing$ for all $i \neq j$). Then, the set $A_1 \cup A_2 \cup \cdots \cup A_k$ is again finite, and

$$|A_1 \cup A_2 \cup \cdots \cup A_k| = |A_1| + |A_2| + \cdots + |A_k|.$$

Theorem 6.1.12 (Difference rule)**.** Let T be a subset of a finite set S. Then:
(a) The set T is finite, and its size $|T|$ is $|T| \leq |S|$.
(b) We have $|S \setminus T| = |S| - |T|$.
(c) If $|T| = |S|$, then $T = S$.

Theorem 6.1.13 (Product rule for k sets)**.** Let $A_1, A_2, \ldots, A_k$ be k finite sets. Then, the set

$$\begin{aligned} & A_1 \times A_2 \times \cdots \times A_k \\ & = \{\text{all } k\text{-tuples } (a_1, a_2, \ldots, a_k) \text{ with } a_i \in A_i \text{ for each } i \in [k]\} \end{aligned}$$

is again finite and has size

$$|A_1 \times A_2 \times \cdots \times A_k| = |A_1| \cdot |A_2| \cdot \cdots \cdot |A_k|.$$

From these basic rules, various other rules can be derived, such as:

Theorem 6.1.14. Let A and B be two finite sets (not necessarily disjoint). Then, $A \cup B$ is again finite, and we have

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

(See §6.1.3 in the notes for more details here.)

6.2. Redoing some proofs rigorously

(I'll be brief here; see §6.2 in the notes.)

We now have the tools to redo some of our above informal proofs formally.

6.2.1. Integers in an interval

Proposition 6.2.1. Let $a, b \in \mathbb{Z}$ be such that $a \leq b + 1$. Then,

$$|[a, b]| = b - a + 1.$$

Proof. We have a bijection

$$\begin{aligned} f : [a, b] &\to [b - a + 1] , \\ x &\mapsto x - (a - 1) . \end{aligned}$$

Thus, the set $[a, b]$ is isomorphic to $[b - a + 1]$, and so has size $b - a + 1$. □

6.2.2. Counting all subsets

Recall:

Theorem 6.2.2. Let $n \in \mathbb{N}$. Then,

$$(\text{\# of subsets of } [n]) = 2^n.$$

Proof. We induct on n.

Base case ($n = 0$)*:* This is saying that (# of subsets of $[0]$) $= 2^0$. But this is clear, since the set $[0]$ is empty and thus has only 1 subset (itself).

Induction step: We proceed from $n - 1$ to n. So let n be a positive integer. We assume (as IH) that the theorem holds for $n - 1$ instead of n. Our goal is to prove that it holds for n.

So our IH says that

$$(\text{\# of subsets of } [n-1]) = 2^{n-1}.$$

Our goal is to prove that

$$(\text{\# of subsets of } [n]) = 2^{n}.$$

We define

- a **red set** to be a subset of $[n]$ that contains n;
- a **green set** to be a subsets of $[n]$ that does not contain n.

Each subset of $[n]$ is either red or green but not both. Hence,

$$\{\text{subsets of } [n]\} = \{\text{red sets}\} \cup \{\text{green sets}\}.$$

Therefore,

$$\begin{aligned} |\{\text{subsets of } [n]\}| &= |\{\text{red sets}\} \cup \{\text{green sets}\}| \\ &= |\{\text{red sets}\}| + |\{\text{green sets}\}| \end{aligned}$$

by the sum rule for two sets (since the sets $\{\text{red sets}\}$ and $\{\text{green sets}\}$ are disjoint). In other words,

$$(\text{\# of subsets of } [n]) = (\text{\# of red sets}) + (\text{\# of green sets}).$$

Now let us count the red sets and the green sets separately.

- The green sets are easy: They are just the subsets of $[n-1]$. So

  $$(\text{\# of green sets}) = (\text{\# of subsets of } [n-1]) = 2^{n-1}$$

  (by the IH).
- The problem of counting the red sets can be reduced to counting the green sets: Indeed, we claim that the sets $\{\text{red sets}\}$ and $\{\text{green sets}\}$ are isomorphic. To prove this, we need to find a bijection between these two sets. We observe that the maps

  $$\begin{aligned} \text{ins}_n : \{\text{green sets}\} &\to \{\text{red sets}\}, \\ G &\mapsto G \cup \{n\} \end{aligned}$$

  and

  $$\begin{aligned} \text{rem}_n : \{\text{red sets}\} &\to \{\text{green sets}\}, \\ R &\mapsto R \setminus \{n\} \end{aligned}$$

are well-defined and mutually inverse (i.e., we have $\operatorname{ins}_n(\operatorname{rem}_n(R)) = R$ for each red set R, and we have $\operatorname{rem}_n(\operatorname{ins}_n(G)) = G$ for each green set G), and thus are bijective. So the sets $\{\text{red sets}\}$ and $\{\text{green sets}\}$ are isomorphic. Therefore, the bijection principle says that

$$\left|\{\text{red sets}\}\right| = \left|\{\text{green sets}\}\right|.$$

In other words,

$$(\text{\# of red sets}) = (\text{\# of green sets}) = 2^{n-1}.$$

Combining what we have shown,

$$\begin{aligned}(\text{\# of subsets of } [n]) &= \underbrace{(\text{\# of red sets})}_{=2^{n-1}} + \underbrace{(\text{\# of green sets})}_{=2^{n-1}} \\ &= 2^{n-1} + 2^{n-1} = 2 \cdot 2^{n-1} = 2^n,\end{aligned}$$

as desired. So the induction step is complete, and the theorem proved. □

More generally:

Theorem 6.2.3. Let $n \in \mathbb{N}$. Let S be any n-element set. Then,

$$(\text{\# of subsets of } S) = 2^n.$$

Proof. Informally, this follows from the previous theorem, since we can rename the n elements of S as $1, 2, \ldots, n$.

Formally, instead of renaming, we have to use a bijection: Since S is an n-element set, there exists a bijection $\alpha : S \to [n]$. Consider this α. Then,

$$\begin{aligned}\alpha_* : \{\text{subsets of } S\} &\to \{\text{subsets of } [n]\}, \\ T &\mapsto \{\alpha(t) \mid t \in T\}\end{aligned}$$

is a bijection as well. For instance, if $\alpha(\text{cat}) = 1$ and $\alpha(\text{dog}) = 2$ and $\alpha(\text{rat}) = 3$, then $\alpha_*(\{\text{cat}, \text{rat}\}) = \{1, 3\}$.

Why is α_* a bijection? Because the map

$$\begin{aligned}\left(\alpha^{-1}\right)_* : \{\text{subsets of } [n]\} &\to \{\text{subsets of } S\}, \\ T &\mapsto \left\{\alpha^{-1}(t) \mid t \in T\right\}\end{aligned}$$

is inverse to α_* (easy to verify).

So the bijection principle yields

$$\left|\{\text{subsets of } S\}\right| = \left|\{\text{subsets of } [n]\}\right|.$$

In other words,

$$(\text{\# of subsets of } S) = (\text{\# of subsets of } [n]) = 2^n$$

by the previous theorem. □

6.2.3. Counting all k-element subsets

Recall:

> **Theorem 6.2.4** (combinatorial interpretation of BCs)**.** Let $n \in \mathbb{N}$, and let k be any number. Let S be an n-element set. Then,
>
> $$(\text{\# of } k\text{-element subsets of } S) = \binom{n}{k}.$$

Our informal proof of this fact used the colors red, green and blue. We can formalize this proof in the same way as we did for the proof of $(\text{\# of subsets of } S) = 2^n$. See §6.2.3 in the notes for details.

> **Corollary 6.2.5.** Let $n \in \mathbb{N}$. Then,
>
> $$\sum_{k=0}^{n} \binom{n}{k} = 2^n.$$

Proof. We have

$$\begin{aligned} 2^n &= (\text{\# of subsets of } [n]) \\ &= \sum_{k=0}^{n} \underbrace{(\text{\# of } k\text{-element subsets of } [n])}_{=\binom{n}{k}} \\ &\qquad \text{(by the sum rule for } k \text{ sets)} \\ &= \sum_{k=0}^{n} \binom{n}{k}. \end{aligned}$$

□

6.3. Lacunar subsets

6.3.1. Definition

Another type of objects we can count are the so-called **lacunar subsets** (aka **sparse subsets**). Here is their definition:

Definition 6.3.1. A set S of integers is said to be **lacunar** if it contains no two consecutive integers (i.e., if there is no $i \in \mathbb{Z}$ such that both i and $i+1$ are in S).

For example, $\{\text{all odd integers}\}$ is lacunar, as is $\{1,3,5,10,27\}$ or $\{0,2\}$ or any singleton set $\{s\}$ or the empty set $\varnothing$. On the other hand, $\{\text{prime numbers}\}$ is not lacunar, since 2 and 3 are consecutive and both in this set. Also, $\{2,3\}$ and $\{1,4,5,9\}$ are not lacunar.

Now we can ask ourselves some questions: For given $n \in \mathbb{N}$,

1. how many lacunar subsets does the set $[n] = \{1,2,\ldots,n\}$ have?
2. how many k-element lacunar subsets does $[n]$ have for a given $k \in \mathbb{N}$?
3. what is the largest size of a lacunar subset of $[n]$?

In this section, we shall answer all three.

6.3.2. The maximum size of a lacunar subset

We start with the third question:

Proposition 6.3.2. Let $n \in \mathbb{N}$. The maximum size of a lacunar subset of $[n]$ is $\left\lfloor \dfrac{n+1}{2} \right\rfloor = \left\lceil \dfrac{n}{2} \right\rceil$.

Proof. The set

$$\{\text{all odd numbers in } [n]\} = \{1,3,5,7,\ldots\} \cap [n]$$

is a lacunar subset of $[n]$, and has size $\left\lfloor \dfrac{n+1}{2} \right\rfloor = \left\lceil \dfrac{n}{2} \right\rceil$. Thus, the size $\left\lfloor \dfrac{n+1}{2} \right\rfloor = \left\lceil \dfrac{n}{2} \right\rceil$ is attainable.

It remains to show that this size is the largest possible – i.e., that if L is a lacunar subset of $[n]$, then

$$|L| \leq \left\lfloor \frac{n+1}{2} \right\rfloor .$$

To prove this, we fix a lacunar subset L of $[n]$. We shall first show that $|L| \leq \dfrac{n+1}{2}$. Once this is proved, $|L| \leq \left\lfloor \dfrac{n+1}{2} \right\rfloor$ will follow (since $|L|$ is an integer).

Let L^+ be the set $\{\ell + 1 \mid \ell \in L\}$. Then, the sets L and L^+ have the same size, but are disjoint (since L is lacunar, so that $\ell \in L$ entails $\ell + 1 \notin L$). So the sum rule yields

$$\begin{aligned} \left|L \cup L^+\right| &= |L| + \left|L^+\right| = |L| + |L| \qquad (\text{since } L \text{ and } L^+ \text{ have the same size}) \\ &= 2 \cdot |L| . \end{aligned}$$

But $L \cup L^+$ is a subset of $[n+1]$. Thus, $|L \cup L^+| \leq |[n+1]| = n+1$. In other words, $2 \cdot |L| \leq n+1$, so that $|L| \leq \dfrac{n+1}{2}$. And as we said, this leads to $|L| \leq \left\lfloor \dfrac{n+1}{2} \right\rfloor$, since any integer that is $\leq \dfrac{n+1}{2}$ must also be $\leq \left\lfloor \dfrac{n+1}{2} \right\rfloor$.

(See the notes for a different proof.) □

6.3.3. Counting all lacunar subsets of $[n]$

Now let us get to our first question: We want to count all the lacunar subsets of $[n]$.

Theorem 6.3.3. For any $n \geq -1$, we have

$$(\text{\# of lacunar subsets of } [n]) = f_{n+2},$$

where $f_0, f_1, f_2, \ldots$ are the Fibonacci numbers ($f_0 = 0$ and $f_1 = 1$ and $f_i = f_{i-1} + f_{i-2}$).

Here, we agree that $[-1] = \varnothing$ (and more generally: $[k] = \varnothing$ for all $k < 0$).

Proof. For any integer $n \geq -1$, let us set

$$\ell_n := (\text{\# of lacunar subsets of } [n]).$$

So we must prove that

$$\ell_n = f_{n+2} \qquad \text{for each } n \geq -1.$$

We have $\ell_{-1} = 1$ (since the set $[-1] = \varnothing$ has only one lacunar subset, namely $\varnothing$ itself) and $f_{-1+2} = f_1 = 1$. So our claim holds for $n = -1$. Similarly, our claim holds for $n = 0$.

Let us next show the following:

Claim 1: We have $\ell_n = \ell_{n-1} + \ell_{n-2}$ for each integer $n \geq 1$.

Proof of Claim 1. Fix an integer $n \geq 1$. We declare a subset of $[n]$ to be

- **red** if it contains n, and
- **green** if it does not contain n.

Then,

$$\begin{aligned}\ell_n &= (\text{\# of lacunar subsets of } [n]) \\ &= (\text{\# of red lacunar subsets of } [n]) + (\text{\# of green lacunar subsets of } [n]).\end{aligned}$$

The green lacunar subsets of $[n]$ are just the lacunar subsets of $[n-1]$. Thus,

$$\begin{aligned}&(\text{\# of green lacunar subsets of } [n]) \\ &= (\text{\# of lacunar subsets of } [n-1]) = \ell_{n-1}.\end{aligned}$$

Now let us count the red lacunar subsets of $[n]$. If R is a red lacunar subset of $[n]$, then R contains n, so that R does **not** contain $n-1$ (since R is lacunar), and therefore $R \setminus \{n\}$ is a lacunar subset of $[n-2]$. Thus, we obtain a map

$$\begin{aligned}\operatorname{rem}_n : \{\text{red lacunar subsets of } [n]\} &\to \{\text{lacunar subsets of } [n-2]\}, \\ R &\mapsto R \setminus \{n\}.\end{aligned}$$

Conversely, if L is a lacunar subset of $[n-2]$, then $L \cup \{n\}$ is a red lacunar subset of $[n]$. Thus, we obtain a map

$$\begin{aligned}\operatorname{ins}_n : \{\text{lacunar subsets of } [n-2]\} &\to \{\text{red lacunar subsets of } [n]\}, \\ L &\mapsto L \cup \{n\}.\end{aligned}$$

These two maps rem_n and ins_n are mutually inverse (since one removes the n and the other inserts an n). So they are bijections, and thus the bijection principle yields

$$|\{\text{red lacunar subsets of } [n]\}| = |\{\text{lacunar subsets of } [n-2]\}|.$$

In other words,

$$\begin{aligned}&(\text{\# of red lacunar subsets of } [n]) \\ &= (\text{\# of lacunar subsets of } [n-2]) = \ell_{n-2}.\end{aligned}$$

Now recall that

$$\begin{aligned}\ell_n &= \underbrace{(\text{\# of red lacunar subsets of } [n])}_{=\ell_{n-2}} + \underbrace{(\text{\# of green lacunar subsets of } [n])}_{=\ell_{n-1}} \\ &= \ell_{n-2} + \ell_{n-1} = \ell_{n-1} + \ell_{n-2}.\end{aligned}$$

This proves Claim 1. □

Claim 1 does not directly show that the sequence $(\ell_{-1}, \ell_0, \ell_1, \ell_2, \ldots)$ is the (shifted) Fibonacci sequence $(f_1, f_2, f_3, \ldots)$; it only shows that the two sequences satisfy the same recurrence (each entry after the first two is the sum of two previous entries). However, we also know that the first two entries of both sequences are the same ($\ell_{-1} = 1 = f_1$ and $\ell_0 = 1 = f_2$). So the two sequences have the same starting values and the same recurrence. Thus, they are equal. (Strictly speaking, this is a strong induction.) □

6.3.4. Counting all k-element lacunar subsets of $[n]$

Let us now address the remaining question about lacunar subsets: how many are there of a given size k ?

Theorem 6.3.4. Let $n \in \mathbb{Z}$ and $k \in \mathbb{N}$ be such that $k \leq n+1$. Then,

$$(\text{\# of } k\text{-element lacunar subsets of } [n]) = \binom{n+1-k}{k}.$$

For instance, for $n = 7$ and $k = 3$, this yields

$$(\text{\# of 3-element lacunar subsets of } [7]) = \binom{7+1-3}{3} = \binom{5}{3} = 10.$$

Note that the condition $k \leq n+1$ in the theorem was chosen to be as liberal as possible (even allowing silly cases such as $k = n+1$). We cannot drop it entirely, since for $k > n+1$ the LHS is still 0 but the RHS becomes nonzero again.

It is possible to prove the theorem by induction, using a red/green dichotomy. This requires an annoying case distinction because of the $k \leq n+1$ condition.

Here is a nicer proof. We will use the following very basic fact:

Proposition 6.3.5. Let $k \in \mathbb{N}$. Let S be a k-element set of integers. Then, there exists a unique k-tuple $(s_1, s_2, \ldots, s_k)$ of integers satisfying

$$S = \{s_1, s_2, \ldots, s_k\} \qquad \text{and} \qquad s_1 < s_2 < \cdots < s_k.$$

This is just saying that any finite set of integers can be uniquely listed in increasing order (smallest element, second-smallest element, and so on). There is a proof by induction, but we skip it since the fact is so self-evident.

We abbreviate the statement

$$S = \{s_1, s_2, \ldots, s_k\} \qquad \text{and} \qquad s_1 < s_2 < \cdots < s_k$$

as

$$S = \{s_1 < s_2 < \cdots < s_k\}.$$

For example,

$$\{\text{odd integers between 0 and 10}\} = \{1 < 3 < 5 < 7 < 9\}.$$

Proof of the theorem. Let $m = n+1-k$. Then, $m \geq 0$ since $k \leq n+1$. Hence $[m]$ is an m-element set.

Now, if $S = \{s_1 < s_2 < \cdots < s_k\}$ is a k-element lacunar subset of $[n]$, then $\overleftarrow{S}$ will mean the set

$$\{s_i - (i-1) \mid i \in [k]\} = \{s_1,\ s_2 - 1,\ s_3 - 2,\ \ldots,\ s_k - (k-1)\}.$$

In other words, $\overleftarrow{S}$ is obtained from S by

- leaving the smallest element of S unchanged,
- decreasing the second-smallest element by 1,
- decreasing the third-smallest by 2,
- and so on.

We call $\overleftarrow{S}$ the **compression** of S. We can easily show that $\overleftarrow{S}$ is a k-element subset of $[m]$. Indeed,

$$\overleftarrow{S} = \{s_1 < s_2 - 1 < s_3 - 2 < \cdots < s_k - (k-1)\}$$

(the inequalities follow from the lacunarity of S: it gives $s_i < s_{i+1} - 1$ and thus $s_i - (i-1) < s_i - i$) shows that $\overleftarrow{S}$ is a k-element set. Moreover, $\overleftarrow{S}$ is a subset of $[m]$ since its highest element is

$$\underbrace{s_k}_{\leq n} - (k-1) \leq n - (k-1) = n - k + 1 = m.$$

Thus can define a map

$$\text{compress} : \{k\text{-element lacunar subsets of } [n]\} \to \{k\text{-element subsets of } [m]\},$$
$$S \mapsto \overleftarrow{S}.$$

Conversely, if $T = \{t_1 < t_2 < \cdots < t_k\}$ is a k-element subset of $[m]$, then $\overrightarrow{T}$ shall mean the set

$$\{t_i + (i-1) \mid i \in [k]\} = \{t_1, \ t_2 + 1, \ t_3 + 2, \ \ldots, \ t_k + (k-1)\}.$$

We can easily see that $\overrightarrow{T}$ is a k-element lacunar subset of $[n]$. So we obtain a map

$$\text{expand} : \{k\text{-element subsets of } [m]\} \to \{k\text{-element lacunar subsets of } [n]\},$$
$$T \mapsto \overrightarrow{T}.$$

It is straightforward to show that these two maps compress and expand are mutually inverse. So they are both bijections, and thus the bijection principle yields

$$\begin{aligned}
&(\#\text{ of } k\text{-element lacunar subsets of } [n]) \\
&= (\#\text{ of } k\text{-element subsets of } [m]) \\
&= \binom{m}{k} \qquad \text{(by the combinatorial interpretation of BCs)} \\
&= \binom{n+1-k}{k} \qquad (\text{since } m = n+1-k).
\end{aligned}$$

□

6.3.5. A corollary

Corollary 6.3.6. Let $n \in \mathbb{N}$. Then, the Fibonacci number f_{n+1} is

$$f_{n+1} = \sum_{k=0}^{n} \binom{n-k}{k} = \binom{n-0}{0} + \binom{n-1}{1} + \cdots + \binom{n-n}{n}.$$

Proof. Renaming n as $n+1$, we can restate this as

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

Now, the LHS is the # of all lacunar subsets of $[n]$, whereas the RHS is the # of all k-element lacunar subsets of $[n]$, summed over all $k \in \{0, 1, \ldots, n+1\}$. By the sum rule, the two numbers are equal. □

6.4. Selections

We now come back to a class of problems we have posed at the start of Chapter 4: Given an n-element set S, how many ways are there to select k elements from S (where n and k are fixed)?

This is an ambiguous question, since it leaves us the following to decide:

1. Do we want k arbitrary elements or k distinct elements? (“selections with/without replacement”)

2. Does the order of these k elements matter or not? (So should we think of “1, 2” and “2, 1” as two different selections?)

So we have four different counting problems.

6.4.1. Unordered selections without replacement

We begin with the case in which we want to select k distinct elements, and the order does not matter. So we are selecting a k-element subset of S. We know the answer:

Theorem 6.4.1. Let $n \in \mathbb{N}$, and let k be any number. Let S be an n-element set. Then,

$$(\text{\# of } k\text{-element subsets of } S) = \binom{n}{k}.$$

6.4.2. Ordered selections without replacement

Now let us care about the order (but still require the k elements to be distinct). Then we are counting not the k-element **subsets** but the **k-tuples**. However, we don't want **all** the k-tuples, but only the ones that consist of **distinct** elements. I shall call such k-tuples **injective**:

Definition 6.4.2. Let $k \in \mathbb{N}$. A k-tuple $(i_1, i_2, \ldots, i_k)$ is said to be **injective** if all its k entries are distinct (i.e., if $i_a \neq i_b$ for all $a \neq b$).

So $(6, 1, 3)$ is injective, but $(6, 1, 6)$ is not. So we want to count the injective k-tuples.

Definition 6.4.3. Let S be a set, and let $k \in \mathbb{N}$. Then, S^k means the Cartesian product $\underbrace{S \times S \times \cdots \times S}_{k \text{ times}}$; it consists of the k-tuples of elements of S.

Theorem 6.4.4. Let $n, k \in \mathbb{N}$. Let S be an n-element set. Then,

$$\begin{aligned}\left(\# \text{ of injective } k\text{-tuples in } S^k\right) &= n\,(n-1)\,(n-2)\cdots(n-k+1) \\ &= k! \cdot \binom{n}{k}.\end{aligned}$$

Example 6.4.5. The # of injective 3-tuples in $\{1, 2, 3, 4, 5, 6\}^3$ is $6 \cdot 5 \cdot 4 = 120$. Two of these tuples are $(2, 5, 4)$ and $(1, 5, 3)$.

Proof idea. How can we construct an injective k-tuple $(i_1, i_2, \ldots, i_k) \in S^k$? Here is an algorithm:

1. First, we choose its first entry i_1. There are n options (since $|S| = n$).
2. Next, we choose its second entry i_2. There are $n - 1$ options (since i_2 must be $\neq i_1$ in order for the k-tuple to be injective).
3. Next, we choose its third entry i_3. There are $n - 2$ options (since i_3 must be distinct from both i_1 and i_2, and these two conditions are not the same since $i_1 \neq i_2$).
4. And so on. We choose all entries from first to last, and each time, for the i-th entry, we have $n - i + 1$ options.

So we have n options at the first step, $n - 1$ options at the second step, $n - 2$ at the third step, and so on, all the way until the last (kth) step, at which we have $n-k+1$ options. Any combination of these options gives a different k-tuple. So the # of outcomes (i.e., injective k-tuples) is the **product** of all these numbers, that is,

$$n\left(n-1\right)\left(n-2\right)\cdots\left(n-k+1\right).$$

This can also be written as $k!\cdot\dbinom{n}{k}$ because $\dbinom{n}{k}=\dfrac{n\left(n-1\right)\left(n-2\right)\cdots\left(n-k+1\right)}{k!}$.

The method we used here is the so-called "**dependent product rule**". Informally, it says that if we perform a multi-step construction, in which we have

- exactly n_1 options in step 1,
- exactly n_2 options in step 2,
- $\ldots$,
- exactly n_k options in step k,

then the entire construction can be performed in $n_1 n_2 \cdots n_k$ many different ways.

In the notes, I explain how this can be re-encoded as an induction proof. □

6.4.3. Ordered selections with replacement

Now for the third problem: How many ways are there to select k elements from an n-element set, if the order matters and distinctness is not required?

This is just counting the k-tuples of elements of this set. The number is n^k, by the product rule.

6.4.4. Unordered selections with replacement

Now to the fourth and final problem: How many ways are there to select k elements from an n-element set, if the order does not matter and distinctness is not required?

The answer is $\dbinom{k+n-1}{k}$. For the proof (sketch), see §6.6.5 in the notes.