1st Year Mathematics
Imperial College London

Autumn 2014

Foundations of Analysis

Lectured by:
Prof. A. CORTI

Humbly Typed by:
Karim BACCHUS

$$Q.E.D.$$

Syllabus

An introductory course involving basic material, which will be widely used later.

The Very Beginning

Sets, Logic and basic proofs. The Archimedean Axiom. The principle of mathematical induction, strong induction, smallest element axiom.

Number Systems

Integers, rational numbers, real numbers, decimal expansions for rationals and reals. Inequalities. The Complex Numbers, De Moivre's Theorem. Polynomials.

Number Theory

Division theorem, Euclids algorithm. Prime Numbers and the Fundamental Theorem of Arithmetic. Arithmetic modulo n. Systems of congruences and the Chinese Remainder Theorem. Euler Phi Function, Fermat's Little Theorem. Public Key Cryptography.

Counting

Permutations and combinations. Inclusion-Exclusion Principle. The Binomial Theorem. Equivalence relations.

Functions

Injective, Surjective and Bijective Functions, mappings between functions. Countability of Sets, Cantor's Diagonal Argument.

Foundations of Analysis

Supremum and Infimums of sets. The Completeness axiom for $\mathbb{R}$. Introduction to limits of sequences.

Appropriate books

M. Liebeck *A Concise Introduction to Pure Mathematics.*
K. Houston *How to Think Like a Mathematician.*
E. Hurst and M. Gould *Bridging the Gap to University Mathematics.*

Contents

1 The Very Beginning

> A human who does not understand the difference between strong induction
> and usual induction is a sad and demeaning spectacle.
>
> *- Alessio Corti*

Sets

A set S is a collection of objects (called the *elements* of the set)

> **Example 1.1.** A way to specify a set is to list the objects (between curly brackets):
>
> $$S = \{1, 3, 7\}$$
>
> The order of elements is unimportant, as is repetition:
>
> $$\{1, 2\} = \{2, 1\} = \{1, 1, 2\}$$

I say $S_1 \subset S_2$ (S_1 *contained* in S_2) if every element of S_1 is also an element of S_2. I can write this as a *statement*: $x \in S_1 \implies x \in S_2$

I say $S_1 = S_2$ if $S_1 \subset S_2$ & $S_2 \subset S_1$. Elements can be sets: $S = \{1, 2, \{1, 2\}\}$.
But there is one thing you are never allowed to do:

> **Axiom 1.2** (Foundation Axiom). $S \notin S$

More things about sets:

- $a \notin S$ "a is not an element of S"
- $S_1 \cup S_2$ "S_1 union S_2" $= \{x \mid x \in S_1 \text{ or } x \in S_2 \text{ (or both)}\} = \{x : x \in S_1 \text{ or } x \in S_2\}$
- $S_1 \cap S_2$ "S_1 intersection S_2" $= \{x \mid x \in S_1 \text{ and } x \in S_2\}$
- $S_1 \backslash S_2$ "S_1 take away S_2" $= \{x \mid x \in S_1 \& x \notin S_2\}$
- $S_1 \triangle S_2$ "symmetric difference" $= \{x \mid x \in S_1 \text{ or } x \in S_2 \text{ but not both }\}$
  $= (S_1 \cup S_2) \backslash (S_1 \cap S_2)$

When you reason about sets & other mathematical objects, it is useful to draw pictures:
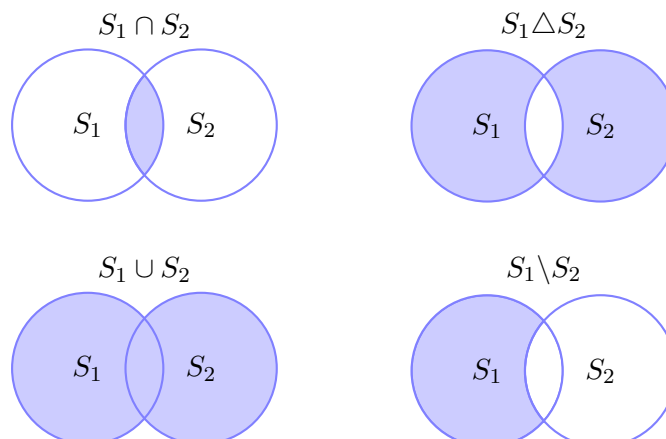
Sometimes (often) it is not practical to list all the elements of a set:

> **Examples 1.3.**
>
> $\mathbb{Z} =$ set of integers $= \{0, +1, -1, 2, -2, 3, -3, \dots\}$
>
> $\mathbb{N} =$ set of natural numbers $= \{0, 1, 2, 3, \dots\} = \{n \in \mathbb{Z} \mid n \geq 0\}$
>
> $\mathbb{Q} =$ set of rational numbers $= \{x \mid x = \frac{p}{q}, \ p \in \mathbb{Z}, q \in \mathbb{N}\backslash\{0\}\}$
>
> $\mathbb{R} =$ set of real numbers
>
> $\mathbb{C} =$ set of complex numbers
>
> $$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$$

$\emptyset$ denotes the *empty set.*

Rules governing the empty set - for every set A:
- $\emptyset \subset A$
- $\emptyset \cup A = A$
- $\emptyset \cap A = \emptyset$

One ambition of maths is to construct everything from first principles. One generally agreed thing is that first principles includes set theory. For instance:

> **Example 1.4.** An approach to constructing $\mathbb{N}$:
>
> $\underline{0} = \emptyset$ is a number, literally.
> $\underline{1} = \{\emptyset\}$
> $\underline{2} = \{\emptyset, \{\emptyset\}\}$
> $\underline{3} = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$
>
> If $\underline{n}$ is a number, then also $\underline{n} + \underline{1} = \underline{n} \cup \{\underline{n}\}$. Now you have all of $\mathbb{N}$!

Q: What does it mean when a natural number is a set of empty sets?

Corti: You are confused, that is the definition! You can never object to a definition. I guess this example shows not only what to do with the empty set, but also if you take going back to first principles a little too far, it becomes something we don't want to do.

Elements of Propositional Calculus

A proof is a chain of statements linked by logical implications (deductions) that establishes the truth of the last statement in the chain.

> **Examples 1.5.** Examples of Statements (aka "propositions")
>
> (i) $n = 3$
>
> (ii) $n^2 - 2n - 3 = 0$
>
> (iii) $x^2 > 5$
>
> (iv) If n^2 is odd then n is odd

(v) If $x \in \mathbb{R}$, $x \geq 0$ then there exists $y \in \mathbb{R}$ such that $y^2 = x$

Implications

Let P & Q be statements.

$$P \implies Q \text{ means any one of:} \begin{cases} \text{If } P \text{ is true, then } Q \text{ is true} \\ \text{If } P \text{ then } Q \\ Q \text{ if } P \\ P \text{ true only if } Q \text{ is true} \\ Q \text{ false} \implies P \text{ false} \\ P \text{ is sufficient for } Q \\ Q \text{ is necessary for } P \end{cases}$$

e.g. $m = 3 \implies m^2 - 2n - 3 = 0$

$$P \iff Q \text{ means any one of:} \begin{cases} P \implies Q \text{ \& } Q \implies P \\ P \text{ if and only if } P \\ P \text{ is necessary and sufficient for } Q \\ P \text{ iff } Q \end{cases}$$

This is just language, don't try to understand it. It's just mathematical English. A computer can be taught that any of those strings of words are completely equivalent, without having to understand anything, and so will you.

Negations

If P is a statement, there is a statement $\overline{P}$ (non-P) such that P is true $\iff$ $\overline{P}$ is false.

$P \implies Q$ is equivalent to $\overline{Q} \implies \overline{P}$ (the basics of proofs by contradiction)

e.g. $(x = 2) \implies (x^2 < 5)$ is equivalent to $(x^2 \geq 5) \implies (x \neq 2)$

Quantifiers

> **Definition.** The symbol $\forall$ means "for all" (the "universal quantifier")
>
> The symbol $\exists$ means "there exists" (the "existential" quantifier)
>
> ($\exists_1$, or $\exists!$ means "there exists a *unique...*")

So instead of writing

$$\text{"If } x \in \mathbb{R}, \ x \geq 0 \text{ then there is } y \in \mathbb{R} \text{ such that } y^2 = x\text{"}$$

I write

$$\text{"}\forall x \in \mathbb{R}, \ (x \geq 0) \implies (\exists y \in \mathbb{R} : y^2 = x)\text{"}$$

In the course of the proof one is allowed to "call up" statements that have been proved previously and axioms (statements that are generally accepted and not proved).

The "propositional calculus" is a very strict system where:

(i) You are only allowd a given (small) list of symbols...

$(\forall, \exists, ", ", \text{ brackets } (), "|", x, \implies, \iff, \text{"and" } \wedge (\&), \text{"or"} \vee, \emptyset ...)$

(ii) There are very precise rules for what constitutes a "well-formed" (generalised) statement, e.g. all brackets have to close

(iii) There are precise rules governing the use of $\implies$

Example 1.6. Prove: for $n \in \mathbb{N}$ n^2 is odd $\implies n$ is odd.

In the proof the following will be taken for granted.

- n even iff n is divisible by 2 i.e. $n = 2k$ for some $k \in \mathbb{N}$
- n is off iff n is not even
- n is odd iff $n = 2k + 1$ for some $k \in \mathbb{N}$

First a wrong proof:

Suppose that n is odd.
Then $n = 2k + 1$ (some k)
Then $n^2 = (2k + 1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1$ which is odd.   ■

Sometimes it's harder to prove $\bar{q} \implies \bar{p}$ than directly - I think it's more psychological; some people are a bit more oppositional, they are angry with the world, and they start each proof by contradiction! I think psychologically it's much better to prove things directly. But in this case $\bar{q} \implies \bar{p}$ is much easier:

Now a correct proof...

The statement I want to prove is equivalent to: n is even $\implies n^2$ is even

Proof. Suppose n is even, then $n = 2k$ (for some $k \in \mathbb{N}$)
Hence $n^2 = (2k)^2 = 4k^2 = 2(2k^2)$ is also even.   ■

Note: ■ means I just finished the proof[1], same as QED, //, □, ■ etc.

Definition. Suppose that for all $n \in \mathbb{N}$ you are given a set A_n.

Then we define the infinite unions and intersections:

$$\bigcup_{n=1}^{\infty} A_n = \{x \mid x \in A_n \text{ for some } n \in \mathbb{N}\}$$

---

[1] I prefer using the black rectangle for all my notes; squares frighten me.

$$\bigcap_{n=1}^{\infty} A_n = \{x \mid x \in A_n \text{ for all } n \in \mathbb{N}\}$$

The Open and Closed Intervals:

$$[a, b] = \{t \in \mathbb{R} \mid a \leq t \leq b\}$$

$$(a, b) = \{t \in \mathbb{R} \mid a < t < b\}$$

$$[a, b) = \{t \in \mathbb{R} \mid a \leq t < b\}$$

etc. with $a \leq b$ with $a, b \in \mathbb{R}$.

N.B. if $a, b = \pm\infty$, we must always use an open interval for the $\pm\infty$ since $\infty \notin \mathbb{R}$.

The Archimedean Axiom

Axiom 1.7 (Archimedean)**.** Given any $A \in \mathbb{R}$, $\exists m \in \mathbb{N}$ such that $m \geq A$.
$[\forall A \in \mathbb{R} \ \exists n \in \mathbb{N} \mid n \geq A]$

This statement is equivalent to: $\forall A \in \mathbb{R}$, $\exists m \in \mathbb{N} \mid m > A$.

The "grammar" that rules how these sentences are put together is very strict. In particular the *order* of the "words" is extremely important. For instance: $\exists n \in \mathbb{N} : \forall A \in \mathbb{R} \ n > A$ still makes "sense". But it is *not* the Archimedean axiom. It is an *utterly false* statement.

Example 1.8. Prove that:

(a) $\bigcup_{n=1}^{\infty} [0, 1 - \frac{1}{m}) = [0, 1)$

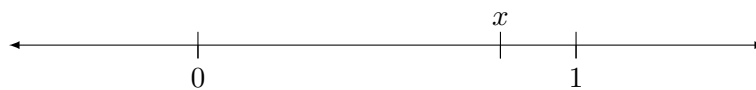(b) $\bigcap_{n=1}^{\infty} (1 - \frac{1}{m}, 1 + \frac{1}{m}) = \{1\}$

Part (a)

Picture & Idea: If n is very large, then $[0, 1 - \frac{1}{m}]$ is almost the interval $[0, 1)$

Proof. Note this $\subset$ is obvious. If $x \in \cup_{n=1}^{\infty} [0, 1 - \frac{1}{m})$ then $\exists m \in \mathbb{N}$ such that $x \in [0, 1 - \frac{1}{m})$. But $[0, 1 - \frac{1}{m}) \subset [0, 1)$ so $x \in [0, 1)$.

We need to show $\supset$ now to complete the proof. Suppose $x \in [0, 1)$:

Then $d = 1 - x$, we then show that we can find an n such that $\frac{1}{n} < d$ (i.e. an n "large enough" so that $n > \frac{1}{d}$, since that implies $x \in [0, 1 - \frac{1}{n})$).

To show that if $0 \le x < 1$, then $\exists n$ such that $0 \le x < 1 - \frac{1}{n}$. We choose n such that $n > \frac{1}{1-x}$ [invoking the Archimedean axiom]. Then

$$1 - x > 0 \; : \; (1-x)n > 1$$
$$\implies 1 - x > \frac{1}{n}$$
$$\implies 1 - \frac{1}{n} > x$$

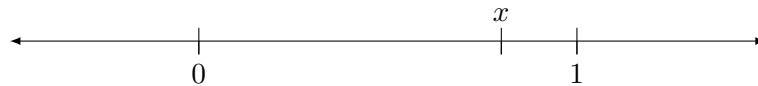which is exactly what I had to show for (a). ∎

Part (b): $\bigcap_{n=1}^{\infty}(1 - \frac{1}{m}, 1 + \frac{1}{m}) = \{1\}$

Proof. It is obvious that $1 \in \bigcap_{n=1}^{\infty}(1 - \frac{1}{m}, 1 + \frac{1}{m})$, since $1 - \frac{1}{n} < x < 1 + \frac{1}{n} \quad \forall n \in \mathbb{N}$. Hence RHS $\subset$ LHS.

To complete the proof we thus need to show that LHS $\subset$ RHS:

$$\{x \; : \; 1 - \frac{1}{n} < x < 1 + \frac{1}{n} \; \forall n \in \mathbb{N}\} \subset \{1\}$$

Picture:

It seems blindingly obivous that $(x <, 1 + \frac{1}{n} \quad \forall n) \implies (x \le 1) \quad (*)$

I prove instead the equivalent to statement $(*)$: $x > 1 \implies \exists n$ such that $x \ge 1 + \frac{1}{n}$

Choose n such that $n > \frac{1}{x-1}$ (invoking the Archimedean axiom) $x - 1 > 0$, so

$$n(x-1) > 1 \implies x - 1 > \frac{1}{n}$$
$$\implies x > 1 + \frac{1}{n}$$

There is a similar claim whose proof I gave to you for $(x > 1 - \frac{1}{n} \; \forall n) \implies (x \ge 1)$.

Since both LHS $\subseteq$ RHS, and RHS $\subseteq$ LHS, $\bigcap_{n=1}^{\infty}(1 - \frac{1}{m}, 1 + \frac{1}{m}) = \{1\}$ ∎

Proposition 1.9. - *Let A, B, C, Ω be sets:*   Lecture 3

(i) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

(ii) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

(iii) *If $A, B, \subset \Omega$, Define $A^C = \Omega \backslash A$ (the complement of A). Then:*

$$A^C \cap B^C = (A \cup B)^C$$
$$(A \cap B)^C = A^C \cup B^C$$

Proof. (i) First let us show $\subset$:

Let $x \in A \cap (B \cup C)$. $x \in A$ and $x \in B \cup C$, so

Either $x \in A$ and $x \in B \implies x \in A \cap B$

Or $x \in A$ and $x \in C \implies x \in A \cap C$.

So

Either $x \in A \cap B$

Or $x \in A \cap C$

that is, $x \in (A \cap B) \cup (A \cap C)$. This shows "$\subset$".

Second, let us show "$\supset$":

Let $x \in (A \cap B) \cup (A \cap C)$, then

Either $x \in A \cap B \implies x \in A$ and $x \in B \implies x \in A$ and $x \in B \cup C$.

Or $x \in A \cap C \implies x \in A$ and $x \in C \implies x \in A$ and $x \in B \cup C$.
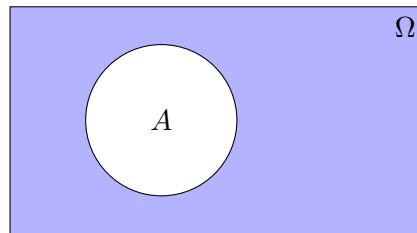
So in both cases $x \in A \cap (B \cup C)$. The two inclusions together imply (i). ∎

Proof. (ii). I could do this in a similar way to (i). Instead, I show that $(i) \implies (ii)$ purely formally:
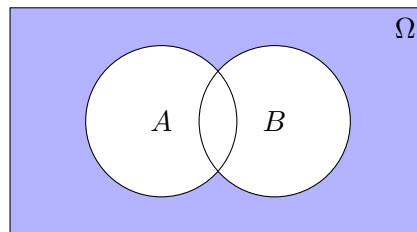
$$
\begin{aligned}
(A \cup B) \cap (A \cup C) &= ((A \cup B) \cap A) \cup ((A \cup B) \cap C) \\
&= ((A \cap A) \cup (B \cap A)) \cup ((A \cap C) \cup (B \cap C)) \\
&= [A \cup (B \cap A) \cup (C \cap A)] \cup (B \cap C) \\
&= A \cup (B \cap C) \quad \blacksquare
\end{aligned}
$$

Now for (iii), sketch the general pictures:

Picture of A^C:

Picture of $(A \cup B)^C$:

Can you "see" that $(A \cup B)^C = A^C \cap B^C$?

Proof. (iii) We show "$\subset$":

Suppose $x \in A^C \cap B^C$. Then $x \notin A$ and $x \notin B \implies x \notin A \cup B$.

Let's show $\supset$:

Suppose $x \notin A \cup B$, then $x \notin A$ and $x \notin B$. So $x \in A^C$ and $x \in B^C$. So $x \in A^C \cap B^C$. ∎

Exercise: Similarly prove on your own that $(A \cap B)^C = A^C \cup B^C$.

Mathematical Induction

Suppose you want to show

$$P_n : 1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}$$

(P_n is a proposition depending on $n \in \mathbb{N}$)

Of course P_0 is true. Indeed P_0 just says $0 = 0$. Suppose that by some way or another somebody hands you the proof of P_k for some particular k. Then you should be happy:

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

so we have shown $P_k \implies P_{k+1}$.

> **Definition** (Principle of Mathematical Induction)**.** Suppose given infinitely many propositions P_n, one for each $n \in \mathbb{N}$. Suppose
>
> (i) P_0 is true
>
> (ii) $\forall k \in \mathbb{N}, (P_k \implies P_{k+1})$ is true
>
> Then P_n is true for all $n \in \mathbb{N}$.

Question: Can the principle of mathematical induction be considered a "mathematical statement" i.e. can we prove it. Yes!

> **Axiom 1.10** (Smallest Element Axiom)**.** Suppose $\phi \neq S \subset \mathbb{N}$. Then S has a smallest element.

Proposition 1.11. *Smallest element axiom $\implies$ mathematical induction.*

Proof. By contradiction. Let $S = \{n \in \mathbb{N} \mid \overline{P_N}\}$. Assume by contradiction $S \neq \phi$. Then S has a smallest element $k \in S$. There are two cases:

Either $k - 1 \in S$. Then k was not the smallest element: a contradiction.

Or $k - 1 \notin S$. This means P_{k-1}. But $P_{k-1} \implies P_k$. So P_k. But then $k \notin S$, also a contradiction.

<div align="right">∎</div>

We will see the converse: mathematical induction $\implies$ smallest element axiom.

> **Example 1.12.** Define the Fibonacci Sequence inductively as
>
> - $F_0 = 0$ and $F_1 = 1$
> - $F_n = F_{n-1} + F_{n-2} \quad (n \geq 2)$

We prove by induction that:

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

In the proof. I want to do something like this:

$$F_{n+1} = F_n + F_{n-1}$$

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

Proof. Let

$$Q_n : \left(F_n = \frac{1}{\sqrt{5}} \left[\left(\frac{1 + \sqrt{5}}{2} \right)^n - \left(\frac{1 - \sqrt{5}}{2} \right)^n \right] \right)$$

I apply induction to $P_n : (Q_n \ \& \ Q_{n-1})$.

At this point we need some algebra:

$$F_{n+1} = \frac{1}{\sqrt{5}} \left(\frac{1 + \sqrt{5}}{2} \right)^{n-1} \left[1 + \frac{1 + \sqrt{5}}{2} \right] - \frac{1}{\sqrt{5}} \left(\frac{1 - \sqrt{5}}{2} \right)^{n-1} \left[1 + \frac{1 - \sqrt{5}}{2} \right]$$

Focus on :

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

$$\implies \left(\frac{1 + \sqrt{5}}{2} \right)^2 = \frac{6 + 2\sqrt{5}}{4}$$

and similarly

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

So indeed

$$F_{n+1} = \frac{1}{\sqrt{5}} \left[\left(\frac{1 + \sqrt{5}}{2} \right)^{n+1} - \left(\frac{1 - \sqrt{5}}{2} \right)^{n+1} \right]$$

So what did we do? We proved $(Q_k \ \& \ Q_{k-1}) \implies Q_{k+1}$. But then obviously $(Q_k \ \& \ Q_{k-1}) \implies (Q_{k+1} \ \& \ Q_k)$. i.e. $P_k \implies P_{k+1}$. P_1 is obviously true, so by induction P_n and also Q_n is true $\forall n$. ∎

Warning. Look at how you set up your proof; you're using *both* Q_n and Q_{n-1} to prove Q_{n+1}, but this is NOT mathematical induction which is only $P_k \implies P_{k+1}$. You have to set up P_k as well. I say this, as usually about 50 or 100 have to resit M1F in September - that's just a fact of life. Last summer the exam asked you to prove induction for this formula and a lot of people (I'm saying $> 50\%$) didn't tell me what P_n is. Do you know how many marks they got, for pages and pages of algebra? Zero.

Definition (Strong principle of induction). Suppose Q_n ($n \in \mathbb{N}$) and assume

   (i) Q_0 is true, and

   (ii) $\forall n \geq 1, (Q_k \text{ for } k < n) \implies Q_n$

then Q_n is true for all n.

Proof. Apply usual induction with $P_n = (Q_n, k \leq n)$. ∎

As an application I show the existence of prime factorisation (uniqueness is much harder and I will treat it later in the course).

Definition. For $n \in N$, $n \neq 0, 1$, we say n is *reducible* if $\exists a, b \in \mathbb{N}$ such that $n = a \cdot b$, $a \neq 1$ and $b \neq 1$. n is *irreducible* (sometimes also called *prime*) if n is not reducible.

E.g. $p = 2, 3, 5, 7, 11, 13, 17, 19, \dots$ are all irreducible.

Theorem 1.13

Every $n \in \mathbb{N}, n \neq 0, 1$ is the product of irreducibles.

Proof. Uses the strong principle of induction with

$$Q_n : n = 0 \text{ or } n = 1 \text{ or } n \neq 0, 1 \text{ is the product of irreducibles.}$$

Q_0, Q_1 are true. Assume now that $n \neq 0, 1$. I prove $(Q_k, k < n) \implies Q_n$.

Either: n is irreducible, in this case obviously we're OK

Or: $n = a \cdot b$ is reducible, with $a \neq 1, b \neq 1$.. But then $a < n, b < n$, so by our inductive assumption Q_a and Q_b are both true: a, b are both products of irreducibles. hence

$$n = \underbrace{\text{product of irreducibles}}_{a} \times \underbrace{\text{product of irreducibles}}_{b}$$

is also the product of irreducibles. ∎

Next I complete a circle of ideas:

$$\text{Usual induction}$$

$$\text{Strong induction} \longleftrightarrow \text{Smallest element axiom}$$
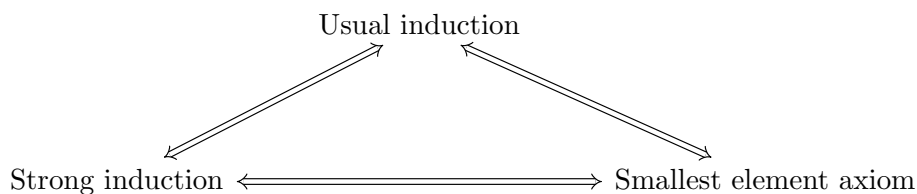
Proposition 1.14. *Strong induction* $\implies$ *smallest element axiom.*

Proof. Let $\phi \neq S \subset \mathbb{N}$. We want to show that S has a smallest element. I use strong induction with

$$Q_n : (\text{If } n \in S \text{ then } S \text{ has a smallest element})$$

If Q_n is true for all n, then I'm done since $S \neq \phi$. Then $\exists n \in S$ and the true of Q_n will imply that S has a smallest element.

(i) Q_0 is clearly true: indeed if $0 \in S$ then 0 is the smallest element of S because it is the smallest element of $\mathbb{N}$

(ii) I need to show that $(Q_k$ for $k < n) \implies Q_n$.

   If n is the smallest element of S then I'm done. Otherwise n is not the smallest element of S. This means $\exists k \in S$, but then Q_k tells me that S has a smallest element. By strong induction, all Q_n are true and then I'm done with this proof.

■

Q: Why have you included 0 in $\mathbb{N}$?

Corti: It's a convention, some people do, some don't. I'm happy to include zero. The perspective I take is somewhat higher - there are many mathematical structures, like a semi-group where you can add things and there is a neutral element, if you don't have a neutral element things becomes shit. Of course it was a big deal in human history when humans decided that 0 is a number. The place in history where this first came up was in those Indian Sanskrit poems, the Bhagavad Gita. Inside there, in Sanskrit texts of the 8th Century, people studied what we would call the Pell equation, and it included a cyclic method to solve the equation - if you allowed 0 as a number it just becomes much clearer. I completely don't care about applications of maths like accounting and bullshit like that. 0 helps with mathematical research.

Contrapositive

How do I form the negation of a mathematical statement? Given P, how to form $\overline{P}$ (non-P)?

Rule No. 1:
$$(\forall x \in A, Q(x)) = P$$
$$\implies (\exists x \in A, \overline{Q(x)}) = \overline{P}$$

Rule No. 2:
$$(\exists x \in A, Q(x)) = P$$
$$\implies (\forall x \in A, \overline{Q(x)}) = \overline{P}$$

Remark 1.15. An element $a \in A$ such that $\overline{Q(a)}$ is called a *counter example* to the statement $(\forall x \in A, Q(x)) = P$. Indeed the very existence of this "example" $a \in A$ shows that P is false.

Example 1.16. A typical exam question: True or false? Find a proof or a counterexample:

$$n \in \mathbb{N} \text{ irreducible} \implies \exists a, b \in \mathbb{N} \text{ with } n = a^2 + b^2$$

A: It is false, that is, the negation:

$$\exists n \in \mathbb{N} \text{ irreducible such that } \forall a, b \in \mathbb{N}, n \neq a^2 + b^2$$

is true.

Proof. We find the contrapositive applying our rules:

$$\forall n \in \{n \in \mathbb{N} \text{ irreducible}\} \, (\exists (a,b) \in \mathbb{N} \times \mathbb{N}(n = a^2 + b^2))$$

Rule No. 1:

$$\exists n \in \{n \in \mathbb{N} \text{ irreducible}\} \, \overline{(\exists (a,b) \in \mathbb{N} \times \mathbb{N}(n = a^2 + b^2))}$$

$$= \exists n \in \{n \in \mathbb{N} \text{ irreducible}\} \, (\forall (a,b) \in \mathbb{N} \times \mathbb{N}\overline{(n \neq a^2 + b^2)})$$

Rule No. 2 (obvious rule):

$$= \exists n \in \{n \in \mathbb{N} \text{ irreducible}\} \, (\forall (a,b) \in \mathbb{N} \times \mathbb{N}(n \neq a^2 + b^2))$$

Indeed take $n = 3$. Then $\forall a, b$, $3 \neq a^2 + b^2$. (Exercise: Prove this!) $\blacksquare$

2 Number Systems

> It really seems almost obscene to go to such a higher layer of pure thought to discuss something as basic as this; but you're mathematicians, you probably look like real weirdos to all your friends, so you can rise to this.
>
> *- Alessio Corti*

Today I take $\mathbb{N}, \mathbb{Z}$ for granted and discuss in some depth $\mathbb{Q}$ (and, a bit, $\mathbb{R}$).

$\mathbb{Q}$, the Rationals

Basically, elements of $\mathbb{Q}$ (the rational numbers) are "things" that you write in the form: $x = m/n$, where $m \in \mathbb{Z}$ and $m \in \mathbb{N}, n \neq 0$. There is a catch, and it is that x cannot be uniquely written in this form: $2/3" = "4/6" = "6/9$. This is not a *major* problem: $m = cm'$, $n = cn'$ with $c = \mathrm{hcf}(m, n)$ then the expression for x: $x = m'/n'$ is unique. But it is a small problem nonetheless.

> **Definition.** The *cartesian product* of two sets, A, B is
> $$A \times B = \{(a, b) \mid a \in A, b \in B\}$$

Consider the Cartesian product $\mathbb{Z} \times (\mathbb{Z} \backslash \{0\})$. An element is an ordered pair (m, n) where $m \in \mathbb{Z}$, $n \in \mathbb{Z} \backslash \{0\}$. (think of (m, n) as a "pedantic" way to start thinking about the fraction m/n.)

> **Definition.** (m, n) is equivalent to (M, N), written as $(m, n) \sim (M, N)$. If $\exists a, b \in \mathbb{Z} \backslash \{0\}$ such that $am = bM$ and $am = bN$.

Properties of equivalence:

  (i) $x \sim x$

  (ii) $x \sim y \implies y \sim x$

  (iii) $(x \sim y)$ and $(y \sim z) \implies x \sim z$

Exercise: Prove this

> **Definition.** If $(m, n) \in \mathbb{Z} \times (\mathbb{Z} \backslash \{0\})$ then the *equivalence class* of (m, n) is the set
> $$[m, n] = \{(M, N) \mid (m, n) \sim (M, N)\}$$

Alternative notation: $[m, n] = m/n$

> **Definition.**
> $$\mathbb{Q} = \{[m, n] \mid (m, n) \in \mathbb{Z} \times (\mathbb{Z} \backslash \{0\})\}$$

Elements $x \in \mathbb{Q}$ are written as "equivalence classes" m/n and now it makes perfect sense to write $2/3 = 4/6 = 6/9$.

Binary operations on $\mathbb{Q}$

Suppose (p, q) and $(m, n) \in \mathbb{Z} \times \mathbb{Z} \backslash \{0\}$. Then

$$(p, q) \widetilde{+} (m, n) = (pn + qm, qn + pm)$$
$$(p, q) \widetilde{\times} (m, n) = (pm, qn)$$

Suppose $(p, q) \sim (P, Q)$ and $(m, n) \sim (M, N)$. Then

$$(p, q) \widetilde{+} (m, n) \sim (P, Q) \widetilde{+} (M, N)$$
$$(p, 1) \widetilde{\times} (m, n) \sim (P, Q) \widetilde{\times} (M, N)$$

This shows that I can use $\widetilde{+}, \widetilde{\times}$ to define operations of equivalence classes:

$$p/q + m/n = pn + qm/qn$$
$$p/q \times m/n = pm/qn$$

Axiom 2.1 (Axioms of $K = \mathbb{Q}, K = \mathbb{R}$)**.**

 (i) $a + b = b + a \quad \forall a, b \in K$ ($+$ is *commutative*)

 (ii) $a \times b = b \times a \quad \forall a, b \in K$

 (iii) $a + (b + c) = (a + b) + c$ ($+$ is *associative*)

 (iv) $a \times (b \times c) = (a \times b) \times c$

 (v) $a \times (b + c) = (a \times b) + (a \times c)$ ($\times$ is *distributive* over $+$)

 (vi) $\exists 0 \in K : a + 0 = a \quad \forall a \in K$

 (vii) $\exists 1 \in K, 0 \neq 1 : a \cdot 1 = a \quad \forall a \in K$

 (viii) $\forall a \in K, \exists (-a) \in K$ such that $a + (-a) = 0$

 (ix) $\forall a \in K \backslash \{0\} (= K^\times), \exists (a^{-1})$ such that $a \times (a^{-1}) = 1$.

$\mathbb{R}$ satisfies an additional axiom

 (x) Completeness axiom (I will talk about this later in the course)

Notation: $a + (-b) = a - b$, and $a \times (b^{-1}) = a/b$.

Axiom 2.2 (Trichotomy axiom)**.** The binary relations $<, =, >$ are such that

$$\forall x \in K \begin{cases} \text{either} & x > 0 \\ \text{or} & x = 0 \\ \text{or} & x < 0 \end{cases}$$

and only one of these conditions holds.

We then have:

$$x > 0 \iff -x < 0 \quad \forall x \in K$$
$$x > 0 \implies a + x > a \quad \forall a \in K$$
$$x > 0 \implies a \times x > 0 \quad \forall a > 0$$

These axioms are not on the exam - it's just an intellectual exercise. We can deduce anything about $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ from just these axioms. (BUT it can be awkward sometimes).

Examples 2.3.

(1) $0 \times x = 0 \quad \forall x \in K$

Proof. First note that:
$$(0 + 1) \cdot x = 1 \cdot x = x$$
$$0 \cdot x + 1 \cdot x = 0 \cdot x + x$$

So $x = 0 \cdot x + x \quad \forall x$

$$(-x) + x = (-x) + (0 \cdot x + c)$$
$$= (-x) + (x + 0 \cdot x) \qquad \text{[Axiom (i)]}$$
$$= (-x + x) + 0 \cdot x \qquad \text{[Axiom (iii)]}$$
$$= 0 + 0 \cdot x \qquad \text{[Axiom (viii)]}$$
$$= 0 \cdot x + 0 \qquad \text{[Axiom (i)]}$$
$$= 0 \cdot x \qquad \text{[Axiom (vi)]}$$

So $0 = 0 \cdot x$. ∎

(2) $\forall x \in K = \mathbb{Q}$ or $\mathbb{R}$, we have either $x = 0$ or $x^2 > 0$ (but never both)

Proof. There are 3 cases:

  (i) $x = 0$ - nothing to do

  (ii) $x > 0 \implies x \cdot x > 0$ (not much to do)

  (iii) $x < 0 \implies -x > 0 \implies (-x) \cdot (-x) > 0$. All I need is to show that $(-x)(-x) = x \cdot x$

First I claim that $-x = (-1) \times x$ for all x. Indeed:

$$0 = 0 \cdot x = (1 + (-1))x = x + (-1) \cdot x$$
$$-x + 0 = -x = -x + x + (-1)$$
$$= 0 + (-1) \cdot x$$
$$= (-1) \cdot x$$

Secondly, I claim $(-1) \cdot (-1) = 1$ (Exercise!), then I can finish (2):

$$(-x) \cdot (-x) = (-1) \cdot x \cdot (-1) \cdot x$$
$$= (-1) \cdot (-1) \cdot x \cdot x$$
$$= 1 \cdot x \cdot x = x \cdot x$$

∎

Example 2.4. Show that $\sqrt{6} - \sqrt{2} > 1$

Remark: If $A, B \in \mathbb{R}$, & $A, B \geq 0$, then $A > B \iff A^2 > B^2$
(*Exercise:* show this from the axioms)

Write $\sqrt{6} - \sqrt{2} > 0$; indeed $\sqrt{6} > \sqrt{2} \iff (\sqrt{6})^2 > (\sqrt{2})^2 \iff 6 > 2$

So
$$
\begin{aligned}
&\sqrt{6} - \sqrt{2} > 1 \\
&\iff (\sqrt{6} - \sqrt{2})^2 > 1^2 \\
&\iff 6 + 2 - 4\sqrt{3} > 1 \\
&\iff 7 > 4\sqrt{3} \\
&\iff (\frac{7}{4})^2 > (\sqrt{3})^2 \\
&\iff \frac{49}{16} > 3 \\
&\iff 49 > 48 \\
&\iff 1 > 0
\end{aligned}
$$

which follows from the axioms.

Example 2.5. Is $\sqrt{2} + \sqrt{3} > \sqrt{10}$?

$$
\begin{aligned}
&\sqrt{2} + \sqrt{3} > \sqrt{10} \\
&\iff (\sqrt{2} + \sqrt{3})^2 > (\sqrt{10})^2 \\
&\iff 5 + 2\sqrt{6} > 10 \\
&\iff \sqrt{6} > \frac{5}{2} \\
&\iff 6 > \frac{25}{4} \\
&\iff 24 > 25 \\
&\iff 0 > 1, \text{ which is false. So the answer is no.}
\end{aligned}
$$

Example 2.6. Solve $x > \dfrac{2}{x+1}$ for $x \in \mathbb{R}$ [$x + 1 = 0 \implies$ RHS makes no sense]

Cases:

   (i) $x + 1 > 0$

  (ii) $x + 1 < 0$

Starting with (i):

$$x(x+1) > 2 \implies x^2 - x - 2 > 0 \implies (x-1)(x+2) > 0$$

Recall $x > -1$, so case (i) $\implies x \in (1, \infty) = \{x \mid 1 < x\}$

Next (ii): $x(x+1) < 2 \implies x^2 - x - 2 < 0 \implies (x-1)(x+2) < 0$

Then $x \in (-2, 1) \cap (-\infty, -1) = (-2, -1)$

So our final answer is: $\boxed{x \in (-2, -1) \cup (1, \infty)}$

> **Definition** (Absolute Value Function)**.**
>
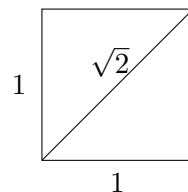> $$|x| = \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0 \end{cases}$$
>
> (so $|x| \geq 0 \quad \forall x \in K$)

Exercise: Solve for $x \in \mathbb{R}$: $|x+4| < |x-5|$, do it properly!
(hint: $|x| < A \iff -A < x < A$)

$\mathbb{R}$, the Real Numbers

Lecture 9   A remarkable thing: $\mathbb{Q} \neq \mathbb{R}$. For example: $\sqrt{2} \notin \mathbb{Q}$.

Do we know for sure that $\sqrt{2}$ exists? i.e. $\exists x \in \mathbb{R}$ such that $x \geq 0$ and $x^2 = 2$. We will answer this analyst-style later in the course. The geometer-style makes this obvious:

I can prove: $x \in \mathbb{Q} \implies x^2 \neq 2$.

Proof. Define $x = p/q$ with $p \in \mathbb{Z}$ and $q \in \mathbb{Z}$ are both are not even.

Suppose for contradiction that $x^2 = 2$. Then $p^2 = 2q^2$. p must be even, so $p = 2p'$ and $4p'^2 = 2q^2$. Then $q^2 = 2p'^2$. So q is also even: a contradiction. ∎

Exercise: $\sqrt{3} \notin \mathbb{Q}$ (3-adic argument using divisibility by 3)

Claim: $\sqrt{2} + \sqrt{3} \notin \mathbb{Q}$.

Proof. Assume for contradiction $r = \sqrt{2} + \sqrt{3} \in \mathbb{Q}$. Then

$$r^2 = (\sqrt{2} + \sqrt{3})^2 = 5 + 2\sqrt{6} \in \mathbb{Q}$$
$$\implies \sqrt{6} = \frac{r^2 - 5}{2} \in \mathbb{Q}$$

but $\sqrt{6} \notin \mathbb{Q}$: If $\sqrt{6} = p/q$, we may assume p, q are not both even, then $p^2 = 6q^2$, so $2 \mid p$ and $4p'^2 = 6q^2$. Then $3q^2 = 2p'^2 \implies q$ is also even, a contradiction. ∎

Exercise: Let $x \in \mathbb{R}$. Prove there is a unique $N \in \mathbb{Z}$ such that $N \leq x < N + 1$. N is called the *round down* of x and is denotes $N : \lfloor x \rfloor$

[hint: assume $x \geq 0$, then consider $S = \{n \in \mathbb{N} \mid n > x\} \subset \mathbb{N}$. Note by the Archimedean axiom S has a smallest element. Claim: $\lfloor x \rfloor = -1+$ smallest element (S).]

Decimal Expansions of Reals

You are all familiar with the

$$\frac{1}{9} = 0.1111\cdots = 0.\overline{1}$$

How can we really make sense of this?

$$
\begin{aligned}
\frac{1}{9} &= 0.1 + 0.01 + 0.001 + 0.0001 + \ldots \\
&= \frac{1}{10} + \frac{1}{100} + \frac{1}{1000} + \frac{1}{10000} + \ldots \\
&= ? \\
&= \frac{1}{10}\left(1 + \frac{1}{10} + \left(\frac{1}{10}\right)^2 + \left(\frac{1}{10}\right)^3 + \ldots\right) \\
&= \frac{1}{10}\cdot\frac{1}{1-\frac{1}{10}} = \frac{1}{10}\frac{10}{9} = \frac{1}{9}
\end{aligned}
$$

Remark 2.7.

$$1 + x + x^2 + \cdots + x^n = \frac{1 - x^{n+1}}{1 - x}$$

If $-1 < x < 1$ then

$$
\begin{aligned}
1 + x + x^2 + \cdots + x^n + \ldots &= \lim_{n\to\infty}\frac{1 - x^{n+1}}{1 - x} \\
&= \frac{1}{1 - x}
\end{aligned}
$$

For now let's agree that all of this makes sense.

Theorem 2.8

$\forall x \in \mathbb{R}, \ \exists! \, n \in \mathbb{Z} : n \le x < n + 1$. We write $n = \lfloor x \rfloor$, the *integer part* of x. The *fractional part* of x is $\{x\} = x - \lfloor x \rfloor$, where $0 \le \{x\} < 1$.

Lecture 10

Corollary 2.9 (Euclid's Algorithm)**.** *Let $p \in \mathbb{Z}$, $q \in \mathbb{N}\backslash\{0\}$. Then $\exists a \in \mathbb{Z}$, $r \in \mathbb{N}$ with $p = qa + r$ and $0 \le r < q$.*

Proof. Take $a = \lfloor\frac{p}{q}\rfloor$. Then $\frac{p}{q} = a + f$ where $f = \{\frac{p}{q}\}$. Let $r = qf$. Note that $r \in \mathbb{Z}$; indeed $r = qf = p - qa \in \mathbb{Z}$. Also $0 \le f < 1 \implies 0 \le r = qf < q$. ∎

We will come back to this later in the course. Today I use this to understand decimal expansions. We consider decimal expansions:

$$x = a_0.a_1a_2a_3\ldots \qquad (*)$$

where $a_0 \in \mathbb{Z}$ and $a_1, a_2, a_3, \ldots$ are digits $0 \le a_i \le 9$ for all i.

I take $(*)$ to mean:

$$x = a_0 + \frac{a_1}{10} + \frac{a_2}{10^2} + \cdots + \frac{a_n}{10^n} + \ldots$$

> **Theorem 2.10**
>
> Every $x \in \mathbb{R}$ has a decimal expansion.

Proof. Let $x \in \mathbb{R}$. We can construct the decimal expansion of x:

$\exists M \in \mathbb{Z} : M \leq x < M + 1$ (M is the *round down* of x, denoted $\lfloor x \rfloor$).

We start with $a_0 = M = \lfloor x \rfloor$ and write:

$$[a_0, a_0 + 1) = [a_0, a_1 + \frac{1}{10}) \cup [a_0 + \frac{1}{10}, a_0 + \frac{2}{10}) \cup \cdots \cup [a_0 + \frac{9}{10}, a_0 + 1)$$

Clearly there is a unique $a_1 \in \{0, 1, \ldots, 9\}$ such that $x \in [a_0, \frac{a_1}{10} + a_0 + \frac{a_1+1}{10})$. This constructs a_1. Continue to do this for a_2 etc. We have a procedure that inductively determines all of the decimal expansion of x. ∎

Example 2.11. $2/7 = 0.\overline{285714}$, using long division.

In summary let $x \in \mathbb{R}$. Then $x = a.a_1 a_2 a_3 a_4 \ldots$ with $a, a_k \in \mathbb{Z}$ and $0 \leq a_k \leq 9$, where $a = \lfloor x \rfloor$ and $a_k (k \geq 1)$ are defined inductively, starting from $f_0 = \{x\}$ by $a_{k+1} = \lfloor 10 f_k \rfloor$ where $f_k = \{10 f_{k-1}\}$.

> **Theorem 2.12**
>
> Let $x \in \mathbb{R}$. Then $x \in \mathbb{Q} \iff$ the decimal expansion of x is eventually periodic,
> i.e. $x = a.a_1 \ldots a_n \overline{b_1 \ldots b_k}$

Example 2.13. $x = 0.1010010001 \cdots \notin \mathbb{Q}$.

Point to take home: the theorem provides a very large supply of explicit, easily described irrational numbers.

Proof. $\impliedby$: If the decimal expansion is eventually periodic, then $x \in \mathbb{Q}$:

$$x = a.a_1 \ldots a_n + \frac{1}{10^n}(0.\overline{b_1 \ldots b_k})$$

It is enough to show that $y = 0.\overline{b_1 \ldots b_k} \in \mathbb{Q}$:

$$y = b_1 \ldots b_k \left(\frac{1}{10^k} + \frac{1}{10^{2k}} + \frac{1}{10^{3k}} + \cdots \right)$$

$$= \frac{b_1 \ldots b_k}{10^k} \frac{1}{1 - \frac{1}{10^k}} \in \mathbb{Q}.$$

Next I prove $\implies$:

Assume $x = p/q \in \mathbb{Q}$, where $p \in \mathbb{Z}, q \in \mathbb{Z}, q \geq 1$. Then the expansion is eventually periodic:

$x = p/q \in \mathbb{Q}$. Let $a = \lfloor x \rfloor$, $f = \{x\}$, so by the division theorem $p = aq + r$ where $r = qf$.

Inductively we calculate:

$a_1 = \lfloor 10f \rfloor, f_1 = \{10f\}$, so $10r = a_1q + r_1$ where $r_1 = qf_1$.

$a_2 = \lfloor 10f_1 \rfloor, f_2 = \{10f_1\}$, so $10r_1 = a_2q + r_2$ where $r_2 = qf_2$.

$$\vdots$$

$a_k = \lfloor 10f_{k-1} \rfloor, f_k = \{10f_{k-1}\}$, so $10r_{k-1} = a_kq + r_k$ where $r_k = qf_k$.

$$\vdots$$

In other words: write $p = aq + r$, then define $r_k (k \geq 1)$, a_k inductively by: $10r_{k-1} = a_kq + r_k$ $\quad 0 \leq r_k \leq 9$. (the process / construction guarantees that $0 \leq a_k \leq 9$). Because $0 \leq r_k < 9$ for all k, at some point $r_n = r_m$, $n > m$. This means that $x = a.a_1 \ldots \overline{a_m \ldots a_{n-1}}$, that is the decimal expansion is eventually periodic. $\blacksquare$

Non-uniqueness of decimal expansions:

$$0.9999\ldots = 0.\overline{9} = 1$$
$$= \frac{9}{1} + \frac{9}{100} + \ldots$$
$$= \frac{9}{10}\left(1 + \frac{1}{10} + \left(\frac{1}{10}\right)^2 + \cdots + \left(\frac{1}{10}\right)^n + \ldots\right)$$
$$= \frac{9}{1} \cdot \frac{1}{1 - \frac{1}{10}} = \frac{9}{10} \times \frac{10}{9} = 1$$

Proposition 2.14. *Suppose $x \in \mathbb{R}$ has two different decimal expansions. Then these are as follows:*
$$x = a_0.a_1a_2 \ldots a_n\overline{9}$$
$$= a_0.a_1a_2 \ldots (a_n + 1) \quad (0 \leq a_n \leq 8)$$

Proof. Suppose the two expansions are:
$$x = a_0.a_1a_2 \ldots a_na_{n+1} \ldots$$
$$= a_0.a_1a_2 \ldots b_nb_{n+1} \ldots$$

with $a_n < b_n$. Then:
$$x = a_0.a_1a_2 \ldots a_na_{n+1} \ldots$$
$$\leq a_0.a_1a_2 \ldots a_2999\ldots$$
$$= a_0.a_1a_2 \ldots (a_n + 1)000\ldots$$
$$\leq a_0.a_1a_2 \ldots b_nb_{n+1} \ldots$$
$$= x$$

So all $\leq$ must have been $=$ and this shows the proposition. $\blacksquare$

Exercise: $\forall x, y \in \mathbb{R}, x < y$ show that

(i) $\exists z \in \mathbb{Q} : x < z < y$

(ii) $\exists z \notin \mathbb{Q} : x < z < y$

Rational Powers

Later in the year you will see the proof of the following:

> **Theorem 2.15: Existence of n-th root**
>
> $\forall n \in \mathbb{N}\backslash\{0\}, \ \forall x \in \mathbb{R}, \ x \geq 0, \ \exists! \, y \in \mathbb{R}, \ y \geq 0 : y^n = x.$

Notation: $y = \sqrt[n]{x} = x^{\frac{1}{n}}$.

Today we take the theorem for granted and draw some consequences.

> **Definition.** For $r = p/q \geq 0$ and $x \in \mathbb{R}$, $x \geq 0$ we define $x^r = \left(x^{\frac{1}{q}}\right)^p$.

Remark 2.16. x^r is well-defined. Recall $p/q = p'/q'$. Iff $\exists m, n$ such that $mp = np'$ and $mq = nq'$. To show x^r is well defined, it is enough to show that $(x^{\frac{1}{q}})^p = (x^{\frac{1}{mq}})^{mp}$.

We'd like to do $(x^{\frac{1}{q}})^p = x^{\frac{mp}{mq}} = x^{\frac{p}{q}} = (x^{\frac{1}{mq}})^{mp}$, but this is using precisely the thing that we want to prove.

Proof.

$$\left(x^{\frac{1}{q}}\right)^p = \left(\left(x^{\frac{1}{mq}}\right)^m\right)p$$

$$\iff x^{\frac{1}{q}} = \left(x^{\frac{1}{mq}}\right)^m$$

$$\iff x = \left(x^{\frac{1}{mq}}\right)^{mq}$$

and this is true by definition of $x^{\frac{1}{mq}}$ ∎

Properties: Suppose $\alpha, \beta \in \mathbb{Q}_{\geq 0}$, and $x, y \in \mathbb{R}_{\geq 0}$. Then

   (i) $x^\alpha x^\beta = x^{\alpha + \beta}$

  (ii) $x^\alpha y^\alpha = (xy)^\alpha$

 (iii) $(x^\alpha)^\beta = x^{\alpha \beta}$

Proof. (i)

Assume the result for integer powers.

$$x^\alpha x^\beta = \left(x^{\frac{1}{n}}\right)^m \left(x^{\frac{1}{q}}\right)^p$$

$$= \left(x^{\frac{1}{nq}}\right)^{mq} \left(x^{\frac{1}{nq}}\right)^{np}$$

$$= \left(x^{\frac{1}{nq}}\right)^{mq+np}$$

by what we just did. By results on integer powers this is $x^{\frac{mq+np}{nq}} = x^{\alpha+\beta}$. ∎

Exercise: (ii) and (iii).

$\mathbb{C}$, the Complex Numbers

Consider
$$\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} := \{(a,b) : a \in \mathbb{R}, b \in \mathbb{R}\}$$

(the symbol :=, "colon equal" means that the thing on the left is defined to be the thing on the right. Similarly =:)

> **Definition.** We define operations on $\mathbb{R}^2$:
> $$(a,b) + (c,d) := (a + c, b + d)$$
> $$(a,b) \cdot (c,d) := (ac - bd, ad + bc)$$
>
> The two operations satisfy all of the axioms for $K = \mathbb{R}, \mathbb{Q}$ except anything to do with $<, >$. We call $\mathbb{R}^2$ together with $+$ and $\cdot$ as $\mathbb{C}$.
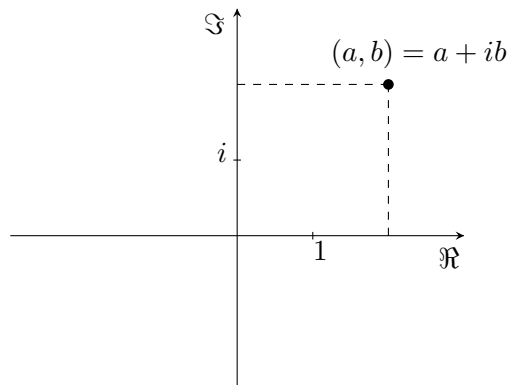
We think of $\mathbb{R}$ as a subset of $\mathbb{C}$:
$$\mathbb{R} \ni x \longmapsto (x, 0) \in \mathbb{C}$$

we write $i = (0, 1)$. Then we can write every element $(a, b) \in \mathbb{C}$ uniquely as
$$\begin{aligned}(a, b) &= (a, 0) + (0, b) \\ &= (a, 0) + (0, b)(0, 1) \\ &= (a, 0) + i(0, b) \\ &= a + ib\end{aligned}$$

and we call this the Cartesian form of the complex numbers (a, b).

You picture the complex numbers on the Cartesian plane:

Note that addition in $\mathbb{C}$ is the addition of vectors in $\mathbb{R}^2$.

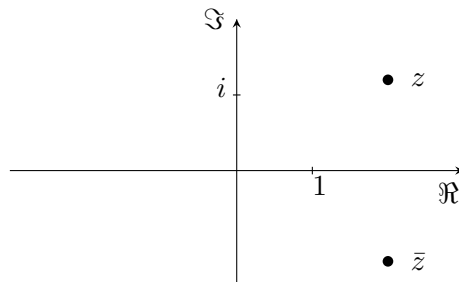Now $i^2 = (0, 1)^2 = (0, 1) \cdot (0, 1) = (-1, 0) = -1$, so using the properties (axioms) we have:
$$\begin{aligned}(a + ib)(c + id) &= ac + ibc + aid + ibid && \text{(distributive property)} \\ &= ac + i^2 bd + ibc + iad && \text{(commutative property)} \\ &= (ac - bd) + i(bd + ad) && \text{(associative property)}\end{aligned}$$

This is familiar and it motivates the definition we started out with.

More terminology and notation

If $z = a + ib \in C$ $(a, b \in \mathbb{R})$, then we denote $\bar{z} := a - ib$ as the *complex conjugate* of z.

Conjugation is reflection in the real axis.

If $z = a + ib \in \mathbb{C}$ $(a, b \in \mathbb{R})$ then:

- $a =$ real part of z, denoted $Re(z)$
- $b =$ imaginary part of z, denoted $Im(z)$

So $z = Re(z) + i\, Im(z)$. Notice that
$$
\begin{cases}
Re(z) & = \dfrac{z + \bar{z}}{2} \\[2mm]
Im(z) & = \dfrac{z - \bar{z}}{2i}
\end{cases}.
$$

We denote $|z| := \sqrt{a^2 + b^2} = \sqrt{z\bar{z}}$ as the *absolute value* of $z \in \mathbb{C}$. (note that if $z = a + ib$, then $z\bar{z} = (a + ib)(a - ib) = a^2 + b^2$.)

Properties of the absolute value of complex numbers

(i) $|z| \geq 0$ and $|z| = 0 \iff z = 0$

(ii) If $\lambda \in \mathbb{R}$, then $|\lambda z| = |\lambda||z|$

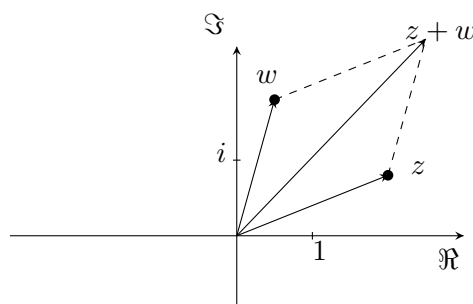(iii) $|z + w| \leq |z| + |w|$ (triangle inequality)

Why is (iii) called the triangle inequality?

First understand $z + w$ in terms of the parallelogram law:

$$
(a + ib) + (a' + ib') = a + a' + i(b + b')
$$

i.e.

$$
\begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} a' \\ b' \end{pmatrix} = \begin{pmatrix} a + a' \\ b + b' \end{pmatrix}
$$

Secondly, we have an interpretation: $|z| = \text{distance}(0, z) = \text{distance}$. The triangle inequality is the obvious fact that the two sides of any triangle (on the Euclidean plane) must be greater than or equal to the other side.
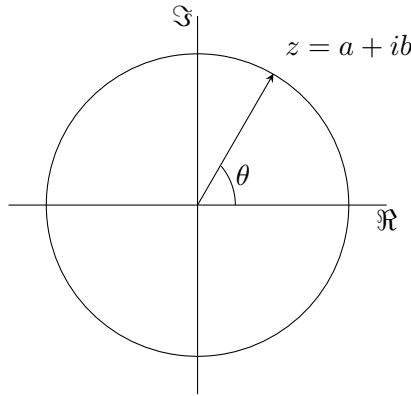
Also note: $|z| \geq 0$ and 0 iff $z = 0$. So if $z \neq 0$, then $1 = \dfrac{z\bar{z}}{|z|^2}$. That is, $z^{-1} = \dfrac{\bar{z}}{|z|^2}$.

> **Example 2.17.** Express $\dfrac{1+i}{1-2i}$ in Cartesian form:
>
> $$\frac{1+i}{1-2i} + \frac{1+i}{1-2i} \times \frac{1+2i}{1+2i}$$
> $$= \frac{(1-2) + (2+1)i}{5}$$
> $$= \frac{-1}{5} + \frac{3i}{5}$$

Polar form of complex numbers

(Motivation: the polar form allows us to neatly visualise complex multiplication much in the same way as the Cartesian form allows us to visualise complex addition.)

z lies on a circle of radius $r = |z|$ and $\theta = $ angle from $\mathbb{R}$ to z.

The polar form of z is in terms of data (r, θ).

> **Definition** (Euler's form)**.** If $\theta \in \mathbb{R}$ then we write
>
> $$e^{i\theta} := \cos\theta + i\sin\theta$$

How can we motivate this?

(i) It just works: $e^{i(\theta_1 + \theta_2)} = e^{i\theta_1} + e^{i\theta_2}$ because

$$e^{i(\theta_1+\theta_2)} := \cos(\theta_1 + \theta_2) + i\sin(\theta_1 + \theta_2)$$
$$= (\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2) + i(\sin\theta_1\cos\theta_2 + \sin\theta_2\cos\theta_1)$$
$$= (\cos\theta_1 + i\sin\theta_1)(\cos\theta_2 + i\sin\theta_2)$$

In otherwords, the trig identities allow us to generalise the exponential function to a function of a complex variable:

$$e^z = e^{x+iy} := e^x(\cos y + i\sin y)$$

while preserving the usual property

$$e^{z_1+z_2} = e^{z_1}e^{z_2}$$

(ii) There are other considerations - you may be familiar with the infinite series for the exponential; if $x \in \mathbb{R}$, then
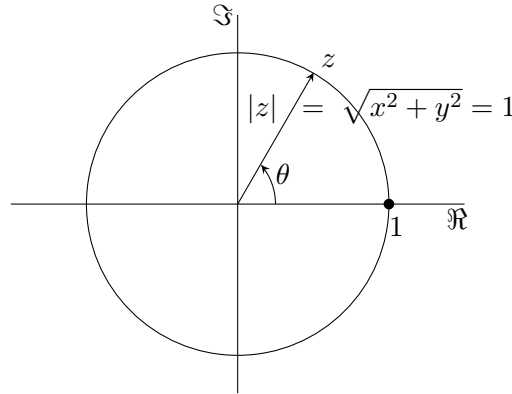
$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

Could it just be that this makes sense for complex numbers?

$$e^{i\theta} = 1 + i\theta - \frac{\theta^2}{2!} - \frac{i\theta^3}{3!} + \frac{\theta^4}{4!} + \dots$$
$$= \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \dots\right) + i\left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \dots\right)$$
$$= \cos\theta + i\sin\theta.$$

I gave some motivation for the formula $e^{i\theta} = \cos\theta + i\sin\theta$. Please now accept this.

Thus, $e^{i\theta} = z$ is the complex no. z with $|z| = \cos^2\theta + \sin^2\theta = 1$ and angle θ from the $\mathbb{R}$-axis to z:

Then every $z = x + iy \in \mathbb{C}$ can be written uniquely if $z \neq 0$ in the polar form $z = re^{i\theta}$ for $r = |z| = \sqrt{x^2 + y^2} \in \mathbb{R}_{\geq 0}$ and θ as an angle $\in (-\pi, \pi]$, where I call r the *modulus* of z and $\theta = arg(z)$, the *argument*.

Conversion formulas:
$$\begin{cases} x & = r\cos\theta \\ y & = r\sin\theta \\ r & = \sqrt{x^2 + y^2} \\ \tan\theta & = \frac{y}{x} \end{cases}$$

Note: our formula for θ does not nail it down uniquely; we need to make sure that $\theta \in (-\pi, \pi]$ is in the correct quadrant.
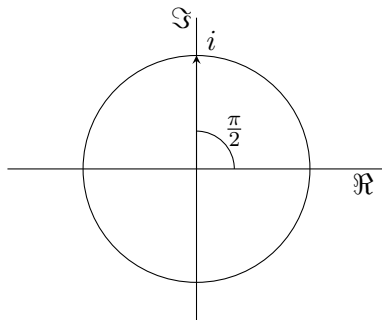

Theorem 2.18: De Moivre's Formula

If $z_1 = r_1 e^{i\theta_1} \in \mathbb{C}$ and $z_2 = r_2 e^{i\theta_2} \in \mathbb{C}$ then
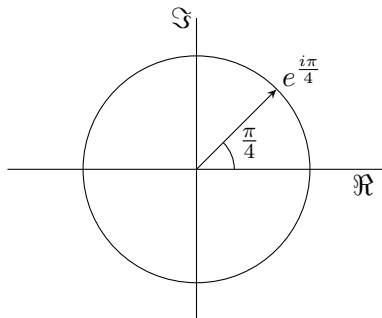
$$z_1 \cdot z_2 = r_1 r_2 e^{i(\theta_1 + \theta_2)}$$

Basically: modulus' multiplies; arguments add.

This allows us to visualise multiplication by $z = r^{i\theta}$: $\mathbb{C} \ni w \longrightarrow z \cdot w \in \mathbb{C}$. This is the linear transformation of $\mathbb{C} = \mathbb{R}^2$ which inflates everything by factor r and rotates everything through angle θ.
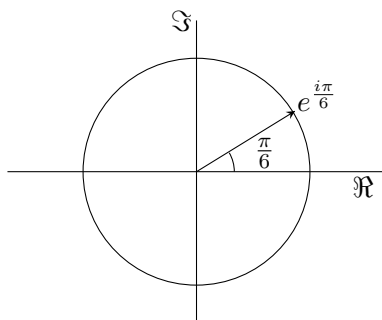
Examples 2.19. Everyone should know these complex numbers:

$$i = e^{\frac{i\pi}{2}}$$

$$\frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2} = e^{\frac{i\pi}{4}}$$

$$\frac{\sqrt{3} + i}{2} = e^{\frac{i\pi}{6}}$$

What's $e^{\frac{i\pi}{12}}$?

Note that

$$\frac{1}{4} - \frac{1}{6} = \frac{3 - 2}{12} = \frac{1}{12}$$

by De Moivre's formula:

$$
\begin{aligned}
&= \frac{\sqrt{2} + i\sqrt{2}}{\sqrt{3} + i} \times \frac{\sqrt{3} - i}{\sqrt{3} + i} \\
&= \frac{(\sqrt{6} + \sqrt{2}) + i(-\sqrt{2} + \sqrt{6})}{4} \\
&= \frac{\sqrt{6} + \sqrt{2}}{4} + i\frac{-\sqrt{2} + \sqrt{6}}{4}
\end{aligned}
$$

Complex roots of 1

Lecture 13   Consider the equation

$$
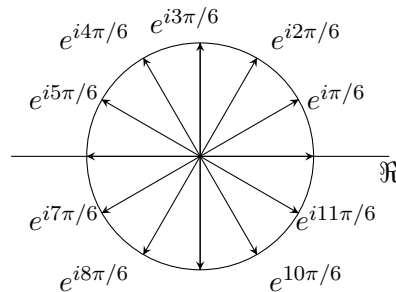z^n = 1 \quad \text{for } z \in \mathbb{C}
$$

Write $z = re^{i\theta}$, then we want $z^n = r^n e^{in\theta} = 1$. That is,

$$
\begin{aligned}
r^n &= 1 \\
\implies r &= 1 \text{ and } n\theta = 2k\pi \text{ for some } k \in \mathbb{Z} \\
\implies \theta &= \frac{2\pi}{n} \cdot R \text{ for some } R \in \mathbb{Z}
\end{aligned}
$$

How many solutions are there?

$r = 1$, so z is on the circle of radius 1.

I draw the picture for $n = 12$:

The solutions to $z^n = 1$ ($z \in \mathbb{C}$) are the complex numbers

$$
z \in \{z = e^{i\frac{2\pi}{n}k} \mid k = 0, 1, 2, \ldots, n - 1\}
$$

They are the vertices of a regular n-gon inscribed in the unit circle with a vertex at 1. Another way to say this: Let $w = e^{i\frac{2\pi}{n}}$. Then the solutions of $z^n = 1$ are $1, w, w^2, w^3, \ldots, w^{n-1}$.

Example 2.20. Solve for z, $z^n = 3$.

We want $z = re^{i\theta}$ such that $z^n = r^n e^{in\theta} = 3 \iff r^n = 3$ and $e^{in\theta} = 1$.

So $z^n = 3 \iff$
$$
z \in \{ \sqrt[n]{3} w^k \mid w = e^{\frac{2\pi i}{n}} \ k = 0, 1, \ldots, n - 1\}
$$

Polynomial equations

> ### Theorem 2.21: Fundamental Theorem of Algebra
>
> Any polynomial with complex coefficients of degree $n \geq 1$:
>
> $$p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0$$
>
> $(a_0, a_1, \ldots, a_n \in \mathbb{C}$, and $a_n \neq 0)$
>
> has a root $\lambda \in \mathbb{C}$.

In other words, there $\exists \lambda \in \mathbb{C}$ such that $p(\lambda) = 0$.

Proof. In the 2nd year Complex Analysis course.

One of the most basic things in maths you learn early on is to solve the quadratic equation. The Fundamental Theorem of Algebra says that any equation of order n has such a result. It is an existential result, it doesn't tell you how to find them.

There is a formula for $n = 3$ (next week), it's very complicated, and strangely useless in a deep kind of way. Then there is a formula for $n = 4$, which I won't tell you as it takes half an hour to write down. For a long time people tried to find formulas for $n = 5$, until in the 19th century Galois proved there is no such formula.

Corollary 2.22. *Let*
$$p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0$$

($a_n \neq 0$) be a complex polynomial of degree n. There are $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{C}$ such that

$$p(z) = a_n(z - \lambda_1)(z - \lambda_2) \ldots (z - \lambda_n)$$

(the equation $p(z) = 0$ has n solutions $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{C}$ - perhaps counted with multiplicity).

Proof. This is based on the *long division algorithm* for polynomials.

Consider any $\lambda \in \mathbb{C}$: I can write:

$$p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0$$
$$= a_n(z - \lambda)(z^{n-1} + b_{n-2} z^{n-2} + \cdots + b_0) + r$$

where r is a constant. So
$$p_n(z) = (z - \lambda)q_{n-1}(z) + r$$

Plug in $z = \lambda$: $p_n(\lambda = r)$, so $\lambda \in \mathbb{C}$ is a root of p, i.e. $p(\lambda) = 0 \iff r = 0$.

So let $\lambda_1 \in \mathbb{C}$ be a root of $p(z)$ (such a thing exists by the fundamental theorem); then

$$p_n(z) = (z - \lambda_1)q_{n-1}(z)$$

where $q_{n-1}(z)$ is a polynomial of degree $n - 1$.

By induction on n, $\exists \lambda_2, \ldots, \lambda_n$ such that

$$q_{n-1}(z) = a_n(z - \lambda_2) \ldots (z - \lambda_n)$$
$$\implies p_n(z) = a_n(z - \lambda_1)(z - \lambda_2) \ldots (z - \lambda_n) \quad \blacksquare$$
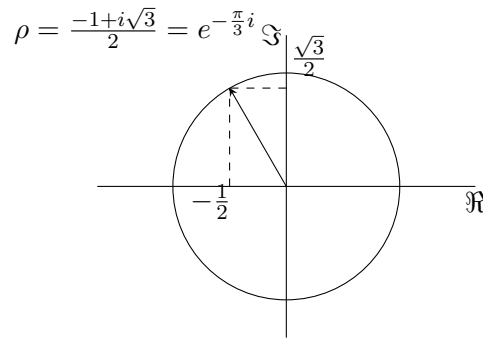
Example 2.23. Split the polynomial:

$$p(z) = z^7 - z^6 - 2z^4 + 2z^3 + z - 1$$

into linear factors and hence find all the roots of p.

We note that $\lambda = 1$ is a root. So $z - 1$ divides $p(z)$ exactly. Normally we'd do long division, but in this case I can "see" my way though:

$$p(z) = +z^6(z-1) - 2z^3(z-1) + (z-1)$$
$$= (z-1)(z^6 - 2z^3 + 1)$$
$$= (z-1)(z^3 - 1)^2$$

To split $z^3 - 1$, we need the cube roots of 1:

$$\rho = \frac{-1 + i\sqrt{3}}{2} = e^{-\frac{\pi}{3}i}$$

$$\rho^2 = \bar{\rho} = \frac{-1 - i\sqrt{3}}{2} = e^{\frac{-2\pi i}{3}}$$
$$= (z-1)^3(z-\rho)^2(z-\bar{\rho})^2$$
$$= (z-1)^3(z-\rho)^2(z-\rho^2)^2$$

So the roots are $z = 1$, $z = \rho$ and $z = \bar{\rho}$, with multiplicities of 3, 2 and 2 respectively.

Example 2.24 (Long division of polynomials).

$$
\begin{array}{r}
z^2 + 3z\ + 7 \\
z - 2\overline{)\ z^3\ + z^2\ + z\ + 1} \\
\underline{-z^3 + 2z^2\quad\quad} \\
3z^2\ + z \\
\underline{-3z^2 + 6z\quad} \\
7z\ + 1 \\
\underline{-7z + 14} \\
15
\end{array}
$$

Conclusion:

$$z^3 + z^2 + z + 1 = (z-2)(z^2 + 3z + 7) + \underbrace{15}_{\text{remainder}}$$

$$\frac{z^3 + z^2 + z + 1}{z - 2} = \underbrace{z^2 + 3z + 7}_{\text{quotient}} + \frac{15}{z - 2}$$

Corollary 2.25 (For polynomials with real coefficients)**.** *Let*

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0$$

be a polynomial with real coefficients (that is $a_0, a_1, \ldots, a_n \in \mathbb{R}$).

Then $p(x)$ can be written as a product of linear (i.e. degree 1) and quadratic (i.e. degree 2) polynomials with real coefficients.

Proof. Left as an exercise. You will need to use two things:

(i) If $\lambda \in \mathbb{C}$ is a root of $p(x)$, then $\bar{\lambda}$ is also a root. ($\overline{z+w} = \bar{z} + \bar{w}$, and $\overline{z \cdot w} = \bar{z} \cdot \bar{w}$)

(ii) If $\lambda \in \mathbb{C}$ then

$$\begin{aligned}
(x - \lambda)(x - \bar{\lambda}) &= x^2 - (\lambda + \bar{\lambda})x + \lambda\bar{\lambda} \\
&= x^2 - 2Re(\lambda)x + |\lambda|^2
\end{aligned}$$

this is a polynomial with real coefficients.

Roots and Coefficients of Algebraic Equations

$$(z - \lambda_1)(z - \lambda_2) \ldots (z - \lambda_n) = z^n + a_{n-1} z^{n-1} + \cdots + a_0$$

where

$$a_{n-1} = -(\lambda_1 + \lambda_2 + \cdots + \lambda_n)$$

$$a_{n-2} = \sum_{i<j} \lambda_i \lambda_j$$

$$a_{n-3} = - \sum_{i<j<k} \lambda_i \lambda_j \lambda_k$$

$$\vdots$$

$$a_0 = (-1)^n \lambda_1 \ldots \lambda_n$$

[Message to take home: the coefficients are simple expressions of the roots.]

> **Example 2.26.** Find a cubic with roots $\lambda_1 = 1$, $\lambda_2 = 1 + i$, $\lambda_3 = 1 - i$.
>
> The cubic will be
> $$z^3 + a_2 z^2 + a_1 z + a_0$$
> where
> $$\begin{aligned}
> a_2 &= -(\lambda_1 + \lambda_2 + \lambda_3) \\
> &= -(1 + 1 + i + 1 - i) = -3
> \end{aligned}$$
>
> $$\begin{aligned}
> a_1 &= \lambda_1 \lambda_2 + \lambda_1 \lambda_3 + \lambda_2 \lambda_3 \\
> &= 1 + i + 1 - i + (1 + i)(1 - i) \\
> &= 4
> \end{aligned}$$
>
> $$\begin{aligned}
> a_0 &= -\lambda_1 \lambda_2 \lambda_3 \\
> &= -(1 + i)(1 - i) \\
> &= -2
> \end{aligned}$$
>
> So the cubic is $z^3 - 3z^2 + 4z - 2$.

Example 2.27. Suppose $ax^2 + bx + c$ has roots λ, $\mu \in \mathbb{C}$. Find a quadratic equation with roots λ^2, μ^2.

Rewrite the quadratic as

$$x^2 + \frac{b}{a}x + \frac{c}{a}$$

then $\lambda + \mu = -\frac{b}{a}$, $\lambda\mu = \frac{c}{a}$

$$\lambda^2 + \mu^2 = (\lambda + \mu)^2 - 2\lambda\mu$$

$$= \frac{b^2}{a^2} - 2\frac{c}{a}$$

$$\lambda^2\mu^2 = \frac{c^2}{a^2}$$

The quadratic polynomial is then

$$x^2 - (\tfrac{b^2}{a^2} - \tfrac{2c}{a})x + \tfrac{c^2}{a^2}$$

Clearing the denominators:

$$a^2x^2 + (2ca - b^2)x + c^2$$

* The Cubic Equation *

Everybody remembers that $ax^2 + bx + c = 0$ has roots

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

I want to tell you a similar formula for cubics:

$$Ax^3 + Bx^2 + Cx + D = 0$$

Recall how to derive the quadratic formula (for solving $ax^2 + bx + c = 0, a \neq 0$).

First divide by a

$$x^2 + \frac{b}{a}x + \frac{c}{a} = 0$$

The first step was to "complete the square":

$$= x^2 + 2\frac{b}{2a}x + \left(\frac{b}{2a}\right)^2 + \frac{c}{a} - \left(\frac{b}{2a}\right)^2$$

$$= \left(x + \frac{b}{2a}\right)^2 + \frac{c}{a} - \left(\frac{b}{2a}\right)^2 = 0$$

Solving for x, we get:

$$x + \frac{b}{2a} = \pm\sqrt{\left(\tfrac{b}{2a}\right)^2 - \tfrac{c}{a}}$$

$$= \pm\frac{\sqrt{b^2 - 4ac}}{2a}$$

$$\implies x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Suppose now that we have a cubic equation:

$$x^3 + Ax^2 + Bx + C = 0$$

The first step is to "complete the cube". Note that

$$(x + \alpha)^3 = x^3 + 3x^2\alpha + 3x\alpha^3 + \alpha^3$$

I want to see a "cube" where $3\alpha = A$:

$$x^3 + Ax^2 + Bx + C = x^3 + 3\tfrac{A}{3}x^2 + 3 \cdot \tfrac{A^2}{9}x + \tfrac{A^3}{27} + Bx + C - \tfrac{A^2}{3}x - \tfrac{A^3}{27}$$

$$= (x + \tfrac{A}{3})^3 + (B - \tfrac{A^2}{3})x + C - \tfrac{A^3}{27}$$

$$= (x + \tfrac{A}{3})^3 + (B - \tfrac{A^2}{3})(x + \tfrac{A}{3}) + C - \tfrac{A^3}{27} - (B - \tfrac{A^2}{3})\tfrac{A}{3}$$

So if we substitute $y = x + \frac{A}{3}$ we have an equation for y of the form

$$y^3 + 3py + 2q = 0$$

(p and q are easily computed from A, B, C).

Completing the cube did not quite solve the problem, but it made it simpler. [Interesting fact: the cubic formula was discovered by Cardano in the early 1500s]

Now we study:

$$y^3 + 3py + 2q = 0$$

We are looking for a solution in the form of $y = u + v$ (Why Prof. Corti? I don't know, it just works...)

$$y^3 = (u + v)^3 = u^3 + v^3 + 3uv(u + v)$$

On the one hand

$$y^3 = -3py - 2q$$

on the other hand

$$y^3 = +3uvy + (u^3 + v^3)$$

One way this could happen is if:

$$\begin{cases} u \cdot v & = -p \\ u^3 + v^3 & = -2q \end{cases} \implies \begin{cases} u^3 \cdot v^3 & = -p^3 \\ u^3 + v^3 & = -2q \end{cases}$$

So u^3, v^3 are the two roots of the quadratic equation

$$z^2 + 2qz - p^3 = 0$$

We find u^3, v^3 with the quadratic formula:

$$u^3, v^3 = -q \pm \sqrt{q^2 + p^3}$$

ad at the end of the day, we obtain :

$$y = u + v = \sqrt[3]{-q + \sqrt{q^2 + p^3}} + \sqrt[3]{-q - \sqrt{q^2 + p^3}}$$

Note: we seem to have $3 \times 3 = 9$ solutions! (??) The key point is $u \cdot v = -p$. Once you find u, determine v uniquely from this.

Example 2.28. A typical (but nice) example: Solve for y

$$y^3 - y + 1 = 0$$

Using Cardano's formula:

$$y = \sqrt[3]{-\frac{1}{2} + \sqrt{\frac{1}{4} - \frac{1}{27}}} + \sqrt[3]{-\frac{1}{2} - \sqrt{\frac{1}{4} - \frac{1}{27}}}$$

$$= \sqrt[3]{-\frac{1}{2} + \frac{1}{6}\sqrt{\frac{23}{3}}} + \sqrt[3]{-\frac{1}{2} - \frac{1}{6}\sqrt{\frac{23}{3}}}$$

This is a real root of the equation.

Fact: You can't simplify this expression in any substantial way.

There are 2 complex roots:

$$\rho\sqrt[3]{-\frac{1}{2} + \frac{1}{6}\sqrt{\frac{23}{3}}} + \rho^2\sqrt[3]{-\frac{1}{2} - \frac{1}{6}\sqrt{\frac{23}{3}}}$$

$$\rho^2\sqrt[3]{-\frac{1}{2} + \frac{1}{6}\sqrt{\frac{23}{3}}} + \rho\sqrt[3]{-\frac{1}{2} - \frac{1}{6}\sqrt{\frac{23}{3}}}$$

where $\rho = e^{\frac{2\pi i}{3}} = \frac{-1 + i\sqrt{3}}{2}$.

Example 2.29. A non-typical example - solve

$$y^3 - 6y - 40 = 0$$

Calculating with the formula, we find that

$$y = \sqrt[3]{20 + 14\sqrt{2}} + \sqrt[3]{20 - 14\sqrt{2}}$$

Here something very special happens:

$$\sqrt[3]{20 \pm 14\sqrt{2}} = 2 \pm \sqrt{2}$$

(this cube root is a "mirage")

I verify this:
$$(2 + \sqrt{2})^3 = 8 + 3 \times 4\sqrt{2} + 3 \times 2 \times 2 + 2\sqrt{2}$$
$$= 20 + 14\sqrt{2}$$

So in fact the formula gives:

$$y = 2 + \sqrt{2} + 2 - \sqrt{2} = 4$$

The explanation for the simplified $\sqrt[3]{20 + 14\sqrt{2}} = 2 + \sqrt{2}$ relies precisely on the fact that the equation has a root $y \in \mathbb{Q}$. This is complicated enough that I cannot tell

you the whole story, as it's a long story. Even when you do Galois Theory you focus on other examples. The reason is that the other two roots are complex conjugated, and you need to extend your field more (??). These guys in the 1500s had absolutely no right to discover this formula as they didn't have the right equipment to deal with it, and the truth is neither do you.

It is easy to find rational roots without using the cubic formula:
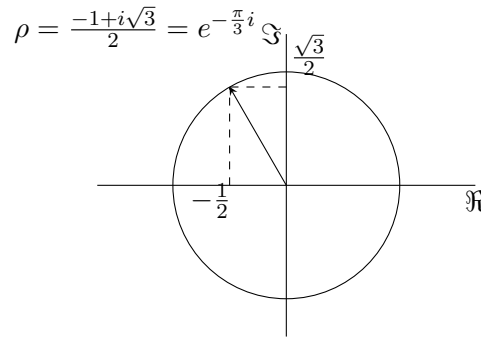
Proposition 2.30. *Suppose the equation*

$$a_0 y^m + a_1 y^{n-1} + \cdots + a_n = 0 \ (a_i \in \mathbb{R})$$

has a root $y = \frac{p}{q} \in \mathbb{Q}$, *then* $p \mid a_n$ *and* $q \mid a_0$.

(Message: to find rational roots, we only need to sift though finitely many candidates)

Example 2.31. A typical (but not so nice, but interesting) example .

Suppose you try to find a nice expression for $x = \cos\frac{\pi}{9}$ by solving a cubic using the cubic formula.

$$\rho = \frac{-1+i\sqrt{3}}{2} = e^{-\frac{\pi}{3}i}$$

This leads to a cubic equation for $x = \cos\frac{\pi}{9}$:

$$4x^3 - 3x - \frac{1}{2} = 0$$

The cubic formula:

$$x = \sqrt[3]{\frac{1}{16} + \frac{i}{16}\sqrt{3}} + \sqrt[3]{\frac{1}{16} - \frac{i}{16}\sqrt{3}}$$
$$= \frac{1}{2}\left(\sqrt[3]{\frac{1}{2} + \frac{i}{2}\sqrt{3}} + \sqrt[3]{\frac{1}{2} - \frac{i}{2}\sqrt{3}}\right)$$

I am faced with the same problem that I started out with! Namely, I was looking for an expression for

$$z = \sqrt[3]{\frac{1}{2} + \frac{i\sqrt{3}}{2}}$$

The cubic formula didn't help!

The fact here is: there is no "nice" expression for $x = \cos\frac{\pi}{9}$ (i.e. all expressions involving any number of radicals of rational numbers).

In some sense, the best you can do with $x = \cos \frac{\pi}{9}$ is to remember that it is a solution of the cubic equation:

$$4x^3 - 3x - \frac{1}{2} = 0$$

What is the key difference between the two "typical" examples (2.27 and 2.30)?

Assuming $p, q \in \mathbb{R}$ and even better, $p, q \in \mathbb{Q}$, it happens that:

$$q^2 + p^3 > 0 \iff \text{cubic formula gives nice expression for roots}$$
$$\iff \text{1 real root \& 2 complex conjugate roots}$$

$$q^2 + p^3 < 0 \iff \text{cubic formula doesn't help much}$$
$$\iff \text{3 real roots.}$$

3 Introduction to $\mathbb{N}$

> If you don't get this, get a tattoo that you can see, not where your boyfriend/girlfriend can see.
>
> *- Alessio Corti*

Highest Common Factor

Theorem 3.1: Divison Algorithm

$\forall a, b \in \mathbb{N}$, $b \neq 0$, $\exists q, r \in \mathbb{N}$ such that

$$a = qb + r \quad \text{with } 0 \leq r < b$$

(q is the *quotient*, r is the *remainder*.)

You probably met this in elementary school. Psychoanalysis is based on this - you have to regress back to your childhood to understand what the natural numbers are. We're not going to go all the way back to childhood, well maybe, but not to pre-birth and your past lives...

Remark 3.2. I've already proved this. The proof was based on taking the $\lfloor \frac{a}{b} \rfloor = q$. In turn this rested on the Archimedean axiom in $\mathbb{R}$. This is "unclean" to use $\mathbb{R}(\mathbb{Q})$ to prove something about $\mathbb{N}$. I now do a proof in $\mathbb{N}$:

Proof.
$$S = \{a - kb \mid k \in \mathbb{N}, 0 \leq a - kb\} \subset \mathbb{N}$$

$S \neq \emptyset$ because $k = 0$ gives $a \in S$. Therefore S has a smallest element. Clearly $\exists q \in \mathbb{N}$ such that $a = qb + r$.

To finish, I show that $r < b$. Suppose for a contradiction $r \geq b$. Then

$$a - (q+1)b = (a - qb) - b = r - b \geq 0$$

So $r - b \in S$ and $r - b < r$. So r was not the smallest element. This contradiction finishes the proof. ∎

Exercise: Prove the following variant:

$$\forall a, b \in \mathbb{Z}, b \neq 0, \exists q, r \in \mathbb{Z} \text{ such that } a = qb + r \text{ with } 0 \leq |r| \leq \frac{|b|}{2}$$

> **Definition.** Let $c, a \in \mathbb{Z}$. I say that c *divides* a (and I write $c \mid a$) if
>
> $$\exists k \in \mathbb{Z} : a = k \cdot c$$

> **Definition.** For $a, b \in \mathbb{Z}$ the *highest common factor*, $\text{hcf}(a, b)$ is
>
> $$d = \text{hcf}(a, b) = \max\{c : c \mid a \ \& \ c \mid b\}$$

Algorithm 3.3 (Computing hcf). *Note: $hcf(a, b) = hcf(\pm a, \pm b) = hcf(b, a)$, so I'm fine to always assume $a, b \in \mathbb{N}$ and $a \geq b$.*

From our division algorithm, we can write $a = qb + r$ with $0 \leq r < b$.

Claim: $(c \mid a$ and $c \mid b) \iff (c \mid b$ and $c \mid r)$.

Proof. $\implies$: $c \mid a$ and $c \mid b$, so we can write $a = k_1 c$, $b = k_2 c$. Then

$$r = a - qb = k_1 c - q k_2 c$$
$$= (k_1 - q k_2)c$$

So $c \mid r$. The other direction is similar. ∎

Recall:

> **Definition.** Let $p \in \mathbb{N} \backslash \{0, 1\}$. p is *irreducible* iff
>
> $$p = nm \implies n = 1 \text{ or } m = 1$$
>
> $(n, m \in \mathbb{N})$

> **Definition.** Let $p \in \mathbb{N} \backslash \{0, 1\}$. p is *prime* iff:
>
> $$(p \mid ab \implies p \mid a \text{ or } p \mid b)$$

Proposition 3.4. *p is irreducible $\iff$ p is prime.*

Proof. Suppose p is prime. Let $p = n \cdot m$. Clearly $p \mid n \cdot m$, so $p \mid n$ or $p \mid m$. Assume wlog that $p \mid n$, so $\exists k$ such that $n = kp$, and then

$$p = nm = p \cdot (km)$$
$$\implies km = 1$$
$$\implies m = 1$$

This shows that p is irreducible.

Now suppose that p is irreducible. If $p \mid ab$, let $c = hcf(p, a)$. Then $c \mid p$, so either $c = p \implies p \mid a$ or $c = 1$. So assume $c = hcf(p, a) = 1$. Then $\exists xy \in \mathbb{Z}$ such that $px + ay = 1$. So $b = pxb + aby \implies p \mid b$. ∎

Remark 3.5. The claim follows from $c \mid A, c \mid B \implies c \mid A + B$.

The claim means that
$$hcf(a, b) = hcf(b, r)$$

This is it! $a \geq b$ and $b > r$ means that the pair (b, r) is simpler (smaller) than the pair (a, b), so I may assume inductively that I can handle computing $hcf(b, r)$.

Theorem 3.6: Bezout's Lemma

Let $d = hcf(a, b)$. Then
$$\exists\, x, y \in \mathbb{Z} : ax + by = d$$

This is a REALLY important fact.

Corollary 3.7. *Suppose $c \mid a$, $c \mid b \implies c \mid \mathrm{hcf}(a, b)$.*

Proof. First remark that $c \mid A, B \in \mathbb{Z} \implies c \mid AB$.

Then $c \mid a$, $c \mid b \implies c \mid ax$, $c \mid by \implies c \mid ax + by \implies c \mid \mathrm{hcf}(a, b)$. ∎

Proof of Theorem. Write $a = bq + r$. By induction $\exists\, x', y'$ such that

$$\begin{aligned}
\mathrm{hcf}(a, b) = \mathrm{hcf}(b, r) &= bx' + ry' \\
&= bx' + (a - bq)y' \\
&= ay' + b(x' - qy')
\end{aligned}$$

So let $x = y'$, $y = x' - qy'$ to obtain the theorem. ∎

Example 3.8. Compute $\mathrm{hcf}(5817, 1428) = d$, and find $x, y \in \mathbb{Z}$ such that $5817x + 1428y = d$. Lecture 17

Using Euclid's algorithm:

$$\begin{aligned}
5817 &= 4 \times 1428 + 105 \\
1428 &= 13 \times 105 + 63 \\
105 &= 1 \times 63 + 42 \\
63 &= 1 \times 42 + 21 \\
42 &= 2 \times 21 + 0
\end{aligned}$$

So $21 = \mathrm{hcf}(5817, 1428)$.

We can then find x, y such that $5817x + 1428y = d$ by working backwards:

$$\begin{aligned}
21 &= -42 + 63 \\
&= -105 + 2 \times 63 \\
&= -105 + 2 \times (-13 \times 105 + 1428) \\
&= -27 \times 105 + 2 \times 1428 \\
&= 27(-5817 + 4 \times 1428) + 2 \times 1428 \\
&= -27(5817) + 110(1428)
\end{aligned}$$

So I can take $x = -27$, $y = 110$.

Definition. $a, b \in \mathbb{Z}$ are *co-prime* iff $\mathrm{hcf}(a, b) = 1$.

Proposition 3.9. *Suppose $a, b \in \mathbb{Z}$ are co-prime. Then*

(i) $a \mid c$, $b \mid c \implies ab \mid c$

(ii) $a \mid bc \implies a \mid c$

Proof. By Bezout's Lemma, $\exists\, x, y \in \mathbb{Z}$ such that $ax + by = 1$. Let me show (i):

Consider $axc + byc = c$. We can see that

$$\begin{aligned}
a \mid c &\text{ so } \exists k_1 \text{ such that } c = ak_1 \\
\& \, b \mid c &\text{ so } \exists k_2 \text{ such that } c = bk_2
\end{aligned}$$

$$\implies abxk_2 + abyk_1 = c$$
$$\implies ab(xk_2 + y_k1) = c$$

so $ab \mid c$, showing (i).

To show part (ii) again: $a \times c + byc = c \implies a \mid (a \times c) + (byc)$, so $a \mid c$. ∎

Terminology. We know irreducible $\iff$ prime. From now on I just use the word prime.

Theorem 3.10: Fundamental Theorem of Arithmetic

Every $n \in \mathbb{N}$ can be written uniquely in the form:

$$n = p_1^{a_1} p_2^{a_2} \ldots p_r^{a_r}$$

where $p_1 < p_2 < \cdots < p_r$ are prime and $a_i \in \mathbb{N}$.

Proof. We already know that n can be written in this form (we proved this using the fact that primes are irreducible.)

We need to show uniqueness. Suppose:

$$n = p_1^{a_1} p_2^{a_2} \ldots p_r^{a_r}, \quad p_1 < p_2 < \cdots < p_r \text{ primes}$$
$$= q_1^{b_1} q_2^{b_2} \ldots q_s^{b_s}, \quad q_1 < q_2 < \cdots < q_s \text{ primes}$$

We said in the definition that p being prime means $p \mid ab \implies p \mid a$ or $p \mid b$, but in fact:

$$p \mid a_1 a_2 \ldots a_m \implies p \mid a_i \text{ for some } i \in \{1, \ldots, p-1\}$$

From the first expression $p_1 \mid n$. From the second expression, we get $p_1 \mid q_i$ for some i. So $q_i = p_1 c$ and then because q_i is irreducible, $c = 1$, so in fact $p_1 = q_i$. Dividing through by p_1:

$$\frac{n}{p_1} = p_1^{a_1 - 1} p_2^{a_2} \ldots p_r^{a_r}$$
$$= q_1^{b_1} \ldots q_i^{b_i - 1} \ldots q_s^{b_s}$$

But $\frac{n}{p_1} < n$, so I could assume (by strong induction) that the two decompositions of $\frac{n}{p_1}$ are identical $\implies$ the two decompositions for n were also identical. ∎

Corollary 3.11. *Suppose that $a = \Pi q_i^{r_i}$ and $b = \Pi p_i^{s_i}$. Then*

$$\mathrm{hcf}(a, b) = \Pi p_i^{\min\{r_i, s_i\}}$$

Note: This is not a good way to compute $\mathrm{hcf}(a, b)$! The good way is to use Euclid's algorithm.

Theorem 3.12: Euclid

There are infinitely many primes.

Proof. Suppose for a contradiction that the set of all primes is finite, and that :

$$p_1, p_2, \ldots, p_r$$

is the list of all primes. Consider

$$N = p_1 p_2 \ldots p_r + 1$$

Claim: $\mathrm{hcf}(N, p_i) = 1$ for all i; that is $p_i \nmid N$.

Indeed:

$$
\begin{aligned}
1 &= N - (p_1 \ldots \hat{\mathbf{p}}_\mathbf{i} \ldots p_r)p_i \\
&= Nx + p_i y
\end{aligned}
$$

with $x, y \in \mathbb{Z} \implies \mathrm{hcf}(N, p_i) = 1$ because $c \mid N$ and $c \mid p_i \implies c \mid 1$.

But then the prime decomposition of N must involve primes other than $p_1, \ldots, p_r$. This contradicts the assumption that $p_1, \ldots, p_r$ has the list of all primes. ∎

Exercise: Show that there are infinitely many primes of the form $4k + 1$ and also of the form $4k + 3$. (One of the two is a lot harder than the other one)

Applications of the fundamental theorem

> **Definition.** $n \in \mathbb{N}$ is a square if $\exists k \in \mathbb{N}$ such that $n = k^2$.

Example 3.13. $n \in \mathbb{N} \ \sqrt{n} \in \mathbb{Q} \implies n$ is a square (i.e. $k = \sqrt{n} \in \mathbb{N}$).

Idea: every prime divides n an even number of times.

Proof. Suppose $\sqrt{n} = \frac{a}{b} \in \mathbb{Q}$, with $a, b \in \mathbb{N}$ and $\mathrm{hcf}(a, b) = 1$. Then $a^2 = nb^2$.

A small digression. Suppose p is prime and $N \in \mathbb{Z}$. We define

$$\mathrm{ord}_p N = \max\{k \in \mathbb{N} \mid p^k \mid N\}$$

So $N = \pm \prod_{p \text{ prime}} p^{\mathrm{ord}_p N}$.

Note: $\mathrm{ord}_p(N_1 N_2) = \mathrm{ord}_p N_1 + \mathrm{ord}_p N_2$.

Back to the example. Let p be prime.

$$
\begin{aligned}
\mathrm{ord}_p(a^2) &= 2\,\mathrm{ord}_p a \\
&= \mathrm{ord}_p n b^2 = \mathrm{ord}_p n + 2\,\mathrm{ord}_p b \\
\implies \mathrm{ord}_p n &= 2(\mathrm{ord}_p a - \mathrm{ord}_p b)
\end{aligned}
$$

i.e. $p \mid n$ an even number of times. ∎

Example 3.14. The equation

$$(2y)^2 = x^3 + 1$$

has no solutions for $x, y \in \mathbb{Z}$, other than $x = -1$, $y = 0$.

Proof.

$$x^3 = (2y)^2 - 1 = (2y - 1)(2y + 1)$$

Note: $\mathrm{hcf}(2y - 1, 2y + 1) = 1$.

Indeed $2 = (2y+1) - (2y-1)$, so if $c \mid 2y+1$ and $c \mid 2y-1 \implies c \mid 2$, so hcf $\mid 2$, so hcf $= 1$ or 2. But both numbers are odd, so hcf $= 1$.

Lemma 3.15. $\text{hcf}(2y-1, 2y+1) = 1$ *and* $(2y-1)(2y+1)$ *is a cube* $\implies$ *both* $2y-1$ *and* $2y+1$ *are cubes.*

Proof of Lemma. Suppose $p \mid x^3$. Then

$$
\begin{aligned}
0 < \text{ord}_p x^3 &= 3(\text{ord}_p x) \\
&= \text{ord}_p(2y-1) + \text{ord}_p(2y+1)
\end{aligned}
$$

because $\text{hcf}(2y-1, 2y+1) = 1$, so p cannot divide both. So either

$$\text{ord}_p(2y-1) = 0 \implies 3(\text{ord}_p x) = \text{ord}_p(2y+1)$$

or

$$\text{ord}_p(2y+1) = 0 \implies 3(\text{ord}_p x) = \text{ord}_p(2y-1)$$

This shows the lemma.

What are the cubes?

$$-1, 0, 1, 8, 27, 64, 125, \ldots$$

Manifestly then $2y-1 = -1$ and $2y+1 = 1 \implies x = -1, y = 0$. ∎

Example 3.16. Suppose that all months have 30 days. Show that some month must start on a Monday.

The whole point is $\text{hcf}(7, 30) = 1$:

$$
\begin{aligned}
30 &= 7 \times 4 + 2 \\
7 &= 2 \times 3 = 1 \\
1 &= 7 - 3 \times (30 - 4 \times 7) \\
&= 13 \times 7 - 3 \times 30
\end{aligned}
$$

So

$$\boxed{3 \times 30 = 13 \times 7 - 1}$$

Every 3 months I start exactly one day earlier. So in the worst case I need $3 \times 6 = 48$ months before I have worked through all days of the week.

The Least Common Multiple

Definition. Let $a, b \in \mathbb{Z}$. We define

$$\text{lcm}(a, b) = \text{ smallest} > 0 \text{ integer divisible by both } a \text{ and } b$$

Note: A linguistic wrinkle: nothing is divisble by 0, hence $\text{lcm}(a, 0)$ is undefined.

Proposition 3.17. *Suppose k is divisble by a, b. Then k is divisible by $\text{lcm}(a, b)$*

I will prove this together with the following:

Proposition 3.18. *Write $d = \mathrm{hcf}(a, b)$. $b = db'$, $a = da'$, then $lcm(a, b) = da'b'$.*

Before even getting into the proof, not that $\mathrm{hcf}(a'b') = 1$. [If $c = \mathrm{hcf}(a', b')$, then $dc \mid da' = a$ and $dc \mid db' = b \implies c = 1$]

Proof. Set $m = da'b'$. Note that $m = ab' = a'b$, so m is divisible by both a and b.

First step: If k is divisible by a, b then k is divisible by m.

$$k = ap = bq$$
$$= a'dp = b'dq$$

Thus $a'p = b'q$, so $a' \mid b'q$ and $\mathrm{hcf}(a'b') = 1 \implies a' \mid q$ (by something that we proved earlier). So we can write $q = a'c$, and $k = bq = ba'c = da'b'c = mc$, meaning k is divisible by m.

Second Step: Let $S = \{k \mid k > 0,\ k \text{ is divisible by } a, b\}$

Claim: $lcm(a, b) = min\, S$ We know that $m \in S$. We also know $k \in S \implies k$ is divisible by m. In particular $k \in S \implies k \geq m$. So indeed $m = min\, S$ and then $m = lcm(a, b)$. ∎

Equations of the form $ax + by = c$

Remark 3.19. The equation has a solution $\iff \mathrm{hcf}(a, b) \mid c$.

Proof. Indeed suppose $\exists x, y \in \mathbb{Z}$ such that $ax + by = c$. Then

$$d = \mathrm{hcf}(a, b),\ d \mid ax,\ d \mid by \implies d \mid ax + by = c$$

Vica-versa: Suppose $d \mid c$, i.e. $c = dc'$. We know there are $x', y' \in \mathbb{Z}$ such that $ax' + by' = d$ by Bezout's Lemma.

Multiply them by c':
$$a(x'c') + b(y'c') = dc' = c$$

So $x = x'$, $y = y'c$ solve the original equation. ∎

Assuming $d \mid c$, let us describe all the solutions:

Remark 3.20. Suppose x_0, y_0 is a solution:

$$ax_0 + by_0 = c$$

Then
$$ax + by = c \iff a(x - x_0) + b(y - y_0) = 0$$
$$\iff x', y' \in \mathbb{Z} \text{ is a solution of}$$
$$\text{the homogeneous equation } ax' + by' = 0$$

Remark 3.21. $ax' + by' = 0 \iff x' = tb', y' = -ta'$ for some $t \in \mathbb{Z}$, where again $d = \mathrm{hcf}(a, b)$ and $a = a'd$, $b = b'd$.

Proof. Indeed $k = ax' = -b'$ is divisible by a and b. So k is divisible by $lcm(a, b) = a'b'd$, so $k = ta'b'd = ax'$. Thus $x' = t'$ and similarly $y' = -ta'$. ∎

Example 3.22. Solve for $x, y \in \mathbb{Z}$:

$$5x + 11y = 1$$

$\mathrm{hcf}(5, 11) = 1$, so the equation is soluble. In general, a solution can always be found by Euclid's algorithm. In this case we can just "eyeball" the solution: $x_0 = -2, y_0 = 1$ solved the equation as $-2 \times 5 + 1 \times 11 = 1$.

We solve the homogeneous equation

$$5x' + 11y' = 0$$

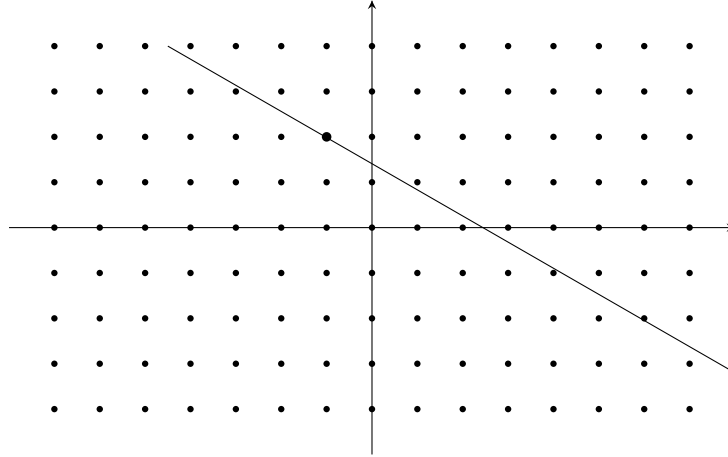$$\Longleftrightarrow \; x' = 11t, \; y' = -5t, \; \text{some } t \in \mathbb{Z}$$

The set of all solutions to the original equation is

$$\{(x, y) = (-2 + 11t, 1 - 5t) \mid t \in \mathbb{Z}\}$$

Sometimes $a, b, c > 0$ are integers and we want to solve $ax + by = c$ for $x, y > 0$.

Example 3.23. *When is $5x + 11y = c$ soluble for $x, y \in \mathbb{N}$?*

Here is a picture of $\mathbb{Z}^2 \subset \mathbb{R}^2$:

Is there a solution in the positive quadrant $\mathbb{N}^2$?

We know $\exists$ a solution $x_0, y_0 \in \mathbb{Z}^2$. We also know every other $x, y \in \mathbb{Z}^2$ is of the form:
$$x_0 + 11y, y_0 - 5t \quad t \in \mathbb{Z}$$
$$= (x_0, y_0) + t(11, -5)$$

For sure we can find a solution $x, y \in \mathbb{N}^2$ if

$$\|(11, -5)\| \leq \|(\tfrac{c}{5}, \tfrac{c}{11})\|$$
$$\Longrightarrow \sqrt{11^2 + 5^2} \leq \sqrt{\tfrac{c^2}{5^2} + \tfrac{c^2}{11^2}}$$

So we're OK if $55 \leq c$. One can do better though!

Recall that the equation $ax \equiv 1 \pmod{m}$ is soluble for $x \iff \mathrm{hcf}(a, m) = 1$.

Modular Arithmetic

> **Definition.** $a, b \in \mathbb{Z}$ $n \in \mathbb{N}$. a is *congruent* to b mod n:
>
> $$a \equiv b \pmod{n}$$
>
> if $n \mid a - b$.

Remark 3.24. $a \equiv a \pmod{n}$. $a \equiv b \iff b \equiv a \pmod{n}$. If $a \equiv b, b \equiv c \implies a \equiv c \pmod{n}$.

In modular arithmetic, we are really jut interested in classes:

$$[a] = \{b \mid a \equiv b \pmod{n}\}$$

Remark 3.25. $\forall a, \exists! b$ such that $a \equiv b \pmod{n}$ and $0 \le b < n$.

This is just the division algorithm: $a = qn + b$ $0 \le b < n$, so $a \equiv b$.

So in fact I can , if I want, think of the set of classes mod n:

$$\{[a] \mid a \in \mathbb{Z}\} = \{0, 1, 2, \dots, n-1\}$$

Notation. $\mathbb{Z}\backslash n\mathbb{Z} = $ the set of classes.

Proposition 3.26. *If $a \equiv b \pmod{n}$ and $c \equiv d \pmod{n}$, then:*

$$a + c \equiv b + d \pmod{n}$$
$$a \cdot c \equiv b \cdot d \pmod{n}$$

i.e. it makes sense to add and multiply $\pmod{n}$.

There is no point trying to remember this proof, because it's just too easy:

Proof. $n \mid a - b$, i.e. $\exists k : a - b = n$. Similarly $\exists h : c - d = hm$.

$$
\begin{aligned}
(a + c) - (b + d) &= (a - b) + (c - d)\\
&= kn + hn = (k + h)n
\end{aligned}
$$

So $a + c \equiv b + d \pmod{n}$.

Similarly:

$$
\begin{aligned}
(a - b)(c + d) &= kn(c + d)\\
&= ac - bd + ad - bc\\
&= ac - bd + ad - ac + ac - bc\\
&= ac - bd + a(d - c) + c(a - b)\\
&= ac - bd - ahn + ckn\\
ac - bd &= ahn - ckn + kn\\
&= n[ah - ck + k(0)]
\end{aligned}
$$

So $ac \equiv bd \pmod{n}$. ∎

Thus, addition and multiplication are defined in $\mathbb{Z}/n\mathbb{Z}$. It makes sense to say

$$[a] + [c] = [a + c]$$
$$[a][c] = [a \cdot c]$$

Demonstration. $\mathbb{Z}/5\mathbb{Z} = \{0, 1, 2, 3, 4\}$

Addition table mod 5:

$+_5$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

Multiplication table mod 5:

$\times_5$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Note: $\forall a \in \mathbb{Z}/5\mathbb{Z}$, $a \neq 0, \exists b \in \mathbb{Z}/5\mathbb{Z} : ab = 1$. This is not true in $\mathbb{Z}/6\mathbb{Z}$ e.g. $2 \cdot 3 = 0$ (mod 6), since 6 is not prime; this will be proved in M1P2 ALGEBRA I next term.

Some examples to familiarise ourselves with modular arithmetic.

Example 3.27. What is the remainder when I divide 7 into 15^{99}?

It is $= 1$, since $15 \equiv 1 \pmod 7$, so $15^{99} \equiv 1 \pmod 7$, i.e. $15^{99} = k \cdot 7 + 1$, some k.

Example 3.28. What is the smallest residue of $2^{100} \pmod{15}$?

Mucking about: $2^4 = 16 \equiv 1 \pmod{15}$, so $2^{100} = (2^4)^{25} \equiv 1^{25} = 1 \pmod{15}$.

Example 3.29. Compute $5^{67} \pmod{14}$ [method: successive squaring]

$$67 = 64 + 3 = 64 + 2 + 1$$
$$= 2^6 + 2 + 1$$

(in binary $67 = 1000011$). I compute $5^{64} \pmod{14}$ by successive squaring:

$$5^1 \equiv 5 \pmod{14}$$
$$5^2 = 25 \equiv -3 \pmod{14}$$
$$5^4 \equiv 9 \equiv -5 \pmod{14}$$
$$5^8 \equiv -3 \pmod{14}$$
$$5^{16} \equiv -5 \pmod{14}$$
$$5^{32} \equiv -3 \pmod{14}$$
$$5^{64} \equiv -5 \pmod{14}$$

So
$$5^{67} = 5^{64} \cdot 5^2 \cdot 5$$
$$\equiv (-5) \times (-3) \times 5$$
$$= 1 \cdot 5 \equiv 5 \pmod{14}$$

Example 3.30. For every $n \in \mathbb{N}$, $5n + 3$ is not a square. i.e. $\nexists k \in \mathbb{N}$ such that $k^2 = 5n + 3$.

Proof. It may be quite hard to do the proof directly. The statement actually says: there is no $k \in \mathbb{Z}$ such that $k^2 \equiv 3 \pmod{5}$.

There are 5 cases to discuss. Let $k \in \mathbb{Z}$, then one of the following 5 cases occurs:

   (i) $k \equiv 0 \pmod{5} \implies k^2 \equiv 0 \pmod{5}$

  (ii) $k \equiv 1 \pmod{5} \implies k^2 \equiv 1 \pmod{5}$

 (iii) $k \equiv 2 \pmod{5} \implies k^2 \equiv 4 \pmod{5}$

 (iv) $k \equiv 3 \pmod{5} \implies k^2 \equiv 4 \pmod{5}$

  (v) $k \equiv 4 \pmod{5} \implies k^2 \equiv 1 \pmod{5}$

So k^2 is never $\equiv 3 \pmod{5}$, i.e. $\nexists k$ such that $k^2 = 5n + 3$. ∎

Proposition 3.31. *Fix $m \in \mathbb{N}, a \in \mathbb{Z}$. The equation $ax \equiv 1 \pmod{m}$ is soluble for $x \in \mathbb{Z}$ (modulo m) iff $hcf(a, m) = 1$.*

Proof.
$$ax \equiv 1 \pmod{m}$$
$$\iff \exists y \in \mathbb{Z} : ax + my = 1$$
$$\iff \mathrm{hcf}(a, m) = 1.$$

∎

Example 3.32. The equation

$$x^2 + 5y^2 = 3z^2 \tag{$*$}$$

has no solution for $x, y, z \in \mathbb{Z}$. (except the trivial one where $x = y = z = 0$).

Proof. Note that if $x, y, z \in \mathbb{Z}$ is a solution and $k \in \mathbb{Z}$ then kx, ky, kz is also a solution.

So assume for contradiction that x, y, z is a nontrivial solution. I may assue $\mathrm{hcf}(x, y, z) = 1$. Now consider $(*)$ mod 5:

$$x^2 \equiv 3z^2 \pmod{5}$$

Exercise: $\implies$ both $x, z \equiv 0 \pmod{5}$. (idea again: 3 is not a square modulo 5)

Going back to the original equation $(*)$, we have $x = 5x'$, $z = 5z'$.

Substituing into $(*)$

$$25x'^2 + 5y^2 = 3 \times 25z'^2$$
$$\implies y^2 = 5(15z^{12} - 5x^{12})$$
$$\equiv 0 \pmod 5$$
$$\implies y \equiv 0 \pmod 5$$

Then $5 \mid x$, $5 \mid z$ and $5 \mid y$, a contradiction. ∎

Lecture 21

Definition.

$$\mathbb{Z}/m\mathbb{Z}^\times = \{\text{multiplicatively invertible elements of } [a] \in \mathbb{Z}/m\mathbb{Z}\}$$
$$= \{[a] \in \mathbb{Z}/m\mathbb{Z} \mid \exists [x] \text{ s.t. } [a][x] = [1]\}$$
$$= \{[a] \mid \text{hcf}(a, b) = 1\}$$

Note: $\text{hcf}(a, m) = \text{hcf}(a + km, m)$ for all $k \in \mathbb{Z}$. So the statement $\text{hcf}(a, m) = 1$ only depends on $[a] \pmod m$.

For a finite set A, I will denote $|A|$ as the number of elements of A. Next week I will find formulas for $\phi(n) = |(\mathbb{Z}/n\mathbb{Z})^\times|$. For now just note the following:

If $n = p$ is prime, then for all $a \in \mathbb{Z}$: either $p \mid a \implies a \equiv 0 \pmod m$, or $\text{hcf}(p, a) = 1 \implies a \in (\mathbb{Z}/p\mathbb{Z})$. So $(\mathbb{Z}/p\mathbb{Z})^\times = \mathbb{Z}/p\mathbb{Z}/\{0\}$. In other words, every non-zero $a \in \mathbb{Z}/p\mathbb{Z}$ is *multiplicatively invertible*.

The operations $+, \times$ on $\mathbb{Z}/p\mathbb{Z}$ satisfy the same axioms for $k = \mathbb{Q}, \mathbb{R}, \mathbb{C}$. We say that $\mathbb{Z}/p\mathbb{Z}$ is a *field.* (whereas e.g. $\mathbb{Z}$ is just a *ring*). When we insist that $\mathbb{Z}/p\mathbb{Z}$ is a *field*, then we call if $\mathbb{F}_p$. Note also that $\phi(p) = p - 1$.

Theorem 3.33: Fermat's Little Theorem

Suppose that p is prime, $a \not\equiv 0 \pmod p$. Then

$$a^{p-1} \equiv 1 \pmod p$$

If you know about groups, then the proof is very quick:

Proof. $(\mathbb{Z}/p\mathbb{Z})^\times, \times = G$, is a group with $p - 1 = n$ elements. It is a general fact (Lagrange's theorem) that if G is a finite group, and $|G| = n$, then $g^n = e \; \forall g \in G$. ∎

I give a proof that's good enough for you:

Proof. Notation: If $n \in \mathbb{Z}$ then $\bar{n}$ = the smallest residue mod p, i..e $0 \leq \bar{n} < p$ and $n \equiv \bar{n} \pmod p$. (formally we write $n = kp + r$, $0 \leq r < p$, so then $r = \bar{n}$)

Consider $S = \{1, 2, 3, \ldots, p - 1\}$. Next consider $S_n = \{\bar{a}, \overline{2a}, \overline{3a}, \ldots, \overline{(p-1)a}\}$.

Claim. $S_n = S$.

All I need to check is that

$$1 \leq k_1 < k_2 \leq p - 1 \implies \overline{k_1 a} \neq \overline{k_2 a}$$

or in fact $k_1 a \not\equiv k_2 a \pmod p$. But this is completely obvious:

$$\leq k_1 < k_2 \leq p - 1 \implies 0 < k_2 - k_1 < p - 1$$

So

$$p \nmid k_2 - k_1 \text{ and } p \nmid a \implies p \nmid (k_2 - k_1)a$$
$$\iff (k_2 - k_1)a \not\equiv 0 \pmod{p}$$
$$\iff k_1 a \not\equiv k_2 a \pmod{p}$$

Next:

$$1 \cdot 2 \cdot 3 \cdots (p-1) = (p-1)!$$
$$= \overline{a} \, \overline{2a} \, \overline{3a} \ldots \overline{(p-1)a}$$
$$= (p-1)! \, a^{p-1} \pmod{p}$$

and all we need to observe is that

$$(p-1)! \not\equiv 0 \pmod{p}$$

Indeed $p \nmid 1$, $p \nmid 2, \ldots, p \nmid (p-1) \implies p \nmid (p-1)!$, so I can divide though by $(p-1)!$ mod p and conclude that

$$a^{p-1} \equiv 1 \pmod{p} \qquad \blacksquare$$

Theorem 3.34: Chinese Remainder Theorem

Suppose $\text{hc}(m_1, m_2) = 1$. Then for all a_1, a_2 the system of equations

$$\begin{cases} x & \equiv a_1 \pmod{m_1} \\ x & \equiv a_2 \pmod{m_2} \end{cases}$$

has a unique solution mod $m_1 m_2$.

Proof. Existence: We want $y_1, y_2 \in \mathbb{Z}$ such that

$$a_1 + y_1 m_1 = a_2 + y_2 m_2$$

so then I can put $x = a_1 + y_1 m_1$.

I need to solve

$$y_1 m_1 - y_2 m_2 = a_2 - a_1$$

for $y_1, y_2 \in \mathbb{Z}$, and this is OK since $\text{hcf}(m_1, m_2) = 1$.

Uniqueness: Suppose we have two solutions $x_1, x_2 \in \mathbb{Z}$, then

$$x_1 - x_2 \equiv a_1 - a_1 = 0 \pmod{m_1}$$
$$\equiv 0 \pmod{m_2}$$

So $m_1 \mid x_1 - x_2$ and $m_2 \mid x_1 - x_2$. But $\text{hcf}(m_1, m_2) = 1$, so $\text{lcm}(m_1, m_2) = m_1 m_2 \mid x_1 - x_2$. So $x_1 \equiv x_2 \pmod{m_1 m_2}$. $\blacksquare$

In other words: (recall if A, B are sets then we denote the cartesian product as $A \times B = \{(a, b) \mid a \in A, \, b \in B\}$)

$$\mathbb{Z}/m_1 m_2 \mathbb{Z} = \mathbb{Z}/m_1 \mathbb{Z} \times \mathbb{Z}/m_2 \mathbb{Z}$$

if $\text{hcf}(m_1, m_2) = 1$. I think of this as a kind of *holography*.

RSA Cryptography

Application of Fermat's Little Theorem (p prime, $a \not\equiv 0 \pmod{p} \implies a^{p-1} \equiv 1 \pmod{p}$, e.g. $2^4 = 16 \equiv 1 \pmod{5}$)

RSA = Riverst, Shamir, Alderman (MIT '77).

It was discovered earlier by Clifford Cocks (GCHQ '73)

Public key cryptography: You can publish the encryption key publicly, so anyone can send you a message - only you can decrypt the message.

Procedure 3.35 (RSA Encryption)**.**

(i) *Choose 2 big prime numbers p, q. (in practice, we want $|p - q$ to be big)*

(ii) *Compute $N = pq$ - make N public.*

(iii) *Compute $(p - 1)(q - 1)$ - keep this secret*

(iv) *Choose e with $\mathrm{hcf}(e, (p - 1)(q - 1)) = 1$. Make e public*

(v) *Compute d such that (from Euclid's algorithm) $de \equiv 1 \pmod{(p-1)(q-1)}$. Keep d secret.*

Write down your message. Turn it into a number $0 \leq x < N$. Compute $y = x^3 \pmod{N}$. y is the encrypted version of x, and now send y, the encrypted message.

How do we get x back from y?

> **Theorem 3.36**
>
> $x \equiv y^d \pmod{N}$

Proof. $N = pq$, $p \neq q$. We have $e, d : de \equiv 1 \pmod{(p - 1)(q - 1)}$.

Claim: $x^{de} \equiv 1 \pmod{N}$.

note that $de = 1 + m(p - 1)(q - 1)$, so $x^{de} = x \left(x^{(p-1)(q-1)} \right)^m$.

$$x^{p-1} \equiv \begin{cases} 1 (x \neq 0 \pmod{p}) & \pmod{p} \\ 0 (x \equiv 0 \pmod{p}) & \pmod{p} \end{cases}$$

So

$$x^{de} = x(x^{p-1})^{m(q-1)}$$

$$\equiv \begin{cases} x(1)^{m(q-1)} \text{ if } x \not\equiv 0 \pmod{p} \\ x(0)^{m(q-1)} \text{ if } x \equiv 0 \pmod{p} \end{cases}$$

$$\equiv x \pmod{p}.$$

Similarly $x^{de} \equiv x \pmod{q}$.

So p and q both divide $x^{de} - x$. Also $\mathrm{hcf}(p, q) = 1$.

Claim: This implies that $N = pq$ divides $x^{de} - x$.

This follows from

Lemma 3.37. *If $\mathrm{hcf}(a, b) = 1$, $a \mid m$, $b \mid m \implies ab \mid m$.*

Proof. If $b \mid ac$ and $\mathrm{hcf}(a, b) = 1$, then $b \mid c$.
($\exists x, y$ such that $ax + by = 1$, so $axc + byc = c$, i.e. $x(ac) + yc(b) = c$)

By assumption, $a \mid m$, so $m = ac$ for some c. Also $b \mid m = ac$, so $b \mid c$, say $c = bk$. So $m = ac = (ab)k$ is divisible by ab. ∎

Recap: Choose p, q, so $N = pq$. Choose e with $\mathrm{hcf}(e, (p-1)(q-1)) = 1$. Find d such that $de \equiv 1 \pmod{(p-1)(q-1)}$. Take a message x with $0 \le x < N$.

Encrypt: $y \equiv x^e \pmod N$. Decrypt: $x \equiv y^d \pmod N$.

> **Example 3.38** (Silly example)**.** $p = 5, q = 7, e = 7$. $N = pq = 35$.
>
> Message: $x = 20 \pmod{35}$.
>
> $y = 20^7 \pmod{35}$. We calculate 20^7 by repeated squaring: $20^7 = 20 \cdot 20^2 \cdot 20^4$, so
>
> $$20^2 = 400 \equiv 50 \pmod{35}$$
> $$\equiv 15 \pmod{35}$$
>
> $$20^4 \equiv (15)^2 = 225 = 210 + 15$$
> $$\equiv 15 \pmod{35}$$
>
> $$20^7 = 20 \cdot 20^2 \cdot 20^4$$
> $$\equiv 20 \cdot 15 \cdot 15$$
> $$\equiv 20 \cdot 15$$
> $$\equiv 20 \pmod{35}$$
>
> $[20 \cdot 15 - 20 = 20 \cdot 14 = 4 \cdot 5 \cdot 2 \cdot 7 \equiv 0 \pmod{35}]$
>
> So $y = 20$ - the same as the original message!
>
> To "decrypt" this, first work out d:
>
> $$de \equiv 1 \pmod{(p-1)(q-1)} = 24$$
> $$\implies 7d \equiv 1 \pmod{24}$$
>
> We do this by Euclid's algorithm:
>
> $$24 = 3 \cdot 7 + 3$$
> $$7 = 2 \cdot 3 + 1$$
>
> So working backwards:
> $$1 = 7 - 2 \cdot 3$$
> $$= 7 - 2(24 - 3 \cdot 7)$$
> $$= 7 \cdot 7 - 2 \cdot 24$$
>
> So $7 \cdot 7 \equiv 1 \pmod{24}$. So $d = 7$. Claim: $x \equiv y^d \pmod N$, so $x \equiv 20^7 \equiv 20 \pmod{35}$ as we saw before.

Notice that our private key was $N = 35 = 7 \times 5$; 5 and 7 are very close to each other, which makes 35 easy to factor with:

Fermat Factorisation.

If p, q are very close, you can find them by looking at squares.

Idea: Use $a^2 - b^2 = (a+b)(a-b)$.

Take $p = a + b, q = a - b \implies a = \frac{p+q}{2}, b = \frac{p-q}{2}$.

We want to write $N = a^2 - b^2$. Then bo back to $p = a + b, q = a - b$.

So take N, and compute $N + 1, N + 4, N + 9, \ldots$ and check if any of them are squares.

> **Examples 3.39.**
>
> (i) In our example, $35 + 1 = 36 = 6^2$, take $a = 6, b = 1$. Then $p = 6 + 1 = 7, q = 6 - 1 = 5$.
>
> (ii) $n = 187$. Calculate $N + 1 = 188, N + 4 = 191, N + 9 = 196 = 14^2$, so $a = 14, b = 3$ and $p = 17, q = 11$.

4 Counting

> I don't know how to count.
>
> *- Alessio Corti*

Combinations

If S is a set, I'll write $|S|$ = number of elements in S ($\in \mathbb{N} \cup \infty$) [the *order* of S]

Counting: the art of finding formulas for $|S|$ for different types of S.

> **Example 4.1.** How many 4-digit PINs are there? (e.g. 0741)
>
> There are 10 choices for the 1st digit, 10 choices for the 2nd digit etc. So the number of combinations is:
> $$10 \times 10 \times 10 \times 10 = 10,000$$

> **Example 4.2.** How many odd 2-digit PINs with 2 distinct digits are there?
>
> $$\underbrace{9}_{\text{choices for the 1st digit}} \times \underbrace{5}_{\text{choices for the 2nd (1,3,5,7,9)}} = 45$$
>
> Key point: Choose the second digit first.

First result:

No. of ways of ordering the numbers $1, 2, 3, \ldots, n$ (equally: ordering the objection of a set of n elements)
$$= n \times (n-1) \times (n-2) \times \cdots \times 1 = n!$$

Binomial coefficients

Notation.
$$\binom{n}{r} = \text{``}n \text{ choose } r\text{''}$$
$$= \text{No. of subsets of order } r \text{ of } \{1, 2, 3, \ldots, n\}$$

Proposition 4.3.
$$\binom{n}{r} = \frac{n!}{r!(n-r)!}$$

Proof. There are $n!$ factorial ways of ordering $1, 2, 3, \ldots, n$. We do the ordering by choosing r objects first, ordering those, and then ordering the remaining $n-r$ objects.

$$n! = \binom{n}{r} r!(n-r)!$$

∎

Note: we could take it as convention that $\binom{n}{0} = 1$.

Proposition 4.4.

$$(x + y)^n = x^n + nx^{n-1}y + \binom{n}{2}x^{n-2}y^2 + \binom{n}{3}x^{n-3}y^3 + \cdots + y^n$$

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

Proof. Write out

$$(x + y)^n = \underbrace{(x + y) \cdot (x + y) \cdot (x + y) \ldots (x + y)}_{n \text{ times}}$$

When you expand n times you select that the term $x^{n-k}y^k$ by choosing y from k of the brackets and there are $\binom{n}{k}$ ways to do that. ∎

Excercise: Prove, by induction or otherwise, that

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

Multinomial Coefficients

Definition. Let $S = \{1, 2, \ldots, n\}$ (or any set of order n). A *partition* of S is a collection $S_1, S_2, \ldots, S_k$ of subsets of S such that every $x \in S$ belongs to one and only one of these subsets.

Another way to say this is that $S = S_1 \cup S_2 \cup \cdots \cup S_k$ and $S_i \cap S_j = \phi$ for $i \neq j$.

Notation. $S = S_1 \coprod S_2 \coprod S_3 \coprod \cdots \coprod S_k$ ($\coprod$ is for disjoint union). An *ordered partition* is a partition where the sets $S_1, \ldots, S_k$ are ordered.

Example 4.5. For $n = 8$, the following are two different *ordered* partitions, though they're actually the same partition:

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

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

Definition. The no of ordered partitions of $\{1, 2, 3, \ldots, n\}$ into k subsets $S_1, S_2, \ldots, S_k$ of orders $r_1, r_2, \ldots, r_k$ is

$$\binom{n}{r_1, r_2, \ldots, r_k}$$

Proposition 4.6.

$$\binom{n}{r_1, r_2, \ldots, r_k} = \frac{n!}{r_1! r_2! \ldots r_k!}$$

Proof. Order the elements $1, 2, \ldots, n$ by first making an ordered partition $S_1, S_2, \ldots, S_k$ and then order the elements of S_1, then order the elements of $S_2, S_3, \ldots$:

$$n! = \binom{n}{r_1, r_2, \ldots, r_k} r_1! r_2! \ldots r_k!$$

∎

Remark 4.7. $r_1 + r_2 + \cdots + r_k = n$ and $\binom{n}{r, n-r} = \binom{n}{r}$

Example 4.8. Suppose 9 students are to be assigned to 3 projects S_1, S_2, S_3 which need 4, 2 and 3 students respectively. *In how many ways can the students be assigned to the projects?*

$$
\begin{aligned}
\binom{9}{4, 3, 2} &= \frac{9!}{4! 3! 2!} \\
&= \frac{5 \cdot 6 \cdot 7 \cdot 8 \cdot 9}{2 \cdot 6} \\
&= 4 \cdot 5 \cdot 7 \cdot 9 = 1260
\end{aligned}
$$

Theorem 4.9: Multinomial Theorem

$$
(x_1 + x_2 + \cdots + x_k)^n = \sum_{\substack{r_1, \ldots, r_k \geq 0 \\ r_1 + r_2 \cdots + r_k = n}} \binom{n}{r_1, r_2, \ldots r_k} x_1^{r_1} x_2^{r_2} \ldots x_k^{r_k}
$$

Proof. Expand the product:

$$
\underbrace{(x_1 + x_2 + \cdots + x_k) \cdot (x_1 + x_2 + \cdots + x_k) \ldots (x_1 + x_2 + \cdots + x_k)}_{n}
$$

Score $x_1^{r_1} x_2^{r_2} \ldots x_k^{r_k}$ by choosing x_1 from r_1 brackets, x_2 from r_2 brackets etc. Then there are precisely $\binom{n}{r_1, r_2, \ldots, r_k}$ ways to do that. ∎

Example 4.10. Compute the constant coefficient c_n of

$$
\left(x + y + z + \frac{1}{xyz} \right)^n
$$

Note that

$$
c_n = \begin{cases} 0 & \text{if } 4 \nmid n \\ \binom{4k}{k, k, k, k} & \text{if } n = 4k \end{cases}
$$

$$
\begin{aligned}
x^{r_1} y^{r_2} z^{r_3} \left(\frac{1}{xyz} \right)^{r_4} = \frac{x^{r_1} y^{r_2} z^{r_3}}{x^{r_4} y^{r_4} z^{r_4}} &= 1 \\
\iff r_4 = r_1 = r_2 = r_3 = k
\end{aligned}
$$

This monomial appears with coefficient $\binom{n}{k, k, k, k}$ when $4k = n$.
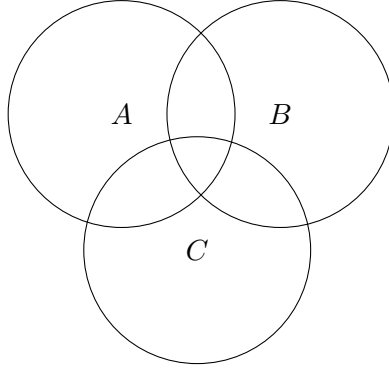
Inclusion-Exclusion Principle

It's obvious that

$$
|A \cup B| = |A| + |B| - |A \cap B|
$$

Similarly

$$
|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|
$$

Proof by picture:

When I write $|A| + |B| + |C|$ I count twice $x \in A \cap B$ (and $A \cap C, B \cap C$) I better take them out: $-|A \cap B| - |A \cap C| - |B \cap C|$. This creates a new problem with $x \in A \cap B \cap C$. It was counted 3 times the first time around. I also took it out 3 times. So I better add those back again: $+|A \cap B \cap C|$.

There is a more general version involving an arbitrary (finite) no. of sets:

Theorem 4.11: Inclusion-Exclusion Principle

Let A_i be subsets of a finite set X, for $1 \leq i \leq n$. Then

$$|\bar{A}_1 \cap \cdots \cap \bar{A}_n| = |X| - \sum_i |A_i| + \sum_{i<j} |A_i \cap A_j| - \cdots + (-1)^n |A_1 \cap \cdots A_n|.$$

Example 4.12. $S = \{n \mid 0 \leq n < 360, \text{hcf}(n, 360) = 1\}$. *What is* $|S|$? Consider

$$S^C \{n \mid 0 \leq n < 360 : \text{hcf}(n, 360) \neq 1\}$$

$360 = 2^3 \times 3^2 \times 5$, so $S^C = A_2 \cup A_3 \cup A_5$ where $A_k = \{u \mid 0 \leq n < 360 \ \& \ k \mid n\}$.

$$|S^C| = |A_2| + |A_3| + |A_5| - |A_6| - |A_{10}| - |A_{15}| + |A_{30}|$$
$$= 180 + 120 + 72 - 60 - 36024 + 12$$
$$= 262$$
$$\implies |S| = 360 - 264 = 96$$

This was the complicated way to use it as an illustration of the exclusion-inclusion principle. There was a simpler way to dos it, using the Chinese Remainder Theorem:

To give a number $n : 0 \leq n < 360 \iff$ To give 3 numbers $\begin{cases} 0 \leq a < 8 \\ 0 \leq b < 9 \\ 0 \leq c < 5 \end{cases}$

$$\begin{cases} n \equiv a \pmod 8 \\ n \equiv b \pmod 9 \\ n \equiv c \pmod 5 \end{cases}$$

$$\text{hcf}(n, 360) = 1 \iff \begin{cases} \text{hcf}(a, 8) = 1 \\ \text{hcf}(b, 9) = 1 \\ \text{hcf}(c, 5) = 1 \end{cases}$$

There are

$$\begin{array}{ll} 4 \text{ choices for } a & (1, 3, 5, 7) \\ 6 \text{ choices for } b & (1, 2, 4, 5, 7, 8) \\ 4 \text{ choices for } c & (1, 2, 3, 4) \end{array}$$

So $|S| = 4 \times 6 \times 4 = 96$.

Recall that

$$\mathbb{Z}/n\mathbb{Z}^{\times} = \{n \in \mathbb{Z}/n\mathbb{Z} \mid \exists x \in \mathbb{Z}/m\mathbb{Z} : ax \equiv 1 \pmod{n}$$
$$= \{[a] \in \mathbb{Z}/m\mathbb{Z} \mid \text{hcf}(a, n) = 1\}$$

$$\phi(n) = |\mathbb{Z}/n\mathbb{Z}^{\times}|$$

Proposition 4.13.

$$\phi(n) = n \prod_{p \mid n} \left(1 - \frac{1}{p} \right)$$

Proof. Step 1: ϕ is a multiplicative function, that is:

$$\text{hcf}(n, m) = 1 \implies \phi(n, m) = \phi(n)\phi(m)$$

Indeed by the Chinese Remainder Theorem,

$$\text{the system } \begin{cases} x & \equiv a \pmod{n} \\ x & \equiv b \pmod{m} \end{cases}$$

has a unqiue solution mod nm and (exercise!)

$$\text{hcf}(x, nm) = 1 \iff \text{hcf}(a, n) = 1 \,\& \, \text{hcf}(b, m) = 1$$

So $(\mathbb{Z}/nm\mathbb{Z})^{\times} = (\mathbb{Z}/n\mathbb{Z})^{\times} \times (\mathbb{Z}/m\mathbb{Z})^{\times}$ and $\phi(nm) = \phi(m)\phi(n)$
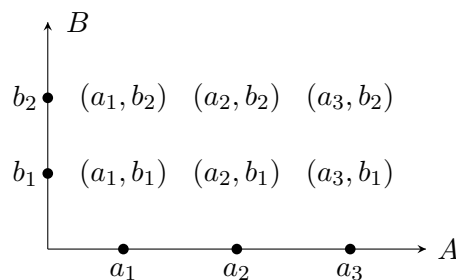
Step 2: If p is prime, then $\phi(p^a) = p^a - p^{a-1}$.

Indeed: $\mathbb{Z}/p^a\mathbb{Z}^{\times} = \{0 \le n < p \mid p \mid n\}^C$. ∎

Relations

Recall the Cartesian product of sets, $A \times B = \{(a, b,) \mid a \in A, b \in B\}$.

A picture: For $A = \{a_1, a_2, a_3\}$ $B = \{b_1, b_2\}$

The construction is called the Cartesian product by analogy with the construction of the 2-plane by Descartes, as the set of ordered pairs (a, b):

$$\text{Plane} = \{(a, b) : a \in \mathbb{R}, \, b \in \mathbb{R}\}$$
$$= \mathbb{R} \times \mathbb{R} = \mathbb{R}^2$$

The notation $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$ is by analogy with $r \times r = r^2$, where r is a number.

> **Definition.** A *relation* $\mathcal{R}$ on a set S is a subset $\mathcal{R} \subset S \times S$.
>
> We write $a \sim_{\mathcal{R}} b \iff (a, b) \in \mathcal{R}$.

Remember that it's an *ordered* pair; suppose you considered a relation between people: you could get x loves y as a relation, which is *not* automatically reflexive or transitive etc. The friend of your friend is not automatically your friend. In fact a friend of my best friend could be my best enemy. If x loves y it does not follow that y loves x - I'm sure you will have experienced this. And finally, most strangely, but true, it's not a fact that x loves x.

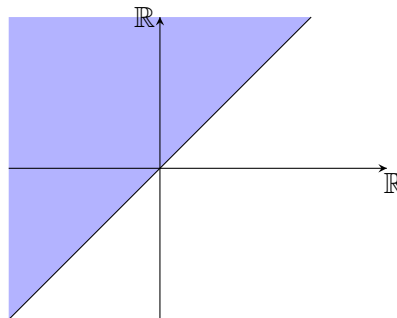**Notation.** The *diagonal* is $\triangle_S = \{(s, s) : s \in S\} \subset S \times S$.
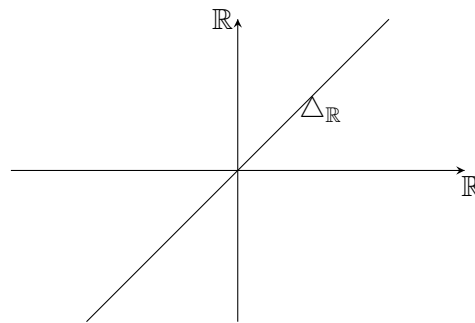
> **Examples 4.14.**
> (i) $\triangle_{\mathbb{R}} \subset \mathbb{R} \times \mathbb{R}$:
>
>
>
> This isn't a particularly interesting relation.
> (ii) $\mathcal{R} = \{(x, y) \mid x \leq y\} \subset \mathbb{R}^2$ [a relation on $\mathbb{R}$]
>
>
>
> $x \sim_{\mathcal{R}} y \iff x \leq y$

Warning. If $x \sim_{\mathcal{R}} y$, it does not follow that $y \sim_{\mathcal{R}} x$.

E.g. $\mathcal{R} = S \times S \backslash \triangle_S$. $(a, b) \in \mathcal{R} \iff a \neq b$.

Let me just ask this strange question:

Example 4.15. *How many relations are there on the set $S = \{1, 2\}$?*
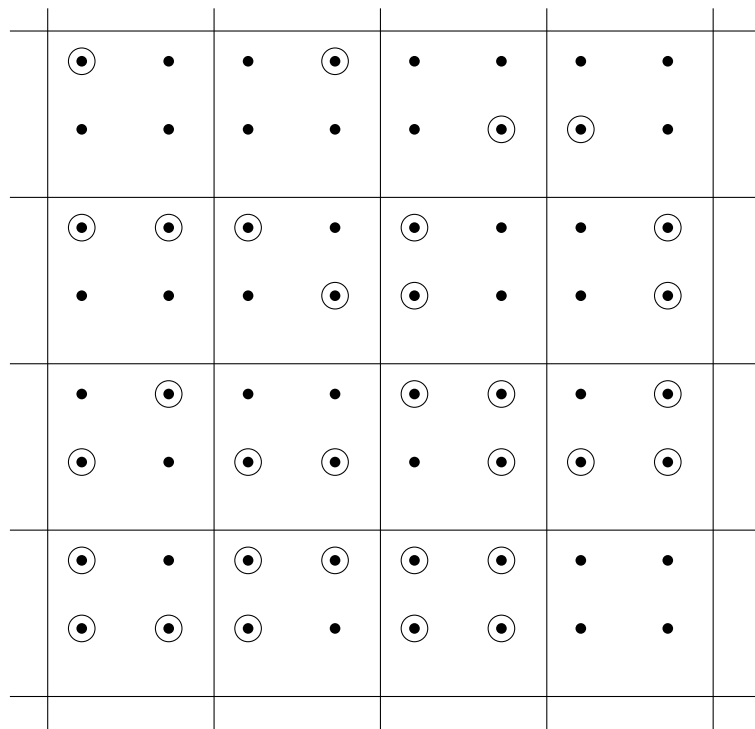
Remarks 4.16.

1. A finite set T has exactly $2^{|T|}$ subsets.

Let S be a set $|S| = n$. Let $P(s) = \{T \mid T \subset S\}$. Then $|P(s)| = 2^n$. Let $P_k(s) = \{T \mid T \subset S \& |T| = k\}$. Then we know $|P_k(s)| = \binom{n}{k}$. Then

$$|P(s)| = \sum k = 0^n \binom{n}{k} = 2^n$$

2. If S_1, S_2 are finite sets then $|S_1 \times S_2| = |S_1||S_2|$.

Consider now $S = \{1, 2\}$. A relation is a subset of $S \times S$, so here are $2^{|S \times S|} = 2^4 = 16$ relations on $\{1, 2\}$. Let us picture them all:

I'm always surprised to see people who take in inside themselves a foreign object, like a new mathematical idea for example... You have to find a way to take them inside yourself, otherwise it's pointless to be here. One way is to do lots of examples, to think of your own examples and to draw pictures.

Definition (Properties of relations).

(i) A relation $\mathcal{R}$ on S is *reflexive* if $\forall a \in S$, $a \sim_{\mathcal{R}} a$.

(ii) A relation is *symmetric* if $\forall a, b \in S, a \sim_{\mathcal{R}} b \implies b \sim_{\mathcal{R}} a$

(iii) A relation is *transitive* if $\forall a, b, c \in S, a \sim_{\mathcal{R}} b, b \sim_{\mathcal{R}} c \implies a \sim_{\mathcal{R}} c$.

Examples 4.17.

(i) $\mathcal{R}$ on $\mathbb{R} : \mathcal{R} = \{(x, y) \in \mathbb{R}^2 : x \neq y\}$.

This is symmetric ($x \neq y \implies y \neq x$). It is not reflective or transitive.

(ii) $\mathcal{R} = \{x, y \mid x < y\} \subset \mathbb{R}^2$ a relation on $\mathbb{R}$.

It is transitive, but not reflexive or symmetric.

(iii) $\mathcal{R} \{x, y \mid x \leq y\} \subset \mathbb{R}^2$ a relation on $\mathbb{R}$. This is reflexive, transitive but not symmetric.

Exercise For every $T \subset \{(i), (ii), (iii)\}$. Find some S, some $\mathcal{R}$ on S such that $\mathcal{R}$ has property $(*) \iff (*) \in T$.

[hint: see what you can do with 16 relations on $\{1, 2\}$]

Definition. $\mathcal{R} \subset S \times S$ is an *equivalence relation* iff $\mathcal{R}$ satisfies $(i), (ii), (iii)$.

Lecture 30   Recap: A relation on a set S is a subset $\mathcal{R} \subset S \times S$. $\mathcal{R}$ is an *equivalence relation* if:

(i) Reflexive: $\forall a \in S, (a, a) \in \mathcal{R}$

(ii) Symmetric: $(a, b) \in \mathcal{R} \iff (b, a) \in \mathcal{R}$

(iii) Transitive: $(a, b) \in \mathcal{R}$ and $(b, c) \in \mathcal{R} \implies (a, c) \in \mathcal{R}$.

Examples 4.18.

1. $(a, b) \in \mathcal{R} \iff a \equiv b \pmod{n}$
   (fix $n \in \mathbb{N}$, a relation on $\mathbb{Z}$)

2. $\mathcal{R}$ on $S = \mathbb{Z} \times (\mathbb{Z}/\{0\})$
   $((p, q), (N, M)) \in \mathcal{R} \iff \exists a, b \in \mathbb{Z}/\{0\}$ such that $pa = Nb$ and $qa = Mb$

Definition. If $a \in S$, the *class of a* is:

$$[a] = \{b \in S \mid (a, b) \in R\}$$

In Example 4.15.1: $[a] =$ the "class of a mod n"

In Example 4.15.2: $[(p, q)] = \frac{p}{q} \in \mathbb{Q}$

In both examples we have a good way to "picture" the set of classes, which are $\mathbb{Z}/n\mathbb{Z}$ and $\mathbb{Q}$ respectively.

Notation. If $R \subset S \times S$ is an *equivalence relation* on S then we write:

$$a, b \in S \ a \sim b \iff (a, b) \in R$$

We can then re-write (i), (ii), (iii) as:

(i) $a \sim a \quad \forall a \in S$

(ii) $a \sim b \iff b \sim a$

(iii) $a \sim b, b \sim a \implies a \sim c$

Remarks 4.19.

• If $a \sim b$ then $[a] = [b] \subset S$

• For all $a, b \in S$ then either $[a] = [b]$ or $[a] \cap [b] = \emptyset$ (obviously not both)

Proof. $[a] = \{b \mid a \sim b\} \subset S$.

Suppose $a \sim b$. Then $b \in [a]$. I show that $[b] \subset [a]$. Suppose $c \in [b]$. Then $b \sim c$. But $a \sim b, b \sim c \implies a \sim c \implies c \in [a]$. Similarly since $b \sim a$, $[a] \subset [b]$, so then $[a] = [b]$.

Now suppose $c \in [a] \cap [b]$.
This means $a \sim c$ and $b \sim c$, and then by what we just did $[a] = [c] = [b]$. ∎

> **Definition** (Quotient Set)**.** If $\mathcal{R}$ is an equivalence relation on S, then we denote the quotient set:
> $$S/\mathcal{R} = \{[a] \mid a \in S\}$$
> $$= \text{ the set of classes}$$

In Example 1, $\mathcal{R}$ is on $\mathbb{Z}$ where $a \sim b \iff a \equiv b \pmod{n}$. Then $\mathbb{Z}/\mathcal{R} = \mathbb{Z}/n\mathbb{Z}$.

In Example 2, $\mathbb{Z} \times (\mathbb{Z}\backslash\{0\})/\mathcal{R} = \mathbb{Q}$. This is the foundation upon which the rational numbers depend on, which is higher - Maths is a tree that grows downward.

Remark 4.20. What makes it easier to work with examples 1 and 2 is that every class has a *distinguished element*. In example 1, every $[a]$ has a unique element $\bar{a}$ such that $0 \le \bar{a} < n$. So we think of the quotient set

$$\mathbb{Z}/\mathcal{R}\{k \mid 0 \le k < n\}$$

In example 2, every $[(p,q)]$ has a unique element (p', q') such that $\text{hcf}(p', q') = 1$. This *is* the rational number $p'/q' \in \mathbb{Q}$ in lowest terms.

Recall that a *partition* of set S is a set $\{S_i \mid i \in I\}$ of subsets of S such that :

(i) $S = \bigcup_{i \in I} S_i$
(ii) $\forall i, j \in I$ either $S_i = S_j$ or $S_i \cap S_i = \emptyset$

> **Theorem 4.21**
>
> To give an equivalence relation on S is the same as to give a partition of S.

Proof. Suppose that $\{X_i \mid i \in I\}$ is a partition. Define the relation

$$\mathcal{R} = \{(a, b) \mid \exists i \in I \text{ such that } a, b, \in S_i\} \subset S \times S$$

Exercise: verify that $\mathcal{R}$ is an equivalence relation.

Suppose $\mathcal{R} \subset S \times S$ is an equivalence relation. Then $\{[a] \mid a \in S\}$ is a partition of S.

Exercise: verify that this set is a partition. Persuade yourself that the two constructions are inverses of each other:

$$\mathcal{R} \subset S \times S \longrightarrow \text{ form } \{[a] \mid a \in\}$$
$$\longrightarrow \text{ form } \{\mathcal{R}' = \{(a, b) \mid \exists c \in S \ (a, b) \in [c]\}$$
$$\implies \mathcal{R}' = \mathcal{R}$$

Start with
$$\{S_i \mid i \in I\} \longrightarrow \text{ form } \mathcal{R} = \{(a, b) \mid \exists i \in I \ a, b \in S_i\}$$
$$\longrightarrow \{[a] \mid a \in S\}$$
$$\text{then } \{[a] \mid a \in S\} = \{S_i \mid i \in I\} \quad ∎$$

5 Functions

It is amazing how economists get Nobel prizes by saying that when there is enough demand, the shit will turn up.

- Alessio Corti

Definition. Let X, Y be sets. A *function*, $f : X \to Y$, "f from X to Y" is a subset: $\Gamma_f \subset X \times Y$, sometimes called "the graph of f", such that

$$\forall x \in X, \, \exists! \, y \in Y : (x, y) \in \Gamma_f$$

This unique y is denoted $y = f(x)$ and is called the *value* of f at x.

The usual definition for a function is a "thing" that $\forall x \in X$ spits out a unique $y \in Y$ called $f(x)$.

Terminology. X is the *domain* of f. Y is the *range* of f. The *image* of f is the set

$$f(X) = \{y \in Y \mid \exists x \in X \text{ with } f(x) = y\}$$

Warning. Do not confuse the *image* with the *range*.

Example 5.1. The function $f : \mathbb{R} \to \mathbb{R}$ such that

$$f(x) = x^2$$

$(\Gamma_f = \{(x, y) \in \mathbb{R} \times \mathbb{R} \mid y = x^2\})$

We have range$(f) = \mathbb{R}$, and image$(f) = \mathbb{R}_{\geq 0} = \{y \mid y \geq 0\}$

There are two ways to make a new function from an old one.
Suppose $f : X \to Y$ is a function.

(i) Restriction of the domain.

If $A \subset X$, then $f_{|A} : A \to Y$ is the function such that $f_{|A}(a) = f(a)$ for all $a \in A$ (f is *restricted* to $A \subset X$)

(ii) Restriction of the range.

If im$(f) \subset B \subset Y$ then I can define a function $g : X \to B$ such that $g(x) = f(x)$ for all $x \in X$.

Examples 5.2.

(i) The function $g : \mathbb{R} \backslash \{0\} \to \mathbb{R}$ such that $g(x) = x^2$. Then $g = f_{|\mathbb{R} \backslash \{0\}}$ where f is as in Example 5.1.

(ii) The function $g : \mathbb{R} \to \mathbb{R}_{\geq 0}$, $g(x) = x^2$ is obtained from the function f in Example 5.1 by restricting its range.

(iii) The function $g : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ such that $g(x) = x^2$ is obtained from Example 5.1 by restricting the domain and the range.

Definition. $f : X \to Y$ is *injective* (a.k.a. 1-to-1) iff:

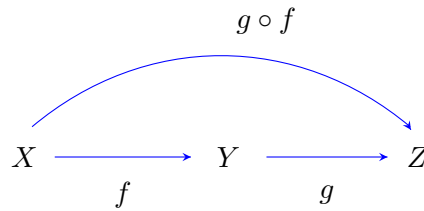$$\forall x_1, x_2 \in X, \; x_1 \neq x_2 \implies f(x_1) \neq f(x_2)$$

f is *surjective* (a.k.a. onto) iff: $\text{im}(f) = Y$, or in otherwords:

$$\forall y \in Y, \; \exists x \in X : f(x) = y.$$

Such an x need not be unique.

Definition. Let X, Y, Z be sets. Let $f : X \to Y$ and $g : Y \to Z$ be functions. Then the composition $g \circ f$ (read as "g composed f") is the function $g \circ f : X \to Z$ such that $\forall x \in X, \; (g \circ f)(x) = g(f(x))$.

Draw a picture:

Definition. Let X be a set. The *identity function* is the function

$$\text{id}_X : X \to X \text{ such that } \text{id}_X(x) = x \; \forall x \in X$$

Definition. Let $f : X \to Y$ be a function. Then f is *invertible* iff $\exists \, g : Y \to X$ such that $g \circ f = \text{id}_X$ and $f \circ g = \text{id}_Y$. Such a g is (necessarily unique) called the *inverse* of f and denoted $g = f^{-1}$.

Examples 5.3.

(i) $f : \mathbb{R} \to \mathbb{R}$ such that $f(x) = x^2$.

$g : \mathbb{R}_{\geq 0} \to \mathbb{R}$ such that $g(y) = \sqrt{y}$.

Note: $f \circ g = \mathbb{R}_{\geq 0} \to \mathbb{R}$. $f \circ g(y) = (\sqrt{y})^2 = y \neq \text{id}$ of anything.

On the other hand, it does not make sense to do $g \circ f$ because $\text{range}(f) = \mathbb{R} \neq \text{domain}(g) = \mathbb{R}_{\geq 0}$

(ii) Take $f : \mathbb{R} \to \mathbb{R}_{\geq 0}$ such that $f(x) = x^2$ and $g : \mathbb{R}_{\geq 0} \to \mathbb{R}$ such that $g(y) = \sqrt{y}$.

Now $f \circ g : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}, \; \forall y \in \mathbb{R}_{\geq 0}$. Now I can say $f \circ g = \text{id}_{\mathbb{R}_{\geq 0}}$.

I am also now allowed to take $g \circ f : \mathbb{R} \to \mathbb{R}$. Then $g \circ f(x) = \sqrt{x^2} = |x|$. This is not the same as $\text{id}_{\mathbb{R}}$.

The function g is not the inverse of f and vice-versa. In fact f is not invertible (g is neither).

(iii) Take $f : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ $f(x) = x^2$ and $g : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ $g(y) = \sqrt{y}$.

Now $f \circ g(y) = (\sqrt{y})^2 = y$. $g \circ f(x) = \sqrt{x^2} = x$. So $f \circ g = \mathrm{id}_{\mathbb{R}_{\geq 0}}$ and $g \circ f = \mathbb{R} \geq 0$. So $g = f^{-1}$ and $f = g^{-1}$.

Theorem 5.4

A function $f : X \to Y$ is invertible $\iff$ f is injective and surjective.

Proof. $\implies$:

f is injective: Suppose $x_1 \neq x_2$ and let $y_1 = f(x_1)$, $y_2 = f(x_2)$. Then $g(y_1) = x_1$ and $g(y_2) = x_2$, so $y_1 \neq y_2$.

f is surjective: Let $y \in Y$. Then $f \circ g(y) = y$. So $x = g(y) \in X$ is an element such that $f(x) = y$.

$\impliedby$:

Indeed $g : Y \to X$ such that $g \circ f = \mathrm{id}_X$ and $f \circ g = \mathrm{id}_Y$. Take

$$\Gamma_g = \{(y, x) \in Y \times X \mid (x, y) \in \Gamma_f\} \subset Y \times X$$

Γ is a function. Indeed

$$\forall y \in Y, \exists x \in X \text{ such that } (y, x) \in \Gamma_g$$

i.e. $y = f(x)$ because f is surjective. Moreover, this $x \in X$ is unique because f is injective. I leave it to you to show that $g \circ f = \mathrm{id}_X$ and $f \circ g = \mathrm{id}_Y$. ∎

Definition. A function f is *bijective* $\iff$ it is injective and surjective.

The last theorem also means bijective $\iff$ invertible.

Lecture 32 **Proposition 5.5.** *Let X, Y be finite sets and $f : X \to Y$ a function.*

$$
\begin{aligned}
f \text{ injective} &\implies |X| \leq |Y| \\
f \text{ surjective} &\implies |X| \geq |Y| \\
f \text{ bijective} &\implies |X| = |Y|
\end{aligned}
$$

Proof. Suppose $X = \{x_1, x_2, x_3, \ldots, x_m\}$. If f is injective, then $f(x_1), \ldots, f(x_m) \in Y$ are n distinct elements of Y. So $|Y| \geq n$.

Similarly if f is surjective and $Y = \{y_1, \ldots, y_m\}$ then there exists elements (necessarily distinct) $x_1, x_2, \ldots, x_m \in X$ with $f(x_i) = y_i$ for $i = 1, \ldots, n$. This shows $|X| \geq n$.

Finally if f is bijective, it's both injective and surjective, so $|X| \leq |Y|$ and $|Y| \leq |X| \implies |X| = |Y|$. ∎

Definition. Let Y, X be sets. The *power set* is

$$Y^X := \{f : X \to Y \mid f \text{ is a function}\}$$

This notation is designed to reflect this:

Lemma 5.6. *If X, Y are finite sets, then :*

$$|Y^X| = |Y|^{|X|}$$

Proof. Say $X = \{x_1, \ldots, x_m\}$. Then you have for a function $f : X \to Y$. There are

$$|Y| \text{ choices for } f(x_1)$$
$$|Y| \text{ choices for } f(x_2)$$
$$\vdots$$
$$|Y| \text{ choices for } f(x_n)$$

So overall:

$$|Y|^{|X|} = \underbrace{|Y| \times |Y| \times \cdots \times |Y|}_{n \text{ times}} \text{ choices for } f.$$

∎

Application: Remember that I told you: If S is a finite set, then S has $2^{|S|}$ subsets.
Let $Z = \{T : T \subset S\}$. Let $2 = \{0, 1\}$.

Claim. There is a "canonical" (natural) bijective function:

$$F : Z \to 2^S$$

Proof. Define F as the mapping of

$$T \longmapsto \chi_T : S \to \{0, 1\}$$

where χ_T is defined as

$$\chi_T(s) = \begin{cases} 1 & \text{if } s \in T \\ 0 & \text{if } s \notin T \end{cases}$$

Then F^{-1} is defined as the mapping of:

$$\chi : S \to \{0, 1\} \longmapsto T = \{s \in S \mid \chi(s) = 1\}$$

It is quite clear that F and F^{-1} are inverses of each other, in other words that F is invertible and so bijective. The whole thing implies that
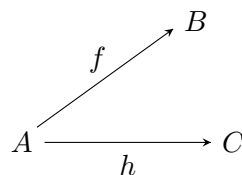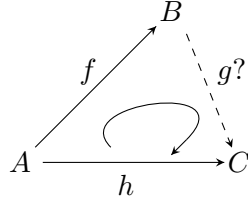
$$|Z| = |2^S| = 2^{|S|}$$

∎

Function maps

Question. Let A, B, C be sets.

Suppose we're given functions $f : A \to B$ and $h : A \to C$:
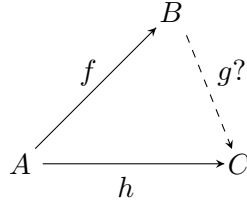
Is there a function $g : B \to C$, such that $h = g \circ f$?

Basic Point: Not always!

Example 5.7.

g does not exist here, as g would have to map b to both c_1 and c_2 - impossible! Formally:

Proof. Suppose that g existed:

Let $a_1, a_2 \in A$ and assume $f(a_1) = f(a_2)$. Then

$$h(a_1) = g(f(a_1)) = g(f(a_2)) = h(a_2) \qquad \blacksquare$$

Theorem 5.8

g exists if and only if

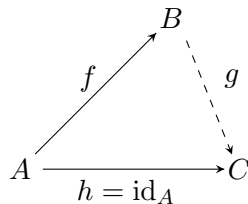$$\forall a_1, a_2 \in A, f(a_1) = f(a_2) \implies h(a_1) = h(a_2)$$

Proof. $\implies$: We already did this.

$\impliedby$: Choose any element $c \in C$. Suppose $b \in B$.

(i) If $b \notin im(f)$ then I declare $g(b) = c$

(ii) If $b \in im(f)$ then $\exists a \in A : f(a) = b$. In this case I declare $g(b) = h(a)$.

This function is well defined, i.e. independent of the choice of $a \in A$ such that $f(a) = b$. Indeed if $a' = A$ and $f(a') = b$ then $h(a) = h(a')$ $\qquad \blacksquare$
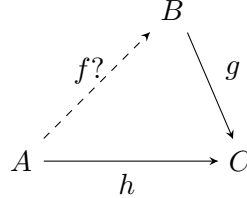
A special case: Take $h = \mathrm{id}_A$

i.e. $h(a) = a$ for all a. Then $\exists g : B \to A$ such that $g \circ f = \mathrm{id}_A$ if and only if $\forall a_1, a_2 \in A, f(a_1) = f(a_2) \implies a_1 = a_2$. i.e. $\exists g : B \to A$ such that $g \circ f = \mathrm{id}_A$ iff f is injective. Such a g is called a *left inverse of f*.

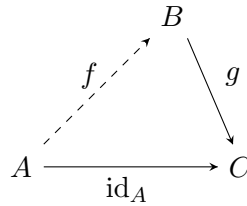We found: f has a left inverse iff f is injective.

A variation on the theme:

Let A, B, C be sets. Let there be functions $g : B \to C$, $h : A \to C$.

Is there $f : A \to B$ such that $g \circ f = h$?

Exercise: Show that f exists $\iff im(h) \subset im(g)$.

Special case: Take $h = \mathrm{id}_A$.

There exists f such that $g \circ f = \mathrm{id}_A \iff A \subset im(g)$ i.e. $A = im(g0$, i.e. g is surjective. Such an f is called a *right inverse* of g.

We found: g has a right inverse iff g is surjective.

Exercise: Suppose $f : A \to B$ is injective and surjective:

Then f has a left inverse. Prove that this left inverse is unique and it also has a right inverse, therefore it is the inverse of f. Also do the opposite:

Then f has a right inverse. Prove that this right inverse is unique, and it is also a left inverse, therefore it is the inverse of f.

Countability

> **Definition.** A set S is *countable* iff there exists a bijective function $f : \mathbb{N} \to S$.

Informally, this means that I can "list" all the elements of S:

$$S = \{s_1, s_2, s_3, s_4, s_5, \dots\}$$

more formally, $s_n = f(n)$.

N.B. A countable set is always infinite!

Proposition 5.9. *Suppose $S \subset \mathbb{N}$ is infinite. Then S is countable.*

Informally: Take $s_1 = minS$. s_2 is the next element: $s_2 = minS\backslash\{s_1\}$; s_3 as the next element: $s_3 = minS\backslash\{s_1, s_2\}$ and so on. This is making a list of all the elements of S. More formally:

Proof. Define $f : \mathbb{N} \to S$ inductively as follows:

- $f(1) = minS$
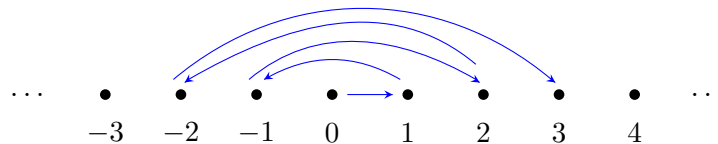
- Assume $f(1), \ldots, f(n-1)$ is defined already.

    Then define
    $$f(n) = minS\backslash\{f(1), \ldots, f(n-1)\}$$

One an prove by induction that this procedure does define a function. One does need a bit of work to show this function is bijective. ∎

Proposition 5.10. *$\mathbb{Z}$ is countable.*

Idea:

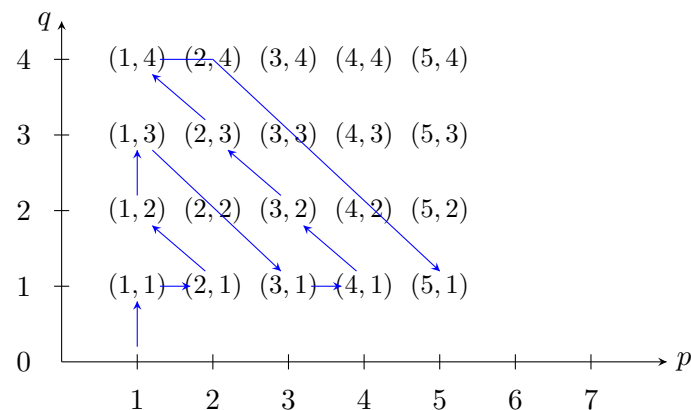*Proof.* Define a bijection $f : \mathbb{N} \to \mathbb{Z}$ by declaring:

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

Exercise: Show that f is indeed bijective. ∎

Proposition 5.11. *$\mathbb{Q}$ is countable.*

Sketch of the usual proof: Consider $\mathbb{Q}_+$ and then arrange pairs $(p, q) \in \mathbb{N}^2$ in a square:

There are two disadvantages to this: You have to cross out pairs which aren't in the lowest terms, and then it is difficult to write down an explicit formula for the kth element. This is just not good enough to write down a bijective formula from $\mathbb{N}$ to $\mathbb{Q}$. I'm going to do something a bit more rigorous:

Proof. Let me show first that

$$\mathbb{Q}_+ = \{x = m/n \in \mathbb{Q} : x > 0\}$$

is countable: Define $f : \mathbb{Q}_+ \to \mathbb{N}$ as $f(m/n) \to 2^n 3^m.$[1] f is injetive. f is not surjective, but it is surjective to $S = im(f)$, so $\mathbb{Q}_+$ is a bijective correspondence with $S \subset \mathbb{N}$. By Proposition 6.1, S is countable:

$$\mathbb{Q}_+ \longrightarrow S \longrightarrow \mathbb{N}$$

So $h = g \circ f : \mathbb{Q}_+ \to \mathbb{N}$ is bijective.

To finish off, we prove as an exercise the following: If A is countable, and B is countable, then $A \cup B$ is countable. This implies that $Q = Q_- \cup \{0\} \cup \mathbb{Q}_+$ is countable   ∎

Theorem 5.12

$\mathbb{R}$ is uncountable.

There are different degrees of infinities. This sort of bothers me a little bit, but maybe I shouldn't go there.

Proof. (Cantor's Diagonal Argument) Suppose for a contradiction that you could "list" all the real numbers. Let this be the list:

$$x_1 = a_1 \cdot a_{11}\, a_{12}\, a_{13}\, a_{14} \ldots$$
$$x_2 = a_2 \cdot a_{21}\, a_{22}\, a_{23}\, a_{24} \ldots$$
$$x_3 = a_3 \cdot a_{31}\, a_{32}\, a_{33}\, a_{34} \ldots$$
$$x_4 = a_4 \cdot a_{41}\, a_{42}\, a_{43}\, a_{44} \ldots$$
$$\vdots$$
$$x_m = a_m \cdot a_{m1}\, a_{m2}\, a_{m3}\, a_{m4} \ldots$$

Here $a_1, a_2, a_3 \in \mathbb{Z}$ and $a_{11}, a_{12}, \ldots, a_{ij} \in \{0, 1, 2, \ldots, 9\}$.

We produce a real number $x \in \mathbb{R}$ which is not on the list:

   (i) Pick $b_1 \in \mathbb{Z}$ such that $b_1 \neq a_1$

  (ii) Pick $c_1 \in \{0, 1, \ldots, 8\}$ such that $c_1 \neq a_{21}$

 (iii) Pick $c_2 \in \{0, 1, \ldots, 8\}$ such that $c_2 \neq a_{32}$

$$\vdots$$

  (n) Pick $c_n \in \{0, 1, \ldots, 8\}$ such that $c_n \neq a_{n+1\,n}$

$$\vdots$$

---

[1]Or just use $f(m/n) \to 2^m(2n - 1)$ which is already a bijection onto $\mathbb{N}$.

Notice: We don't allow 9, so we don't end up with the number ending Let $x = b_1 \cdot c_1 c_2 \ldots c_n \ldots$.

Then x is not on the list. $x \neq$ every number on the list (at least one digit in the decimal expansion is different!) ∎

I'm not a big fan of countability and uncountability, so this is all I'm going to say about this topic. Philosophically I am uncomfortable with there being an infinite set strictly bigger than the set of natural numbers. I have a hard time making this ontological commitment to a higher degree of infinity than the smallest I need to accept. There are ways to justify this position mathematically, but they have some catches, and that wouldn't be the mainstream view so I'm not going to go there.

6 Foundations of Analysis

> If you really understood this course, you won't need to do any revision. Just
> relax and have some beer.
>
> *- Alessio Corti*

The Completeness Axiom

Definition. Let $\emptyset \neq S \subset \mathbb{R}$. Any element $x \in S$ is the *maximum* of S:

$$x = \max S \text{ if } \forall y \in S, \ y \leq x$$

Similarly $x \in S$ is the *minimum* of S

$$x = \min S \text{ if } \forall y \in S, \ x \leq y$$

It makes sense to call it *the* maximum as there can only be one.

Remark 6.1. If $\emptyset \neq S \subset \mathbb{R}$ then let $\phi \neq -S = \{-x \mid x \in S\} \subset \mathbb{R}$, $\min S = -\max(-S)$.

Warning. Make sure $S \neq \emptyset$, as the empty set has no maximum.

Examples 6.2.

   (i) $S = \{x \mid x \geq 0\}$ then S has no maximum, $\min S = 0$.

  (ii) $S = (0, 1)$ We would like to say that 1 is the max, but this is not true because $1 \notin S$. So S has no maximum. S has no min either. Clearly $x = 1$ is an "important" number for the set $(0, 1)$. The next few definitions address this.

Definition. $\emptyset \neq S \subset \mathbb{R}$ is *bounded above* if

$$\exists M \in \mathbb{R} \text{ s.t. } \forall x \in S, \ x \leq M$$

such an M is called an *upper bound* for S.

S is *bounded below* if

$$\exists M \in \mathbb{R} \text{ such that } \forall x \in S, M \leq x$$

Such an M is called a *lower bound*.

S is *bounded* if S is *bounded above* and *below*.

Equivalently S is bounded if

$$\exists R > 0 \text{ such that } \forall x \in S, -R \leq x \leq R$$

or equivalently

$$\exists R > 0 \text{ such that } \forall x \in S, |x| \leq R$$

Examples 6.3. $S = \{x \in \mathbb{R} \mid x \geq 0\}$ is bounded below but not above.

$S = \{x \in \mathbb{R} \mid x \leq 0\}$ is bounded above, not below.

$S = (0,1) = \{x \in \mathbb{R} \mid 0 < x < 1\}$ is bounded (above and below).

Definition. Suppose $\emptyset \neq S \subset \mathbb{R}$ is bounded above. The *least upper bound* of S a.k.a. the "sup" (*supremum*) of S, is the smallest of all upper bounds (assuming that it exists!), i.e.

$$x = \sup S = \min T \text{ where } \emptyset \neq T = \{y \mid y \text{ is an upperbound for } S\}$$

Similarly the *greatest lower bound* of S, or "inf" (*infimum*) is the largest of all lower bounds (assuming that it exists!) i.e.

$$x = \inf S = \max T \text{ where } \emptyset \neq T = \{y \mid y \text{ is a lower bound for } S\}$$

Example 6.4. $S = (0,1)$. Then $T = \{y \mid y \text{ is an upperbound for } S\}$.

T has a minimum: $1 = \min T = \sup(0,1)$. Similarly $0 = \inf(0,1)$.

Axiom 6.5 (Completeness of $\mathbb{R}$)**.** Suppose that $\emptyset \neq S \subset \mathbb{R}$ is bounded above. Then S has a supremum.

This is an article of faith about the real numbers. There are very strange things that I would like to discuss that I won't tell you as I don't want to scare your innocent minds. You may want to use set theory to construct the real numbers and prove that the completion axiom is a theorem there. But what's worse is that if you study logic carefully, then you cannot state this axiom in "first order logic", but I'm not going to go there; it's something that bothers me but it shouldn't bother you. Just believe it.

Consequence: If $\emptyset \neq S \subset \mathbb{R}$ is bounded below then S has an inf. Indeed: S is bounded below $\implies -S$ is bounded above and $\inf S = -sup(-S)$.

Exercise. There is a real number $x \in \mathbb{R}$ such that $x^2 = 2$. In other words, $\sqrt{2}$ exists as a real number.

Hint:

(i) Let $r = sup\{x \in \mathbb{R} \mid x^2 < 2\}$. You next try to show $r^2 = 2$. To do that the following characterisation of sup S may help.

(ii) $L = \sup S$ iff

    a) L is an upperbound ($\forall x \in S$, $x \leq L$)

    b) $\forall \epsilon > 0, \epsilon \in \mathbb{R}, \exists x \in S$ such that $L - \epsilon < x$

Lecture 27   So I introduced sup S for $\emptyset \neq S \subset \mathbb{R}$ bounded above:

$$sup S = \min\{t \in \mathbb{R} : t \text{ an upper bound for } S\}$$
$$= \min\{t \in \mathbb{R} \mid \forall x \in S, x \leq t\}$$

Criterion to help recognising sup S

$L = sup S$ if and only if

(1) L is an upper bound for S, i.e. $\forall x \in S, x \leq L$

(2) $\forall \epsilon > 0, \exists x \in S$ such that $L - \epsilon < x$¿

I think the idea is easy to understand:

Think: ϵ is *small*, then $L - \epsilon$ is a little bit smaller than L; then (2) is just saying that $L - \epsilon$ is not an upperbound for S.

[Recall that M is an upperbound for $S \iff \forall x \in S, x \leq M$. M is not an upperbound for $S \iff \exists x \in S$ such that $M < x$]

Equivalent formulation of (2) is

$$\forall \epsilon > 0,\ L - \epsilon \text{ is not an upper bound for } S$$

also equivalent

(2†) $\forall L', L' < L, L'$ is not an upper bound for S.

Proof of the Criterion. Suppose that L is sup S. Then (1) is obvious (the min of a set by definition belongs to that set.)

(2) is done by the definition of min: If L' is an upper bound $\implies L \leq L'$, or equivalently L' is not an upperbound $\impliedby L' < L$. This shows that (2†) holds, hence (2) holds.

For the converse, suppose that (1) and (2†) hold. I want to show $L = \sup S$. i.e. $L = \min T$, where $T = \{t \in \mathbb{R} \mid t$ an upper bound of S. (1) is saying $L \in T$. (2†) is saying $L' \in T \implies L \leq L'$. i.e. indeed $L = \min T$. ∎

Application:

> **Theorem 6.6: Existence of Square-Root**
>
> $\forall y \in \mathbb{R}, y \geq 0, \exists! x \in \mathbb{R}, x \geq 0$ such that $x^2 = y$.

Remark 6.7.

(1) $x = \sqrt{y}$. This shows the existence of the square-root function.

(2) Using the same ideas, you can show $\forall n \geq 2, n \in \mathbb{N}$:

$$\forall y \in \mathbb{R}, y \geq 0, \exists! x \in \mathbb{R}, x \geq 0 \text{ such that } x^m = y$$

(i.e. the existence of the n-th root function.)

Proof. We are supposed to do this using sup.

Fix $y \geq 0$. Let $x = sup\{t \geq 0 \mid t^2 \leq y\}$. In order for this to make sense, we need to check that $T = \{t \geq 0 \mid t^2 \leq\}$ is bounded above. [$0 \in T$, so $T \neq \emptyset$]

Indeed $M = \max\{1, y\}$ is an upper bound for T. (think about this as an exercise).

Claim. $x^2 = y$.

Step 1: $x^2 \leq y$. I want to use property (2) of the criterion for sup. If x' is a little smaller than x then x' is not an upperbound for T. So

$$\exists t \in T : x' < t \implies x'^2 < y^2 \leq y.$$

So again: $\forall x', x' < x \implies x'^2 \le y$. In other words, $\forall \epsilon > 0, (x - \epsilon)^2 \le y$

$$\implies x^2 - 2\epsilon \le (x - \epsilon)^2 = x^2 - 2\epsilon x + \epsilon^2 \le y$$

i.e. $\forall \epsilon > 0$:

$$x^2 \le y + 2\epsilon x \le y + 2M\epsilon$$

Exercise: Let $x_1, x_2 \in \mathbb{R}, A > 0$. Suppose $\forall \epsilon > 0, x_1 \le x_2 + A\epsilon$. Then $x_1 \le x_2$.

Once you've done this, you will have shown $x^2 \le y$ i.e. Step 1.

Step 2: $y \le x^2$. We take this back to property (1) of the criterion for sup. x is an upper bound for $T = \{t \mid t^2 \le y\}$.

Suppose for contradiction $x^2 < y$. Then I claim x is not an upper bound for T, i.e.

$$\exists \epsilon > 0 \text{ such that } x + \epsilon \le T$$

i.e.

$$\exists \epsilon > 0 \text{ such that } (x + \epsilon)^2 < y$$

In other words:

$$\text{If } x^2 < y, \text{ then } \exists \epsilon > 0 \text{ such that } (x + \epsilon)^2 < y$$

Compute

$$(x + \epsilon)^2 = x^2 + 2\epsilon x + \epsilon^2$$

provided that I choose my ϵ to be ≤ 1, this is $\le x^2 + 2\epsilon x + \epsilon = x^2 + \epsilon(2x + 1)$.

Formally, I choose:

$$0 < \epsilon < \min \left\{ 1, \frac{y - x^2}{2x + 1} \right\}$$

then I claim $(x + \epsilon)^2 < y$ (and this finishes the proof).

Indeed:

$$\begin{aligned}
(x + \epsilon)^2 &= x^2 + 2\epsilon x + \epsilon \\
&\le x^2 + 2\epsilon x + \epsilon \\
&= x^2 + \epsilon(2x + 1) \\
&< x^2 + \frac{y - x^2}{2x + 1} 2x + 1 = y
\end{aligned}$$

■

* Sequences and Limits *

The rest of these notes are unexamined in M1F (but are in M1P1, where it's recovered).

> **Definition.** A *sequence* of real numbers is an infinite list of real numbers:
>
> $$a_1, a_2, a_3, \ldots, a_n, \ldots$$
>
> one for each positive integer n.

Notation. We denote a sequence $(a_n)_{n \in \mathbb{N}^\times}$, or just (a_n).

Note the "round" parentheses $(,)$ to distinguish from the notation for sets that uses $\{,\}$ curly brackets.

Remark 6.8. A sequence is a function $a : \mathbb{N}^\times \to \mathbb{R}$ with $a(n) = a_n$.

Examples 6.9.

(i) $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots, \frac{1}{n}, \ldots$

$$a_n = \frac{1}{n}$$

(ii) $1, 0, 1, 0, 1, 0, 1, \ldots$

$$a_n = \frac{1 - (-1)^n}{2} = \begin{cases} 1 \text{ if } n \text{ is odd} \\ 0 \text{ if } n \text{ is even} \end{cases}$$

(iii) $1, 1, 2, 3, 5, 8, \ldots$ - the Fibonacci sequence.

The easiest way to define this is by induction:

$$\begin{cases} a_1 & = a_2 = 1 \\ a_n & = a_{n-1} + a_{n-2} \quad \text{for } n \geq 3 \end{cases}$$

(iv) $1, 1 + \frac{1}{2}, 1 + \frac{1}{2} + \frac{1}{3}, \ldots$.

$$a_n = \sum_{k=1}^{n} \frac{1}{k}$$

(v) $\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \ldots$

$$a_n = \frac{n}{n+1}$$

I want to define the notion of a limit of a sequence (a_n) as $n \to \infty$. ("n tends to ∞")

A little preparation

Here $l \in \mathbb{R}, \epsilon > 0$. The statement $|x - l| < \epsilon$ means "x is within ϵ of l", another way to say this is :

$$-\epsilon < x - l < \epsilon \iff l - \epsilon < x < l + \epsilon$$

in other words $x \in$ open window of radius ϵ centred at l.

> **Definition.** Let (a_n) be a sequence and $l \in \mathbb{R}$. I say that $a_n \to l$ ("a_n tends to l") as $n \to \infty$ and I write this as $\lim_{n\to\infty} a_n = l$ iff:
>
> $$\forall \epsilon > 0, \exists N \in \mathbb{N} \text{ such that } n \geq N \implies |a_n - l| < \epsilon$$

Intuitively: If n is large, then a_n is very close to l.

> **Example 6.10.** $a_n = \frac{1}{n} \to 0$ as $n \to \infty$
>
> *Proof.* Fix $\epsilon > 0$.
>
> Let N be any integer such that $N > \frac{1}{\epsilon}$. (we know such N exists by the Archimedean axiom.) If $n \geq N$ then
>
> $$|a_n - 0| = \frac{1}{n} \leq \frac{1}{N} < \epsilon$$

■

Example 6.11. Let $a_n = \dfrac{1-(-1)^n}{2}$, i.e. $a_n = \begin{cases} 1 \text{ if } n \text{ is odd} \\ 0 \text{ if } n \text{ is even} \end{cases}$

This sequence tends to *no limit*.

This is rather awkward to prove directly from the definition, becuase you would have to show

$$\forall l \in \mathbb{R}, \; \exists > 0 \forall N, \exists n \geq N \text{ s.t. } |a_n - l| \geq \epsilon$$

But you try.

Proposition 6.12 (Uniqueness of limits). *Let (a_n) be a sequence. Suppose $l_1 = \lim_{n\to\infty} a_n$ and $l_2 = \lim_{n\to\infty} a_n$. Then $l_1 = l_2$.*

Proof. Fix $\epsilon > 0$. By the definition of a limit

$$\exists N_1 \text{ such that } n \geq N_1 \implies |a_n - l_1| < \epsilon/2$$

$$\exists N_2 \text{ such that } n \geq N_2 \implies |a_n - l_2| < \epsilon/2$$

Consider $N = max\{N_1, N_2\}$. If $n \geq N \implies |a_n - l_1| < \epsilon/2$ and $|a_n - l_2| < \epsilon/2$.

Now look at this:
$$\begin{aligned} |l_1 - l_2| &= |l_1 - a_n + a_n - l_2| \\ &\leq |a_n - l_1| + |a_n - l_2| \\ &< \epsilon/2 + \epsilon/2 = \epsilon \end{aligned}$$

So $|l_1 - l_2| < \epsilon \implies |l_1 - l_2| = 0$, i.e. $l_1 = l_2$. ■

Strategy to prove $a_n \to l$ directly

Step 1: Fix $\epsilon > 0$.

Step 2: Compute $|a_n - l|$

STEP 3: Solve (for n) the inequality $|a_n - l| < \epsilon$.

Step 4: Choose $N \in \mathbb{N}$ such that $n \geq N \implies |a_n - l| < \epsilon$

Step 5: Turn this "upside down" and write your own proof.

Example 6.13. Show $a_n = \dfrac{n+5}{n+1} \to 1$.

Proof. Fix $\epsilon > 0$. Let $N = \lceil \frac{4}{\epsilon} - 1 \rceil + 1$. Now if $n \geq N$, then

$$n > \frac{4}{\epsilon} - 1 \implies n + 1 > \frac{4}{\epsilon}$$

$$\implies \frac{4}{n+1} = |a_n - l| < \epsilon.$$

■

Felina. Show that $a_n = \dfrac{n^2 - 5}{n^3 - 50} \to 0$

Rough work. We want to prove that $a_n \to 0$. Compute:

$$|a_n - l| = \frac{|n^2 - 5|}{|n^3 - 50|}$$

It is going to be awkward to solve for n:

$$\frac{|n^2 - 5|}{|n^3 - 50|} < \epsilon$$

We need to modify STEP 3 and 4:

$(2\frac{1}{2})$ Find $b_n \geq 0$ and $n_0 \in \mathbb{N}$ such that $n \geq n_0 \implies |a_n - l| \leq b_n$.

(3') Solve for n, $b_n < \epsilon$.

(4') Choose $N \geq n_0 \in \mathbb{Z}$ such that $n \geq N \implies b_n < \epsilon$.

(5) Turn this upside down and write your proof.

$(2\frac{1}{2})$ Choose b_n, now $|a_n - l| = \left| \dfrac{n^2 - 5}{n^3 - 50} \right|$. I choose $b_n = \frac{2}{n}$.

I need n_0 such that $\left| \dfrac{n^2 - 5}{n^3 - 50} \right| \leq b_n$ for all $n \geq n_0$.

$n_0 \geq 3$, so numerator is positive. We want $n^3 - 50 \geq \frac{n^3}{2}$.

Note: If $n \geq 4$ then both the numerator $n^2 - 5 > 0$ and the denomiator $n^3 - 50 > 0$. I am definitely going to take $n_0 \geq 4$, so

$$\left| \frac{n^2 - 5}{n^3 - 50} \right| = \frac{n^2 - 5}{n^3 - 50}$$

and then I want to fix

Now we're ready to put everything together into a proof:

Proof. Fix $\epsilon > 0$. Let $N = \max\{\lceil \frac{2}{\epsilon} \rceil + 1, 5\}$.

Claim: If $n \geq 5$, then $|a_n - l| \leq 2/n$.

Proof of Claim.

$$|a_n - l| = \frac{|n^2 - 5|}{|n^3 - 50|} = \frac{n^2 - 5}{n^3 - 50}$$

$$\leq \frac{n^2}{n^3/2} = 2/n \;//$$

by the claim, if $n \geq N$, then

$$|a_n - l| \leq \frac{2}{n} < \frac{2}{2/\epsilon} = \epsilon \hspace{3cm} \blacksquare$$

- End of Foundations of Analysis -