Proofs So Far

This document will contain all the theorems and proofs we have covered in the course so far. If it is ever more than a day out of date, please email the instructor to ask for an update.

Contents

1 Week 1

1.1 e is irrational

We prove that e is irrational. Our proof is inspired by a proof attributed to Joseph Fourier, from the textbook "Mélanges d'Analyse Algébrique et de Géométrie" by Janot de Stainville (1815), pp. 340–341.

Lemma 1.1 (Geometric series)**.** *If z is a real number and $|z| < 1$, then*

$$1 + z + z^2 + \cdots = \frac{1}{1-z}.$$

Proof. For any positive integer n, let S_n denote the partial sum,

$$S_n = 1 + z + z^2 + \cdots + z^n.$$

Multiplying S_n by $1 - z$ yields

$$(1-z)S_n = (1 + z + \cdots + z^n) - (z + z^2 + \cdots + z^{n+1}) = 1 - z^{n+1}.$$

Therefore,

$$S_n = \frac{1 - z^{n+1}}{1-z}.$$

Since $|z| < 1$, we have $\lim_{n\to\infty} z^{n+1} = 0$. It follows that

$$\lim_{n\to\infty} S_n = \frac{1}{1-z}. \qquad \square$$

Corollary 1.2. *If t is a real number and $|t| > 1$, then*

$$\frac{1}{t} + \frac{1}{t^2} + \frac{1}{t^3} + \cdots = \frac{1}{t-1}.$$

Proof. A consequence of lemma 1.1, obtained by multiplying each side by z, is

$$z + z^2 + z^3 + \cdots = \frac{z}{1-z}. \tag{1.1}$$

If $|t| > 1$, then setting $z = 1/t$ ensures $|z| < 1$, and substituting into (1.1) yields the result. $\square$

Lemma 1.3. *If $r > 0$, then*

$$\frac{1}{r+1} < \frac{1}{r+1} + \frac{1}{(r+1)(r+2)} + \frac{1}{(r+1)(r+2)(r+3)} + \cdots < \frac{1}{r}.$$

Proof. The inequality on the left follows from the fact that all terms besides $1/(r+1)$ are positive. For the inequality on the right, observe that $(r+1)(r+2) > (r+1)^2$ and $(r+1)(r+2)(r+3) > (r+1)^3$, etc.. Taking reciprocals, we find

$$\begin{aligned}\frac{1}{r+1} + \frac{1}{(r+1)(r+2)} + \frac{1}{(r+1)(r+2)(r+3)} + \cdots &< \frac{1}{r+1} + \frac{1}{(r+1)^2} + \frac{1}{(r+1)^3} + \cdots \\ &= \frac{1}{(r+1)-1} = \frac{1}{r} \quad \text{(applying corollary 1.2).}\ \square\end{aligned}$$

Theorem 1.4. *The number e is irrational.*

Proof. If e were rational, then there would exist an integer $n > 1$ such that $ne \in \mathbb{Z}$. In this case, define

$$x = \frac{1}{0!} + \frac{1}{1!} + \cdots + \frac{1}{(n-1)!} + \frac{1}{n!}, \quad y = \frac{1}{(n+1)!} + \frac{1}{(n+2)!} + \frac{1}{(n+3)!} + \cdots,$$

so that $e = x + y$. Multiplying by $n!$ yields

$$n!e = n!x + n!y. \tag{1.2}$$

Since ne is an integer, $n!e$ must also be an integer. Cancelling denominators, we find

$$n!x = \frac{n!}{0!} + \frac{n!}{1!} + \frac{n!}{2!} + \cdots + \frac{n!}{(n-1)!} + \frac{n!}{n!} \in \mathbb{Z}.$$

Equation (1.2) now implies that $n!y$ is an integer. Note that $n!y > 0$ since its factors $n!$, and y are positive. On the other hand, the previous lemma bounds $n!y$:

$$\begin{aligned} n!y &= \frac{n!}{(n+1)!} + \frac{n!}{(n+2)!} + \frac{n!}{(n+3)!} + \cdots \\ &= \frac{1}{n+1} + \frac{1}{(n+1)(n+2)} + \frac{1}{(n+1)(n+2)(n+3)} + \cdots \\ &< \frac{1}{n} < 1 \quad (\text{since } n > 1). \end{aligned}$$

But one cannot have $n!y \in \mathbb{Z}$ and $0 < n!y < 1$. Hence e cannot be rational. □

1.2 Pythagoras's theorem (Proof 1)

The famous theorem of Pythagoras. The proof uses four copies of the triangle, whose hypotenuses form the sides of a square, so that the triangles lie inside the square.

Theorem 1.5 (Pythagoras)**.** *Suppose a right triangle has hypotenuse of length c and other sides of length a and b. Then $c^2 = a^2 + b^2$.*

Proof. The internal angles of a triangle sum to π, but the triangle has a right angle, so the two acute angles sum to $\pi/2$. Arranging four congruent copies of the triangle as shown in figure 1 ensures that each internal angle of the outermost quadrilateral is the sum of one of the acute angles from one triangle, and the other acute angle from another triangle. Hence the outermost quadrilateral is a rectangle. As the triangles are congruent, the sides of the outermost quadrilateral are equal in length, so the outer quadrilateral is a square.

As the triangles are right, the internal angles of the innermost quadrilateral are also right. Moreover, the sides of the innermost quadrilateral are each of length $(b - a)$, so it is square.

The outer square has area c^2. The inner square has area $(b - a)^2$. Each triangle has area $\frac{1}{2}ab$. Hence

$$c^2 = b^2 - 2ab + a^2 + 4\left(\frac{1}{2}ab\right).$$

The result follows. □

1.3 Pythagoras's theorem (Proof 2)

The famous theorem of Pythagoras. The proof uses four copies of the triangle, whose hypotenuses form the sides of a square, so that the triangles lie outside the square.

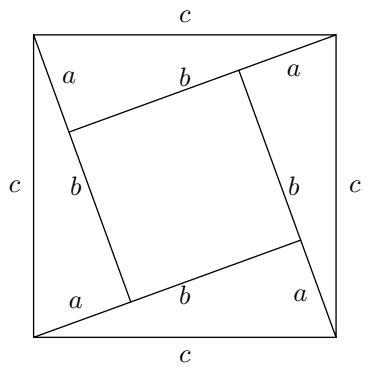


Figure 1: Four congruent copies of the triangle arranged so that their hypotenuses form the sides of a square that contains the triangles.

Theorem 1.6 (Pythagoras)**.** *Suppose a right triangle has hypotenuse of length c and other sides of length a and b. Then $c^2 = a^2 + b^2$.*

Proof. The internal angles of a triangle sum to π, but the triangle has a right angle, so the two acute angles sum to $\pi/2$. Let us arrange four triangles, congruent to the original, as shown in figure 2, so that the innermost quadrilateral is a square. Note that the innermost quadrilateral is guaranteed to be a rhombus, as its sides must have equal length, so insisting it be a square is permissible.

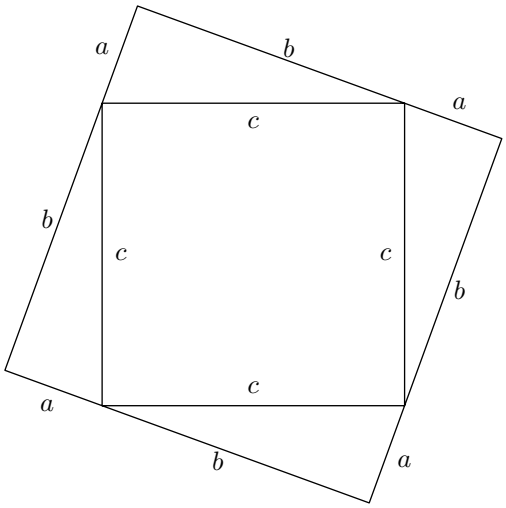


Figure 2: Four congruent copies of the triangle arranged so that their hypotenuses form the sides of a square that excludes the triangles.

The outer perimeter appears to be a quadrilateral, but we must argue that no angle is made at the meeting points of the two triangles that make up each side. Indeed, summing the angles on the interior, we arrive at $\pi/2$ from the right angle of the innermost quadrilateral, and $\pi/2$ from the sum of the two different acute angles from the congruent triangles, for a total of π; no angle is made. The outermost quadrilateral (as we may now certainly call it) has every angle right, because the triangles are right, and every side of length $(b + a)$, because the triangles are congruent. Hence the outermost quadrilateral is also a square.

The outer square has area $(a + b)^2$. The inner square has area c^2. Each triangle has area $\frac{1}{2}ab$.

Hence

$$c^2 + 4\left(\frac{1}{2}ab\right) = b^2 + 2ab + a^2.$$

The result follows. □

2 Week 2

2.1 Existence of an irrational number

We show that there is at least one irrational number. (It is good to be sure!)

Lemma 2.1. *There is a positive real number z such that $z^2 = 2$.*

Proof. Consider a right isosceles triangle, in which the sides of equal length have length 1. By Pythagoras' theorem, the square of the length of the hypotenuse is $1^2 + 1^2 = 2$. Any length must be a real number, so there is a real number whose square is 2. □

Lemma 2.2. *If x is an integer, and x^2 is even, then x is even.*

Proof. If x is odd, then $x = 2n + 1$ for some integer n. It follows that

$$x^2 = (2n + 1)^2 = 4n^2 + 4n + 1 = 2(2n^2 + 2n) + 1.$$

But this is evidently an odd number. Hence the square of an odd number is odd. Therefore, if the square of an integer is even, the number itself must be even. □

Theorem 2.3. *If $z^2 = 2$, then z is irrational.*

Proof. Suppose there were a rational number whose square was 2. Such a rational number could be expressed as a/b for integers a and b. Furthermore, by reducing the fraction, we can assume that a and b are *not both* even.

The equality $(a/b)^2 = 2$ implies that $a^2 = 2b^2$. Since the right side is even, the left side is even, and thus a is even by lemma 2.2.

Therefore, there exists an integer c such that $a = 2c$. The equality $a^2 = 2b^2$ now implies that $(2c)^2 = 2b^2$. Expanding and dividing by 2 yields $2c^2 = b^2$. Since the left side is even, the right side is even, and so b is even by lemma 2.2 again.

We find that a and b must both be even, contradicting the reducedness of the fraction a/b. Hence there cannot be a rational number whose square is 2.

By lemma 2.1, there is a real number whose square is 2. So z is irrational. □

2.2 Divergence of the harmonic series (Proof 1)

We follow the original proof of Nicole Oresme.

Theorem 2.4. *The harmonic series diverges:*

$$\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots = \infty.$$

Proof. Define S_c to be the sum of the terms of the series starting with the reciprocal of 2^c up to (but not including) the reciprocal of 2^{c+1}:

$$\begin{aligned}
S_0 &= \frac{1}{2^0} = \frac{1}{1}, \\
S_1 &= \frac{1}{2^1} + \frac{1}{2^1 + 1} = \frac{1}{2} + \frac{1}{3}, \\
S_2 &= \frac{1}{2^2} + \frac{1}{2^2 + 1} + \frac{1}{2^2 + 2} + \frac{1}{2^2 + 3} = \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7}.
\end{aligned}$$

We may bound each S_c,

$$\begin{aligned} S_c &= \frac{1}{2^c} + \frac{1}{2^c+1} + \frac{1}{2^c+2} + \cdots + \frac{1}{2^{c+1}-1}, \quad \text{(with } 2^c \text{ terms)} \\ &> \frac{1}{2^{c+1}} + \frac{1}{2^{c+1}} + \frac{1}{2^{c+1}} + \cdots + \frac{1}{2^{c+1}}, \quad \text{(with } 2^c \text{ terms)} \\ &= 2^c \cdot \frac{1}{2^{c+1}} = \frac{1}{2}. \end{aligned}$$

Let T_n denote the sum of the first $2^n - 1$ terms of the harmonic series. Then

$$T_n = S_0 + S_1 + \cdots + S_{n-1} > n \cdot \frac{1}{2},$$

since there are n quantities in the list $S_0, S_1, \ldots, S_{n-1}$ and each is greater than $1/2$. It follows that the harmonic series eventually exceeds $n/2$ for every integer n, and so it diverges. □

2.3 Divergence of the harmonic series (Proof 2)

Here we prove the divergence of the harmonic series, using an idea of Mengoli (1625–1686), and following the treatment of "When Less is More: Visualizing Basic Inequalities," by Claudi Alsina and Roger B. Nelson, pp. 11–12.

Lemma 2.5 (Mengoli's inequality)**.** *For any real number $r > 1$,*

$$\frac{1}{r-1} + \frac{1}{r} + \frac{1}{r+1} > \frac{3}{r}.$$

Proof. We consider the outermost terms (on the left side) first.

$$\frac{1}{r-1} + \frac{1}{r+1} = \frac{r+1+r-1}{r^2-1} = \frac{2r}{r^2-1}.$$

Since $r^2 - 1 < r^2$, we find that

$$\frac{1}{r-1} + \frac{1}{r+1} = \frac{2r}{r^2-1} > \frac{2r}{r^2} = \frac{2}{r}.$$

Adding $1/r$ to both sides, we find

$$\frac{1}{r-1} + \frac{1}{r} + \frac{1}{r+1} > \frac{2}{r} + \frac{1}{r} = \frac{3}{r}.$$

□

Theorem 2.6. *The harmonic series diverges:*

$$\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots = \infty.$$

Proof. Suppose to the contrary that the series converged to a real number S,

$$S = \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots.$$

Grouping the terms three by three, we can rewrite this as

$$S = 1 + \left(\frac{1}{3-1} + \frac{1}{3} + \frac{1}{3+1}\right) + \left(\frac{1}{6-1} + \frac{1}{6} + \frac{1}{6+1}\right) + \cdots.$$

The lemma implies the bound

$$S > 1 + \frac{3}{3} + \frac{3}{6} + \frac{3}{9} + \cdots = 1 + \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \cdots = 1 + S.$$

But no real number S can satisfy $S > 1 + S$. □

2.4 Divergence of the harmonic series (Proof 3)

We follow the proof of Johan Bernoulli.

Lemma 2.7 (Johan Bernoulli)**.**

$$\frac{1}{1\times 2}+\frac{1}{2\times 3}+\frac{1}{3\times 4}+\frac{1}{4\times 5}+\cdots=1.$$

Proof. We will see how to prove this lemma later. For now, you can use it without proof. □

Theorem 2.8. *The harmonic series diverges:*

$$\frac{1}{1}+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\cdots=\infty.$$

Proof. Suppose the harmonic series sums to a real number, S.

By Bernoulli's lemma 2.7,

$$\begin{aligned}\frac{1}{2}+\frac{1}{6}+\frac{1}{12}+\frac{1}{20}+\frac{1}{30}+\cdots&=1\\ \frac{1}{6}+\frac{1}{12}+\frac{1}{20}+\frac{1}{30}+\cdots&=\frac{1}{2}\\ \frac{1}{12}+\frac{1}{20}+\frac{1}{30}+\cdots&=\frac{1}{3}\\ \frac{1}{20}+\frac{1}{30}+\cdots&=\frac{1}{4}\end{aligned}$$

Summing these equations, the right hand side is equal to S. Summing the left hand side by summing down each (finite-length) column, and then summing the columns, we obtain

$$\frac{1}{2}+\frac{2}{6}+\frac{3}{12}+\frac{4}{20}+\frac{5}{30}+\cdots=\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\frac{1}{6}+\cdots=S-1.$$

So $S=S-1$, but no real number has this property. □

2.5 Divergence of the harmonic series (Proof 4)

We follow the proof of Leo Goldmakher.

Lemma 2.9. *If $n\in\mathbb{N}$, then* $\dfrac{1}{2n-1}+\dfrac{1}{2n}>\dfrac{1}{n}$.

Proof.

$$\frac{1}{2n-1}+\frac{1}{2n}-\frac{1}{n}=\frac{1}{2n-1}-\frac{1}{2n}=\frac{2n-(2n-1)}{2n(2n-1)}=\frac{1}{2n(2n-1)}>0,$$

as $n>\frac{1}{2}$. Adding $\frac{1}{n}$ to both sides of the inequality establishes the lemma. □

Theorem 2.10. *The harmonic series diverges:*

$$\frac{1}{1}+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\cdots=\infty.$$

Proof. Suppose the harmonic series sums to a real number, S. Then

$$\begin{aligned} S &= \frac{1}{2} + \frac{1}{1} + \left(\frac{1}{2(2)-1} + \frac{1}{2(2)}\right) + \left(\frac{1}{2(3)-1} + \frac{1}{2(3)}\right) + \cdots \\ &> \frac{1}{2} + \frac{1}{1} + \quad \frac{1}{2} \quad + \quad \frac{1}{3} \quad + \cdots \\ &= \frac{1}{2} + S. \end{aligned}$$

But this cannot be true for any real number. Hence our original supposition was incorrect, and the harmonic series does not sum to a real number. □

2.6 Divergence of the harmonic series (Proof 5)

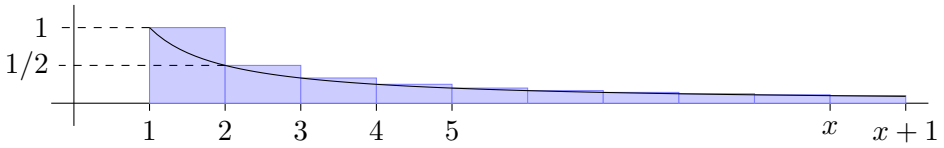
We use a geometric proof, comparing the series to the logarithm function.

Define the function $H : \mathbb{Z} \to \mathbb{R}$ by

$$H(x) = \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{x}.$$

Theorem 2.11. *For all natural numbers x, $\log(x+1) < H(x) < \log(x) + 1$.*

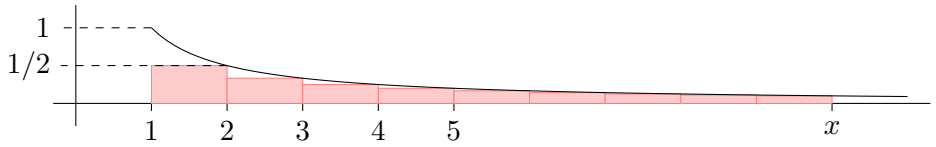
Proof. By definition, $\log(x+1) = \int_1^{x+1} \frac{1}{t}\,\mathrm{d}t$, illustrated by the area under the curve $y = 1/t$ below.



On the other hand, the harmonic sum $H(x)$ equals the area of the blue rectangles. Since this area contains the area under the curve, we find that

$$H(x) > \log(x+1). \tag{2.1}$$

One may also compare using rectangles below the curve.



The area of the red rectangles equals $\frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{x} = H(x) - 1$. Since this area is contained in the area under the curve, we find that

$$\log(x) > H(x) - 1. \tag{2.2}$$

The theorem follows immediately from the two estimates (2.1) and (2.2). □

Corollary 2.12. *The harmonic series diverges:*

$$\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots = \infty.$$

Proof. Theorem 2.11 tells us that the partial sums $H(x)$ of the harmonic series are greater than $\log(x+1)$, for every positive integer x. Hence

$$\lim_{x\to\infty} H(x) > \lim_{x\to\infty} \log(x+1) = \infty.$$ □

2.7 The sum of odd squares

We give a formula for the sum of the first n odd square integers.

The theorem below is about the sum of the squares of the first n odd natural numbers. How do we know this is the same as the sum of the first n odd squares?

Theorem 2.13. *If $n \in \mathbb{N}$, then*

$$1^2 + 3^2 + \cdots + (2n-1)^2 = \frac{n}{3}(4n^2 - 1).$$

Proof. For $n = 1$, there is only one term on the left, and $1^2 = \frac{1}{3}(4(1)^2 - 1) = 1$ is true.

Now suppose that the result is true for a particular value $k \in \mathbb{N}$. Then

$$\begin{aligned} 1^2 + 3^2 + \cdots + (2k-1)^2 + (2(k+1)-1)^2 &= \frac{k}{3}\left(4k^2 - 1\right) + (2k+1)^2 \\ &= \frac{k}{3}(2k-1)(2k+1) + (2k+1)^2 \\ &= \frac{2k+1}{3}\left(2k^2 + 5k + 3\right). \end{aligned}$$

On the other hand

$$\begin{aligned} \frac{k+1}{3}\left(4(k+1)^2 - 1\right) = \frac{k+1}{3}\left(4k^2 + 3 + 8k\right) &= \frac{k+1}{3}(2k+1)(2k+3) \\ = \frac{2k+1}{3}(k+1)(2k+3) &= \frac{2k+1}{3}\left(2k^2 + 5k + 3\right), \end{aligned}$$

proving the statement for $k + 1$.

Hence the statement is true for all $n \in \mathbb{N}$. □

2.8 Telescoping sums and series

We prove the telescoping sum formula and a few corollaries

A telescoping sum is one in which part of each term cancels with part of the subsequent term, in such a way that we can reduce a sum with very many terms to a sum with very few terms. To make an argument along these lines rigourously, we have to use induction.

Theorem 2.14. *For a sequence $(a_k)_{k\in\mathbb{N}}$,*

$$\sum_{k=1}^{n} (a_{k+1} - a_k) = a_{n+1} - a_1.$$

Proof. If $n = 1$, then the sum has only one term, and that term is equal to $a_2 - a_1$, which is the right hand side.

Now suppose that the result holds for some $n \in \mathbb{N}$. Then

$$\sum_{k=1}^{n+1} (a_{k+1} - a_k) = \sum_{k=1}^{n} (a_{k+1} - a_k) + a_{n+2} - a_{n+1} = a_{n+1} - a_1 + a_{n+2} - a_{n+1} = a_{n+2} - a_1.$$

Hence, by induction, the result holds for all $n \in \mathbb{N}$. □

Corollary 2.15. *For a sequence $(a_k)_{k\in\mathbb{N}}$, with limit a,*

$$\sum_{k=1}^{\infty} (a_{k+1} - a_k) = a - a_1.$$

Proof. By the definition of a series,

$$\sum_{k=1}^{\infty}(a_{k+1}-a_k)=\lim_{n\to\infty}\sum_{k=1}^{n}(a_{k+1}-a_k)\,.$$

Applying theorem 2.14, we learn

$$\sum_{k=1}^{\infty}(a_{k+1}-a_k)=\lim_{n\to\infty}(a_{n+1}-a_1)=\lim_{n\to\infty}(a_{n+1})-a_1=a-a_1.$$

□

The following is Johan Bernoulli's lemma, which we saw in our third proof of the divergence of the harmonic series. We are now ready to prove it.

Proposition 2.16.

$$\sum_{k=1}^{\infty}\frac{1}{k(k+1)}=1.$$

Proof. Defining $a_k=1/(k(k+1))$ and $b_k=-1/k$, we observe that

$$a_k=\frac{1}{k(k+1)}=\frac{1}{k}-\frac{1}{k+1}=\left(-\frac{1}{k+1}\right)-\left(-\frac{1}{k}\right)=b_{k+1}-b_k.$$

Hence, by corollary 2.15,

$$\sum_{k=1}^{\infty}a_k=\sum_{k=1}^{\infty}b_{k+1}-b_k=\lim_{n\to\infty}b_n-b_1=0-(-1)=1.$$ □

2.9 The identity of Nicomachus (Proof 1)

We prove the identity commonly called "Nicomachus' Theorem." This is named for Nicomachus of Gerasa, a Neo-Pythagorean living around 100CE, who wrote a fascinating arithmetic book. Of historical interest: Nicomachus never actually wrote the identity named for him, but he did observe the following:

$$1=1,\quad 3+5=8,\quad 7+9+11=27,\quad 13+15+17+19=64,\quad \ldots,$$

which is closely related. Summing consecutive odd numbers in this way yields the cubes, and one may deduce "Nicomachus' Theorem" from this observation. Instead of following the method described above, we will make a different inductive argument. You may also be interested in this rather wonderful graphical proof: `http://blogs.mathworks.com/loren/2010/03/04/nichomachuss-theorem/`

Lemma 2.17. *If n is a natural number, then* $\displaystyle\sum_{k=1}^{n}k=\frac{n(n+1)}{2}$.

Proof. Let us denote the sum in question

$$S(n)=1+2+3+\cdots+(n-2)+(n-1)+n.$$

Then $S(1)=1=\frac{2}{2}=\frac{1(1+1)}{2}$.

Suppose that $S(n)=n(n+1)/2$, for some particular n. Then

$$S(n+1)=\frac{n(n+1)}{2}+(n+1)=(n+1)\left[\frac{n}{2}+1\right]=\frac{n+1}{2}[n+2].$$

Hence, by induction, $S(n)=n(n+1)/2$ for all natural numbers n. □

Lemma 2.18. *If n is a natural number, then* $\displaystyle\sum_{k=1}^{n} k^3 = \frac{n^2(n+1)^2}{4}$.

Proof. If $n = 1$, then both sides of the equation are equal to 1. Now suppose that the identity holds for a particular value n. Then

$$\sum_{k=1}^{n+1} k^3 = \frac{n^2(n+1)^2}{4} + (n+1)^3 = \frac{(n+1)^2}{4}\left[n^2 + 4(n+1)\right] = \frac{(n+1)^2(n+2)^2}{4}.$$

Hence, by induction, the result holds for all $n \in \mathbb{N}$. □

Corollary 2.19 (The identity of Nicomachus)**.** *If n is a natural number, then* $\displaystyle\sum_{k=1}^{n} k^3 = \left(\sum_{j=1}^{n} j\right)^2$.

Proof. The result is immediate from lemmas 2.17 and 2.18. □

2.10 The identity of Nicomachus (Proof 2)

We prove the identity commonly called "Nicomachus' Theorem." This is named for Nicomachus of Gerasa, a Neo-Pythagorean living around 100CE, who wrote a fascinating arithmetic book. Of historical interest: Nicomachus never actually wrote the identity named for him, but he did observe the following:

$$1 = 1, \quad 3 + 5 = 8, \quad 7 + 9 + 11 = 27, \quad 13 + 15 + 17 + 19 = 64, \quad \ldots,$$

which is closely related. Summing consecutive odd numbers in this way yields the cubes, and one may deduce "Nicomachus' Theorem" from this observation. Instead of following the method described above, we will make a different inductive argument. You may also be interested in this rather wonderful graphical proof: `http://blogs.mathworks.com/loren/2010/03/04/nichomachuss-theorem/`

Lemma 2.20. *If n is a natural number, then* $\displaystyle\sum_{k=1}^{n-1} k = \frac{n(n-1)}{2}$.

Proof. Let us denote the sum in question

$$S(n-1) = 1 + 2 + 3 + \cdots + (n-2) + (n-1).$$

Then $S(1) = 1 = \frac{2}{2} = \frac{1(1+1)}{2}$.

Suppose that $S(n-1) = (n-1)n/2$, for some particular n. Then

$$S(n) = \frac{(n-1)n}{2} + n = n\left[\frac{n-1}{2} + 1\right] = \frac{n}{2}[n+1].$$

Hence, by induction, $S(n-1) = (n-1)n/2$ for all natural numbers n. □

Theorem 2.21 (The identity of Nicomachus)**.** *If n is a natural number, then* $\displaystyle\sum_{k=1}^{n} k^3 = \left(\sum_{j=1}^{n} j\right)^2$.

Proof. When $n = 1$, the left side equals 1 and the right side equals 1.

Now suppose that that $n > 1$ and the identity of Nichomachus has been proved for $n - 1$. Then we compute

$$\begin{aligned}\left(\sum_{i=1}^{n} i\right)^2 &= \left(n + \sum_{i=1}^{n-1} i\right)^2 = n^2 + 2n\left(\sum_{i=1}^{n-1} i\right) + \left(\sum_{i=1}^{n-1} i\right)^2 \\ &= n^2 + 2n\frac{n(n-1)}{2} + \left(\sum_{i=1}^{n-1} i\right)^2 \quad \text{(by the lemma)} \\ &= n^2 + 2n\frac{n(n-1)}{2} + \sum_{j=1}^{n-1} j^3 \quad \text{(by the inductive hypothesis)} \\ &= n^2 + n^3 - n^2 + \sum_{j=1}^{n-1} j^3 \\ &= \sum_{j=1}^{n} j^3.\end{aligned}$$

By induction, it follows that $\left(\sum_{i=1}^{n} i\right)^2 = \sum_{j=1}^{n} j^3$ for all positive integers n. □

3 Week 3

3.1 Throdder squares

We show that no square integers take the form $3k + 2$.

Definition 3.1. An integer n is called *threven* if it can be written in the form $n = 3k$ for another integer k. If $n = 3k + 1$ for another integer k, then n is called *throdd*. If $n = 3k + 2$ for integer k, then n is called *throdder*.

For this proof, you may assume that every integer is precisely one of threven, throdd or throdder.

Theorem 3.2. *There are no throdder squares. Moreover, if the square of an integer is threven, then the integer is threven.*

Proof. If n is threven, then, for some $k \in \mathbb{Z}$, $n = 3k$. So $n^2 = 9k^2 = 3(3k^2)$, so n^2 is threven. If n is throdd, then, for some $k \in \mathbb{Z}$, $n = 3k + 1$. So $n^2 = 9k^2 + 6k + 1 = 3(3k^2 + 2k) + 1$, so n^2 is throdd. If n is throdder, then, for some $k \in \mathbb{Z}$, $n = 3k + 2$. So $n^2 = 9k^2 + 12k + 4 = 3(3k^2 + 4k + 1) + 1$, so n^2 is throdd. So all square numbers must be threven or throdd, not throdder. The only way to produce a threven square is by starting with a threven number and squaring it, because the other cases both produce throdd squares. □

3.2 $\sqrt{3}$ is irrational

We establish the existence of another irrational number.

Theorem 3.3. *There is a positive irrational number x for which $x^2 = 3$.*

Proof. We already know there is a real number z for which $z^2 = 2$. Constructing a right triangle with sides of length 1 and z adjacent to the right angle, Pythagoras' theorem tells us the square of the length of the hypotenuse is 3. As lengths are positive real numbers, $x \in \mathbb{R}^+$.

Now suppose that $x \in \mathbb{Q}$. Then there are integers p, q, which are not both divisible by 3, such that $x = p/q$. (Otherwise, divide p and q by 3 until at least one is not divisible by 3.) Hence $3q^2 = p^2$, so p^2 is divisible by 3. But then (using the theorem on Throdd squares) p is also divisible by 3, so $p = 3k$ for some $k \in \mathbb{Z}$. Hence $3q^2 = 9k^2$, so $q^2 = 3k^2$, and q^2 is divisible by 3. But then q is also divisible by 3 (using the theorem on Throdd squares again). This contradicts p and q being not both divisible by 3. Hence our supposition was incorrect, and $x \notin \mathbb{Q}$.

As $x \in \mathbb{R} \setminus \mathbb{Q}$, x is irrational. □

3.3 A bound on the tail of the factorial

We find a bound on the tail of the factorial. That is, we bound the ratio $n!/r!$ for nonnegative integers $n \geqslant r$.

Theorem 3.4. *Suppose $n, k \in \mathbb{N}$ and $n \geqslant k - 1$. Then $n! \geqslant (k-1)!k^{n-k+1}$.*

Proof. Throughout this proof, let k be any natural number. If $n = k - 1$, then the claim is $(k-1)! \geqslant (k-1)!k^0$, which is true as $k \in \mathbb{N}$. Now suppose that the theorem holds for some particular $n \geqslant k$. Then

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

But $n + 1 \geqslant k$, so

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

Hence, by induction, $n! \geqslant (k-1)!k^{n-k+1}$ for all integer $n - 1 \geqslant k$. As k could have been any natural number, this proof holds for all k. □

3.4 A formulaic definition of the Fibonacci numbers

We already saw the Fibonacci sequence defined using a recurrence relation. Sometimes a recurrence relation can be transformed into an explicit formulaic definition. We see how to do that here.

Definition 3.5. The *golden ratio* is the real number $\phi = \dfrac{1+\sqrt{5}}{2} \approx 1.618$.

The conjugate of the golden ratio is the real number $\bar{\phi} = \dfrac{1-\sqrt{5}}{2}$.

Lemma 3.6. *The golden ratio and its conjugate both satisfy* $\phi^2 = \phi + 1$.

Proof. Squaring ϕ yields

$$\phi^2 = \left(\frac{1+\sqrt{5}}{2}\right)^2 = \frac{1+2\sqrt{5}+5}{4} = \frac{3+\sqrt{5}}{2}.$$

On the other hand,

$$\phi + 1 = \frac{1+\sqrt{5}}{2} + 1 = \frac{3+\sqrt{5}}{2}.$$

The proof for $\bar{\phi}$ is nearly identical. □

Theorem 3.7. *For every natural number n, the Fibonacci numbers satisfy* $F_n = \dfrac{\phi^n - \bar{\phi}^n}{\sqrt{5}}$.

Proof. When $n = 1$, $F_n = 1$ by definition, and the right side also evaluates to 1. When $n = 2$, $F_n = 1$ by definition, and $\phi^2 - \bar{\phi}^2 = (\phi + 1) - (\bar{\phi} + 1) = \sqrt{5}$, so the right side equals 1 too.

Now, define $f(n) = (\phi^n - \bar{\phi}^n)/\sqrt{5}$. We compute

$$\begin{aligned} f(n+2) &= \frac{\phi^{n+2} - \bar{\phi}^{n+2}}{\sqrt{5}} = \frac{\phi^n\phi^2 - \bar{\phi}^n\bar{\phi}^2}{\sqrt{5}} \\ &= \frac{\phi^n(\phi+1) - \bar{\phi}^n(\bar{\phi}+1)}{\sqrt{5}} \\ &= \frac{\phi^{n+1} - \bar{\phi}^{n+1} + \phi^n - \bar{\phi}^n}{\sqrt{5}} \\ &= \frac{\phi^{n+1} - \bar{\phi}^{n+1}}{\sqrt{5}} + \frac{\phi^n - \bar{\phi}^n}{\sqrt{5}} = f(n+1) + f(n). \end{aligned}$$

We have shown that $f(n+2) = f(n+1) + f(n)$.

Now let us suppose that, for some particular n, $f(n+1) = F_{n+1}$ and $f(n) = F_n$. Then

$$f(n+2) = f(n+1) + f(n) = F_{n+1} + F_n = F_{n+2}.$$

We already have the base case(s) $f(1) = F_1$, $f(2) = F_2$ so, by induction, $f(n) = F_n$ for all natural numbers n. □

3.5 The AMGM inequality

We prove the arithmetic/geometric mean inequality (AMGM inequality). Our proof follows the method of Cauchy, but the more modern treatments of "When Less is More: Visualizing Basic Inqualities," by Claudi Alsina and Roger B. Nelson, pp. 6–8, and "The Cauchy-Schwarz Master Class," by J. Michael Steele, Chapters 1 and 2. The original proof by Augustin-Louis Cauchy is contained in

his "Cours d'analyse de l'École Royale Polytechnique, première partie, Analyse algébrique," (1821), pp. 457-459.

We will prove that the arithmetic mean is at least as large as the geometric mean, for any list of positive real numbers. Moreover, the arithmetic mean and geometric mean coincide only in the case where the numbers are all equal!

Let AMGM(n) be the statement "for every n-tuple of positive real numbers $(x_1, x_2, \ldots, x_n)$,

$$\sqrt[n]{\prod_{i=1}^{n} x_i} \leqslant \frac{1}{n}\sum_{i=1}^{n} x_i,$$

with equality iff $x_1 = x_2 = \ldots = x_n$".

Lemma 3.8. *If x and y are positive real numbers, then*

$$\sqrt{xy} \leqslant \frac{x+y}{2},$$

with equality only if $x = y$. That is, AMGM(2).

Proof. Consider the quantity $\Delta = \sqrt{x} - \sqrt{y}$. As the square of a real number, $\Delta^2 \geqslant 0$. Also, $\Delta = 0$ if and only if $\sqrt{x} = \sqrt{y}$, which occurs if and only if $x = y$. Expanding Δ^2, we find

$$0 \leqslant \Delta^2 = \left(\sqrt{x} - \sqrt{y}\right)^2 = x - 2\sqrt{xy} + y,$$

with equality if and only if $x = y$. Rearranging terms and dividing by 2, we find that

$$\frac{x+y}{2} \geqslant \sqrt{xy},$$

with equality if and only if $x = y$. □

To warm up for the next stage in the proof, we prove the AMGM inequality for four variables. Consider four positive real numbers x, y, u, v. The two-variable AMGM inequality implies

$$\sqrt[4]{xyuv} = \sqrt{\sqrt{xy}\sqrt{uv}} \leqslant \frac{\sqrt{xy} + \sqrt{uv}}{2},$$

with equality if and only if $\sqrt{xy} = \sqrt{uv}$.

Again by the two-variable AMGM inequality, we have

$$\sqrt{xy} \leqslant \frac{x+y}{2} \text{ and } \sqrt{uv} \leqslant \frac{u+v}{2},$$

with equalities if and only if $x = y$ and $u = v$, respectively. We use this to continue,

$$\begin{aligned}\sqrt[4]{xyuv} &\leqslant \frac{\sqrt{xy} + \sqrt{uv}}{2} \\ &\leqslant \frac{\frac{x+y}{2} + \frac{u+v}{2}}{2} \\ &= \frac{x+y+u+v}{4},\end{aligned}$$

with equality if and only if $\sqrt{xy} = \sqrt{uv}$ and $x = y$ and $u = v$.

The conjunction of these three equalities implies that $\sqrt{x^2} = \sqrt{u^2}$ and so $x = u$ (both are required to be positive). Thus the three equalities $\sqrt{xy} = \sqrt{uv}$, $x = y$, $u = v$, are equivalent to the equalities $x = y = u = v$. We find the four-variable AMGM inequality,

$$\sqrt[4]{xyuv} \leqslant \frac{x+y+u+v}{4}$$

with equality if and only if $x = y = u = v$.

We can use the same method to prove the AMGM inequality for 8 variables, 16 variables, 32 variables, etc.. To write the proof more formally:

Lemma 3.9. *For any natural number n, the statements* AMGM(n) *and* AMGM(2) *imply the statement* AMGM($2n$).

Proof. Assume the statements AMGM(n) and AMGM(2), i.e., the AMGM inequality for n variables and for 2 variables (which we have already proven). To prove the AMGM inequality for $2n$ variables, consider $2n$ positive real numbers and call them $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$. By AMGM(2), we find

$$\sqrt[2n]{\prod_{i=1}^{n} x_i \prod_{i=1}^{n} y_i} = \sqrt{\sqrt[n]{\prod_{i=1}^{n} x_i}\sqrt[n]{\prod_{i=1}^{n} y_i}} \leqslant \frac{\sqrt[n]{\prod_{i=1}^{n} x_i} + \sqrt[n]{\prod_{i=1}^{n} y_i}}{2},$$

with equality if and only if $\prod x_i = \prod y_i$.

Now we apply AMGM(n) to continue,

$$\begin{aligned}\sqrt[2n]{\prod_{i=1}^{n} x_i \prod_{i=1}^{n} y_i} &\leqslant \frac{\sqrt[n]{\prod_{i=1}^{n} x_i} + \sqrt[n]{\prod_{i=1}^{n} y_i}}{2} \\ &\leqslant \frac{\frac{1}{n}\sum_{i=1}^{n} x_i + \frac{1}{n}\sum_{i=1}^{n} y_i}{2} = \frac{x_1 + \cdots + x_n + y_1 + \cdots + y_n}{2n},\end{aligned}$$

with equality if and only if $\prod_{i=1}^{n} x_i = \prod_{i=1}^{n} y_i$ and $x_1 = \cdots = x_n$ and $y_1 = \cdots = y_n$. But these equality conditions are equivalent to $x_1 = \cdots = x_n = y_1 = \cdots = y_n$. □

Since we have proven AMGM(2), the previous lemma implies AMGM(4) and AMGM(8), etc.. Therefore, we have proven AMGM(2^c) for every natural exponent c. The next lemma fills in the gaps.

Lemma 3.10. *For any integer $n \geqslant 3$, the statement* AMGM(n) *implies the statement* AMGM($n-1$).

Proof. Assume AMGM(n). Consider a list of positive real numbers $x_1, \ldots, x_{n-1}$. Define x_n to be their arithmetic mean,

$$x_n = \frac{x_1 + \cdots + x_{n-1}}{n-1}.$$

Applying AMGM(n) to the list $x_1, \ldots, x_n$ yields

$$\begin{aligned}\sqrt[n]{x_1 \cdots x_{n-1} x_n} &\leqslant \frac{x_1 + \cdots + x_{n-1} + x_n}{n} \\ &= \frac{(x_1 + \cdots + x_{n-1})\left(1 + \frac{1}{n-1}\right)}{n} \\ &= \frac{x_1 + \cdots + x_{n-1}}{n-1} = x_n,\end{aligned}$$

with equality if and only if $x_1 = x_2 = \cdots = x_{n-1} = x_n$. Multiplying both sides by $x_n^{-1/n}$, we find

$$\sqrt[n]{x_1 \cdots x_{n-1}} \leqslant x_n^{1-1/n} = x_n^{\frac{n-1}{n}},$$

with equality if and only if $x_1 = x_2 = \cdots = x_{n-1} = x_n$.

Finally, raising both sides to the $n/(n-1)$ power, we find

$$\sqrt[n-1]{x_1 \cdots x_{n-1}} \leqslant x_n,$$

with equality if and only if $x_1 = x_2 = \cdots = x_{n-1} = x_n$. Since x_n is the arithmetic mean of $x_1, \ldots, x_{n-1}$, this is the AMGM inequality, i.e. the statement AMGM$(n-1)$. □

Taken together, these lemmas imply the AMGM inequality.

Theorem 3.11. *For every natural number n, and every list of positive real numbers $x_1, x_2, \ldots, x_n$,*

$$\sqrt[n]{\prod_{i=1}^{n} x_i} \leqslant \frac{1}{n} \sum_{i=1}^{n} x_i,$$

with equality if and only if all numbers are equal, $x_1 = x_2 = \cdots = x_n$.

Proof. When $n = 1$, the statement is tautology. When $n = 2$, the statement is proven in lemma 3.8. Lemma 3.9 implies the statement whenever n is a power of 2. Finally, lemma 3.10 implies the statement for all remaining positive integers n. □

3.6 Scaled Fibonacci numbers

We prove a bound on a scaled version of the Fibonacci numbers

Theorem 3.12. *Suppose a sequence is defined inductively using the recurrence relation*

$$a_{n+2} = \frac{a_{n+1}}{2n+1} + \frac{a_n}{2n-1}.$$

Then, provided the initial terms a_0, a_1 are both 1, every term a_n is no greater than 1.

Proof. Suppose $a_0 = a_1 = 1$. Then $a_2 = 0 \leqslant 1$ and $a_3 = 1 \leqslant 1$.

Now suppose that, for some $n \geqslant 2$, $0 \leqslant a_n, a_{n+1} \leqslant 1$. Then, because $n \geqslant 2$, a_{n+2} is the sum of two nonnegative ratios, so $a_{n+2} \geqslant 0$. Moreover

$$\begin{aligned} a_{n+2} &= \frac{a_{n+1}}{2n+1} + \frac{a_n}{2n-1} \\ &\leqslant \frac{1}{2n+1} + \frac{1}{2n-1} \\ &= \frac{2n-1+2n+1}{(2n+1)(2n-1)} \\ &= \frac{4n}{4n^2-1}. \end{aligned}$$

But $4n^2 - 1 > 4n$ for all natural numbers $n \geqslant 2$. So the ratio is less than one, and $a_{n+2} < 1$.

Hence, by induction on n, $a_n \leqslant 1$ for all $n \in \mathbb{N}$. □

4 Week 4

4.1 A divisibility theorem

We show that the numbers one less than each power of 4 are each divisible by 3.

Theorem 4.1. *If n is a nonnegative integer, then $3 \mid (4^n - 1)$.*

Proof. If $n = 0$, then $4^n - 1 = 0$, and $3 \mid 0$.

Now suppose that, for some nonegative integer n, $3 \mid (4^n - 1)$. Then $3 \mid 4(4^n - 1)$ also. But $4(4^n - 1) = 4^{n+1} - 4 = [4^{n+1} - 1] - 3$, so $4^{n+1} - 1$ is also divisible by 3.

Hence, by induction, $3 \mid (4^n - 1)$ for all $n \in \mathbb{N}^0$. □

4.2 A moving sum of fractions

This result gives us a lower bound on the sum of the reciprocals of $(n + 1)$ up to $2n$.

Theorem 4.2. *For all integers $n \geqslant 2$,*

$$\sum_{k=1}^{n} \frac{1}{n+k} \geqslant \frac{7}{12}.$$

Proof. $\frac{1}{3} + \frac{1}{4} = \frac{7}{12}$, so the result holds with equality for $n = 2$.

Now asume that, for some particular number $n \geqslant 2$, the result holds. Consider the sum

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

The first term is positive. By the inductive hypothesis, the other terms sum to at least $\frac{7}{12}$.

Hence, by induction, the theorem holds. □

4.3 Every natural number has a prime factorisation

We prove that every positive integer can be factored into prime numbers.

Definition 4.3. A natural number p is called *prime* if $p > 1$ and the only factors of p are 1 and p itself.

Lemma 4.4. *If $n \in \mathbb{N}$, then $n = 1$ or n has a prime factor.*

Proof. If $n = 1$, then the statement is true. The statement is true when $n = 2$, since 2 is a prime factor of itself.

Now suppose that $n > 2$, and the lemma has been verified for all smaller cases. If n itself is prime, then n is a prime factor of itself, and there is nothing left to prove. If n is not prime, then n must have a factor m with $1 < m < n$.

Since $1 < m < n$, the inductive hypothesis implies that m has a prime factor p. Since p is a factor of m, and m is a factor of n, we find that p is a factor of n. Hence n has a prime factor.

By induction, every integer greater than 1 has a prime factor. □

Theorem 4.5. *If n is a natural number, then n can be expressed as a product of primes.*

Proof. If $n = 1$, then n is the *empty product*.

So let us assume $n > 1$ and that the theorem has been verified for smaller values. By lemma 4.4, n has a prime factor p, so we may write $n = pm$ for some positive integer m. Since $p \geqslant 2$, we must have $1 \leqslant m < n$.

By the inductive hypothesis, m can be expressed as a product of prime numbers. Since $n = pm$ and p is prime, we find that n can be expressed as a product of prime numbers too.

By induction, every positive integer can be expressed as a product of prime numbers. □

4.4 There are infinitely many primes

We prove that the number of primes is infinite. Our proof is based on Euclid's "Elements" Book IX, Proposition 20.

Lemma 4.6. *If $a, M \in \mathbb{Z}$ and $a > 1$, then a cannot be a factor of both M and $M + 1$.*

Proof. If a is a factor of M and a is a factor of $M+1$, then a is a factor of their difference $(M+1)-M$, which equals 1. But a cannot be a factor of 1, since $a > 1$. □

Theorem 4.7. *The number of primes is greater than any natural number.*

Proof. Suppose that n is a positive integer, and consider a set S consisting of n prime numbers. Write M for the product of all these n prime numbers, and consider the natural number $M + 1$.

Since $M + 1$ is an integer greater than 1, it has a prime factor p. By the lemma, p cannot be a factor of M. It follows that p is a prime number and not a member of the set S. Hence there are more than n prime numbers. □

4.5 The well-ordering principle

The well-ordering principle is an important property of the natural numbers which the integers do not posess. It tells us that there is a canonical place to start listing the natural numbers, because all the others appear by starting at 1 and continuing to increment. Notice that the integers do not have this property; wherever you start in the integers, successive incrementing is insufficient to produce all the integers because there are always some integers less than the starting number.

Definition 4.8. Suppose that S is a set containing only real numbers. If $x \in S$ and, for all $y \in S$, it is true that $x \leqslant y$, then x is called the *least element* of S.

Note that, in definition 4.8, we refer to x as *the* least element, not *a* least element. Why is this? Can you prove a justification?

Not all sets of real numbers have a least element. In particular, $\mathbb{R}$ itself has no least element, and neither does $\mathbb{Z}$. The set of nonnegative rational numbers has a least element 0, but the set of positive rational numbers has no least element. This strange behaviour leads to the invention of the infimum (and supremum) and will become very important towards the end of the module. For now, we concentrate on sets of natural numbers.

The set $\mathbb{N}$ has a least element: 1. The principle of mathematical induction lets us generalise this statement to theorem 4.9

Theorem 4.9. *If $S \subset \mathbb{N}$ and $S \neq \emptyset$, then S has a least element.*

Proof. Suppose $S \subset \mathbb{N}$ and there is no least element of S. Then $1 \notin S$, as otherwise 1 would be the least element of S.

Suppose that, for some particular $m \in \mathbb{N}$, none of $1, 2, \ldots, m$ appears in S. If $m+1 \in S$ then there are no members of S less than $m+1$, so $m+1$ would be the least element of S. Therefore none of $1, 2, \ldots, m+1$ appear in S. Hence, by induction, $\mathbb{N} \cap S = \emptyset$. As $S \subset \mathbb{N}$, it follows that $S = \emptyset$.

This contradicts the hypothesis of the theorem, so there must be a least element of S. □

4.6 The Cauchy-Schwarz inequality

Here we prove the Cauchy-Schwarz inequality, one of the most important inequalities in all of mathematics.

Theorem 4.10. *For all natural numbers n, and all real sequences $(x_j)_{j=1}^n$, $(y_j)_{j=1}^n$,*

$$\left(\sum_{j=1}^{n} x_j y_j\right)^2 \leqslant \left(\sum_{j=1}^{n} x_j^2\right)\left(\sum_{j=1}^{n} y_j^2\right).$$

Proof. Let

$$X = \sqrt{\sum_{j=1}^{n} x_j^2}, \qquad Y = \sqrt{\sum_{j=1}^{n} y_j^2}.$$

If $X = 0$ or $Y = 0$, then the inequality holds trivially, as both sides evaluate to 0. Henceforth assume $X, Y > 0$. By the AM-GM inequality, for each j,

$$\frac{1}{2}\left(\frac{x_j^2}{X^2} + \frac{y_j^2}{Y^2}\right) \geqslant \sqrt{\frac{x_j^2}{X^2}\frac{y_j^2}{Y^2}} = \frac{|x_j y_j|}{XY}.$$

Multiplying both sides by (-1), we obtain

$$-\frac{1}{2}\left(\frac{x_j^2}{X^2}+\frac{y_j^2}{Y^2}\right) \leqslant -\frac{|x_j y_j|}{XY}.$$

Therefore, because for all $\alpha \in \mathbb{R}$ $-|\alpha| \leqslant \alpha \leqslant |\alpha|$,

$$-\frac{1}{2}\left(\frac{x_j^2}{X^2}+\frac{y_j^2}{Y^2}\right) \leqslant -\frac{x_j y_j}{XY} \leqslant \frac{1}{2}\left(\frac{x_j^2}{X^2}+\frac{y_j^2}{Y^2}\right)$$

Summing these inequalities,

$$-1 = -\sum_{j=1}^{n}\frac{1}{2}\left(\frac{x_j^2}{X^2}+\frac{y_j^2}{Y^2}\right) \leqslant \sum_{j=1}^{n}\frac{x_j y_j}{XY} \leqslant \sum_{j=1}^{n}\frac{1}{2}\left(\frac{x_j^2}{X^2}+\frac{y_j^2}{Y^2}\right) = 1.$$

Therefore

$$-XY \leqslant \sum_{j=1}^{n} x_j y_j \leqslant XY.$$

Squaring both sides yields the required inequality. □

Want a challenge? Try going through the proof and figure out when equality occurs to prove the sharpest version of the theorem.

One application of this inequality is the *definition* of the angle between two vectors, in n dimensions. Consider two vectors $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$. (A vector is just a list of real numbers, typically written between parentheses.) The *length* of these vectors is defined by

$$|\mathbf{x}| = \sqrt{x_1^2 + \cdots + x_n^2}, \quad |\mathbf{y}| = \sqrt{y_1^2 + \cdots + y_n^2}.$$

The *dot product* of these vectors is defined by

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

The inequality $\mathrm{CS}(n)$ implies that

$$\mathbf{x} \cdot \mathbf{y} = C|\mathbf{x}||\mathbf{y}|,$$

for some real number C satisfying $|C| \leqslant 1$.

There exists a unique number θ between 0 and π such that $C = \cos(\theta)$. This number θ is the *angle* between the two vectors. It is the unique number between 0 and π satisfying

$$\mathbf{x} \cdot \mathbf{y} = \cos(\theta)|\mathbf{x}||\mathbf{y}|.$$

4.7 The binomial theorem

We prove the binomial theorem, which relates the coefficients in the expansion of $(x+y)^n$ to factorials.

Lemma 4.11. *The binomial coefficients satisfy* $\binom{n}{k} = \binom{n}{n-k}$ *for all* $n, k \in \mathbb{N}^0$ *such that* $0 \leqslant k \leqslant n$.

Proof. We can compute

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

In the last step, we reindexed the sum by replacing k by $n-k$.

Since there is an equality of polynomials, $(x+y)^n = (y+x)^n$, there is an equality of coefficients: $\binom{n}{k} = \binom{n}{n-k}$. □

Lemma 4.12. *The binomial coefficients satisfy the following properties.*

1. $\binom{n}{0} = \binom{n}{n} = 1$ *for all* $n \in \mathbb{N}^0$.
2. *If* $n, k \in \mathbb{N}$, *and* $1 \leqslant k \leqslant n$, *then* $\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}$.

Proof. Observe that $\binom{0}{0} = 1$, since $(x+y)^0 = 1 = 1 \times x^0 y^0$. We have verified the first claim when $n = 0$.

The rest of the proof relies on the following computation, valid for all natural numbers n.

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

On the other hand,

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

Matching coefficients, from left to right, we find three facts (the middle is valid when $1 \leqslant k \leqslant n$):

$$\binom{n}{0} = \binom{n+1}{0}, \qquad \binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}, \qquad \binom{n}{n} = \binom{n+1}{n+1}.$$

Since $\binom{0}{0} = 1$ and $\binom{n}{0} = \binom{n+1}{0}$ for all natural numbers n, induction implies that $\binom{n}{0} = 1$ for all nonnegative integers n. Similarly, since $\binom{0}{0} = 1$ and $\binom{n}{n} = \binom{n+1}{n+1}$ for all nonnegative integers n, induction implies that $\binom{n}{n} = 1$ for all nonnegative integers n. □

Next, we relate these binomial coefficients to the factorial. The proof will rely on the recursive relationship found in the previous proposition.

Theorem 4.13 (Binomial theorem)**.** *For all $k, n \in \mathbb{N}^0$ such that $0 \leqslant k \leqslant n$, we have*

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

Proof. For all $k, n \in \mathbb{N}^0$ such that $0 \leqslant k \leqslant n$, define

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

The theorem may be rephrased as the equality $\binom{n}{k} = C(n,k)$ for all such natural numbers k, n. A straightforward computation demonstrates that $C(n,0) = C(n,n) = 1$, for all $n \in \mathbb{N}$, and so the theorem holds in this case.

To continue the proof, we demonstrate that the function $C(n,k)$ satisfies the same recurrence relation as the binomial coefficients. For all natural numbers n, k such that $1 \leqslant k \leqslant n$, we compute

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

This is enough to prove the theorem. Indeed, the truth of the statement $\binom{n}{k} = C(n,k)$ has been demonstrated when (n,k) belongs to either edge of the (infinite) triangular region in figure 3. Since $\binom{n}{k}$ and $C(n,k)$ satisfy the same recurrence relation, we find that the truth of the statement $\binom{n}{k} = C(n,k)$ along any vertical line in the diagram implies the truth of the statement along the next vertical line to the right. Induction implies that $C(n,k)$ is true for all natural numbers n, k.

□

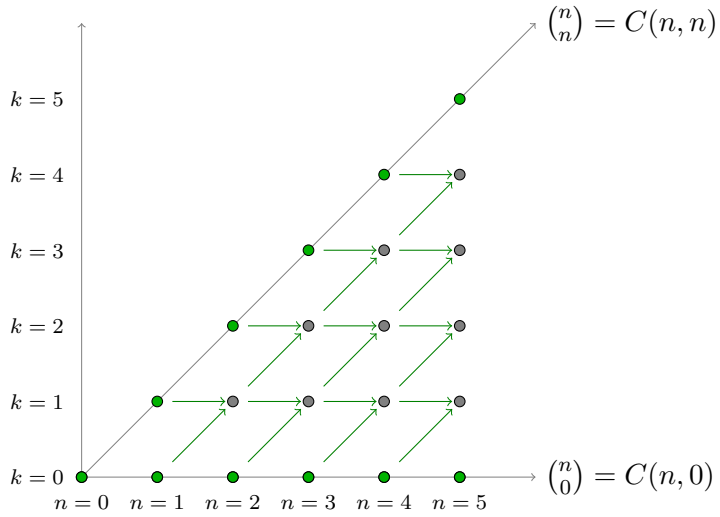


Figure 3: The point with coordinates (n, k) represents the statement $\binom{n}{k} = C(n, k)$. Green dots are proven true at the beginning of the proof. Green arrows represent implications: the proof of each statement marked with a gray dot requires the truth of two statements to its left.

4.8 The quotient-remainder theorem

The famous quotient-remainder theorem is the basis for modular arithmetic (clock arithmetic). At first glance, the quotient-remainder theorem may seem obvious, but it is actually not so easy to prove. This is not the simplest proof, but it goes via another useful result, whose proof demonstrates a different structure for an inductive proof.

Theorem 4.14. *For every $m \in \mathbb{N}$, in any set of m consecutive integers, m divides precisely one of them.*

Proof. Suppose, for a contradiction, that there is an integer k such that none of

$$k, \quad k+1, \quad \ldots, \quad k+m-1$$

are divisible by m. Then

$$k+m, \quad k+m+1, \quad \ldots, \quad k+m+m-1$$

is another set of m consecutive integers of which are none are divisible by m. Moreover, these integers are consecutive to the original integers, so all integers

$$k, \quad k+1, \quad \ldots, \quad k+2m-1$$

are not divisible by m. Hence, by induction, no integers greater than or equal to k can be divisible by m. However, $|k|+1$ is an integer greater than k, so $m(|k|+1)$ is an integer greater than k and divisible by m. Hence our original assumption must have been false, and no such k exists. Therefore, in any set of m consecutive integers, at least one is divisible by m.

Now suppose, for a contradiction, that there is an integer k such that at least two of

$$k, \quad k+1, \quad \ldots, \quad k+m-1$$

are divisible by m. Without loss of generality, we may assume $m \mid k$ (otherwise, increment k by 1 until this holds, and no numbers divisible by m have left the set). Then there is an integer p such that $1 \leqslant p \leqslant m-1$ and $m \mid k$ and $m \mid (k+p)$. But then $m \mid p$, which is impossible, as $m > p$. Hence our assumption was false, and no such k exists. Therefore, at most one number in a set of m consecutive integers is divisible by m. □

Theorem 4.15 (Quotient-remainder theorem)**.** *For any $x \in \mathbb{Z}$ and $m \in \mathbb{N}$, there exists some $r \in \mathbb{Z}$ with $0 \leqslant r \leqslant m-1$ and there exists some $t \in \mathbb{Z}$ such that $x = mt + r$.*

Proof. Suppose the negation. Then there exists an integer x and natural number m such that

$$x-(m-1), \quad x-(m-2), \quad \ldots, \quad x-1, \quad x$$

is a set of m consecutive integers, of which none is divisible by m. This contradicts theorem 4.14, so the quotient-remainder theorem holds. □

4.9 Bézout's lemma

Bézout's lemma is a consequence of the quotient-remainder theorem, and is another fundamental theorem in number theory.

Definition 4.16. Given any two nonzero integers a, b, we define the *greatest common divisor* of a and b to be the largest positive integer d for which both $d \mid a$ and $d \mid b$. We usually denote the greatest common divisor by $\mathrm{GCD}(a, b)$.

The greatest common divisor is also known as the "highest common factor", and various other arrangements of these words.

The greatest common divisor is an essential object in number theory. To calculate the greatest common divisor of a pair of integers, one can use Euclid's algorithm.

Lemma 4.17 (Bézout's lemma)**.** *Suppose that $a, b \in \mathbb{Z} \setminus \{0\}$, and $d = \mathrm{GCD}(a, b)$. Then there exist integers X, Y such that $aX + bY = d$.*

Moreover, if $n \in \mathbb{N}$ also has the property that there exist integers x, y such that $ax + by = n$, then $d \mid n$.

Proof. Choosing $x = y = 1$, we find that the set

$$\Delta = \{n \in \mathbb{N} : \exists\, x, y \in \mathbb{Z} : ax + by = n\}$$

is nonempty. As Δ is a nonempty subset of the natural numbers, it has a least element, which we shall denote d. We aim to show that $d = \mathrm{GCD}(a, b)$, and that d divides all elements of the set Δ. We shall show the latter first.

Associated with d, there are particular integers X, Y for which $aX + bY = d$. Suppose now that $n \in \Delta$ also, with associated integers x, y, so that $ax + by = n$. Suppose further that $n > d$, but $d \nmid n$. By the quotient-remainder theorem, there exists some $t, r \in \mathbb{Z}$, with $0 < r < d$ such that

$$n = td + r.$$

But then

$$\begin{aligned} r &= n - td \\ &= ax + by - t(aX + bY) \\ &= a(x - tX) + b(y - tY), \end{aligned}$$

so $r \in \Delta$ also. But this contradicts our definition of d as the least element of Δ. Therefore, $d \mid n$ for all $n \in \Delta$.

By choosing one of x, y to be 1, and the other to be 0, we see that

$$\begin{aligned} a &= 1a + 0b \in \Delta, \\ b &= 0a + 1b \in \Delta. \end{aligned}$$

Therefore $d \mid a$ and $d \mid b$; d is a common divisor a and b. Now suppose that g is also a common divisor of a and b. Then, for some integers p, q, $a = pg$ and $b = qg$. Hence

$$\begin{aligned} d &= aX + bY \\ &= (pg)X + (qg)Y \\ &= g(pX + qY), \end{aligned}$$

so $g \mid d$. In particular, this means that $g \leqslant d$, so d is the greatest common divisor of a and b. □

4.10 Euclid's lemma

We study various forms of Euclid's lemma. This result is a corollary of Bézout's lemma, and it is in turn the principal tool in the proof of the fundamental theorem of arithmetic.

Definition 4.18. Suppose that a and b are integers. If it happens that 1 is the only natural number which divides both a and b, then we call a and b *coprime*. Another way of saying this is

$$a, b \text{ coprime} \quad \Leftrightarrow \quad \mathrm{GCD}(a, b) = 1.$$

Lemma 4.19 (Euclid's lemma: Gauss's form)**.** *Suppose that $n, a, b \in \mathbb{Z}$. If n and a are coprime, and $n \mid ab$, then $n \mid b$.*

Proof. As $\mathrm{GCD}(n, a) = 1$, Bézout's lemma tells us that there exist integers X, Y such that

$$1 = aX + nY.$$

Multiplying by b, we obtain

$$b = aXb + nYb = n\left(\frac{ab}{n}X + Yb\right).$$

But $n \mid ab$, so the parenthetical quantity on the right hand side is an integer, so $n \mid b$. □

Lemma 4.20 (Euclid's lemma)**.** *Suppose that $p, a, b \in \mathbb{Z}$. If p is prime, and $p \mid ab$, then $p \mid a$ or $p \mid b$.*

Proof. As p is prime, either $p \mid a$ (in which case we are done) or p and a are coprime. In the latter case, by lemma 4.19, $p \mid b$. □

Lemma 4.21 (Euclid's lemma: generalised form)**.** *Suppose that p is prime, and n is a natural number. For any finite sequence of integers $(a_j)_{j=1}^n$, if*

$$p \mid \prod_{j=1}^{n} a_j,$$

then, for some $j \in \{1, 2, \ldots, n\}$, $p \mid a_j$.

Proof. If $n = 1$, then the result holds for $j = 1$. Euclid's lemma 4.20 is the case $n = 2$.

Now suppose that $n \geqslant 2$ and the result holds for any sequence of integers of the particular length $n - 1$, and consider a sequence $(a_j)_{j=1}^n$ for which

$$p \mid \prod_{j=1}^{n} a_j.$$

By Euclid's lemma 4.20, either $p \mid a_n$, or

$$p \mid \prod_{j=1}^{n-1} a_j.$$

Therefore, by the inductive hypothesis, $p \mid a_j$ for some $j \in \{1, 2, \ldots, n\}$.

The result follows, by induction on n. □

4.11 The fundamental theorem of arithmetic

We have already seen that every integer can be factorised into primes. Here we show that this prime factorisation is unique, up to reordering: $12 = 2 \times 2 \times 3 = 2 \times 3 \times 2$, but $12 \neq 5 \times$ [any other prime(s)]. The proof relies upon Euclid's lemma.

Lemma 4.22. *If any two finite products of primes are equal, then the products have the same number of factors and, up to reordering, the factors are the same.*

Proof. For any positive integer N, let $A(N)$ be the statement: "If $n, m \in \mathbb{N}^0$, both $(p_j)_{j=1}^n$ and $(q_j)_{j=1}^m$ are (possibly empty) sequences of primes,

$$\prod_{j=1}^{n} p_j = \prod_{j=1}^{m} q_j,$$

and $n, m \leqslant N$, then $n = m$ and we can reorder the sequence $(q_j)_{j=1}^m$ so that $p_j = q_j$ for each $j \in \{1, 2, \ldots, n\}$." We aim to show, by induction on N, that $A(N)$ is true for all positive integers N.

If $N = 0$, then $0 \leqslant n, m \leqslant 0$ implies $n = m$, and both sequences are empty, so the sequences are the same. Therefore $A(0)$ is true.

Now suppose $A(N)$, for some particular N. Suppose further that $n, m \in \mathbb{N}^0$, both $(p_j)_{j=1}^n$ and $(q_j)_{j=1}^m$ are sequences of primes,

$$\prod_{j=1}^{n} p_j = \prod_{j=1}^{m} q_j,$$

and $n, m \leqslant N + 1$. If $n = 0$ or $m = 0$, then both are 0, and both products are empty, so we may assume further that $n, m \geqslant 1$. Then

$$p_1 \mid \prod_{j=1}^{m} q_j$$

so, by the generalised form of Euclid's lemma, there is some $j \in \{1, 2, \ldots, m\}$ for which $p_1 \mid q_j$. We can reorder the finite sequence $(q_j)_{j=1}^m$ so that $p_1 \mid q_1$. As q_1 is prime, this implies $p_1 = q_1$. Therefore,

$$\prod_{j=2}^{n} p_j = \prod_{j=2}^{m} q_j,$$

and these are products of $n - 1, m - 1 \leqslant N$ primes. By the inductive hypothesis, $n - 1 = m - 1$ and, up to reordering of $(q_j)_{j=2}^m$, the factors match. By reintroducing the first term to each product, we establish that $A(N)$ implies $A(N + 1)$.

Hence, by induction, $A(N)$ holds for all positive integers N. □

Theorem 4.23 (Fundamental theorem of arithmetic)**.** *Every natural number may be written as a product of prime numbers, and that product is unique up to the order of the factors.*

Proof. We already proved the prime factorisation theorem, which established the existence of such a product of primes for each natural number. It remains to show that the uniqueness claim is true.

Suppose that there is a natural number x with two prime factorisations

$$\prod_{j=1}^{n} p_j = x = \prod_{j=1}^{m} q_j.$$

i.e. n and m are (possibly different) natural numbers and all p_j, q_j are primes. By lemma 4.22, these prime factorisations are the same, up to reordering. □

4.12 Binary representation of nonnegative integers

This theorem says that we can write all natural numbers using the binary system. It is worth considering whether we can write all natural numbers using other systems: ternary? hexadecimal? decimal? How would we have to adjust this proof?

Definition 4.24. We say that a sequence is *eventually zero* if all but finitely many of its terms are 0.

Definition 4.25. Each term in a *binary sequence* is either 0 or 1.

Theorem 4.26. *For every $x \in \mathbb{N}^0$, there exists a binary sequence $(a_n)_{n\in\mathbb{N}^0}$, which is eventually zero, such that*

$$x = \sum_{n=0}^{\infty} a_n 2^n.$$

Proof. Note that, because the sequence is eventually zero, the series is in fact a finite sum (just with arbitrarily many terms) so convergence is guaranteed.

The case $x = 0$ is trivially true, with the empty sum given by the sequence in which every term is zero.

Suppose that, for some integer $x \geqslant 0$, for every integer k such that $0 \leqslant k \leqslant x$, there is a binary sequence $(a_{k,n})_{n\in\mathbb{N}^0}$, which is eventually zero, such that

$$k = \sum_{n=0}^{\infty} a_{k,n} 2^n.$$

If $x + 1$ is even, then $\frac{x+1}{2}$ is an integer such that $0 \leqslant \frac{x+1}{2} \leqslant x$ so, by the inductive hypothesis, there exists a sequence $(a_{\frac{x+1}{2},n})_{n\in\mathbb{N}^0}$, which is eventually zero, for which

$$\frac{x+1}{2} = \sum_{n=0}^{\infty} a_{\frac{x+1}{2},n} 2^n.$$

Hence

$$x + 1 = 2\sum_{n=0}^{\infty} a_{\frac{x+1}{2},n} 2^n = \sum_{n=0}^{\infty} a_{\frac{x+1}{2},n} 2^{n+1} = \sum_{n=1}^{\infty} a_{\frac{x+1}{2},n-1} 2^n = \sum_{n=0}^{\infty} a_{x+1,n} 2^n,$$

where the last equality is ensured by defining $a_{x+1,n} = a_{\frac{x+1}{2},n-1}$ for $n \in \mathbb{N}$, and $a_{x+1,0} = 0$. The new sequence is eventually zero because $(a_{\frac{x+1}{2},n})_{n\in\mathbb{N}^0}$ was eventually zero. The new sequence is binary because the original sequence was binary and the only additional term is 0.

If, instead, $x + 1$ is odd, then $\frac{x}{2}$ is an integer such that $0 \leqslant \frac{x}{2} \leqslant x$ so, by the inductive hypothesis, there exists a sequence $(a_{\frac{x}{2},n})_{n\in\mathbb{N}^0}$, which is eventually zero, for which

$$\frac{x}{2} = \sum_{n=0}^{\infty} a_{\frac{x}{2},n} 2^n.$$

Hence

$$x + 1 = 1 + 2\sum_{n=0}^{\infty} a_{\frac{x}{2},n} 2^n = 1 + \sum_{n=0}^{\infty} a_{\frac{x}{2},n} 2^{n+1} = 1 + \sum_{n=1}^{\infty} a_{\frac{x}{2},n-1} 2^n = \sum_{n=0}^{\infty} a_{x+1,n} 2^n,$$

where the last equality is ensured by defining $a_{x+1,n} = a_{\frac{x}{2},n-1}$ for $n \in \mathbb{N}$, and $a_{x+1,0} = 1$. The new sequence is eventually zero because $(a_{\frac{x}{2},n})_{n\in\mathbb{N}^0}$ was eventually zero. The new sequence is binary because the original sequence was binary and the only additional term is 1.

The result follows, by complete induction. □

Check that we have done enough to justify the base; does the step actually work every time we apply it, or is there a problem in one of the first few applications?

Note that we still have not shown that the binary representation of a natural number is unique. How could this be done?

5 Week 5

5.1 Stacking

We analyse a single-player game.

Consider the following one-player game, called *stacking*. The player begins with a score of 0 points and a stack of (at least 1) boxes. On each turn, the player can divide any one stack of $n = a+b$ boxes into two new stacks of a boxes and b boxes, and add ab points to their score. The new stacks do not add to existing stacks, but are completely separate stacks. The game ends when every box is isolated in a stack of height 1.

Theorem 5.1. *Suppose the player begins a game of stacking with n boxes in a single stack. The score at the end of the game will be $n(n-1)/2$, regardless of the player's strategy.*

Proof. Suppose you begin a game with a single stack containing a single box; $n = 1$. Then there is no possible move, and your final score is $0 = 1(1-1)/2$.

Now suppose it is known that, for some n, for all $k \in \mathbb{N}$ with $k \leqslant n-1$, beginning a game with k boxes in a single stack results in a score of $k(k-1)/2$, regardless of strategy. Consider a game which begins with n boxes in a single stack. The first turn in that game results in one stack of a boxes, and another stack of b boxes, and scores ab points, for some $a, b \in \mathbb{N}$. Note that the choice of strategy in the first turn only determines a, b subject to the constraint $a + b = n$. As no boxes may ever be added to stacks, the two stacks may be considered as two new games, which do not interact with one another. But $a, b \leqslant n-1$, so the scores from those two games are known, and independent of strategy. Therefore the game beginning with n boxes in a single stack must score

$$\begin{aligned} ab + \frac{a(a-1)}{2} + \frac{b(b-1)}{2} &= \frac{2a(n-a) + a(a-1) + (n-a)(n-a-1)}{2} \\ &= \frac{n(n-1)}{2}. \end{aligned}$$

Hence, by induction, the theorem holds. □

5.2 Counting the inductive steps in the proof of AMGM

We count how many times we have to apply the lemmata used to prove AM-GM for each n.

In the proof of the arithmetic mean / geometric mean inequality, we used the following lemmata:

Lemma 5.2. AMGM(2).

Lemma 5.3. *If* AMGM(2) *and* AMGM(n), *then* AMGM($2n$).

Lemma 5.4. *If* AMGM(n), *then* AMGM($n-1$).

Theorem 5.5. *Suppose $n \in \mathbb{N}$ and $2 \leqslant n \leqslant 2^p$ for $p \geqslant 1$. We can prove* AMGM(n) *by applying lemmata 5.2, 5.3 and 5.4 a total of no more than $2p-1$ times.*

Proof. Note to presenter / reader: We are doing induction on p, not on n. It may help to visualise n, 2^p, 2^{p+1} on a number line.

If $p = 1$, then $n = 2$. It takes 1 application of lemma 5.2 to prove AMGM(2), and $1 \leqslant 2(1) - 1$.

Suppose that, for some particular $p \geqslant 1$, for all natural k with $2 \leqslant k \leqslant 2^p$, we can prove AMGM($k$) by applying lemmata 5.2, 5.3 and 5.4 a total of no more than $2p-1$ times. Fix some natural n so that $2 \leqslant n \leqslant 2^{p+1}$. If it happens that $n \leqslant 2^p$, then AMGM(n) can be proved in no more than $2p-1$ (which is less than $2(p+1)-1$) applications of the lemmata. Otherwise, $2^p < n \leqslant 2^{p+1}$.

If n is even, then $4 \leqslant 2^p + 2 \leqslant n$, so $2 \leqslant \frac{n}{2} \leqslant 2^p$, so AMGM($\frac{n}{2}$) can be proved using no more than $2p - 1$ applications of the lemmata. By applying lemma 5.3 to AMGM($\frac{n}{2}$), AMGM(n) is proved in one more step, for a total of $2p$ applications, which is less than $2(p + 1) - 1$ applications.

If n is odd, then $2^p + 1 \leqslant n \leqslant 2^{p+1} - 1$, so $2 \leqslant 2^{p-1} + 1 \leqslant \frac{n+1}{2} \leqslant 2^p$, so AMGM($\frac{n+1}{2}$) can be proved using no more than $2p - 1$ applications of the lemmata. Applying lemma 5.3 to AMGM($\frac{n+1}{2}$), AMGM($n + 1$) is proved in one more step. Applying lemma 5.4 to AMGM($n + 1$) yields AMGM(n) in one further step. In this way, AMGM(n) is proved using no more than $2p - 1 + 2 = 2(p + 1) - 1$ applications of the lemmata.

Hence, by induction, it takes no more than a total of $2p - 1$ applications of the lemmata to prove AMGM(n), for $2 \leqslant n \leqslant 2^p$. □

5.3 A game of 'Winner takes all'

We study a game, and find who will win it.

Dave and Mike have decided to play a game of 'Winner takes all'. Before the game begins, each player puts all of their heavy metal CDs into a separate pile, so that there are two piles. A coin is tossed to decide the order of play. On each player's turn, they may take as many CDs as they like from either one pile or the other (but not from both piles) and remove them from the game. On each turn, the player must remove at least one CD from the game. The game is *won* by the player who removes the last CD from the last pile, and lost by the other player. The winning player gets to keep all of the CDs.

Dave very much does not want to lose all his heavy metal CDs.

It seems reasonable to assume that at least one of Dave and Mike owns at least one heavy metal CD, as otherwise there is not much of a game. For $n, k \in \mathbb{N}^0$ with $(n, k) \neq (0, 0)$, let $G(n, k)$ denote a game which begins with one pile of n CDs and one pile of k CDs.

Theorem 5.6. *Player 2 can always win the game $G(n, n)$.*

Proof. In $G(1, 1)$, whichever of the two permissible moves player 1 makes, player 2 will always win with their next turn.

Suppose that player 2 can always win the game $G(k, k)$ for all integers $k = 1, 2, \ldots, n$, and consider the game $G(n + 1, n + 1)$. On their first turn, player 1 must remove p CDs from one pile or the other, for some integer p with $1 \leqslant p \leqslant n + 1$. As the choice of piles does not affect the game, the only consequent choice is the choice of p. If player 1 chooses $p = n + 1$, then player 2 can win the game by removing all the CDs from the other pile. Otherwise, player 2 can remove p CDs from the other pile, and the game is reduced to $G(n + 1 - p, n + 1 - p)$, which is a game that can be won by player 2.

Hence, by complete induction, player 2 can win every game. □

Corollary 5.7. *Player 1 can always win the game $G(n, k)$ if $n \neq k$.*

Proof. The game $G(0, k)$ is equivalent to the game $G(k, 0)$, and is won by player 1 if they remove all the CDs on their first turn. Henceforth, we can assume $n, k \in \mathbb{N}$. On their first turn, player 1 may choose to remove enough CDs from the larger pile so that the piles are equal in size. This reduces the game to $G(n, n)$, but with the players reversed. By theorem 5.6, player 1 (called player 2 in the original theorem) can always win that game. □

5.4 Enumerating the power set

We study the set of subsets of a finite set, drawing a link to binomial coefficients.

Lemma 5.8. *If k is a natural number, and S is a set with $\#S = n \geqslant k$, then*

$$\#\,\mathcal{P}_k(S) = \frac{n-k+1}{k}\#\,\mathcal{P}_{k-1}(S).$$

Proof. We define a new sort of subset of S, called a subcommittee. A *subcommittee* of S consists of a nonempty $C \subset S$ with one element elected as *president*. When describing a subcommittee of a set, we use a "hat" to mark the president. For example, if $S = \{a, b, c, d\}$, one example of a subcommittee of S would be $\{b, c, \hat{d}\}$ (a subset with three elements, with d as president). We write $\hat{\mathcal{P}}_k(S)$ for the set of k-element subcommittees of S.

Define a function $\kappa : \hat{\mathcal{P}}_k(S) \to \mathcal{P}_k(S)$ (we use the Greek letter "κ" (kappa) for "kommunism"), by the rule that when a subcommittee is input, the underlying subset is output. In other words, the president is peacefully dethroned and the elements of the subcommittee are once again equals. For example,

$$\kappa(\{b, c, \hat{d}\}) = \{b, c, d\}.$$

The function κ is k-to-1. In other words, there at k different subcommittees that all map to the same $T \in \mathcal{P}_k(S)$, under the function κ, because any of the k elements of T could have been the president. For example,

$$\kappa(\{\hat{b}, c, d\}) = \kappa(\{b, \hat{c}, d\}) = \kappa(\{b, c, \hat{d}\}) = \{b, c, d\}.$$

Three distinct subcommittees map to the same subset via the function κ.

Next, define a function $\rho : \hat{\mathcal{P}}_k(S) \to \mathcal{P}_{k-1}(S)$ (we use the letter "rho" for "revolution"), by the rule that when a subcommittee is input, the president is thrown out, and the output is the set of remaining elements. For example,

$$\rho(\{b, c, \hat{d}\}) = \{b, c\}.$$

We claim that the function ρ is $(n-k+1)$-to-1. Indeed, how many subcommittees could lead to a given $(k-1)$-element subset $T \subset S$? Such a subcommittee must be composed of the elements of T together with a president (before the revolution), and the president must be an element of $S \setminus T$. Corresponding to the $n-(k-1)$ possible presidents before the revolution, there are $n-(k-1)$ possible subcommittees possible before a revolution led to T. Hence ρ is an $(n-k+1)$-to-1 function.

What we have proven about κ and ρ yield equalities,

$$k\#\,\mathcal{P}_k(S) = \#\hat{\mathcal{P}}_k(S) = (n-k+1)\#\,\mathcal{P}_{k-1}(S).$$

Therefore,

$$\#\,\mathcal{P}_k(S) = \frac{n-k+1}{k}\#\,\mathcal{P}_{k-1}(S). \qquad \square$$

Theorem 5.9. *Suppose that S is a finite set and $n = \#S$. Then $\#\,\mathcal{P}_k(S) = \binom{n}{k}$.*

Proof. When $k = 0$, we find that $\#\,\mathcal{P}_k(S) = 1$ since the only 0-element subset of any set is the empty set.

Now suppose that $1 \leqslant k \leqslant n$ and we know that $\#\,\mathcal{P}_{k-1}(S) = \binom{n}{k-1}$. The binomial theorem tells us that binomial coefficients can be expressed with factorials:

$$\#\,\mathcal{P}_{k-1}(S) = \frac{n!}{(k-1)!(n-k+1)!}.$$

The lemma implies that

$$\begin{aligned}\#\,\mathcal{P}_k(S) &= \frac{n-k+1}{k}\#\,\mathcal{P}_{k-1}(S)\\ &= \frac{n-k+1}{k}\frac{n!}{(k-1)!(n-k+1)!}\\ &= \frac{n!}{k!(n-k)!} = \binom{n}{k},\end{aligned}$$

where we have used the binomial theorem again to justify the last equality.

By induction on k, we find that $\#\,\mathcal{P}_k(S) = \binom{n}{k}$ for all $0 \leqslant k \leqslant n$. □

In the above proof, we see an example of induction within a finite subset of the naturals. Here we are doing induction on k, starting at 0 and going up to n. The number n is an arbitrary parameter, that can be any natural number, but is fixed throughout the proof. So the induction only goes the finite distance $0 \to 1 \to 2 \to \ldots \to n$, but the "finite distance" can be as long as we like; the induction works regardless of what n we choose.

Corollary 5.10. *If S is a finite set, then $\#\,\mathcal{P}(S) = 2^{\#S}$.*

Proof. Let $n = \#S$. The subsets of S have cardinality from 0 to n inclusive, and

$$\#\,\mathcal{P}(S) = \#\,\mathcal{P}_0(S) + \#\,\mathcal{P}_1(S) + \cdots + \#\,\mathcal{P}_n(S).$$

By the previous theorem, we find

$$\#\,\mathcal{P}(S) = \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{n}.$$

By the binomial theorem, this coincides with the expansion of $(1+1)^n$. Hence

$$\#\,\mathcal{P}(S) = (1+1)^n = 2^n = 2^{\#S}.$$ □

5.5 Fermat's little theorem

Fermat may be famous as a mathematician for his "last theorem", but he had plenty of others. This divisibility result is one of the best-known.

Lemma 5.11. *If $pa/b \in \mathbb{Z}$, p is prime and $a, b \in \mathbb{Z}$ with $b \neq 0$, then $p \mid b$ or $a/b \in \mathbb{Z}$.*

Proof. Suppose $p \nmid b$ and $a/b \notin \mathbb{Z}$. Then there exists a prime factor of b of a higher power than that in a. But $p \nmid b$, so p is not that prime factor. So there exists a prime factor of b of a higher power than that in pa. Because prime factorizations of integers are unique, it must be that $pa/b \notin \mathbb{Z}$. We have proved the contrapositive. □

Lemma 5.12. *For any prime number p, and any integer k such that $1 \leqslant k \leqslant p-1$, the binomial coefficient $\binom{p}{k}$ is divisible by p.*

Proof. Consider the binomial coefficient

$$\binom{p}{k} = \frac{p!}{k! \cdot (p-k)!} = \frac{p \cdot (p-1)!}{k(k-1)\cdots 1 \cdot (p-k)(p-k-1)\cdots 1}.$$

All factors in the denominator of the fraction are less than p. Hence the prime decomposition of the denominator involves only prime numbers less than p. Similarly, the prime decomposition of the numerator involves p together with some prime numbers less than p.

Since p occurs in the prime decomposition of the numerator, but not in the prime decomposition of the denominator, and the binomial coefficient is an integer, we find that the binomial coefficient is a multiple of p. □

Theorem 5.13 (Fermat's little theorem)**.** *If x is a natural number and p is a prime number, then $x^p - x$ is a multiple of p.*

Proof. When $x = 0$, $x^p - x = 0$, which is a multiple of p.

Now suppose that the result is proved for a nonnegative integer x: $x^p - x$ is a multiple of p. We can use the binomial theorem to study $(x+1)^p - (x+1)$.

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

By the inductive hypothesis, $x^p - x$ is a multiple of p. By the lemma, all the appearing binomial coefficients are multiples of p. Therefore $(x+1)^p - (x+1)$ is a multiple of p.

By induction, $x^p - x$ is a multiple of p for all natural numbers x. □

5.6 There are two vertices with the same degree

A result about simple graphs.

Theorem 5.14. *Any finite simple graph with at least two vertices has a pair of vertices with the same degree.*

Proof. Suppose that Γ is a graph with n vertices and each vertex has a different degree. Then we can enumerate the vertices so that their order is increasing: $v_0, v_1, v_2, \ldots, v_{n-1}$. The order of each vertex is a nonnegative number, and no greater than the number of other vertices, so it must be that each vertex v_k has order k. In particular v_0 has order 0 (it is no edges to other vertices) and v_{n-1} has order $n-1$ (it has an edge to every other vertex). But those two statements contradict one another. Therefore at least two of the vertices must have the same order. □

5.7 Counting graphs

We count the number of distinct graphs on a set of n vertices.

Theorem 5.15. *Let V be a finite set, and let $n = \#V$. The number of distinct simple graphs with vertex set V is equal to*

$$2^{n(n-1)/2}.$$

Proof. The set of 2-element subsets of V, denoted $\mathcal{P}_2(V)$, has cardinality $\binom{n}{2}$. This binomial coefficient can be evaluated in terms of factorials and simplified:

$$\binom{n}{2} = \frac{n!}{2! \cdot (n-2)!} = \frac{n(n-1)(n-2)\cdots 1}{2 \cdot (n-2)(n-3)\cdots 1} = \frac{n(n-1)}{2}.$$

The edge set of a simple graph is a subset of $\mathcal{P}_2(V)$.

In this way, the edge set is an element of $\mathcal{P}(\mathcal{P}_2(V))$. The number of possible edge-sets can now be computed,

$$\#\,\mathcal{P}(\mathcal{P}_2(V)) = 2^{\#\,\mathcal{P}_2(V)} = 2^{n(n-1)/2}.$$ □

5.8 Counting digraphs

A digraph is an object related to a graph, but in which the "edges" now have an orientation. Each directed edge has an source vertex and a sink vertex.

Definition 5.16. A *digraph* is an ordered pair (V, E), where V is a finite set, and E is a set of ordered pairs of elements of V, called *directed edges*.

Theorem 5.17. *There are* $2^{(\#V)^2}$ *digraphs on the vertex set* V*.*

Proof. Let $n = \#V$. There are n^2 ordered pairs of elements of V, as there are two ordered pairs for each element of $\mathcal{P}_2(V)$, plus one ordered pair for each element of V. The number of different digraphs on V is the number of subsets of the directed edges, which is $\mathcal{P}(\{\text{all possible directed edges}\})$, which has cardinality 2^{n^2}. □

6 Week 6

6.1 $V - E = 1$ for trees

We establish Euler's formula in the special case of a tree. The general form of Euler's formula will be studied later.

Definition 6.1. We say that a path $(v_0, \ldots, v_\ell)$ *backtracks* if it contains a sequence (v_p, v_q, v_p), so that it traverses an edge in one direction and then immediately traverses the same edge in the opposite direction.

In a path that backtracks, by replacing each instance of (v_p, v_q, v_p) with (v_p), we can remove the backtrack. Note that the path may still backtrack, so it is necessary to keep successively removing such instances of backtracking until the path is backtrack-free:

$$(v_0, \ldots, v_p, v_q, v_r, v_q, v_p, \ldots, v_\ell) \mapsto (v_0, \ldots, v_p, v_q, v_p, \ldots, v_\ell) \mapsto (v_0, \ldots, v_p, \ldots, v_\ell).$$

Lemma 6.2. *Suppose the graph $\Gamma = (V, E)$ is a tree, and $v \in V$ has degree 1. Let $V' = V \setminus \{v\}$, and E' be the set of all edges in E excluding the edge incident with v, and define $\Gamma' = (V', E')$. (The process of producing Γ' from Γ is shown in figure 4.) Then Γ' is a tree.*

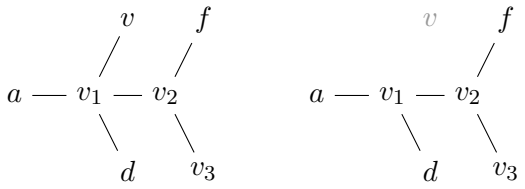


Figure 4: Producing a subgraph $\Gamma' = (V', E')$ by deleting vertex v and edge $\{v, v_1\}$.

Proof. We have to show that Γ' is connected and has no circuits.

If Γ' has a circuit, then that is also a circuit in Γ. But Γ is a tree, so no such circuit exists.

In Γ, there is only one vertex which shares an edge with v; we call that vertex v_1 so that the edge $\{v, v_1\}$ is the only edge deleted in producing Γ' from Γ. Suppose that $x, y$are vertices in Γ'. Then x, y are vertices in Γ, so there is a path in Γ connecting x, y. If that path contains v, then it must contain the backtrack (v_1, v, v_1), so making the path backtrack-free means that it does not pass through v. The backtrack-free path is in Γ'. Therefore Γ' is connected. □

Theorem 6.3. *Suppose that V and E are the vertex and edge sets in a tree, with $\#V \in \mathbb{N}$. Then $\#V = \#E + 1$.*

Proof. We apply induction on the cardinality of V. If $\#V = 1$ (the smallest possible), then there cannot be an edge in the tree. Hence $\#E = 0$ and so $\#V = \#E + 1$.

Now suppose that the theorem is true for all trees with n vertices, for some particular $n \geqslant 1$. Consider a tree Γ with $n + 1$ vertices V, and edge set E. Let $(v_0, v_1 \ldots, v_\ell)$ be a backtrack-free path of maximal length in the tree. Note that such a path exists: since there are no circuits, paths cannot hit the same point twice; since V is finite, no path can pass through more than the finite number $\#V$ vertices. The length of this path is positive, since Γ is a connected graph with more than 1 vertex.

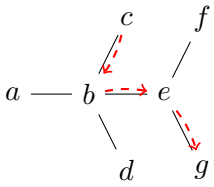


Figure 5: A tree with seven vertices and six edges. Highlighted is a path of maximal length (length 3).

We claim that v_0 has degree 1, i.e., the *only* edge incident with v_0 is the edge $\{v_0, v_1\}$. Indeed, if there were another edge $\{v_0, w\}$, then the maximal-length path $(v_0, \ldots, v_\ell)$ could be extended further to a path $(w, v_0, \ldots, v_\ell)$, contradicting the maximality of the length.

Consider a new graph Γ' with vertex set $V' = V \setminus \{v_0\}$ (all vertices except for v_0), and edge set $E' = E \setminus \{\{v_0, v_1\}\}$. By lemma 6.2, Γ' is a tree. Since Γ' is a tree with n vertices, the inductive hypothesis states that $\#V' = \#E' + 1$. But $\#V = \#V' + 1$ and $\#E = \#E' + 1$, since Γ' was produced by deleting one vertex and deleting one edge. Therefore $\#V = \#E + 1$ too. By induction, we find that $\#V = \#E + 1$ for all finite trees. □

6.2 Planar drawings of connected graphs satisfy $V - E + F = 2$

We prove Euler's famous formula, $V - E + F = 2$, for all planar graphs.

Definition 6.4. A graph is said to be *planar* if it can be drawn in a plane so that no edges cross.

Theorem 6.5. *Let $\Gamma = (V, E)$ be a connected graph (with at least one vertex). Then, in any planar drawing of Γ, the number of faces $\#F$ satisfies the formula*

$$\#V - \#E + \#F = 2.$$

Proof. We apply induction on the number of edges of Γ. If Γ has no edges, then Γ must be the one-point graph (since Γ is connected). Drawing a point in the plane yields $\#V = 1$, $\#E = 0$, and $\#F = 1$. Thus $\#V - \#E + \#F = 2$.

Now assume that $\Gamma = (V, E)$ is a connected graph with at least one edge, and that the theorem has been proven for connected graphs with fewer edges. If Γ has no circuits, then Γ is a tree (since it is assumed connected). In this case $\#V - \#E = 1$ by an earlier theorem. Moreover, a planar drawing of Γ cannot partition the plane into multiple faces, since Γ has no circuits. Hence $\#F = 1$ for any planar drawing of Γ. Hence, for any planar drawing of Γ,

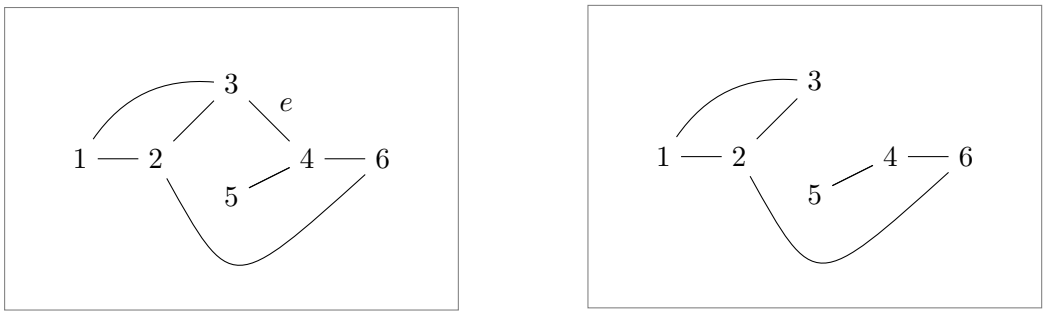
$$(\#V - \#E) + \#F = 1 + 1 = 2.$$

We have proven the theorem in the case where Γ has no circuits.

Otherwise, Γ has at least one circuit. In this case, choose an edge e in the circuit, and let Γ' be the graph obtained by removing e from the edge set of Γ. In other words, define $V' = V$, and define $E' = E \setminus \{e\}$ to obtain the graph $\Gamma' = (V', E')$.

Observe that Γ' is still connected, since the edge e belonged to a circuit.

Given a planar drawing of Γ with $\#F$ faces, we obtain a planar drawing of Γ' by removing the line segment corresponding to the edge e. We claim that the resulting planar drawing of Γ' has $\#F - 1$ faces. Indeed, since e was part of a circuit, the edge e separated two faces in the planar drawing of Γ (by the Jordan curve theorem). Removing the edge e merges those two faces.



Planar drawing of Γ Planar drawing of Γ'

Figure 6: An example in which an edge is removed from a planar graph. The edge e belongs to the circuit 2-3-4-6-2. Removing the edge e merges two faces. Note that faces do not have to have polygonal boundaries!

Since Γ' has 1 fewer edges than Γ, and the drawing of Γ' has 1 fewer faces than a drawing of Γ, we find that

$$\#V - (\#E - 1) + (\#F - 1) = 2.$$

Therefore $\#V - \#E + \#F = 2$.

By induction, we have proven that $\#V - \#E + \#F = 2$ for all planar drawings of connected graphs. □

6.3 K_5 is not planar

We show that the complete graph on 5 vertices is nonplanar, by way of a lemma bounding the number of edges a planar graph may have in terms of the number of vertices it may have.

Recall that the *complete graph* with n vertices is the graph K_n with vertex set $V = \{1, 2, \ldots, n\}$ and every edge possible: $E = \mathcal{P}_2(V)$ is the set of all 2-element subsets of V. The complete graphs on 1, 2, 3, and 4 vertices are planar, as exhibited below.

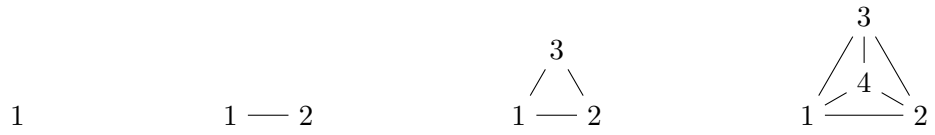


Figure 7: Drawings of K_1, K_2, K_3, and K_4.

Lemma 6.6. *If $\Gamma = (V, E)$ is a connected planar graph with more than two vertices, then $\#E \leqslant 3\#V - 6$.*

Proof. Since we assume that $\#V \geqslant 3$, it follows that $3\#V - 6 \geqslant 3$. Hence the lemma is true when $\#E \leqslant 3$.

Now assume that Γ has more than three edges, and consider a planar drawing of Γ. Let F be the set of faces in this planar drawing. For any face $f \in F$, define $\mathrm{gon}(f)$ to be the number of edges that touch the face f (e.g., $\mathrm{gon}(f) = 3$ if f is a triangular face). Since every edge touches at most two faces in the planar drawing, we find that

$$\sum_{f \in F} \mathrm{gon}(f) \leqslant 2\#E.$$

On the other hand, every face has at least three edges touching it, i.e., for all $f \in F$, $\text{gon}(f) \geqslant 3$. [The reader/presenter should consider why this is true.] Therefore,

$$\sum_{f \in F} \text{gon}(f) \geqslant 3\#F.$$

Putting these two estimates together, we find that $2\#E \geqslant 3\#F$. This estimate, together with Euler's formula, yields the result as follows:

$$\begin{aligned}
& \#V - \#E + \#F = 2 \\
\Rightarrow \quad & 3\#V - 3\#E + 3\#F = 6 \\
\Rightarrow \quad & 3\#V - 3\#E + 2\#E \geqslant 6 \\
\Rightarrow \quad & -\#E \geqslant 6 - 3\#V \\
\Rightarrow \quad & \#E \leqslant 3\#V - 6.
\end{aligned}$$

□

Theorem 6.7. *The complete graph on 5 vertices is not planar.*

Proof. The complete graph on 5 vertices contains $\binom{5}{2} = \frac{5!}{2!\cdot 3!} = 10$ edges. But $10 \leqslant 3(5) - 6 = 9$ is false. By lemma 6.6, the complete graph on 5 vertices is not planar. □

6.4 $K_{3,3}$ is not planar

We show that the complete bipartite graph $K_{3,3}$ is nonplanar, by way of a lemma relating the number of edges, vertices and minimal circuit length for planar graphs.

Definition 6.8. A graph $\Gamma = (V, E)$ is called *bipartite* if the set V can be divided into two disjoint sets U_1 and U_2 such that every edge connects a vertex in U_1 with a vertex in U_2. A bipartite graph is called *complete* if every vertex of U_1 is connected with every vertex of U_2.

Lemma 6.9. *Suppose that $\Gamma = (V, E)$ is a planar graph. Moreover, suppose that $\#E \geqslant \ell > 2$, and Γ has no circuit of length less than ℓ. Then*

$$\#E \leqslant \frac{\ell}{\ell - 2}(\#V - 2).$$

Proof. Consider a planar drawing of Γ, and let F be the resulting set of faces. For any face $f \in F$, define $\mathrm{gon}(f)$ to be the number of edges that touch the face f (e.g., $\mathrm{gon}(f) = 3$ if f is a triangular face). Since every edge touches at most two faces in the planar drawing, we find that

$$\sum_{f \in F} \mathrm{gon}(f) \leqslant 2\#E.$$

Since the graph has no cycle of length less than ℓ, every bounded face must touch at least ℓ edges. Moreover, since $\#E \geqslant \ell$, the unbounded face must touch ℓ edges too [Presenter: can you explain why?]. Hence every face contains at least ℓ edges, and so

$$\sum_{f \in F} \mathrm{gon}(f) \geqslant \ell\#F.$$

Putting these estimates together, we find that

$$\ell\#F \leqslant 2\#E.$$

Using Euler's formula, $\#V - \#E + \#F = 2$, we deduce the estimate

$$\#E \leqslant \frac{\ell}{\ell - 2}(\#V - 2).$$

[The reader/presenter should carry out this algebraic deduction.] □

Theorem 6.10. *The complete bipartite graph $K_{3,3}$ is not planar.*

Proof. The graph $K_{3,3}$ has 6 vertices and 9 edges. Moreover, $K_{3,3}$ has no circuits of length less than 4 (any path alternates between x's and y's, and cannot return to where it starts in fewer than 4 steps). If $K_{3,3}$ were planar, then we would find

$$9 \leqslant \frac{4}{4 - 2}(6 - 2) = 2 \times 4 = 8,$$

a contradiction. Hence $K_{3,3}$ cannot be planar. □

6.5 Subgraphs of planar graphs

We observe that the subgraphs of planar graphs are also planar, and use this to show that certain hierarchies of graphs are nonplanar

Lemma 6.11. *Suppose Γ is a planar graph, and Γ' is a subgraph of Γ. Then Γ' is planar.*

Proof. Consider a planar drawing of Γ. From that drawing, delete all paths that correspond to edges not in Γ'. As no paths have been added, the new drawing cannot have any crossing paths, so it is the drawing of a planar graph. From the new drawing, delete all points that correspond to vertices that do not appear in Γ'. The resulting figure is a drawing of a graph which has precisely the same edges and vertices as Γ', and is a planar drawing, so it is a planar drawing of Γ'. □

Proposition 6.12. *If $q \geqslant 5$, then the complete graph K_q is nonplanar.*

Proof. If $q = 5$, then K_q is nonplanar by an earlier result. Now suppose that K_q is nonplanar and consider the graph $\Gamma = K_{q+1}$. From Γ, form its subgraph Γ' by deleting one vertex and all edges incident to that vertex. Then Γ' has q vertices. Moreover, Γ' has q fewer edges than Γ has edges, so Γ' has

$$\frac{q(q+1)}{2} - q = \frac{q(q-1)}{2}$$

edges. There is only one graph with q vertices and $q(q-1)/2$ edges, and that is K_q. By lemma 6.11, if Γ is planar, then Γ' is planar, but this contradicts the nonplanarity of K_q, so $\Gamma = K_{q+1}$ is nonplanar. Hence, by induction, K_q is nonplanar for all $q \geqslant 5$. □

Proposition 6.13. *If $q \geqslant 3$, then the complete bipartite graph $K_{q,3}$ is nonplanar.*

Proof. If $q = 3$, then $K_{q,3}$ is nonplanar by an earlier result. Now suppose that $K_{q,3}$ is nonplanar and consider the graph $\Gamma = K_{q+1,3}$. Label the vertices of Γ as $x_1, x_2, \ldots, x_{q+1}, y_1, y_2, y_3$ in such a way that no 'x' vertex shares an edge with another 'x' vertex, and no 'y' vertex shares an edge with another 'y' vertex. From Γ, form the subgraph Γ' by deleting the vertex x_{q+1} and all incident edges. Then Γ' is a graph with vertices $x_1, x_2, \ldots, x_q, y_1, y_2, y_3$. For $1 \leqslant j \leqslant q$ and $1 \leqslant k \leqslant 3$, the edge $\{x_j, y_k\}$ is in Γ and was not deleted, so it is in Γ'. Hence Γ' contains $K_{q,3}$ as a subgraph. If $\{\alpha, \beta\}$ is an edge in Γ', then it is an edge in Γ and $x_{q+1} \notin \{\alpha, \beta\}$. Hence $\{\alpha, \beta\} = \{x_j, y_k\}$ for some choice of $1 \leqslant j \leqslant q$ and $1 \leqslant k \leqslant 3$. Therefore Γ' is a subgraph of $K_{q,3}$, and it follows that $\Gamma' = K_{q,3}$. By lemma 6.11, if Γ is planar, then Γ' is planar, but this contradicts the nonplanarity of $K_{q,3}$, so $\Gamma = K_{q+1,3}$ is nonplanar. Hence, by induction, $K_{q,3}$ is nonplanar for all $q \geqslant 3$. □

6.6 Planar complete multipartite graphs

We define multipartite graphs, and fully classify them according to planarity.

Definition 6.14. For integer $n \geqslant 2$, and any choice of natural $p_1, \ldots, p_n$, the *complete n-partite* graph $K_{p_1,p_2,\ldots,p_n}$ is the graph (V, E) with

$$V = \{v_{1,1}, \ldots, v_{1,p_1}, \quad v_{2,1}, \ldots, v_{2,p_2}, \quad \ldots, \quad v_{n,1}, \ldots, v_{n,p_n}\},$$
$$E = \{\{v_{j,x}, v_{k,y}\} : j \neq k,\ 1 \leqslant x \leqslant p_j,\ 1 \leqslant y \leqslant p_k\}.$$

Theorem 6.15. *We adopt the convention that $p_1 \geqslant p_2 \geqslant \ldots \geqslant p_n$ for complete n-partite graphs. The planar complete multipart graphs are:*

$n = 2$: *$K_{p,q}$, for $q \in \{1, 2\}$.*

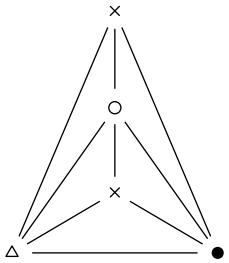


Figure 8: A planar drawing of the graph $K_{2,1,1,1}$.

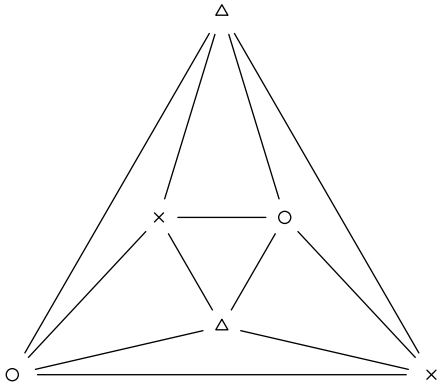


Figure 9: A planar drawing of the graph $K_{2,2,2}$.

$n = 3$**:** $K_{p,1,1}$ *and, for* $r \in \{1, 2\}$, $K_{2,2,r}$.

$n = 4$**:** $K_{p,1,1,1}$ *for* $p \in \{1, 2\}$.

All other complete multipartite graphs are nonplanar.

Proof.
$\boxed{n \geqslant 5}$ If $n \geqslant 5$, then $K_{1,1,1,1,1} = K_5$ is a subgraph of $K_{p_1,p_2,\ldots,p_n}$. But K_5 is nonplanar, so $K_{p_1,p_2,\ldots,p_n}$ is nonplanar.
$\boxed{n = 4}$ The graph $K_{2,1,1,1}$ has the drawing shown in figure 8, so it is planar. The graph $K_{1,1,1,1}$ is a subgraph of $K_{2,1,1,1}$, so is also planar. Every other complete 4-partite graph has $K_{3,1,1,1}$ or $K_{2,2,1,1}$ as a subgraph. But each of these graphs have the nonplanar graph $K_{3,3}$ as a subgraph. Hence no other complete 4-partite graphs are planar.
$\boxed{n = 3}$ The graph $K_{2,2,2}$ has the drawing shown in figure 9, so it is planar. The graph $K_{2,2,1}$ is a subgraph of $K_{2,2,2}$, so is also planar.

The graph $K_{p,1,1}$ has the drawing shown in figure 10, for any p, so it is planar.

All other complete tripartite graphs $K_{p,q,r}$ have both $p \geqslant 3$ and $q + r \geqslant 3$. But then $K_{3,3}$ is a subgraph of $K_{p,q,r}$, so $K_{p,q,r}$ is nonplanar.
$\boxed{n = 2}$ As $K_{p,2}$ is a subgraph of $K_{p,1,1}$, which is planar, $K_{p,2}$ is planar. Hence its subgraph $K_{p,1}$ is also planar. All other complete bipartite graphs $K_{p,q}$ obey $p \geqslant q \geqslant 3$, so have $K_{3,3}$ as a subgraph. Therefore, no other complete bipartite graphs are planar. □

6.7 Planar graphs are 6-colourable

It is a famous theorem that all planar graphs are 4-colourable. Unfortunately, the proof is vastly too complicated for this course. In fact, the proof separates into so many cases that it was only completed

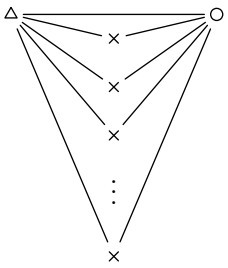


Figure 10: A planar drawing of the graph $K_{p,1,1}$ for $p \in \mathbb{N}$.

with the aid of a computer. Of course, if all planar graphs are 4-colourable, then they are also 5-colourable, 6-colourable, etc. The proof that all planar graphs are 5-colourable is accessible without much more graph theory than you now know, but it is still quite a long proof. Here we give a proof that all graphs are 6-colourable.

Let $\Gamma = (V, E)$ be a graph. A *colouring* of Γ consists of a finite set of colours C (e.g. $C = \{\text{red}, \text{green}, \text{blue}\}$) and a function from V to C (the colouring) such that adjacent vertices do not have the same colour. For example, bipartite graphs can be coloured using only two colours.

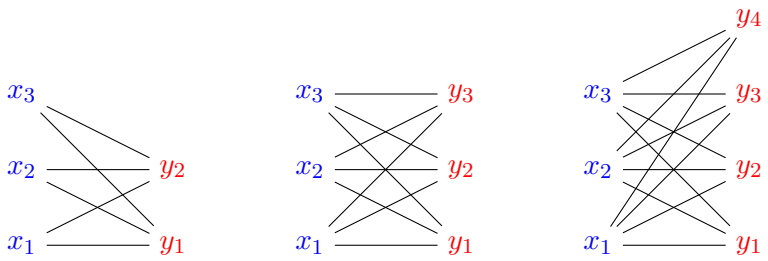
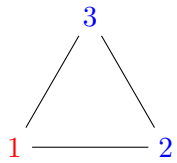
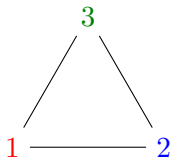


Figure 11: The complete bipartite graphs $K_{3,2}$, $K_{3,3}$, and $K_{3,4}$. We have coloured the graphs using red and blue. Vertices joined by an edge are given different colours

On the other hand, the complete graphs K_3, K_4, K_5, ..., require at least 3, 4, 5, etc., colours for a colouring. [Reader: explain why!]



A forbidden colouring!



An accepted colouring.

Lemma 6.16. *Suppose that $\Gamma = (V, E)$ is a planar graph. Then the average degree of the vertices of Γ is less than 6.*

Proof. By a previous lemma (in the proof that K_5 was not planar), we know that

$$\#E \leqslant 3\#V - 6.$$

On the other hand, counting edges, vertex by vertex, yields

$$2\#E = \sum_{v \in V} \deg(v).$$

Hence we have

$$\frac{1}{2} \sum_{v \in V} \deg(v) \leqslant 3\#V - 6.$$

Multiplying by 2 and dividing by $\#V$ yields

$$\frac{\sum_{v \in V} \deg(v)}{\#V} \leqslant 6 - \frac{12}{\#V}.$$

The left side is the average degree of the vertices of Γ. Since $12/\#V$ is positive, the lemma follows. □

Theorem 6.17. *Every planar graph can be coloured with six colours.*

Proof. We apply induction on the number of vertices. If the graph has 6 or fewer vertices, it can certainly be coloured with six colours (give every vertex its own colour!).

Now suppose that $\Gamma = (V, E)$ is a planar graph with more than 6 vertices, and assume the theorem is proven for planar graphs with fewer vertices. By the lemma, the average degree of a vertex of Γ is less than 6. Hence there exists a vertex $v \in V$ such that $\deg(v) \leqslant 5$.

Let Γ' be the graph obtained by deleting v and all the edges containing v. This is a planar graph with fewer vertices than Γ, and so Γ' can be coloured by six colours. Carry out such a colouring.

Since v was adjacent to at most five other vertices, there exists a colour among the six, unused by the vertices adjacent to v in the colouring of Γ'. Colour v by such an unused colour. The result is a colouring of Γ by six colours.

[Reader / Presenter: Illustrate the process of removing v, colouring everything else, then bringing back v and choosing an appropriate colour.] □

7 Week 7

7.1 Compositions of n

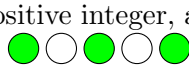
We study the number of different sums of natural numbers that yield a certain result.

Definition 7.1. Let n be a positive integer. A *composition* of n is an ordered sequence of positive integer(s) that sums to n. For example, there are 8 compositions of 4:

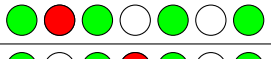
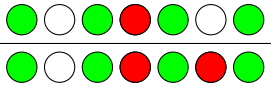
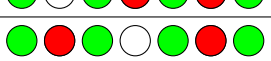
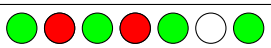
$$4=4, \quad 3+1=4, \quad 1+3=4, \quad 2+2=4,$$
$$2+1+1=4, \quad 1+2+1=4, \quad 1+1+2=4, \quad 1+1+1+1=4.$$

Note that $1+3$ and $3+1$ are different compositions.

Theorem 7.2. *The number of compositions of n equals 2^{n-1}.*

Proof. Let n be a positive integer, and consider a sequence of n green balls, with $n-1$ empty circular slots between them:  .

Let S be the set of $n-1$ circular slots. Any composition of n may be encoded by a subset of S according to the following procedure: for a composition of the form $x_1+\cdots+x_k=n$, place a red ball in the slot after x_1 green balls, after x_1+x_2 green balls, etc.. The encoding is illustrated by the table below.

$4=4$	
$3+1=4$	
$1+3=4$	
$2+2=4$	
$2+1+1=4$	
$1+2+1=4$	
$1+1+2=4$	
$1+1+1+1=4$	

This encoding gives a bijection between the set of compositions of n and the set of subsets of $\{1,\ldots,n-1\}$. The latter set has 2^{n-1} elements, and so there are 2^{n-1} compositions of n. □

7.2 Summing to n

We study nonnegative integer solutions of $x_1+x_2+\cdots+x_k=n$, where n and k are both fixed, and $x_1,\ldots,x_k$ are to be chosen.

Before we prove the theorem below, let's see a few examples. If $k=1$, then we are looking for solutions to the equation $x_1=n$. There's one solution to this silly equation. If $k=2$, then we are looking for ordered pairs of nonnegative integers (x_1,x_2) such that $x_1+x_2=n$. There are $n+1$ possibilities:

$$n+0=n, \quad (n-1)+1=n, \quad (n-2)+2=n, \quad \ldots, \quad 1+(n-1)=n, \quad 0+n=n.$$

If $k=3$, then this gets more difficult. Consider the case $k=3$ and $n=2$; we are looking for ordered triples (x_1,x_2,x_3) such that $x_1+x_2+x_3=2$. Here are the six possibilities:

$$2+0+0=2, \qquad 1+1+0=2, \qquad 1+0+1=2,$$
$$0+2+0=2, \qquad 0+1+1=2, \qquad 0+0+2=2.$$

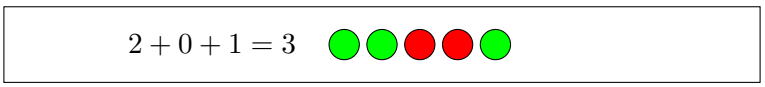
Theorem 7.3. *Suppose $k, n \in \mathbb{N}$ are fixed. Then the number of nonnegative integer solutions $(x_1, \ldots, x_k)$ to the equation*

$$x_1 + x_2 + \cdots + x_k = n$$

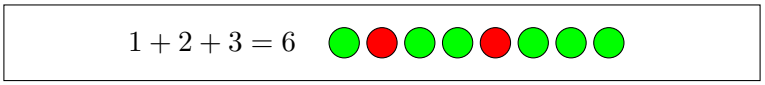
is equal to $\binom{n+k-1}{k-1}$.

Proof. Let k and n be positive integers, and let $(x_1, \ldots, x_k)$ be a nonegative integer solution to the equation $x_1 + \cdots + x_k = n$. We encode this solution with a sequence of green and red balls (or letters G and R will suffice when colors are limited) in the following way: begin with x_1 green balls, then one red ball, then x_2 green balls, then one red ball, then x_3 green balls, then one red ball, etc., until one places the final x_k green balls. In other words, the "plus signs" become red balls, and each x_j becomes x_j green balls. The resulting sequence of balls contains n green balls, and $k-1$ red balls.

For example, the nonegative integer solution $2+0+1=3$, with $x_1 = 2$, $x_2 = 0$, $x_3 = 1$, and $n = 3$, would be encoded by



and the nonegative integer solution $1 + 2 + 3 = 6$, with $x_1 = 1$, $x_2 = 2$, $x_3 = 3$, and $n = 6$, would be encoded by



This encoding is bijective: if one is given a sequence of green and red balls, with $k-1$ red balls, and n green balls, one may convert red balls to plus signs to get a nonegative integer solution to the equation $x_1 + \cdots + x_k = n$. Hence there is a one to one correspondence between the two sets:

- the set of nonegative integer solutions to the equation $x_1 + \cdots + x_k = n$;
- the set of left-to-right arrangements of $k-1$ red balls and n green balls.

The latter set can be enumerated as follows: such an arrangement of $k-1$ red balls and n green balls can be obtained by *beginning* with $n+k-1$ balls in order (labelled $1, 2, 3, \ldots, n+k-1$), and choosing a subset of $k-1$ of them to paint red (the rest to be painted green). This provides a one-to-one correspondence between the latter set and

- the set of $(k-1)$-element subsets of $\{1, 2, 3, \ldots, n+k-1\}$.

We enumerated this set in an earlier proof; it has $\binom{n+k-1}{k-1}$ elements. Thus the set of nonegative integer solutions to the equation $x_1 + \cdots + x_k = n$ has the same number of elements. □

Corollary 7.4. *Suppose $k, n \in \mathbb{N}$. Then the number of **natural number** solutions $(x_1, \ldots, x_k)$ to the equation*

$$x_1 + x_2 + \cdots + x_k = n$$

is equal to $\binom{n-1}{k-1}$, if $n \geqslant k$, and is zero otherwise.

Proof. If $n < k$, then one may not add k natural numbers to obtain n. So we assume $n \geqslant k$ in what follows.

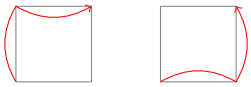
Any natural number solution to $x_1 + x_2 + \cdots + x_k = n$ yields a nonegative integer solution to the equation $y_1 + y_2 + \cdots + y_k = n - k$, by defining $y_i = x_i - 1$ for all i between 1 and n. In the opposite direction, every nonegative integer solution to the equation $y_1 + y_2 + \cdots + y_k = n - k$ gives a natural number solution to $x_1 + x_2 + \cdots + x_k = n$.

Hence there is a bijection between the two solution sets. By the previous theorem, there are $\binom{n-k+k-1}{k-1} = \binom{n-1}{k-1}$ nonegative integer solutions to the equation $y_1 + \cdots + y_k = n - k$. Hence there are the same number of natural number solutions to $x_1 + \cdots + x_k = n$. □

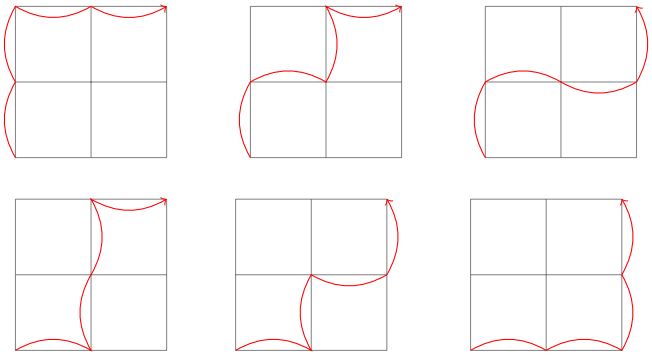
7.3 Lattice paths

We find the number of different paths through a rectangular lattice which take the fewest possible steps.

Before we study this theorem, we illustrate "admissible lattice-paths" with a few examples. If $a = b = 1$, then the grid is a simple square, and there are two admissible lattice paths.



An admissible lattice path proceeds only along the grid lines, and only travels upwards and rightwards. For a more complicated example, consider $a = b = 2$. There are 6 admissible lattice paths.



We are interested in the "admissible" paths in particular, because they are the ones that require the fewest steps to traverse from the bottom-left to top-right corner.

Theorem 7.5. *Let a and b be positive integers. Consider a grid, a units wide and b units tall. The number of admissible lattice-paths from the bottom-left corner to the top-right corner equals $\binom{a+b}{a}$.*

Proof. Let a and b be positive integers, and consider an admissible lattice path from the bottom-left corner to the top-right corner of the a by b grid. Any such path can be encoded by a sequence of letters: "N" for north and "E" for east. For example, in the case $a = 2, b = 2$, the six admissible lattice paths are encoded:

$$\text{NNEE}, \qquad \text{NENE}, \qquad \text{NEEN}, \qquad \text{ENNE}, \qquad \text{ENEN}, \qquad \text{EENN}.$$

For a sequence of Ns and Es to encode an admissible lattice path, the letter "E" must occur a times and the letter "N" must occur b times. Therefore, the number of admissible lattice paths equals the number of strings of Ns and Es, with a Es and b Ns.

The set of such strings can be enumerated as follows: begin with the set of numbers $\{1, 2, \ldots, a+b\}$ from 1 to $a + b$; choose a subset $K \subset \{1, 2, \ldots, a + b\}$ of cardinality a. This determines a string of $a + b$ letters, by placing Ns and Es in order, with Es in each position labeled by K, and Ns in each position not labeled by K. For example,

$a = 2,$	$b = 2,$	$K = \{2, 4\}$	corresponds to NENE;
$a = 3,$	$b = 4,$	$K = \{1, 3, 7\}$	corresponds to ENENNNE.

This gives a bijection between the set of strings of a Es and b Ns and the set of a-element subsets of $\{1, 2, \ldots, a + b\}$. We have already enumerated the latter set; it has $\binom{a+b}{a}$ elements. Therefore the number of admissible lattice paths equals $\binom{a+b}{a}$ too. □

7.4 Elementary properties of bijections

We survey a few elementary properties of bijections, which will come in handy in later proofs.

The first proposition says that if we start with an injection, and shrink its codomain down to just the image, then we get the surjection property.

Proposition 7.6. *Suppose $f : X \to Y$ is an injection and define $g : X \to f(X)$ by $g(x) = f(x)$. Then g is a bijection.*

Proof. If $g(x) = g(y)$, then $f(x) = f(y)$, so $x = y$. Hence g is an injection. By definition, $g(X) = f(X)$, which is the codomain of g, so g is a surjection. □

The injection and surjection parts of the above proof are completely separate. So we could have started with any function, shrunk its codomain to the image, and got a surjection. The resulting surjection would be a bijection if and only if the original function was an injection.

The next proposition allows us to show that two bijections are in fact the same function, even if we don't know whether they agree at one point. (If they might disagree at more than one point, then we would need some extra information to know whether they were the same function. Why is that?)

Proposition 7.7. *Suppose $\xi \in X$, and $f, g : X \to Y$ are bijections. If, for all $x \in X \backslash \{\xi\}$, $f(x) = g(x)$, then $f(\xi) = g(\xi)$.*

Proof. Suppose $f(\xi) \neq g(\xi)$. If there exists an $x \in X \backslash \{\xi\}$ such that $g(x) = f(\xi)$, then also $f(x) = f(\xi)$, which contradicts f being injective. So for all $x \in X \setminus \{\xi\}$, $g(x) \neq f(\xi)$. But also $g(\xi) \neq f(\xi)$, so $f(\xi)$ is not in the image of g, which contradicts g being surjective. Hence our supposition was incorrect, and $f(\xi) = g(\xi)$. □

The following proposition uses two pieces of terminology that may be new: TFAE is a standard abbreviation for "the following are equivalent", a *singleton* is a 1-element set.

Proposition 7.8. *For any function $f : X \to Y$, TFAE:*

1. *The function f is a bijection.*
2. *The preimage of every singleton in f is a singleton. (i.e. for all $\alpha \in \mathcal{P}_1(Y)$, $f^{-1}(\alpha) \in \mathcal{P}_1(X)$.)*

Proof.

(1)⇒(2) Suppose f is a bijection. As f is a surjection, for each $y \in Y$, the preimage of $\{y\}$ in f is nonempty: $\#f^{-1}(\{y\}) \geqslant 1$. As f is an injection, for each $y \in Y$, the preimage of $\{y\}$ in f can have at most one element: $\#f^{-1}(\{y\}) \leqslant 1$. So the preimage of each $\{y\} \subset Y$ in f is a singleton.

(2)⇒(1) Suppose that for all $\alpha \in \mathcal{P}_1(Y)$, $\#f^{-1}(\alpha) = 1$. This means that no two elements in X map to a single $y \in Y$ (otherwise there would be a set containing more than one element that is the preimage of a singleton), so f is an injection. Also, every element in Y is the output of the function evaluated at some element of X (otherwise there would be an empty preimage), so f is a surjection. □

Definition 7.9. Suppose $f : X \to Y$ and $g : U \to V$ are functions, and $Y \subset U$. Then the *composition of g with f*, denoted $g \circ f : X \to Z$, is the function defined by

$$g \circ f(x) = g(f(x)),$$

for all $x \in X$.

Note that composition of functions does not commute (except for very particular choices of functions), so it is important to get it the right way around. Fortunately, this is easy to remember, as g and f appear in the same order in both expressions $g \circ f(x)$ and $g(f(x))$, which, by definition, mean the same thing.

It is easy enough to prove that composition of functions is associative: $(f \circ g) \circ h = f \circ (g \circ h)$.

Proposition 7.10. *Suppose $f : X \to Y$ is a bijection, and define $g : Y \to X$ by*

$$g(y) = \textit{the only element of the singleton } f^{-1}(\{y\}).$$

Then g is a bijection. Moreover, $f \circ g = \mathrm{Id}_Y$ and $g \circ f = \mathrm{Id}_X$. We call g the inverse *of f and usually denote it f^{-1}.*

Proof. By proposition 7.8, for each $y \in Y$, $f^{-1}(\{y\})$ is indeed a singleton, so this is a valid definition of a function.

If $y, y' \in Y$ and $g(y) = g(y')$, then $f^{-1}(\{y\}) = f^{-1}(\{y'\})$, so $\{y\}$ and $\{y'\}$ have the same preimage in f, which means both y and y' are the output of f for the same input. This can only occur if $y = y'$. So g is an injection. As $f(X) \subset Y$, $f^{-1}(Y) \supset X$ [actually they are both equalities, but these are the inclusions that really matter to the argument], so g is a surjection. We have established that the function g is a bijection.

For $x \in X$, by the definition of preimage, the preimage of $\{f(x)\}$ in f certainly contains x, so

$$g \circ f(x) = g(f(x)) = \text{the only element of the singleton } f^{-1}(\{f(x)\}) = x.$$

The definition of the preimage of $U \subset Y$ in f is the set of points $x \in X$ for which $f(x) \in U$. So applying f to any member of the preimage of $\{y\}$ must yield y. Hence, for $y \in Y$,

$$f \circ g(y) = f(g(y)) = f\Big(\text{the only element of the singleton } f^{-1}(\{y\})\Big) = y. \qquad \square$$

Proposition 7.11. *Suppose $f : X \to Y$ and $g : Y \to Z$ are bijections. Then $g \circ f : X \to Z$ is a bijection, and its inverse is $f^{-1} \circ g^{-1} : Z \to X$.*

Proof. As the domain of g is a subset of the codomain of f, composition of functions $g \circ f$ yields a well-defined function.

By the definition of composition, for $\alpha \subset Z$, the preimage of α in $g \circ f$ is obtained by finding the preimage of α in g, and then the preimage of that set in f:

$$[g \circ f]^{-1}(\alpha) = f^{-1}\left(g^{-1}(\alpha)\right).$$

By proposition 7.8, the preimage of any singleton in g is a singleton, and the preimage of any singleton in f is a singleton. Hence the preimage of any singleton in $g \circ f$ is a singleton. Hence, by proposition 7.8 (used in the opposite direction this time), $g \circ f$ is a bijection.

By proposition 7.10, for all $x \in X$, $y \in Y$ and $z \in Z$, $f^{-1} \circ f = \mathrm{Id}_X$, $f \circ f^{-1} = \mathrm{Id}_Y$, $g^{-1} \circ g = \mathrm{Id}_Y$ and $g \circ g^{-1} = \mathrm{Id}_Z$. But

$$\begin{aligned}
\left(f^{-1} \circ g^{-1}\right) \circ (g \circ f) &= f^{-1} \circ \left(g^{-1} \circ g\right) \circ f = f^{-1} \circ \mathrm{Id}_Y \circ f = f^{-1} \circ f = \mathrm{Id}_X, \\
(g \circ f) \circ \left(f^{-1} \circ g^{-1}\right) &= g \circ \left(f \circ f^{-1}\right) \circ g^{-1} = g \circ \mathrm{Id}_Y \circ g^{-1} = g \circ g^{-1} = \mathrm{Id}_Z. \qquad \square
\end{aligned}$$

7.5 Counting permutations

Bijections are very powerful tools for deciding the cardinality of finite sets. But sometimes it can be useful to count sets *of bijections*. We can do this using bijections on the set of bijections! Here we prove the fundamental result that there are $(\#X)!$ bijections from a finite set X to itself.

Definition 7.12. Suppose X is a nonempty finite set. A bijection from X to X is called a *permutation on X*. We denote the set of all permutations on X by $\mathrm{Perm}(X)$. If $a \in X$, then the set of all permutations on X which send a to itself is denoted $\mathrm{Perm}_a(X)$.

Lemma 7.13. *Suppose X is a finite set with $\#X = n > 1$, and $a \in X$. Then $\#\mathrm{Perm}_a(X) = \#\mathrm{Perm}(X \setminus \{a\})$.*

Proof. Define the operator $s : \mathrm{Perm}(X \setminus \{a\}) \to \mathrm{Perm}_a(X)$ by

$$[s\phi](x) = \begin{cases} a & \text{if } x = a, \\ \phi(x) & \text{otherwise.} \end{cases}$$

We want to show that s is a bijection, by showing that s is an injection and a surjection.

If $\phi, \psi \in \mathrm{Perm}(X \setminus \{a\})$ and $\phi \neq \psi$, then there is some $b \in X \setminus \{a\}$ for which $\phi(b) \neq \psi(b)$. But then $[s\phi](b) \neq [s\psi](b)$, so $[s\phi] \neq [s\psi]$. Therefore s is an injection.

Now suppose $f \in \mathrm{Perm}_a(X)$. We define ϕ by $\phi(x) = f(x)$ for each $x \in X \setminus \{a\}$. As $f \in \mathrm{Perm}_a(X)$, $f(X \setminus \{a\}) = X \setminus \{a\}$, so $\phi \in \mathrm{Perm}(X \setminus \{a\})$. Choosing ϕ in this way ensures $[s\phi] = f$; so every f is the output of s for some input ϕ. Therefore s is a surjection.

As s is a bijection, $\#\mathrm{Perm}_a(X) = \#\mathrm{Perm}(X \setminus \{a\})$. □

Lemma 7.14. *Suppose X is a finite set and $a, y, z \in X$. For $f, g \in \mathrm{Perm}_a(X)$, we define $f_y, g_z \in \mathrm{Perm}(X)$ by*

$$f_y(x) = \begin{cases} y & \text{if } x = a, \\ a & \text{if } x = f^{-1}(y), \\ f(x) & \text{otherwise,} \end{cases} \qquad g_z(x) = \begin{cases} z & \text{if } x = a, \\ a & \text{if } x = g^{-1}(z), \\ g(x) & \text{otherwise.} \end{cases} \tag{7.1}$$

Then $f_y = g_z$ only if both $f = g$ and $y = z$.

Proof. Suppose $f, g \in \mathrm{Perm}_a(X)$, $y, z \in X$ and $f_y = g_z$. Then, in particular, $y = f_y(a) = g_z(a) = z$, so $y = z$. As f_y is a bijection, it has an inverse f_y^{-1}, which is also the inverse of g_y. Consider

$$x = f_y^{-1}(g_y(x)) = \begin{cases} f_y^{-1}(y) & \text{if } x = a, \\ f_y^{-1}(a) & \text{if } x = g^{-1}(y), \\ f_y^{-1}(g(x)) & \text{otherwise.} \end{cases}$$

In particular, looking at the case $x = g^{-1}(y)$, this tells us $f_y^{-1}(a) = g^{-1}(y)$. By definition of f_y, $f_y^{-1}(a) = f^{-1}(y)$, so $f^{-1}(y) = g^{-1}(y)$. We use this to rewrite $g_y(x)$ as

$$g_y(x) = \begin{cases} y & \text{if } x = a, \\ a & \text{if } x = f^{-1}(y), \\ g(x) & \text{otherwise.} \end{cases}$$

As $g_y = f_y$, it follows that $g(x) = f(x)$ for all $x \notin \{a, f^{-1}(y)\}$. But $f, g \in \mathrm{Perm}_a(X)$, so $f(a) = g(a)$. As f and g are bijections and they are equal on their whole domain except for one point, they must be equal on their whole domain; $f = g$. □

Lemma 7.15. *Suppose X is a finite set, $a \in X$, and $h \in \mathrm{Perm}(X)$. Then there exist $y \in X$ and $f \in \mathrm{Perm}_a(X)$ such that f_y, as defined by equation (7.1), satisfies $f_y = h$.*

Proof. Suppose we have $h \in \mathrm{Perm}(X)$. We define $f \in \mathrm{Perm}_a(X)$ by

$$f(x) = \begin{cases} a & \text{if } x = a, \\ h(a) & \text{if } x = h^{-1}(a), \\ h(x) & \text{otherwise,} \end{cases}$$

And choose $y = h(a)$. Then

$$f_y(x) = \begin{cases} y & \text{if } x = a, \\ a & \text{if } x = f^{-1}(y), \\ f(x) & \text{otherwise} \end{cases} \qquad = \begin{cases} h(a) & \text{if } x = a, \\ a & \text{if } x = f^{-1}(h(a)), \\ f(x) & \text{otherwise.} \end{cases}$$

By the definition of f, $f^{-1}(h(a)) = h^{-1}(a)$, so

$$f_y(x) = \begin{cases} h(a) & \text{if } x = a, \\ a & \text{if } x = h^{-1}(a), \\ f(x) & \text{otherwise.} \end{cases}$$

For $x \notin \{a, h^{-1}(a)\}$, the definition of f sets $f(x) = h(x)$, so

$$f_y(x) = \begin{cases} h(a) & \text{if } x = a, \\ a & \text{if } x = h^{-1}(a), \\ h(x) & \text{otherwise,} \end{cases}$$

so $f_y(x) = h(x)$. □

Theorem 7.16. *Let X be a nonempty finite set, and let $n = \#X$. Then the number of permutations on X is equal to $n!$.*

Proof. We prove the result by induction; since the theorem is only stated for X nonempty, the base case is that of a one-element set. But when $n = \#X = 1$, there is exactly one bijection from X to itself; since $1! = 1$, the theorem is proven in this case.

Now assume that $\#X = n > 1$, and the theorem holds for sets with $n - 1$ elements. Choose an element $a \in X$. Then, by lemma 7.13 and the inductive hypothesis, $\#\,\mathrm{Perm}_a(X) = \#\,\mathrm{Perm}(X \setminus \{a\}) = (n-1)!$.

For $f \in \mathrm{Perm}_a(X)$ and $y \in X$, define $f_y \in \mathrm{Perm}(X)$ by

$$f_y(x) = \begin{cases} y & \text{if } x = a, \\ a & \text{if } x = f^{-1}(y), \\ f(x) & \text{otherwise.} \end{cases}$$

Then $f_y(a) = f_z(a)$ implies $y = z$, so $f_y = f_z$ only if $y = z$. It follows that $\#\{f_y : y \in X\} = n$. It is tempting to claim that therefore $\#\,\mathrm{Perm}(X) = n \times \#\,\mathrm{Perm}_a(X) = n!$, but there are two issues still to be justified:

1. Do we know that, by beginning with different $f, g \in \mathrm{Perm}_a(X)$ and $y, z \in X$, the functions $f_y, g_z \in \mathrm{Perm}(X)$ are different? In other words, do we risk counting functions in $\mathrm{Perm}(X)$ twice? This might make $n!$ an over-estimate for $\#\,\mathrm{Perm}(X)$.

2. Do we know that, given any $h \in \mathrm{Perm}(X)$, we can find $f \in \mathrm{Perm}_a(X)$ and $y \in X$ such that $f_y = h$? If not, then it might be that $n!$ is an under-estimate for $\#\,\mathrm{Perm}(X)$.

Lemma 7.14 establishes that if $f \neq g$ or $y \neq z$, then $f_y \neq g_z$. This means that the map $(f, y) \mapsto f_y$, from $\mathrm{Perm}_a(X) \times X$ to $\mathrm{Perm}(X)$ is a injection, so $\#\,\mathrm{Perm}(X) \geqslant \#X \times \#\,\mathrm{Perm}_a(X) = n!$.

Lemma 7.15 states that the map $(f, y) \mapsto f_y$ is a surjection from $\mathrm{Perm}_a(X) \times X$ to $\mathrm{Perm}(X)$, so $\#\,\mathrm{Perm}(X) \leqslant \#X \times \#\,\mathrm{Perm}_a(X) = n!$. Hence $\#\,\mathrm{Perm}(X) = n!$.

The theorem follows by induction on n. □

7.6 An easier proof for counting permutations and bijections

We present an easier approach to counting permutations and bijections.

To prove that the number of permutations on X is $\#X!$, we might try to argue as follows:

> Suppose $\#X = n$. We count the permutations on X. The first entry of X must be mapped to a point in X, and there are n ways to do this. The second point in X must be mapped to a point in X different from that to which the first point was mapped, and there are $n - 1$ ways to do this. So far, we have $n \times (n - 1)$ choices for the permutation. For the third point, we have $n - 2$ choices. For the fourth point, there are $n - 3$ choices, and so on, until we arrive at a total of $n!$ choices for the permutation.

But the words "and so on" hide an inductive argument. Let us try to formalise this induction:

> Suppose $\#X = 1$. Then there is only the identity permutation on X. So the theorem holds for $n = 1$.
> Now suppose that, for some particular $n \geqslant 2$, for all sets S with cardinality $n - 1$, $\#\,\mathrm{Perm}(S) = (n - 1)!$. Suppose X is a set with cardinality n, and consider $\mathrm{Perm}(X)$. Any permutation on X must map the first entry of X to a point in X, and there are n ways to do this. Specifically, let us denote by a the first entry of X, and by y our choice of point to which a is mapped. Then $\#(X \setminus \{a\}) = n - 1 = \#(X \setminus \{y\})$. As these are both $(n - 1)$-element sets, our inductive hypothesis guarantees that there are $(n-1)!$ bijections between them. Hence there are $n \times (n-1)!$ permutations on X.
> The result holds by induction.

This is an improvement, but there is an error! The inductive hypothesis does not mention bijections between *different sets*, only permutations, which are bijections from a set to itself. So the proof appears to be useless.

This line of thinking leads us to consider a way we could actually use the inductive hypothesis:

Suppose $\#X = 1$. Then there is only the identity permutation on X. So the theorem holds for $n = 1$.
Now suppose that, for some particular $n \geqslant 2$, for all sets S with cardinality $n-1$, $\#\operatorname{Perm}(S) = (n-1)!$. Suppose X is a set with cardinality n, and consider $\operatorname{Perm}(X)$.
In order to construct a set with only $n-1$ elements, let us remove some element a from X. That leaves us with the set $X \setminus \{a\}$, and we know that $\#\operatorname{Perm}(X \setminus \{a\}) = (n-1)!$. When we "add back in" the point a to the domain and codomain of each of these permutations, it must be that a is mapped to a by each of the "extended permutations". This is the idea behind $\operatorname{Perm}_a(X)$, which can be shown to have the same cardinality as $\operatorname{Perm}(X \setminus \{a\})$.
But this only gets us some of the permutations in $\operatorname{Perm}(X)$; what about the ones that do not leave a constant? Perhaps we can produce them by "swapping" the a in the codomain with another point in the codomain. This is the motivation for introducing the construction f_y. We then have to show that the map "produce f_y from a given f and y" is a bijection, which is the point of the other two lemmata.
Drawing all these ideas together, we are able to complete the inductive step, and apply induction to conclude the theorem.

This proof was successful, but it required several technical lemmata in order to make it work. It would be nice if we could find a more straightforward proof.

Looking back at the the error in our first attempt at an inductive argument, provides a hint at an alternative approach. We needed to use something stronger than our inductive assumption. Instead of "$\#X = (n-1) \Rightarrow \#\operatorname{Perm}(X) = (n-1)!$", we needed to use "$\#X = (n-1) = \#Y \Rightarrow$ there are $(n-1)!$ bijections from X to Y". So why not just make that our inductive assumption? That means we need a slightly different base, but it is not much more complicated. We end up proving the more general theorem that allows us to count bijections, instead of just permutations. But that is not a problem! Permutations are just bijections with codomain equal to domain, so the result for permutations is an immediate corrolary.

Theorem 7.17. *Suppose X and Y are finite sets, with the same cardinality. Then there are $\#X!$ bijections from X to Y.*

Proof. If $\#X = 1 = \#Y$, then there is only one function from X to Y, and it is a bijection.

Now suppose that, for some $n \geqslant 2$, for any sets X, Y with $\#X = n - 1 = \#Y$, there are exactly $(n-1)!$ bijections from X to Y. Suppose A, B are sets with $\#A = n = \#B$, and suppose $a \in A$. Then, for $f : A \to B$ to be a bijection, there are n possibilities for $b = f(a) \in B$. Given such a choice, we define $X = A \setminus \{a\}$ and $Y = B \setminus \{b\}$. But $\#X = n - 1 = \#Y$, so there are $n-1$ bijections from X to Y. Therefore, there are $n \times (n-1)! = n!$ bijections from A to B.

The theorem is proved, by induction. □

Corollary 7.18. *Suppose X is a finite set. Then $\#\operatorname{Perm}(X) = \#X!$.*

Proof. Apply theorem 7.17 with $Y = X$. □

Sometimes, it can be easier to prove the more general theorem first! This is particularly common for induction.

8 Week 8

8.1 There exist bijections between $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{N}^2$

We construct some bijections between sets, in order to show that they, perhaps surprisingly, have the same cardinality.

Theorem 8.1. *Define a function $f : \mathbb{N} \to \mathbb{Z}$ by the following rule.*

$$f(n) = \begin{cases} n/2 & \text{if } n \text{ is even,} \\ \frac{1-n}{2} & \text{if } n \text{ is odd.} \end{cases}$$

Then f is a bijection.

Proof. We begin by proving that f is surjective. If $z \in \mathbb{Z}$, then either $z \leqslant 0$ or $z > 0$. If $z > 0$, then $2z$ is an even natural number, and $z = f(2z)$. Hence z is contained in the image of f. On the other hand, if $z \leqslant 0$, then $1 - 2z$ is a positive odd number, and $f(1 - 2z) = z$. Hence z is contained in the image of f in this case too. Therefore f is surjective.

To see that f is injective, suppose that $x, y \in \mathbb{N}$ and $f(x) = f(y)$. Observe that f sends even natural numbers to positive numbers, and odd natural numbers to non-positive numbers. Hence $f(x) = f(y)$ implies that x and y have the same parity. [Note to presenter/reader: "parity" means "evenness/oddness," for example, the parity of 2 is "even"].

If x and y are both even, then $f(x) = f(y)$ implies that $x/2 = y/2$, so $x = y$. If x and y are both odd, then $f(x) = f(y)$ implies that $(1 - x)/2 = (1 - y)/2$, and so $x + 1 = y + 1$, so $x = y$. Hence, in both cases, $f(x) = f(y)$ implies $x = y$, and therefore f is injective.

[Note to reader/presenter: it is recommended to draw a picture with an arrow-diagram of the function. This makes the proof clearer.] □

Theorem 8.2. *Define a function $f : \mathbb{N} \to \mathbb{N}^2$ by "snaking":*

$$f(1) = (1, 1), \quad f(2) = (2, 1), \quad f(3) = (2, 2), \quad f(4) = (1, 2), \quad \dots$$

as in figure 12. Then f is a bijection.

Proof. Look at the picture.

[Note to reader/presenter: The diagram would suffice as a proof for most mathematicians. A computer scientist might find it interesting to write a formal algorithm to implement the snaking, and carry out a machine-checked proof that every ordered pair of natural numbers is obtained exactly once by the algorithm.] □

Conclusion: Recall that if there is a bijection from a set S to the set $\mathbb{N}$, we say that S is countable, and we write $\#S = \aleph_0$. We have now proven that $\#\mathbb{Z} = \aleph_0$ and $\#\mathbb{N}^2 = \aleph_0$.

8.2 There exists an injective/surjective function from $\mathbb{N}$ to $\mathbb{Q}$.

We aim to show that $\mathbb{N}$ and $\mathbb{Q}$ have the same cardinality. It is not so easy to construct a bijection from $\mathbb{N}$ to $\mathbb{Q}$, so instead we construct two functions: one injective and one surjective. After we have studied it in week 10, the Cantor-Schröder-Bernstein theorem will give us the result we want.

Proposition 8.3. *There is an injection from $\mathbb{N}$ to $\mathbb{Q}$.*

Proof. Let $f : \mathbb{N} \to \mathbb{Q}$ be given by $f(n) = n$. If we were to restrict the codomain of f to its image, we would get the function $\mathrm{Id}_{\mathbb{N}}$, which is certainly an injection. Hence f is an injection. □

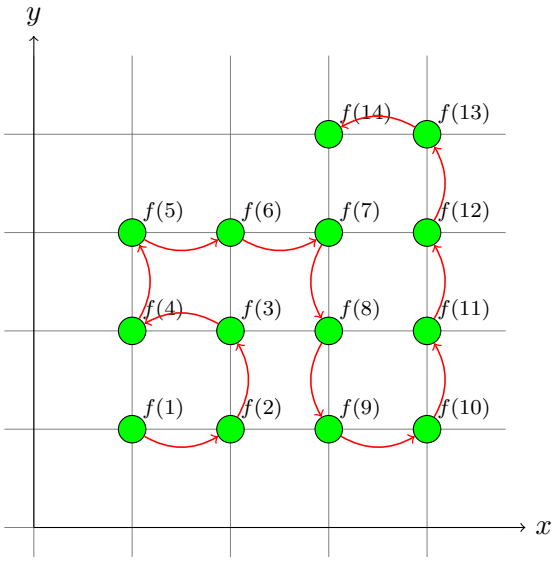


Figure 12: "Snaking" outwards, touching every ordered pair of natural numbers exactly once.

Proposition 8.4. *There is a surjection from* $\mathbb{N}$ *to* $\mathbb{Q}$.

Proof. Look at figure 13 for a bijection $f : \mathbb{N} \to \mathbb{Z}^2$. We defined $\mathbb{Q}$ as the set of answers to the question "by what should I multiply x to get y?" when x and y are both integers, and x is nonzero. Let $g : \mathbb{Z}^2 \to \mathbb{Q}$ be defined by

$$g(x,y) = \begin{cases} \text{the answer to the question "by what should I multiply } x \text{ to get } y\text{?"} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

Then g is a surjection, so $g \circ f$ is a surjection.

[Reader / Presenter: It is clear that $g(x,y) = 0$ for many different (x,y), so g is not a bijection. But even if we restricted the domain of g to get rid of all but one of the points that map to 0, g would still not be a bijection, as $g(2,1) = g(4,2)$. Can you see a geometric interpretation (on figure 13) for the (x,y) that have to be removed from the domain to make g a bijection?] □

8.3 The set of r-tuples of natural numbers is countable.

Using a "snaking" bijection, we showed that the set of ordered pairs of natural numbers is countable. One might manage to produce such a snaking bijection in 3 dimensions for ordered triplets, but it gets a little more difficult to think in 4 or more dimensions. We present an alternative, inductive approach.

In a previous proof (there exist bijections between $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{N}^2$), we proved that

$$\#\mathbb{Z} = \aleph_0 \text{ and } \#\mathbb{N}^2 = \aleph_0.$$

Let r be a natural number, and define $\mathbb{N}^r$ to be the set of r-tuples of natural numbers. For example,

$$(1,3,7) \in \mathbb{N}^3, \text{ and } (2,3,7,4) \in \mathbb{N}^4.$$

There are natural "rewriting" bijections among various sets of tuples. For example, there is a bijection from $\mathbb{N}^2 \times \mathbb{N}^3$ to $\mathbb{N}^5$, sending an ordered pair $((a_1,a_2),(a_3,a_4,a_5))$ to the 5-tuple $(a_1,a_2,\ldots,a_5)$. For any positive integers r,s, define

$$\alpha_{r,s} : \mathbb{N}^r \times \mathbb{N}^s \to \mathbb{N}^{r+s}$$

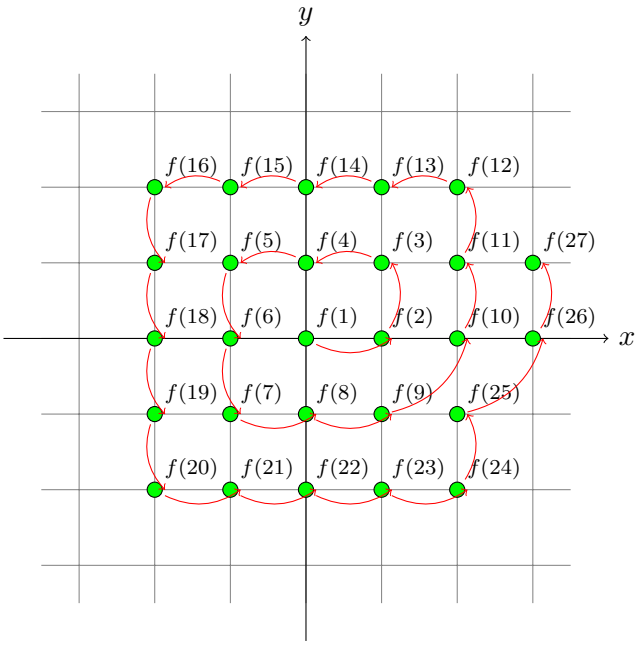


Figure 13: Spiraling outwards, touching every ordered pair of integers exactly once.

to be this rewriting bijection.

For what follows, we require the Cartesian product not only of sets but also of functions. Suppose that A, B, C, D are four sets. Suppose that $f : A \to C$ is a function and $g : B \to D$ is a function. Then we can form a function

$$[f \times g] : (A \times B) \to (C \times D)$$

by the rule

$$[f \times g](a, b) = (f(a), g(b)) .$$

If f and g are injective, then $[f \times g]$ is injective. If f and g are surjective, then $[f \times g]$ is surjective. Hence, if f and g are bijective, then $[f \times g]$ is bijective.

With this background in place, we can prove an important result.

Theorem 8.5. *If r is a natural number, then the set $\mathbb{N}^r$ is countable, so there exists a bijection from $\mathbb{N}$ to $\mathbb{N}^r$.*

Proof. When $r = 1$, the identity function is a bijection from $\mathbb{N}$ to $\mathbb{N}^1 = \mathbb{N}$. When $r = 2$, we have found a bijection $f : \mathbb{N} \to \mathbb{N}^2$ in a previous proof (one may take the "snaking" bijection).

Now suppose that $r > 2$ and assume that there exists a bijection $g : \mathbb{N} \to \mathbb{N}^{r-1}$. Then we can compose four bijections as in the diagram below.

$$\mathbb{N} \to \mathbb{N}^{r-1} \to \mathbb{N}^{r-2} \times \mathbb{N} \to \mathbb{N}^{r-2} \times \mathbb{N}^2 \to \mathbb{N}^r.$$

1. The first bijection, from $\mathbb{N} \to \mathbb{N}^{r-1}$, is g.
2. The second bijection, from $\mathbb{N}^{r-1} \to \mathbb{N}^{r-2} \times \mathbb{N}$, is $\alpha_{r-2,1}^{-1}$.
3. The third bijection, from $\mathbb{N}^{r-2} \times \mathbb{N} \to \mathbb{N}^{r-2} \times \mathbb{N}^2$, is $\mathrm{Id}_{\mathbb{N}^{r-2}} \times f$.
4. The fourth bijection, from $\mathbb{N}^{r-2} \times \mathbb{N}^2 \to \mathbb{N}^r$, is $\alpha_{r-2,2}$.

The composition of these four bijections is a bijection from $\mathbb{N}$ to $\mathbb{N}^r$.

Hence, by induction, $\mathbb{N}^r$ is countable for all $r \in \mathbb{N}$. □

Corollary 8.6. *If r is a natural number, then the set $\mathbb{Z}^r$ is countable.*

Proof. By theorem 8.5, there exists a bijection $f : \mathbb{N} \to \mathbb{N}^r$. By an earlier result, there exists a bijection $g : \mathbb{N} \to \mathbb{Z}$. Therefore, the function $h : \mathbb{Z}^r \to \mathbb{N}$, defined by

$$h = f^{-1} \circ [g^{-1} \times \cdots \times g^{-1}]$$

formulaically, $h(x_1, \ldots, x_r) = f^{-1}(g^{-1}(x_1), \ldots, g^{-1}(x_r))$, is a bijection. □

8.4 Cardinality of countable unions

We study the cardinality of unions of countable sets.

Proposition 8.7. *Suppose J is a countable index set and, for all $j \in J$, A_j is countable. Then*

$$A := \bigcup_{j \in J} A_j$$

is a countable set.

Proof. By countability of each A_j, there exist surjections $f_j : \mathbb{N} \to A_j$. By countability of J, there exists a surjection $g : \mathbb{N} \to J$.

Consider the map $f : \mathbb{N}^2 \to A$ defined by

$$f(x, y) = f_{g(x)}(y).$$

Suppose $a \in A$. Then, for some $j \in J$, $a \in A_j$. Because g is a surjection, there exists $x \in \mathbb{N}$ for which $g(x) = j$ and, because f_j is a surjection, there exists $y \in \mathbb{N}$ such that $f_j(y) = a$. We have justified existence of $(x, y) \in \mathbb{N}^2$ for which $f(x, y) = a$. So f is a surjection.

By countability of $\mathbb{N}^2$, there exists a surjection $h : \mathbb{N} \to \mathbb{N}^2$. But then $f \circ h : \mathbb{N} \to A$ is a surjection. We have shown that A is countable. □

8.5 A set and its power set have different cardinality

We give a famous proof from Georg Cantor's 1891 paper, "Über eine elementare Frage der Mannigfaltigkeitslehre". The original can be easily found online, thanks to the Göttingen archive. The proof is adapted from Cantor's.

Theorem 8.8 (Cantor)**.** *If S is a set, then there does not exist a surjective function from S to its power set $\mathcal{P}(S)$.*

Proof. Suppose that $f : S \to \mathcal{P}(S)$ is a function. Thus, for every *element* $x \in S$, there is a *subset* $f(x) \subset S$. Therefore, f partitions the elements of S into two types of elements:

$$x \text{ is } f\text{-happy if } x \in f(x); \tag{8.1}$$

$$x \text{ is } f\text{-sad if } x \notin f(x). \tag{8.2}$$

Define a new subset $T \subset S$ by the following rule:

$$T = \{x \in S : x \text{ is } f\text{-sad}\}. \tag{8.3}$$

In other words, T is the set of f-sad elements of S.

We claim that T is not an element of the image of f. Indeed, suppose to the contrary that $y \in S$ and $f(y) = T$. It must be that either $y \in T$, or $y \notin T$, but we will show that each leads to a contradiction. If $y \notin T$ then y is f-happy (by (8.3)), and therefore $y \in f(y) = T$ (by (8.1)), and so $y \in T$; so $y \notin T$ is impossible. But if $y \in T$, then y is f-sad (by (8.3)), and therefore $y \notin f(y) = T$ (by (8.2)), and so $y \notin T$; so $y \in T$ is also impossible. Thus we have a contradiction.

Hence there does not exist an element $y \in S$ such that $f(y) = T$. Hence the function $f : S \to \mathcal{P}(S)$ is not surjective. □

8.6 Subsets, power sets and cardinality

We study how cardinality interacts with subsets and power sets.

Recall that we say that two sets are equinumerous if they have the same cardinality. The following proposition says that the operation "take the power set" preserves equinumerosity.

Proposition 8.9. *Suppose* $\#X = \#Y$. *Then* $\#\mathcal{P}(X) = \#\mathcal{P}(Y)$. *Moreover, if* $\#X \leqslant \#Y$, *then* $\#\mathcal{P}(X) \leqslant \#\mathcal{P}(Y)$.

Proof. Suppose there is an injection $f : X \to Y$. We define $g : \mathcal{P}(X) \to \mathcal{P}(Y)$ by

$$g(S) = \{f(x) : x \in S\}.$$

[Note that we could just write $g(S) = f(S)$, with $g(S)$ representing the function g evaluated at input S, and $f(S)$ representing the image of the set S in the function f. This would make for a slick-looking proof, but the notation might be confusing on first read, so we will avoid it!] Suppose that $g(S) = g(T)$. Then for each $x \in S$ there is some $x' \in T$ such that $f(x) = f(x')$. But f is an injection (i.e. $x' = x$), so for each $x \in S$, it must be that $x \in T$. Similarly, for each $x' \in T$, injectivity of f implies $x' \in S$ also. Therefore $S = T$. We have established the latter claim.

Now suppose that f is also surjective, and consider some $V \subset Y$. If $y \in V$, then there is some $x \in X$ such that $f(x) = y$, so there is some $U \subset X$ for which $g(U) = V$. Hence g is a bijection. □

Proposition 8.10. *If* $X \subset Y$, *then* $\#X \leqslant \#Y$.

Proof. Let $f : X \to Y$ be the function given by $f(x) = \mathrm{Id}_X(x)$. Then f is an injection because Id_X is an injection. □

Corollary 8.11. $\#(X \cap Y) \leqslant \#X$.

Proof. Let $Z = X \cap Y$. Then $Z \subset X$. The result follows by proposition 8.10. □

9 Week 9

9.1 The set of real numbers is uncountable

We show that there are uncountably many real numbers. That's far more than there are rational numbers.

Lemma 9.1. *Let $T = \mathcal{P}(\mathbb{N}) \setminus \{x \in \mathcal{P}(\mathbb{N}) : x \text{ is finite}\}$. Then $\#T > \aleph_0$.*

Proof. Let

$$S = \bigcup_{r \in \mathbb{N}} \mathcal{P}_r(\mathbb{N}),$$

so that $T = \mathcal{P}(\mathbb{N}) \setminus S$. We proved earlier that, for each $r \in \mathbb{N}$, $\#\mathbb{N}^r = \aleph_0$. Let $\mathbb{N}^r_{\neq}$ be the set of r-tuples of naturals in which no entry is repeated, so $\mathbb{N}^r_{\neq} \subset \mathbb{N}^r$. The function $g_r : \mathbb{N}^r_{\neq} \to \mathcal{P}_r(\mathbb{N})$ defined by

$$g_r((x_1, x_2, \ldots, x_r)) = \{x_1, x_2, \ldots, x_r\}$$

is a surjection because, for any set $Y = \{y_1, y_2, \ldots, y_r\} \in \mathcal{P}_r(\mathbb{N})$, the tuple $y = (y_1, y_2, \ldots, y_r) \in \mathbb{N}^r_{\neq}$ and $g_r(y) = Y$. But then $\#\mathcal{P}_r(\mathbb{N}) \leqslant \#\mathbb{N}^r_{\neq} \leqslant \#\mathbb{N}^r = \aleph_0$. We proved earlier that a countable union of countable sets is countable, so $\#S \leqslant \aleph_0$.

Suppose $\#T \leqslant \aleph_0$. Then $\mathcal{P}(\mathbb{N}) = T \cup S$ is a union of countable sets, so $\#\mathcal{P}(\mathbb{N}) = \#(T \cup S) \leqslant \aleph_0$. But this contradicts our theorem about the cardinality of the power set of naturals, so $\#T > \aleph_0$. □

Theorem 9.2. *$\#\mathbb{R} > \aleph_0$.*

Proof. Let T be as defined in lemma 9.1, and define $g : T \to \mathbb{R}$ by

$$g(x) = \sum_{n \in x} \frac{1}{2^n}.$$

We will show that g is an injection.

Suppose that $x, y \in T$, $g(x) = g(y)$, and $x \neq y$. Because x and y are different, there must be some $m \in \mathbb{N}$ for which m is an element of one of x and y, but not the other. Pick the least such m and assume, without loss of generality, that $x \ni m \notin y$. Then

$$\begin{aligned}
0 &= g(x) - g(y) \\
&= \sum_{n \in x} \frac{1}{2^n} - \sum_{n \in y} \frac{1}{2^n} \\
&= \left(\sum_{\substack{n \in x: \\ n < m}} \frac{1}{2^n} + \frac{1}{2^m} + \sum_{\substack{n \in x: \\ n > m}} \frac{1}{2^n} \right) - \left(\sum_{\substack{n \in y: \\ n < m}} \frac{1}{2^n} + \sum_{\substack{n \in y: \\ n > m}} \frac{1}{2^n} \right) \\
&= \frac{1}{2^m} + \sum_{\substack{n \in x: \\ n > m}} \frac{1}{2^n} - \sum_{\substack{n \in y: \\ n > m}} \frac{1}{2^n} \\
&\geqslant \frac{1}{2^m} + \sum_{\substack{n \in x: \\ n > m}} \frac{1}{2^n} - \sum_{n=m+1}^{\infty} \frac{1}{2^n} \\
&= \sum_{\substack{n \in x: \\ n > m}} \frac{1}{2^n},
\end{aligned}$$

by the geometric series lemma. But $x \in T$, so there exists some $n > m$ for which $n \in x$. (Actually, there must be infinitely many such n, but we only need one for this argument.) Therefore, for some $n \in \mathbb{N}$, $0 \geqslant 1/2^n$, which is false. Hence no such x and y exist, and g is an injection. □

9.2 $\mathbb{C}$ is a field

In the reading, the operations of "addition" and "multiplication" on $\mathbb{C}$ were introduced, and we claimed it was natural to give these new operations familiar names. Here we show that in fact complex addition and multiplication do behave in the way we expect.

Definition 9.3. A *field* is a triple $(F, +, \cdot)$, where F is a set, and $+$ and $\cdot$ are *binary operators* on F (i.e. they each take two elements of F as input, and each give a single output) in such a way that the following conditions (called the *field axioms*) are obeyed:

1. **Closure under addition**. If $x, y \in F$, then $x + y \in F$.
2. **Commutativity of addition**. If $x, y \in F$, then $x + y = y + x$.
3. **Associativity of addition**. If $x, y, z \in F$, then $x + (y + z) = (x + y) + z$.
4. **Existence of additive identity**. There is a number $0_F \in F$ such that for all $x \in F$, $x + 0_F = x$.
5. **Existence of additive inverses**. For each $x \in F$, there is some $y \in F$ such that $x + y = 0_F$.
6. **Closure under multiplication**. If $x, y \in F$, then $x \cdot y \in F$.
7. **Commutativity of multiplication**. If $x, y \in F$, then $x \cdot y = y \cdot x$.
8. **Associativity of multiplication**. If $x, y, z \in F$, then $x \cdot (y \cdot z) = (x \cdot y) \cdot z$.
9. **Existence of multiplicative identity**. There is number $1_F \in F$ such that for all $x \in F$, $x \cdot 1_F = x$.
10. **Existence of multiplicative inverses**. For each $x \in F \setminus \{0_F\}$, there is some $y \in F$ such that $x \cdot y = 1_F$.
11. **Distributive law**. If $x, y, z \in F$, then $x \cdot (y + z) = (x \cdot y) + (x \cdot z)$.
12. **Zero-one law**. $0_F \neq 1_F$.

In the reading, the symbol $+$ was used to represent 3 different concepts, with very little care:

1. The notation "$x + yi$" is introduced as another way of writing (x, y). This is pure symbolism, with no relation to real addition.
2. The expression
$$(x_1 + y_1 i) + (x_2 + y_2 i) = (x_1 + x_2) + (y_1 + y_2)i \tag{9.1}$$
contains three $+$ signs. Those inside parentheses on the left are used as above, but the $+$ linking the two parenthetical terms on the left is used to represent the new "addition" of complex numbers.
3. The $+$ inside each parentheses on the right of equation (9.1) is used to represent addition of real numbers in the way with which we are familiar. The $+$ linking the two parenthetical terms on the right is another case of symbolism for alternative representation of a complex number.

To minimize such confusion, for the following proof, we will avoid usage 1, writing all complex numbers as ordered pairs $x = (a, b)$. We will distinguish between the other uses by using the usual addition symbol $+$ to represent addition of real numbers, and using the symbol $\oplus$ to represent addition of complex numbers.

For multiplication, we will write ab for the product of real numbers a, b, and $x \cdot y$ for multiplication of complex numbers x, y.

In proving theorem 9.4, we will make full use of the fact that $\mathbb{R}$ is a field.

Theorem 9.4. *Define* $(\mathbb{C}, \oplus, \cdot)$ *as:*

- $\mathbb{C} = \{(a, b) \in \mathbb{R} \times \mathbb{R}\}$
- *If* $(a, b), (c, d) \in \mathbb{C}$*, then* $(a, b) \oplus (c, d) = (a + c, b + d)$*.*
- *If* $(a, b), (c, d) \in \mathbb{C}$*, then* $(a, b) \cdot (c, d) = (ac - bd, ad + bc)$*.*

Then $(\mathbb{C}, \oplus, \cdot)$ *is a field.*

Proof. 1. **Closure under addition**. If $(a, b), (c, d) \in \mathbb{C}$, then $a, b, c, d \in \mathbb{R}$ and $\mathbb{R}$ is closed under addition. Hence $a + c, b + d \in \mathbb{R}$, so $(a, b) \oplus (c, d) = (a + c, b + d) \in \mathbb{C}$.

2. **Commutativity of addition**. If $(a, b), (c, d) \in \mathbb{C}$, then $a, b, c, d \in \mathbb{R}$, and addition in $\mathbb{R}$ is commutative. Hence $a + c = c + a$ and $b + d = d + b$, so $(a, b) \oplus (c, d) = (a + c, b + d) = (c + a, d + b) = (c, d) \oplus (a, b)$.

3. **Associativity of addition**. If $(a, b), (c, d), (e, f) \in \mathbb{C}$, then $a, b, c, d, e, f \in \mathbb{R}$, and addition in $\mathbb{R}$ is associative. Hence

$$\begin{aligned}(a, b) \oplus ((c, d) \oplus (e, f)) &= (a, b) \oplus (c + e, d + f) \\ &= (a + (c + e), b + (d + f)) \\ &= ((a + c) + e, (b + d) + f) \\ &= (a + c, b + d) \oplus (e, f) \\ &= ((a, b) \oplus (c, d)) \oplus (e, f).\end{aligned}$$

4. **Existence of additive identity**. If $(a, b) \in \mathbb{C}$, then $a, b \in \mathbb{R}$, and 0 is an additive identity in $\mathbb{R}$. Hence $(a, b) \oplus (0, 0) = (a + 0, b + 0) = (a, b)$. Hence $(0, 0)$ is an additive identity in $\mathbb{C}$.

5. **Existence of additive inverses**. If $(a, b) \in \mathbb{C}$, then $a, b \in \mathbb{R}$, so they have real additive inverses $-a, -b$, respectively. Using this, we construct $(-a, -b) \in \mathbb{C}$ for which $(a, b) \oplus (-a, -b) = (a - a, b - b) = (0, 0)$ so $(-a, -b)$ is the complex additive inversa of (a, b).

6. **Closure under multiplication**. If $(a, b), (c, d) \in \mathbb{C}$, then $a, b, c, d \in \mathbb{R}$ and $\mathbb{R}$ is closed under multiplication. Hence $ac, bd, ad, bc \in \mathbb{R}$. Also, $\mathbb{R}$ is closed under addition and subtraction, so $(a, b) \cdot (c, d) \in \mathbb{C}$.

7. **Commutativity of multiplication**. If $(a, b), (c, d) \in \mathbb{C}$, then $a, b, c, d \in \mathbb{R}$, and addition and multiplication in $\mathbb{R}$ are commutative. Hence

$$\begin{aligned}(a, b) \cdot (c, d) &= (ac - bd, ad + bc) \\ &= (ca - db, da + cb) \\ &= (c, d) \cdot (a, b).\end{aligned}$$

8. **Associativity of multiplication**. If $(a, b), (c, d), (e, f) \in \mathbb{C}$, then $a, b, c, d, e, f \in \mathbb{R}$, and multiplication in $\mathbb{R}$ is associative and distributive over addition (and subtraction), and addition in $\mathbb{R}$ is associative. Hence

$$\begin{aligned}(a, b) \cdot ((c, d) \cdot (e, f)) &= (a, b) \cdot (ce - df, cf + de) \\ &= (a(ce - df) - b(cf + de), a(cf + de) + b(ce - df)) \\ &= ((ac - bd)e - (ad + bc)f, (ac - bd)f + (ad + bc)e) \\ &= (ac - bd, ad + bc) \cdot (e, f) \\ &= ((a, b) \cdot (c, d)) \cdot (e, f).\end{aligned}$$

9. **Existence of multiplicative identity**. If $(a, b) \in \mathbb{C}$, then $a, b \in \mathbb{R}$, 1 is a multiplicative identity in $\mathbb{R}$, and 0 is an additive identity in $\mathbb{R}$. Hence $(a, b) \cdot (1, 0) = (a - 0, b + 0) = (a, b)$. Hence $(1, 0)$ is a multiplicative identity in $\mathbb{C}$.

10. **Existence of multiplicative inverses**. If $(a, b) \in \mathbb{C}$, then $a, b \in \mathbb{R}$, so $a^2 + b^2 \in \mathbb{R}$. Provided $(a, b) \neq (0, 0)$, $a^2 + b^2 \neq 0$. Hence $a^2 + b^2$ has a real multiplicative inverse, $1/(a^2 + b^2)$. Using closure of $\mathbb{R}$ under multiplication and existence of real additive inverses, we see that $a/(a^2+b^2) \in \mathbb{R}$ and $-b/(a^2 + b^2) \in \mathbb{R}$. We use these facts to construct $(a/(a^2 + b^2), -b/(a^2 + b^2)) \in \mathbb{C}$ which has the property

$$(a, b) \cdot (a/(a^2 + b^2), -b/(a^2 + b^2)) = ((a^2 + b^2)/(a^2 + b^2), (-ab + ba)/(a^2 + b^2)) = (1, 0),$$

hence it is a multiplicative inverse of (a, b).

11. **Distributive law**. If $(a, b), (c, d), (e, f) \in \mathbb{C}$, then $a, b, c, d, e, f \in \mathbb{R}$, multiplication distributes over addition in $\mathbb{R}$, and addition is commutative and associative. Hence

$$\begin{aligned}
(a, b) \cdot ((c, d) \oplus (e, f)) &= (a, b) \cdot (c + e, d + f) \\
&= (a(c + e) - b(d + f), a(d + f) + b(c + e)) \\
&= ((ac - bd) + (ae - bf), (ad + bc) + (af + be)) \\
&= (ac - bd, ad + bc) \oplus (ae - bf, af + be) \\
&= ((a, b) \cdot (c, d)) \oplus ((a, b) \cdot (e, f)).
\end{aligned}$$

12. **Zero-one law**. As $1 \neq 0$ in $\mathbb{R}$, so $(1, 0) \neq (0, 0)$ in $\mathbb{C}$. □

9.3 Equivalent fractions

We define an equivalence on the set of fractions, which encodes the idea of two different fractions representing the same rational number.

Here we take for granted the set of integers, and all familiar properties of addition, subtraction, and multiplication of integers. We also take for granted the "integral domain" property of integers, which states that $xy = 0$ implies $x = 0$ or $y = 0$, for all integers x, y. Equivalently, if $x \neq 0$ and $y \neq 0$ then $xy \neq 0$. This implies the cancellation property: if $x, y, z \in \mathbb{Z}$, and $z \neq 0$, and $xz = yz$, then $x = y$.

A *fraction* will mean an ordered pair $(x, y) \in \mathbb{Z}^2$ such that $y \neq 0$. Let F be the set of fractions. Consider the relation $\sim$ on F given by the rule

$$(x, y) \sim (u, v) \text{ means that } xv = yu.$$

The point of $\sim$ is that if two fractions obey $(x, y) \sim (u, v)$, then they are two expressions of the same rational number. For example, $(2, 4) \sim (12, 24)$, and they both correspond to the rational number that we are used to representing $\frac{1}{2}$.

Theorem 9.5. *The relation $\sim$ is an equivalence relation on the set F.*

Proof. We check that $\sim$ is reflexive, symmetric, and transitive.

Reflexive If $(x, y) \in F$, then we have $xy = yx$ by the commutative property of multiplication. Hence $(x, y) \sim (x, y)$.

Symmetric Suppose that $(x, y) \in F$ and $(u, v) \in F$. If $(x, y) \sim (u, v)$, then $xv = yu$. This implies that $uy = vx$ by the commutative property of multiplication (and symmetry of $=$). Hence $(u, v) \sim (x, y)$.

Transitive Suppose that (x, y) and (u, v) and (r, s) are elements of F. If $(x, y) \sim (u, v)$ and $(u, v) \sim (r, s)$, then we find that $xv = yu$ and $us = vr$. Note that $y \neq 0$ and $v \neq 0$ and $s \neq 0$. We now break the proof into two cases.

If $r = 0$, then the equality $us = vr$ and the nonvanishing of s implies that $u = 0$. From this, the equality $xv = yu$ and the nonvanishing of v implies that $x = 0$. Since $x = 0$ and $r = 0$, we quickly find that $xs = yr$ and so $(x, y) \sim (r, s)$ as required.

If $r \neq 0$, then we multiply the equality $xv = yu$ by rs on both sides to obtain $(xv)(rs) = (yu)(rs)$. The commutative and associative properties of multiplication yield $(xs)(vr) = (yr)(us)$. Since $us = vr$, we find that

$$(xs)(vr) = (yr)(vr)$$

Since $v \neq 0$ and $r \neq 0$, it follows that $vr \neq 0$. Hence, by the cancellation property, $xs = yr$. Therefore $(x, y) \sim (r, s)$ as required.

[Note to reader/presenter: can you find a violation of transitivity if we allowed fractions of the form $(x, 0)$?] □

9.4 Descent of functions

Given a function, whose output is determined only by the equivalence class of its input, we construct a corresponding function which acts upon the equivalence classes themselves.

The "set of equivalence classes" is characterized by what mathematicians call a universal property. This is the most important property of the construction of equivalence classes, and perhaps the most challenging to understand at the beginning.

Consider the following example. Define an equivalence relation on $\mathbb{N}^0$ by the rule "$x \sim y$ means that x and y have the same *parity*", i.e., both x and y are even or both x and y are odd.

There are two equivalence classes for this relation, which we will call E and O:

$$E = \{0, 2, 4, \ldots\}, \qquad O = \{1, 3, 5, \ldots\}.$$

Thus $\mathbb{N}^0/\sim = \{E, O\}$ is a set with two elements.

Whenever S is a set, and $\sim$ is an equivalence relation, there is a natural surjective function $p : S \to S/\sim$ called the *canonical projection*. It is defined by the rule:

$$p(s) = \text{ the equivalence class to which } s \text{ belongs.}$$

We have "projected" each element of S onto the equivalence class to which it belongs. For example, the canonical projection $p : \mathbb{N}^0 \to \mathbb{N}^0/\sim$ satisfies:

$$p(4) = E, \qquad p(11) = O, \qquad p(0) = E, \qquad \ldots.$$

Projection functions are so-called because of a physical realisation. Imagine a lamp is illuminating a surface. If a point was placed between the lamp and surface, then it would cast a shadow at a point on the surface. Moving the point along the line it originally made with the lamp would not move the shadow. This sets up an equivalence relation, $\equiv$, defined by "casts the same shadow on the surface as" on the 3-d space $L \subset \mathbb{R}^3$ between the lamp and the surface. The canonical projection $p : L \to (L/\equiv)$ returns the equivalence class to which any point belongs. That equivalence class can be thought of as being represented by the shadow. So p **projects** L onto a set represented by the shadow points.

The following theorem studies a function whose output is determined only by the equivalence class to which its input belongs. The claim is that we can construct a corresponding function which has domain "the equivalence classes themselves". Before working on the proof of the theorem, it may help to study the application below.

Theorem 9.6. *Suppose that S is a set and $\sim$ is an equivalence relation on S. Suppose that T is a set and $f : S \to T$ is a function. Finally, suppose that, for all $x, y \in S$,*

$$x \sim y \text{ implies that } f(x) = f(y).$$

Then, there exists a unique function $\bar{f} : (S/\sim) \to T$ such that

$$\bar{f} \circ p = f.$$

Proof. Consider an equivalence class $M \in S/\sim$. Define $\bar{f}(M)$ by the following process:

1. Choose an element $s \in M$.
2. Let $\bar{f}(M) = f(s)$.

We claim that the result $\bar{f}(M)$ does not depend on the choice of element $s \in M$. Indeed, if we choose another element $s' \in M$, then $s \sim s'$. It follows that $f(s) = f(s')$, and so the resulting value of $\bar{f}(M)$ does not change whether one chooses s or s'. Hence, we find a *well-defined* function $\bar{f} : (S/\sim) \to T$, regardless of the choices made in the process. Essentially by definition, we have

$$\bar{f} \circ p(s) = f(s) \text{ for all } s \in S.$$

Thus $\bar{f} \circ p = f$ as required.

To prove uniqueness, suppose that $\bar{g} : (S/\sim) \to T$ is another function which satisfies $\bar{g} \circ p = f$. Suppose that $M \in S/\sim$, and pick any $s \in M$. Then we find that

$$f(s) = \bar{g}(p(s)) = \bar{g}(M), \qquad \text{and} \qquad f(s) = \bar{f}(p(s)) = \bar{f}(M).$$

Hence $\bar{f}(M) = \bar{g}(M)$ for all $M \in S/\sim$. Therefore $\bar{f} = \bar{g}$, demonstrating uniqueness. □

When $f : S \to T$ is a function, and $x \sim y$ implies that $f(x) = f(y)$, we say that f *descends to a well-defined* function $\bar{f} : (S/\sim) \to T$. Here is an example, using the set $\mathbb{N}^0$ and the parity equivalence relation $\sim$ from before.

Define a function $\alpha : \mathbb{N}^0 \to \{2, 4\}$ by the rule:

$$\alpha(x) = 3 + (-1)^x.$$

For example, $\alpha(2) = 4$, and $\alpha(5) = 2$.

Then, if $x \sim y$, we find that $\alpha(x) = \alpha(y)$. In other words, if x and y have the same parity, then $\alpha(x)$ and $\alpha(y)$ either both evaluate to 2 or both evaluate to 4. Therefore, the function α *descends* to a well-defined function

$$\bar{\alpha} : \mathbb{N}^0/\sim \to \{2, 4\}.$$

Recall that $\mathbb{N}^0/\sim = \{E, O\}$. The function $\bar{\alpha}$ is given by:

$$\bar{\alpha}(E) = 4, \quad \bar{\alpha}(O) = 2.$$

10 Week 10

10.1 Descent of binary operators

We prove a slightly more general version of the descent of functions theorem. It turns out that, for our purposes, this is a more useful version.

Theorem 10.1. *Suppose S is a set with equivalence $\sim$, $p : S \to S/\sim$ is the canonical projection function, and $f : S \times S \to S$ obeys*

$$x \sim x' \text{ and } y \sim y' \qquad \Rightarrow \qquad f(x,y) \sim f(x',y').$$

Then there exists a unique $\bar{f} : (S/\sim)^2 \to S/\sim$ which satisfies

$$\bar{f} \circ [p \times p] = p \circ f.$$

Proof. Define $\bar{f}$ by $\bar{f}(M,N) = p(f(x,y))$, where $x \in M$ and $y \in N$. To see that $\bar{f}$ is well-defined, consider $x, x' \in M$ and $y, y' \in N$. Then $f(x,y) \sim f(x',y')$, so $p(f(x,y)) = p(f(x',y'))$. Therefore, the particular choice of $x \in M$ and the particular choice of $y \in N$ have no effect on the definition of $f(M,N)$; $\bar{f}$ is a well-defined function.

Now suppose that $\bar{g} : (S/\sim)^2 \to S/\sim$ also satisfies $\bar{g} \circ [p \times p] = p \circ f$. Then, for all $x, y \in S$, $\bar{f}(p(x), p(y)) = \bar{g}(p(x), p(y))$. However, because every equivalence class is nonempty, for all $M, N \in S/\sim$, there are $x, y \in S$ such that $p(x) = M$ and $p(y) = N$. Therefore, for all $M, N \in S$, $\bar{f}(M,N) = \bar{g}(M,N)$. Hence $\bar{f} = \bar{g}$. This shows that $\bar{f}$ is unique. □

10.2 Rational addition

We formally define the rational numbers as equivalence classes of fractions, and show how to add them.

Definition 10.2. The set of rational numbers, denoted $\mathbb{Q}$, is the set of equivalence classes of fractions. That is $\mathbb{Q} = F/\sim$, where F is the set of fractions and $\sim$ is the equivalence we defined earlier on the set of fractions. If (x,y) is a fraction, we let x/y or $\frac{x}{y}$ denote its equivalence class. Thus,

$$\frac{x}{y} = \frac{u}{v} \text{ means that } xv = yu.$$

Definition 10.3. Define a binary operator (a kind of function) $\oplus : F^2 \to F$ by the rule

$$(x,y) \oplus (u,v) = (xv + yu, yv).$$

Note that since $y, v \neq 0$, we also have $yv \neq 0$, so the output of the function $\oplus$ is again an element of F.

Theorem 10.4. *The binary operator $\oplus : F^2 \to F$ descends to a binary operator $\overline{\oplus} : \mathbb{Q}^2 \to \mathbb{Q}$.*

Proof. Suppose that $(x,y) \sim (x',y')$ and $(u,v) \sim (u',v')$ are two pairs of equivalent fractions. We claim that

$$(x,y) \oplus (u,v) \sim (x',y') \oplus (u',v').$$

To prove this, it suffices to show that

$$(xv + yu, yv) \sim (x'v' + y'u', y'v').$$

Since $(x, y) \sim (x', y')$, we know that $xy' = yx'$. Since $(u, v) \sim (u', v')$, we know that $uv' = vu'$. Hence we compute

$$\begin{aligned}(xv + yu)y'v' &= (xy')vv' + (uv')yy' \\ &= (yx')vv' + (vu')yy' \\ &= (x'v' + y'u')yv.\end{aligned}$$

This implies $(xv+yu, yv) \sim (x'v'+y'u', y'v')$ as required. The "descent of a binary operator" theorem implies that $\overline{\oplus}$ is well-defined. □

In this way, we find that the function $\overline{\oplus} : \mathbb{Q}^2 \to \mathbb{Q}$, given by

$$\frac{x}{y} \overline{\oplus} \frac{u}{v} = \frac{xv + yu}{yv},$$

is well defined. The resulting rational number (an equivalence class of fractions) depends only on the rational summands (each being an equivalence class of fractions).

10.3 Construction of $\mathbb{Z}/m\mathbb{Z}$

The integers modulo m are usually represented as the integers $0, 1, \ldots, m-1$, and have modular addition and modular multiplication defined upon them. (Sometimes these operations are called clock arithmetic.) Here we construct sets representing the modular integers as equivalence classes of integers, and also define the modular arithmetic operations as descents of the integer arithmetic operations.

Here we take for granted the set of integers, and all familiar properties of addition, subtraction, and multiplication of integers. We also take for granted the "integral domain" property of integers, which states that $xy = 0$ implies $x = 0$ or $y = 0$, for all integers x, y. Equivalently, if $x \neq 0$ and $y \neq 0$, then $xy \neq 0$. This implies the cancellation property: if $x, y, z \in \mathbb{Z}$ and $z \neq 0$ and $xz = yz$, then $x = y$.

Fix a positive integer m throughout the following; m will be called the *modulus*. Define a relation $\equiv_m$ on $\mathbb{Z}$ by the rule

$$x \equiv_m y \text{ means that } (x - y) \text{ is an integer multiple of } m.$$

In other words, $x \equiv_m y$ if and only if there exists an integer d such that $x - y = dm$. For example, $2 \equiv_7 -12$ since $2 - (-12) = 14$ and 14 is a multiple of 7.

Theorem 10.5. *For any positive integer modulus m, the relation $\equiv_m$ is an equivalence relation.*

Proof. We check that $\equiv_m$ is reflexive, symmetric, and transitive.

Reflexive If $x \in \mathbb{Z}$, then $x - x = 0$, and 0 is a multiple of m (note $0 = 0 \times m$). Thus $x \equiv_m x$.

Symmetric If $x, y \in \mathbb{Z}$ and $x \equiv_m y$, then $x - y = dm$ for some integer d. Hence $y - x = (-d)m$. Therefore $y \equiv_m x$.

Transitive Suppose that $x, y, z \in \mathbb{Z}$ and $x \equiv_m y$ and $y \equiv_m z$. Then there exist integers d and e such that $x - y = dm$ and $y - z = em$. It follows that

$$x - z = (x - y) + (y - z) = dm + em = (d + e)m.$$

Hence $x \equiv_m z$. □

The set of equivalence classes for the relation $\equiv_m$ is often called $\mathbb{Z}_m$ or $\mathbb{Z}/m\mathbb{Z}$. Elements of $\mathbb{Z}_m$ are called *congruence classes modulo m*.

In most mathematical texts, the relation $\equiv_m$ is usually written a little differently. Instead of writing "$x \equiv_m y$" one writes "$x \equiv y \bmod m$". Understanding the arithmetic and algebra of congruence classes is at the foundation of number theory.

Proposition 10.6. *Integer addition, understood as a binary operator $+ : \mathbb{Z}^2 \to \mathbb{Z}$ descends to a binary operator $\overline{\oplus} : \mathbb{Z}_m^2 \to \mathbb{Z}_m$.*

Proof. Suppose that $x \equiv_m x'$ and $y \equiv_m y'$. Then $x - x' = dm$ and $y - y' = em$, for some integers d and e. We find that

$$(x + y) - (x' + y') = (x - x') + (y - y') = dm + em = (d + e)m.$$

Hence $x + y \equiv_m x' + y'$. The "descent of a binary operator theorem" yields the result. □

Therefore, we may "add congruence classes modulo m". Let's interrogate what this means.

The point of showing that a binary operator $+$ descends from a set to the equivalence classes on that set is to show that the new binary operator $\overline{\oplus}$ is "well-defined". But how, exactly, is it defined? The "descent of a binary operator theorem" tells us how: $\alpha \,\overline{\oplus}\, \beta$ is defined to be the equivalence class

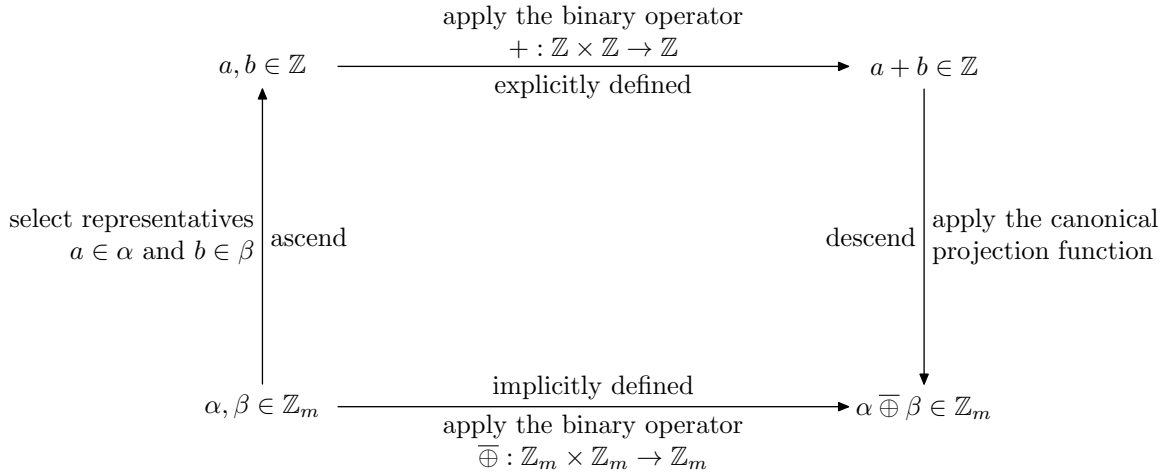


Figure 14: Definition of $\overline{\oplus}$ as the descent of $+$ for $\mathbb{Z}_m$.

to which $a + b$ belongs, whenever $a \in \alpha$ and $b \in \beta$. The descent theorem says that it does not matter how we choose a, b, the result will always be the same. We could say the way to calculate $\alpha \,\overline{\oplus}\, \beta$ is to "ascend" to representatives $a \in \alpha$ and $b \in \beta$, apply the (clearly defined) binary operator $+$ to a and b, then "descend", meaning find the equivalence class to which $a + b$ belongs. Figure 14 summarises the process. In order to traverse the diagram from bottom left to bottom right along the direct path, it is equivalent (by definition!) to traverse the diagram along the indirect path that visits the top left and top right corners on the way.

Proposition 10.7. *Integer multiplication, $\times : \mathbb{Z}^2 \to \mathbb{Z}$ descends to a binary operator $\overline{\otimes} : \mathbb{Z}_m^2 \to \mathbb{Z}_m$.*

Proof. Suppose that $x \equiv_m x'$ and $y \equiv_m y'$. Then $x - x' = dm$ and $y - y' = em$, for some integers d and e. Therefore

$$(x \times y) - (x' \times y') = xy - xy' + xy' - x'y' = x(y - y') + y'(x - x') = xem + y'dm = (xe + y'd)m.$$

Therefore $x \times y - x' \times y'$ is a multiple of m. Hence $x \times y \equiv_m x' \times y'$. We can apply the "descent of a binary operator theorem" to construct $\overline{\otimes} : \mathbb{Z}_m^2 \to \mathbb{Z}_m$. □

Therefore, we may "multiply congruence classes modulo m". The meaning of multiplication of congruence classes parallels the meaning for addition of equivalence classes.

The "descent of a binary operator theorem" tells us that $\alpha \,\overline{\otimes}\, \beta$ is defined to be the equivalence class to which $a \times b$ belongs, whenever $a \in \alpha$ and $b \in \beta$. The descent theorem says that it does not matter how we choose a, b, the result will always be the same. We could say the way to calculate $\alpha \,\overline{\otimes}\, \beta$ is to "ascend" to representatives $a \in \alpha$ and $b \in \beta$, apply the (clearly defined) binary operator $\times$ to a and b, then "descend", meaning find the equivalence class to which $a \times b$ belongs. Figure 15 summarises the process. In order to traverse the diagram from bottom left to bottom right along the direct path, it is equivalent (by definition!) to traverse the diagram along the indirect path that visits the top left and top right corners on the way.

Theorem 10.8. *If m is prime, then $(\mathbb{Z}_m, \overline{\oplus}, \overline{\otimes})$ is a field.*

Proof.

1. **Closure under addition.** If $\alpha, \beta \in \mathbb{Z}_m$, then there exist $a \in \alpha$ and $b \in \beta$ and $a, b \in \mathbb{Z}$. But $a + b \in \mathbb{Z}$ because integers are closed under addition. But, because $\equiv_m$ is an equivalence on $\mathbb{Z}$,

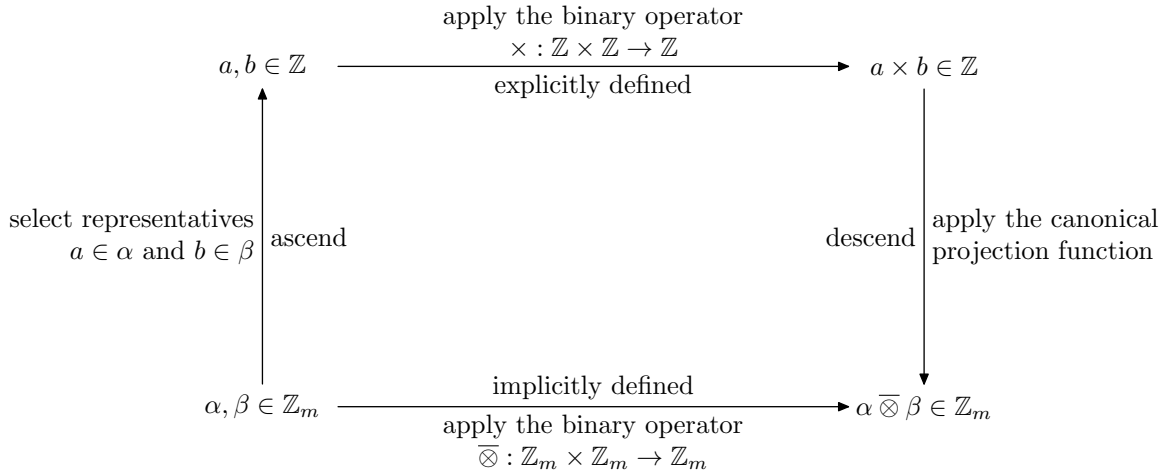


Figure 15: Definition of $\overline{\otimes}$ as the descent of $\times$ for $\mathbb{Z}_m$.

there must be some $\gamma \in \mathbb{Z}_m$ for which $a + b \in \gamma$. Finally, as $\overline{\oplus}$ is the descent of $+$, it holds that $\alpha \mathbin{\overline{\oplus}} \beta = \gamma$.

2. **Commutativity of addition.** If $\alpha, \beta \in \mathbb{Z}_m$, then there exist $a \in \alpha$ and $b \in \beta$ and $a, b \in \mathbb{Z}$. By definition, $\alpha \mathbin{\overline{\oplus}} \beta$ is the equivalence class to which $a + b$ belongs, and $\beta \mathbin{\overline{\oplus}} \alpha$ is the equivalence class to which $b + a$ belongs. But $a + b = b + a$, because integer addition is commutative. So $\alpha \mathbin{\overline{\oplus}} \beta = \beta \mathbin{\overline{\oplus}} \alpha$.

3. **Associativity of addition.** If $\alpha, \beta, \gamma \in \mathbb{Z}_m$, then there exist $a \in \alpha$, $b \in \beta$, $c \in \gamma$ and $a, b, c \in \mathbb{Z}$. By definition, $(\alpha \mathbin{\overline{\oplus}} \beta) \mathbin{\overline{\oplus}} \gamma$ is the equivalence class to which $(a + b) + c$ belongs, and $\alpha \mathbin{\overline{\oplus}} (\beta \mathbin{\overline{\oplus}} \gamma)$ is the equivalence class to which $a + (b + c)$ belongs. But $(a + b) + c = a + (b + c)$, because integer addition is associative. So $(\alpha \mathbin{\overline{\oplus}} \beta) \mathbin{\overline{\oplus}} \gamma = \alpha \mathbin{\overline{\oplus}} (\beta \mathbin{\overline{\oplus}} \gamma)$.

4. **Existence of additive identity.** We define $0_m \in \mathbb{Z}_m$ to be the equivalence class to which the integer 0 belongs. If $\alpha \in \mathbb{Z}_m$, then there exists an integer $a \in \alpha$. But $\alpha \mathbin{\overline{\oplus}} 0_m$ is the equivalence class to which $a + 0$ belongs, and $a + 0 = a$ because 0 is an additive identity for the integers. Hence a also belongs to $\alpha \mathbin{\overline{\oplus}} 0_m$, and $\alpha \mathbin{\overline{\oplus}} 0_m = \alpha$.

5. **Existence of additive inverses.** If $\alpha \in \mathbb{Z}_m$, then there exists $a \in \alpha$, and $a, -a \in \mathbb{Z}$. But $a + (-a) = 0 \in 0_m$. So, defining β to be the equivalence class to which $-a$ belongs, we find $\alpha \mathbin{\overline{\oplus}} \beta = 0_m$, and β is an additive inverse of α.

6–8. Closure under multiplication, commutativity of multiplication, and associativity of multiplication, have proofs very similar to the above proofs of the corresponding properties of addition.

9. **Existence of multiplicative identity.** We define $1_m \in \mathbb{Z}_m$ to be the equivalence class to which the integer 1 belongs. If $\alpha \in \mathbb{Z}_m$, then there exists an integer $a \in \alpha$. But $\alpha \mathbin{\overline{\otimes}} 1_m$ is the equivalence class to which $a \times 1$ belongs, and $a \times 1 = a$ because 1 is an multiplicative identity for the integers. Hence a also belongs to $\alpha \mathbin{\overline{\otimes}} 1_m$, and $\alpha \mathbin{\overline{\otimes}} 1_m = \alpha$.

10. **Existence of multiplicative inverses.** If $\alpha \in \mathbb{Z}_m \setminus \{0_m\}$, then there exists an integer $a \in \alpha$. Because m is prime, if $\mathrm{GCD}(m, a) \neq 1$, then $m \mid a$, in which case $\alpha = 0_m$. Therefore $\mathrm{GCD}(m, a) = 1$. It follows from Bézout's lemma that there are integers x, y for which $ax + my = 1$. So there is an integer x such that $ax - 1$ is an integer multiple of m. In other words, $ax \equiv_m 1$.

Let ξ be the equivalence class to which x belongs. Then $\alpha \mathbin{\overline{\otimes}} \xi = 1_m$, so ξ is a multiplicative inverse of α.

11. **Distributivity of multiplication over addition.** If $\alpha, \beta, \gamma \in \mathbb{Z}_m$, then there exist $a \in \alpha$, $b \in \beta$, $c \in \gamma$ and $a, b, c \in \mathbb{Z}$. By definition, $\alpha \mathbin{\overline{\otimes}} (\beta \mathbin{\overline{\oplus}} \gamma)$ is the equivalence class to which $a \times (b+c)$ belongs, and $(\alpha \mathbin{\overline{\otimes}} \beta) \mathbin{\overline{\oplus}} (\alpha \mathbin{\overline{\otimes}} \gamma)$ is the equivalence class to which $a \times b + a \times c$ belongs. But $a \times (b+c) = a \times b + a \times c$, because integer addition is associative. So $\alpha \overline{\otimes} (\beta \overline{\oplus} \gamma) = (\alpha \overline{\otimes} \beta) \overline{\oplus} (\alpha \overline{\otimes} \gamma)$.

12. **Zero / one property.** If $1_m = 0_m$, then $1 \equiv_m 0$. But then 1 is an integer multiple of m. This is true only for $m = 1$. But m is prime, so $m \geqslant 2$. Therefore $1_m \neq 0_m$. □

10.4 Construction of $\mathbb{Q}$

We construct the rational numbers as equivalence classes of ordered pairs of integers, where the ordered pairs represent what we usually think of as fractions.

Recall that F is the set of fractions, $\sim$ is the equivalence on this set, and the rationals have been defined as $\mathbb{Q} = F/\sim$.

Definition 10.9. Define a function $\otimes : F^2 \to F$ by the rule

$$(x, y) \otimes (u, v) = (xu, yv).$$

Theorem 10.10. *The binary operator $\otimes : F^2 \to F$ descends to a binary operator $\overline{\otimes} : \mathbb{Q}^2 \to \mathbb{Q}$.*

Proof. Suppose that $(x, y) \sim (x', y')$ and $(u, v) \sim (u', v')$ are two pairs of equivalent fractions. By definition, $(x, y) \otimes (u, v) = (xu, yv)$, which is a fraction because the integers are closed under multiplication and two nonzero integers have nonzero product. We claim that

$$(x, y) \otimes (u, v) \sim (x', y') \otimes (u', v').$$

To prove this, it suffices to show that

$$(xu, yv) \sim (x'u', y'v').$$

Since $(x, y) \sim (x', y')$, we know that $xy' = yx'$. Since $(u, v) \sim (u', v')$, we know that $uv' = vu'$. Hence we compute

$$\begin{aligned}(xu)(y'v') &= (xy')(uv') \\ &= (yx')(vu') \\ &= (x'u')(yv).\end{aligned}$$

This implies $(xu, yv) \sim (x'u', y'v')$ as required. The "descent of a binary operator" theorem implies that $\overline{\otimes}$ is well-defined. □

Theorem 10.11. *The triple $(\mathbb{Q}, \overline{\oplus}, \overline{\otimes})$ is a field.*

Proof.

1–3. The arguments for closure under addition, commutativity of addition, and associativity of addition are similar to the equivalent properties of multiplication.

4. **Existence of additive identity.** We define $0_{\mathbb{Q}} \in \mathbb{Q}$ to be the equivalence class to which the fraction $(0, 1)$ belongs. If $\alpha \in \mathbb{Q}$, then there exists a fraction $(a, A) \in \alpha$. But $\alpha \mathbin{\overline{\oplus}} 0_{\mathbb{Q}}$ is the equivalence class to which $(a \times 1 + 0 \times A, A \times 1) = (a, A)$ belongs. So $\alpha \mathbin{\overline{\oplus}} 0_{\mathbb{Q}} = \alpha$.

5. **Existence of additive inverses.** If $\alpha \in \mathbb{Q}$, then there exists a fraction $(a, A) \in \alpha$. But then also $(-a, A) \in F$. But $(a, A) \oplus (-a, A) = (aA - Aa, A^2) = (0, A^2) \sim (0, 1) \in 0_{\mathbb{Q}}$. So, defining β to be the equivalence class to which $(-a, A)$ belongs, we find $\alpha \mathbin{\overline{\oplus}} \beta = 0_{\mathbb{Q}}$, and β is an additive inverse of α.

6. **Closure under multiplication.** If $\alpha, \beta \in \mathbb{Q}$, then there exist fractions $(a, A) \in \alpha$ and $(b, B) \in \beta$ so that $a, A, b, B \in \mathbb{Z}$. It was argued in the above $(a, A) \otimes (b, B) \in F$. But, because $\sim$ is an equivalence on F, there must be some $\gamma \in \mathbb{Q}$ for which $(a, A) \otimes (b, B) \in \gamma$. Finally, as $\overline{\otimes}$ is the descent of $\otimes$, it holds that $\alpha \mathbin{\overline{\otimes}} \beta = \gamma$.

7. **Commutativity of multiplication.** If $\alpha, \beta \in \mathbb{Q}$, then there exist fractions $(a, A) \in \alpha$ and $(b, B) \in \beta$ so that $a, A, b, B \in \mathbb{Z}$. By definition, $\alpha \mathbin{\overline{\otimes}} \beta$ is the equivalence class to which (aA, bB) belongs, and $\beta \mathbin{\overline{\otimes}} \alpha$ is the equivalence class to which (Aa, Bb) belongs. But integer multiplication is commutative, so $\alpha \mathbin{\overline{\otimes}} \beta = \beta \mathbin{\overline{\otimes}} \alpha$.

8. **Associativity of multiplication.** If $\alpha, \beta, \gamma \in \mathbb{Q}$, then there exist fractions $(a, A) \in \alpha$, $(b, B) \in \beta$, and $(c, C) \in \gamma$ so that $a, A, b, B, c, C \in \mathbb{Z}$. By definition, $(\alpha \mathbin{\overline{\otimes}} \beta) \mathbin{\overline{\otimes}} \gamma$ is the equivalence class to which $[(a, A) \otimes (b, B)] \otimes (c, C)$ belongs, and $\alpha \mathbin{\overline{\otimes}} (\beta \mathbin{\overline{\otimes}} \gamma)$ is the equivalence class to which $(a, A) \otimes [(b, B) \otimes (c, C)]$ belongs. But

$$\begin{aligned}
[(a, A) \otimes (b, B)] \otimes (c, C) &= (ab, AB) \otimes (c, C) \\
&= (abc, ABC) \\
&= (a, A) \otimes (bc, BC) \\
&= (a, A) \otimes [(b, B) \otimes (c, C)].
\end{aligned}$$

   So $(\alpha \mathbin{\overline{\otimes}} \beta) \mathbin{\overline{\otimes}} \gamma = \alpha \mathbin{\overline{\otimes}} (\beta \mathbin{\overline{\otimes}} \gamma)$.

9. **Existence of multiplicative identity.** We define $1_{\mathbb{Q}} \in \mathbb{Q}$ to be the equivalence class to which the fraction $(1, 1)$ belongs. If $\alpha \in \mathbb{Q}$, then there exists a fraction $(a, A) \in \alpha$. But $\alpha \mathbin{\overline{\otimes}} 1_{\mathbb{Q}}$ is the equivalence class to which $(a \times 1, A \times 1) = (a, A)$ belongs. So $\alpha \mathbin{\overline{\otimes}} 1_{\mathbb{Q}} = \alpha$.

10. **Existence of multiplicative inverses.** If $\alpha \in \mathbb{Q} \setminus \{0_{\mathbb{Q}}\}$, then there exists a fraction $(a, A) \in \alpha$ for nonzero integers a, A. But then also $(A, a) \in F$. But $(a, A) \otimes (A, a) = (aA, Aa) \sim (1, 1) \in 1_{\mathbb{Q}}$. So, defining β to be the equivalence class to which (A, a) belongs, we find $\alpha \mathbin{\overline{\otimes}} \beta = 1_{\mathbb{Q}}$, and β is a multiplicative inverse of α.

11. **Distributivity of multiplication over addition.** If $\alpha, \beta, \gamma \in \mathbb{Q}$, then there exist fractions $(a, A) \in \alpha$, $(b, B) \in \beta$, and $(c, C) \in \gamma$ so that $a, A, b, B, c, C \in \mathbb{Z}$. By definition, $\alpha \mathbin{\overline{\otimes}} [\beta \mathbin{\overline{\oplus}} \gamma]$ is the equivalence class to which $(a, A) \otimes [(b, B) \oplus (c, C)]$ belongs, and $[\alpha \mathbin{\overline{\otimes}} \beta] \mathbin{\overline{\oplus}} [\alpha \mathbin{\overline{\otimes}} \gamma]$ is the equivalence class to which $[(a, A) \otimes (b, B)] \oplus [(a, A) \otimes (c, C)]$ belongs. But

$$\begin{aligned}
(a, A) \otimes [(b, B) \oplus (c, C)] &= (a, A) \otimes (bC + cB, BC) \\
&= (abC + aBc, ABC) \\
&\sim (abAC + acAB, ABAC) \\
&= (ab, AB) \oplus (ac, AC).
\end{aligned}$$

   So $\alpha \mathbin{\overline{\otimes}} [\beta \mathbin{\overline{\oplus}} \gamma] = [\alpha \mathbin{\overline{\otimes}} \beta] \mathbin{\overline{\oplus}} [\alpha \mathbin{\overline{\otimes}} \gamma]$.

12. **Zero / one property.** If $1_{\mathbb{Q}} = 0_{\mathbb{Q}}$, then $(1, 1) \sim (0, 1)$. But then $1 \times 1 = 1 \times 0$, which is false. Therefore $1_{\mathbb{Q}} \neq 0_{\mathbb{Q}}$. □

11 Week 11

11.1 Convergent rational sequences are Cauchy

We show that if a sequence of rational numbers has a rational limit, then the sequence must be Cauchy. We use the trick of "halving epsilon"

Lemma 11.1 (Triangle inequality)**.** *Suppose* $\alpha, \beta \in \mathbb{Q}$*. Then* $|\alpha + \beta| \leqslant |\alpha| + |\beta|$*.*

Proof. Multiplying numbers from an ordered field by nonnegative numbers preserves order. As both sides of the target inequality are nonnegative, it is equivalent to prove

$$(\alpha + \beta)^2 \leqslant (|\alpha| + |\beta|)^2, \tag{11.1}$$

the right-hand side of which is equal to

$$|\alpha|^2 + 2|\alpha\beta| + |\beta|^2 = \alpha^2 + 2|\alpha\beta| + \beta^2.$$

Expanding the left hand side of inequality (11.1), we get $(\alpha + \beta)^2 = \alpha^2 + 2\alpha\beta + \beta^2$. Finally, observe that $\alpha\beta \leqslant |\alpha\beta|$. □

Proposition 11.2. *Suppose that* $(a_n)_{n\in\mathbb{N}}$ *is a convergent sequence of rational numbers. Then* $(a_n)_{n\in\mathbb{N}} \in C_{\mathbb{Q}}$*.*

Proof. As $(a_n)_{n\in\mathbb{N}}$ converges, there is a rational limit a such that, for any $\delta > 0$, there is some $N = N(\delta) \in \mathbb{N}$ for which if $n > N$ then $|a_n - a| < \delta$. In particular, for any $\varepsilon > 0$, there is some $N = N(\varepsilon/2) \in \mathbb{N}$ for which if $n > N$ then $|a_n - a| < \varepsilon/2$. Fix $\varepsilon > 0$, and suppose $m, n > N(\varepsilon/2)$. Then, using lemma 11.1 with $\alpha = a_m - a$ and $\beta = a - a_n$,

$$|a_m - a_n| = |(a_m - a) + (a - a_n)| \leqslant |a_m - a| + |a - a_n| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Therefore $(a_n)_{n\in\mathbb{N}}$ is Cauchy. □

11.2 Rational Cauchy sequences are bounded

We show that all rational Cauchy sequences are bounded.

Lemma 11.3 (Reverse triangle inequality)**.** *Suppose* $\alpha, \beta \in \mathbb{Q}$*. Then* $|\alpha| - |\beta| \leqslant |\alpha - \beta|$*.*

Proof. By the closure of the rationals under addition and multiplication, $\gamma = \alpha - \beta \in \mathbb{Q}$. By the triangle inequality (an earlier lemma), $|\gamma + \beta| \leqslant |\gamma| + |\beta|$. It follows that $|\alpha| \leqslant |\alpha - \beta| + |\beta|$. □

Proposition 11.4. *Suppose that* $(a_n)_{n\in\mathbb{N}}$ *is a Cauchy sequence of rational numbers. Then there exists* $M \in \mathbb{Q}$ *such that*

$$-M \leqslant a_1, a_2, a_3, \ldots \leqslant M.$$

Proof. We fix $\varepsilon = 1$, so that $\mathbb{Q} \ni \varepsilon > 0$. By definition, there is a natural number N such that, for all $m, n > N$, $|a_m - a_n| < \varepsilon$. In particular, for all $m > N$, $|a_m - a_{N+1}| < \varepsilon$. Hence, by lemma 11.3, $|a_m| < |a_{N+1}| + 1$ holds for all $m > N$. It follows that, for all $m \in \mathbb{N}$,

$$|a_m| < \max\{|a_1|, |a_2|, \ldots, |a_N|, |a_{N+1}| + 1\}.$$

Therefore, a valid choice of M is the right-hand side of the above inequality. □

11.3 The Cauchy equivalence

We establish that the Cauchy "equivalence" is indeed an equivalence relation.

Theorem 11.5. *The relation $\sim$ on $C_{\mathbb{Q}}$ given by*

$$(a_n)_{n\in\mathbb{N}} \sim (b_n)_{n\in\mathbb{N}} \qquad \textit{if} \qquad (a_n - b_n) \to 0 \textit{ as } n \to \infty$$

is an equivalence relation.

Proof. **Reflexivity** If $(a_n)_{n\in\mathbb{N}} \in C_{\mathbb{Q}}$, then $(a_n - a_n) = 0$ for all $n \in \mathbb{N}$, and the sequence which is identically 0 has limit 0.

Symmetry If $(a_n)_{n\in\mathbb{N}} \sim (b_n)_{n\in\mathbb{N}}$, then for each $\varepsilon > 0$ there is some $N(\varepsilon) \in \mathbb{N}$ such that, for all $n > N$, $|a_n - b_n| < \varepsilon$. But $|b_n - a_n| = |a_n - b_n|$, so the same natural number function $N(\varepsilon)$ can be used to show $(b_n - a_n) \to 0$. Therefore $(b_n)_{n\in\mathbb{N}} \sim (a_n)_{n\in\mathbb{N}}$.

Transitivity Suppose that $(a_n)_{n\in\mathbb{N}} \sim (b_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}} \sim (c_n)_{n\in\mathbb{N}}$. Then, for any $\delta > 0$, there exist natural numbers $N_{a,b}(\delta)$ and $N_{b,c}(\delta)$ such that:

$$\text{if } n > N_{a,b}, \text{ then } |a_n - b_n| < \delta,$$

and

$$\text{if } n > N_{b,c}, \text{ then } |b_n - c_n| < \delta.$$

For $\varepsilon > 0$, let $N = \max\{N_{a,b}(\varepsilon/2), N_{b,c}(\varepsilon/2)\}$. This ensures that

$$\text{if } n > N, \text{ then } |a_n - b_n| < \frac{\varepsilon}{2} \text{ and } |b_n - c_n| < \frac{\varepsilon}{2}.$$

It follows that

$$\text{if } n > N, \text{ then } |a_n - b_n| + |b_n - c_n| < \varepsilon.$$

By the triangle inequality, $|a_n - c_n| = |a_n - b_n + b_n - c_n| \leqslant |a_n - b_n| + |b_n - c_n| < \varepsilon$. Hence $(a_n)_{n\in\mathbb{N}} \sim (c_n)_{n\in\mathbb{N}}$. □

11.4 Adding and multiplying rational Cauchy sequences

We aim to construct real numbers as equivalence classes of rational Cauchy sequences. We will eventually have to show that adding and multiplying real numbers produces real numbers (see the field axioms), so it would help to know that adding and multiplying the underlying objects, rational Cauchy sequences, produces more rational Cauchy sequences.

Proposition 11.6. *Suppose that $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ are both rational Cauchy sequences. Then $(a_n + b_n)_{n\in\mathbb{N}}$ is also a rational Cauchy sequence.*

Proof. The rational numbers are closed under addition, so $(a_n + b_n)_{n\in\mathbb{N}}$ is a rational sequence.

Let $\delta > 0$. Then there exist natural numbers $N_a(\delta), N_b(\delta)$ such that

$$\begin{aligned} &\text{if } m, n > N_a(\delta), \text{ then } |a_m - a_n| < \delta, \\ &\text{if } m, n > N_b(\delta), \text{ then } |b_m - b_n| < \delta. \end{aligned}$$

Suppose $\varepsilon > 0$, set $N = \max\{N_a(\varepsilon/2), N_b(\varepsilon/2)\}$, and suppose that $m, n > N$. Then

$$\begin{aligned} |(a_m + b_m) - (a_n + b_n)| &= |(a_m - a_n) + (b_m - b_n)| \\ &\leqslant |a_m - a_n| + |b_m - b_n| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Therefore the sequence $(a_n + b_n)_{n\in\mathbb{N}}$ is Cauchy. □

Proposition 11.7. *Suppose that $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ are both rational Cauchy sequences. Then $(a_n b_n)_{n\in\mathbb{N}}$ is also a rational Cauchy sequence.*

Proof. The rational numbers are closed under multiplication, so $(a_n b_n)_{n\in\mathbb{N}}$ is a rational sequence.

As both of the original sequences are Cauchy, they are both bounded. So we can choose a positive rational number M that is a common bound for both sequences, in the sense that, for all $n \in \mathbb{N}$, $|a_n|, |b_n| < M$. Let $\delta > 0$. Then there exist natural numbers $N_a(\delta), N_b(\delta)$ such that

$$\begin{aligned} &\text{if } m, n > N_a(\delta), \text{ then } |a_m - a_n| < \delta, \\ &\text{if } m, n > N_b(\delta), \text{ then } |b_m - b_n| < \delta. \end{aligned}$$

Suppose $\varepsilon > 0$, set $N = \max\{N_a(\varepsilon/2M), N_b(\varepsilon/2M)\}$, and suppose that $m, n > N$. Then

$$\begin{aligned} |(a_m b_m) - (a_n b_n)| &= |(a_m - a_n)b_m + a_n(b_m - b_n)| \\ &\leqslant |a_m - a_n||b_m| + |a_n||b_m - b_n| \\ &< \frac{\varepsilon}{2M} M + M \frac{\varepsilon}{2M} = \varepsilon. \end{aligned}$$

Therefore the sequence $(a_n b_n)_{n\in\mathbb{N}}$ is Cauchy. □

11.5 Subsequences belong to the same equivalence class

We show that the operation of "taking a subsequence" of a rational Cauchy sequence produces another rational Cauchy sequence which is Cauchy equivalent to the original sequence.

Suppose that we have a sequence

$$(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, \ldots).$$

We can define a new sequence by deleting a few terms

$$(x_1, x_2, \quad x_5, \quad x_7, \quad x_{10}, \ldots).$$

We may delete as many terms as we like, as long as we leave infinitely many terms undeleted. The new sequence is called a *subsequence* of the original sequence. A more formal way of choosing a subsequence is to define a strictly increasing function $r : \mathbb{N} \to \mathbb{N}$ so that the nth term in the subsequence is $x_{r(n)}$. Because the domain of r is $\mathbb{N}$, we often denote $r(n)$ by r_n instead, so the subsequence is written $(x_{r_n})_{n\in\mathbb{N}}$. In the example above, we have

$$r_1 = 1, \quad r_2 = 2, \quad r_3 = 5, \quad r_4 = 7, \quad r_5 = 10, \quad \ldots.$$

Note that, because r is strictly increasing, the terms of the subsequence $(x_{r_n})_{n\in\mathbb{N}}$ are guaranteed to be in the same order as they were in the original sequence.

We are aiming to show that every subsequence of a rational Cauchy sequence is also a rational Cauchy sequence, and belongs to the same Cauchy equivalence class.

Proposition 11.8. *Denote the Cauchy equivalence by $\sim$. Suppose that $(x_n)_{n\in\mathbb{N}}$ is a rational Cauchy sequence, and $(r_n)_{n\in\mathbb{N}}$ is a strictly increasing sequence of natural numbers. Then $(x_{r_n})_{n\in\mathbb{N}}$ is a rational Cauchy sequence and $(x_n)_{n\in\mathbb{N}} \sim (x_{r_n})_{n\in\mathbb{N}}$.*

Proof. As x_{r_n} is a term in the original sequence, it is a rational number. For $\varepsilon > 0$, there exists a natural number $N = N(\varepsilon)$ such that, for all $m, n > N$, $|x_m - x_n| < \varepsilon$. Note that $r_j \geqslant j$ for all natural numbers j, so $r_m, r_n > N$ also. Therefore $|x_{r_m} - x_{r_n}| < \varepsilon$. We have shown that $(x_{r_n})_{n\in\mathbb{N}}$ is a rational Cauchy sequence.

For $\varepsilon > 0$, there exists a natural number N such that, for all $m, n > N = N(\varepsilon)$, $|x_m - x_n| < \varepsilon$. Setting $m = r_n \geqslant n > N$, we have $|x_{r_n} - x_n| < \varepsilon$. Therefore, the sequence $(x_{r_n} - x_n)_{n\in\mathbb{N}}$ converges to 0. Hence $(x_n)_{n\in\mathbb{N}} \sim (x_{r_n})_{n\in\mathbb{N}}$. □