Numbers and Sets

Prof. W.T. Gowers

Michaelmas 2004

LaTeXed by Sebastian Pancratz

These notes are based on a course of lectures given by Prof. W.T. Gowers in Part IA of the Mathematical Tripos at the University of Cambridge in the academic year 2004–2005.

These notes have not been checked by Prof. W.T. Gowers and should not be regarded as official notes for the course. In particular, the responsibility for any errors is mine — please email Sebastian Pancratz (`sfp25`) with any comments or corrections.

Contents

Chapter 1

Introduction

Aim

To introduce the (university-level) mathematical way of thinking, talking, writing, etc.

For example, we shall look at

- precise, formal definitions
- rigorous proofs
- foundational questions.

Numbers

Informal concepts of "number"

- expression of quantity
- 1, 2, 3, 4, ... etc.
- binary operations — things like $+, -, \times, \div$
- a point on the number line.

Number systems

- $\mathbb{N}$ is the set of all *natural numbers*, that is, 1, 2, 3, 4, ... etc.
- $\mathbb{Z}$ is the set of all *integers*, i.e., etc. ..., -3, -2, -1, 0, 1, 2, 3, 4, ... etc.
- $\mathbb{Q}$ is the set of *rational numbers*, i.e., all fractions $\frac{p}{q}$ where p, q are integers and $q \neq 0$.
- $\mathbb{C}$ is the set of all *complex numbers*, i.e., numbers of the form $a + ib$ where a, b are real numbers and $i = \sqrt{-1}$.

Example. A real number that is irrational — $\sqrt{2}$. We shall prove

$$\sqrt{2} \text{ is irrational,}$$

or rather that there are no two positive integers p, q such that

$$\left(\frac{p}{q}\right)^2 = 2 \quad \text{i.e.} \quad p^2 = 2q^2.$$

We use *proof of contradiction*, or *reduction ad absurdum*. That is, assume that the statement to be proved is false, and show that this has impossible consequences.

Proof. Assume that $p^2 = 2q^2$. We may assume also that p and q are not both even, since otherwise keep dividing by 2 until one or the other is odd, e.g. $\frac{140}{96} = \frac{70}{48} = \frac{35}{24}$.

So we find p and q, not both even, such that $p^2 = 2q^2$. But then p^2 is even, so p is even[1]. So write $p = 2r$. Then $(2r)^2 = 2q^2$ so $4r^2 = 2q^2$, $2r^2 = q^2$, so q^2 is even, so q is even. Contradiction. □

This shows that not every polynomial has a root in $\mathbb{Q}$, even if it is sometimes negative and sometimes positive.

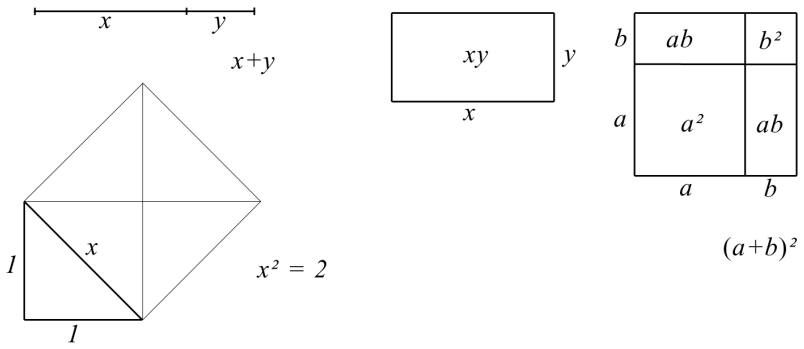

Figure 1.1: Greek way of thinking about numbers.

To "show that $\sqrt{2}$ exists" one must "construct" a larger number system $\mathbb{R}$ and *prove*, given that construction, that there is some x in $\mathbb{R}$ such that $x^2 = 2$.

A real number is *algebraic* if it is a root of some polynomial with integer (or, equivalently, rational) coefficients. E.g. $\sqrt{2}$, since it is a root of $x^2 - 2$, or $\frac{1+\sqrt{5}}{2}$ or $\sqrt[3]{2}$.

Non-algebraic real numbers are called *transcendental*.

Inventing new systems

Bad example,

$$0x = 1, 0\infty = 1$$
$$1 = 0\infty = (0+0)\infty = 0\infty + 0\infty = 1 + 1 = 2$$

[1] p is odd $\implies p = 2r+1 \implies p^2 = 4r^2 + 4r + 1$

Good example (complex numbers),

$$x^2 + 1 = 0$$

The fundamental theorem of algebra states that every non-constant polynomial (with real or complex coefficients) has a root in $\mathbb{C}$.

E.g. $x^5 + x + 1$,

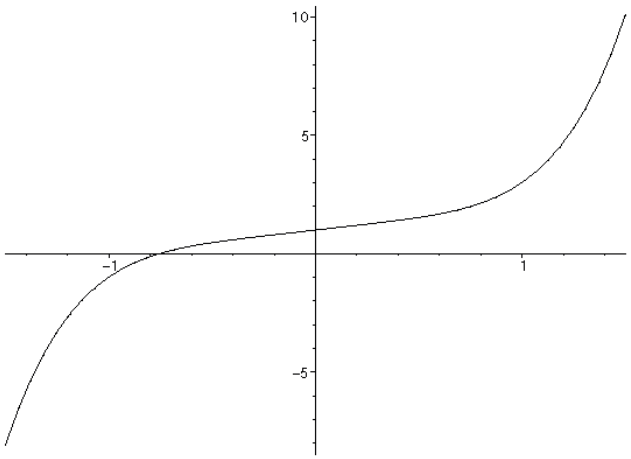

Figure 1.2: Real root of $x^5 + x + 1$.

E.g. $x^{20} + 8x^4 + 1$

Chapter 2

Elementary logic

We shall look at the meanings of the words AND, OR, IMPLIES, NOT, ALL, SOME. We shall use letters to stand for *statements*.

If P and Q are statements, then P and Q is the statement that is true if and only if P is true and Q is true.

P or Q is true if and only if P is true or Q is true (or both). This is sometimes written $P \vee Q$.

P implies Q, sometimes written $P \implies Q$, is true if and only if it is not the case that P is true and Q is false. E.g.

$$\text{`}n \text{ is a prime } > 2\text{'} \implies \text{`}n \text{ is odd'}.$$

Note that the statement

"Let n be a positive integer that is both even and odd. Then $n = 17$."

is true.

Truth tables

The meanings of $\wedge, \vee, \implies$ can be captured in tabular form as follows.

P	Q	$P \wedge Q$	$P \vee Q$	$P \implies Q$
T	T	T	T	T
T	F	F	T	F
F	T	F	T	T
F	F	F	F	T

These can be used to prove logical equivalences. For example, the technique of proof by contradiction relies on the equivalence between $P \implies Q$ and $\neg Q \implies \neg P$ (where $\neg$ is the symbol for NOT). $\neg Q \implies \neg P$ is the *contrapositive* of $P \implies Q$.

P	Q	$P \implies Q$	$\neg P$	$\neg Q$	$\neg Q \implies \neg P$
T	T	T	F	F	T
T	F	F	F	T	F
F	T	T	T	F	T
F	F	T	T	T	T
		↑			↑

To avoid confusion, ¬ should be thought of as short for "it is not the case that".

¬ (Every Cambridge don is over the age of 50.)

means the same as

At least one Cambridge don is under the age of 50 or exactly 50.

Chapter 3

Sets

A *set* is a collection of "objects" usually of the mathematical kind, such as numbers or points in space etc. The objects are called the *elements*. If x is an element of a set A then we write $x \in A$.

3.1 Notation for sets

(i) Listing the elements (inside curly brackets).
$\{1, 2, 5\}$ is the set whose elements are 1, 2 and 5.

(ii) Implied lists.
$\{1, 2, 4, 8, \ldots, 2^{2^{20}}\}$ is the set consisting of all numbers 2^k where $0 \leq k \leq 2^{20}$.

(iii) Definition in terms of *properties*.
$\{n : n - 1$ is a perfect square$\}$
The set of all n such that some property holds. More formally,

$$\{n \in \mathbb{N} : n - 1 \text{ is a perfect square}\}.$$

3.2 Russell's Paradox

It isn't true that every property defines a set. E.g. consider the property "is not an element of itself". A set A has that property iff $A \notin A$. Imagine there were a set $B = \{A : A \notin A\}$. Then, is $B \in B$?

If $B \in B$, then B does not have the property, so $B \notin \{A : A \notin A\}$, so $B \notin B$, so B does have the property, so $B \in \{A : A \notin A\}$ so $B \in B$. Contradiction.

There is no such thing as the set of all sets.

You don't get into difficulties if you take an existing set X and define A inside X by some property:

$$A = \{x \in X : \text{ the property is true for } x\}$$

3.3 Definitions

Definition (Set equality)**.** Two sets A and B are equal if $A \subset B$ and $B \subset A$. I.e., every element of A is an element of B and vice versa.

This is what you use when you come up with two sets and need to show that they are equal.

Definition (Empty set)**.** $\emptyset$, or the *empty set*, is the set with no elements.

Definition (Subset)**.** We write $A \subset B$ (A is a *subset* of B) if every element of A is an element of B.

Notation (Subsets)**.** $A \subset B$ A is a subset of B
$A \subsetneq B$ A is a proper subset of B

Definition (Intersection)**.** The *intersection* of A and B, written $A \cap B$, is $\{x : x \in A \text{ and } x \in B\}$.

Example. $\emptyset \subset \{17\}$

$$\{1,3,5,7,9,\dots\} \cap \{2,4,6,8,\dots\} = \emptyset$$
$$\implies \{1,3,5,7,9,\dots\} \cap \{2,4,6,8,\dots\} \subset \{17\}$$

Definition (Union)**.** Let A and B be sets. The *union* of A and B, written $A \cup B$, is $\{x : x \in A \text{ or } x \in B\}$.

Example. $\{1,3,5\} \cup \{4,5,6\} = \{1,3,4,5,6\}$

Definition (Complement)**.** If you have a designated universal set X and A is a subset of X, then the *complement* of A, written A^C, is $\{x : x \notin A\}$, when it is understood that this means $\{x \in X : x \notin A\}$.

Definition (Difference)**.** If A and B are sets then the *difference* $A \backslash B$ is $\{x \in A : x \notin B\}$. E.g. if $A = \{1,3,5\}$ and $B = \{4,5,6\}$ then $A \setminus B = \{1,3\}$ and $B \setminus A = \{4,6\}$.

Definition (Symmetric difference)**.** The *symmetric difference* of two sets A and B, written $A \triangle B$, is

$$\begin{aligned} &\{x : x \in A \text{ or } x \in B \text{ but not both}\} \\ &= (A \cup B) \setminus (A \cap B) \\ &= (A \setminus B) \cup (B \setminus A) \end{aligned}$$

We can illustrate these concepts on *Venn diagrams*:

3.4 Union and intersection of more than two sets

If $A_1, A_2, \dots, A_n$ are sets then

$$\bigcap_{i=1}^{n} A_i = A_1 \cap A_2 \cap \cdots \cap A_n = \{x : x \in A_i \text{ for every } i\}$$
$$\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \cdots \cup A_n = \{x : x \in A_i \text{ for some } i\}$$

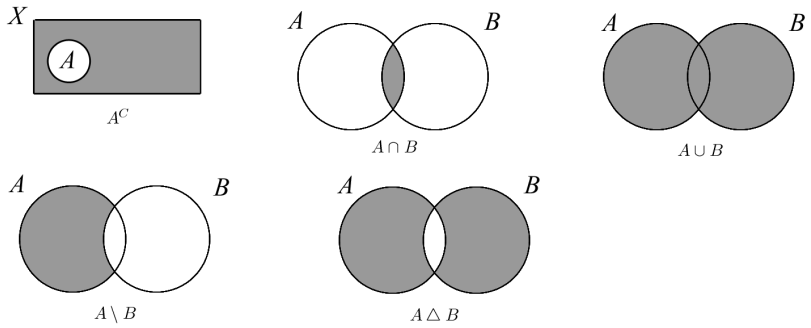

Figure 3.1: Venn diagrams.

Similarly, one can define

$$\bigcap_{i=1}^{\infty} A_i \qquad \text{and} \qquad \bigcup_{i=1}^{\infty} A_i.$$

More generally, if Γ is any set (called an *indexing* set) and for each $\gamma \in \Gamma$ we have a set A_γ then

$$\bigcap_{\gamma\in\Gamma} A_\gamma = \{x : x \in A_\gamma \text{ for every } \gamma \in \Gamma\}$$
$$\bigcup_{\gamma\in\Gamma} A_\gamma = \{x : x \in A_\gamma \text{ for some } \gamma \in \Gamma\}$$

3.5 Some basic facts about sets

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

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

(iii) de Morgan's laws,
$(A \cap B)^C = A^C \cup B^C$
$(A \cup B)^C = A^C \cap B^C$

More generally,
$A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$
$A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$

(iv) $A \triangle (B \triangle C) = (A \triangle B) \triangle C$

(v) The distributive laws,
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

3.6 How to prove set identities

We shall illustrate this with the examples

$$A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$$
$$\text{and } A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C).$$

One approach is via truth tables. (The columns are to be read as "x is an element of".)

A	B	C	$B \cup C$	$A \setminus (B \cup C)$	$A \setminus B$	$A \setminus C$	$(A \setminus B) \cap (A \setminus C)$
T	T	T	T	F	F	F	F
T	T	F	T	F	F	T	F
T	F	T	T	F	T	F	F
T	F	F	F	T	T	T	T
F	T	T	T	F	F	F	F
F	T	F	T	F	F	F	F
F	F	T	T	F	F	F	F
F	F	F	F	F	F	F	F
				↑			↑

To show

$$A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C).$$

Proof. Let $x \in A \setminus (B \cap C)$. Then $x \in A$ but $x \notin B \cap C$. If $x \notin C$ then $x \in A \setminus C$ and therefore $x \in (A \setminus B) \cup (A \setminus C)$. If $x \notin B$ then $x \in A \setminus B$ and therefore $x \in (A \setminus B) \cup (A \setminus C)$. But one or the other of these must be true or else we would have $x \in B \cap C$. So $x \in (A \setminus B) \cup (A \setminus C)$.

Now let $x \in (A \setminus B) \cup (A \setminus C)$. If $x \in A \setminus B$ then $x \in A$ and $x \notin B$. Since $x \notin B$, $x \notin B \cap C$ so $x \in A \setminus (B \cap C)$. Similarly if $x \in A \setminus C$. □

3.7 Functions

Given a set A and another set B, a function f from A to B is a way of assigning to each $x \in A$ an element $y \in B$. We write $y = f(x)$ and say that y is the *image* of x.

A is called the *domain* of f. B is called the *range* of f. We write $f\colon A \to B$ for the statement that f is a function with domain A and range B.

If $y = f(x)$ we also write $f\colon x \mapsto y$ and say x *maps to* y.

Two functions f and g are equal if they have the same domain A, the same range B and $f(x) = g(x)$ for every $x \in A$.

Example. (i) $f\colon \mathbb{R} \to \mathbb{R}, x \mapsto x^2$

(ii) $g\colon \mathbb{R} \to \{x \in \mathbb{R} : x \geq 0\}, x \mapsto x^2$

Let $f\colon A \to B$. If $x \in A$ and $f\colon x \mapsto y$ then we write $y = f(x)$. y is called the *image* of x. x is called *a preimage* of y.

Example.

$$f\colon \mathbb{R} \to \mathbb{R}, x \mapsto x^2$$

Then the image of 6 is 36, and the preimages of 36 are 6 and -6.

If $X \subset A$ then we define

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

This is called the *image* of X. (Note, this has a different meaning of "image" since $X \subset A$ rather than $X \in A$.) $f(A)$ is also called the *image* of f (not the same as the range). [In some books they say "codomain" instead of "range" and "range" instead of "image".]

If $Y \subset B$ then we define

$$f^{-1}(Y) = \{x \in A : f(x) \in Y\}.$$

This is called the *inverse image* of Y. (In general, f^{-1} isn't anything.)

Example.

$$f\colon \mathbb{R} \to \mathbb{R}, x \mapsto x^2$$

The image of $\{x : -1 \leq x < 4\}$ is $\{x : 0 \leq x < 16\}$.
The inverse image of $\{x : -1 \leq x < 4\}$ is $\{x : -2 < x < 2\}$.

Definition (Composition)**.** If $f\colon A \to B$ and $g\colon B \to C$ then we can define a function $g \circ f$, called the *composition* of f with g, by

$$\begin{aligned} g \circ f\colon\ & A \to C \\ & x \mapsto g(f(x)). \end{aligned}$$

Definition (Inverses)**.** If $f\colon A \to B$ and there is a function $g\colon B \to A$ such that

(i) $gf(x) = x$ for every $x \in A$

(ii) $fg(y) = y$ for every $y \in B$

then g is called the *inverse* of f and is denoted by f^{-1}.

By no means all functions have inverses.

Example. Let $\mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$.

(i) $f\colon \mathbb{R}_+ \to \mathbb{R}_+, x \mapsto x^2$
Then the inverse of f is $g\colon \mathbb{R}_+ \to \mathbb{R}_+, x \mapsto \sqrt{x}$.

(ii) $f\colon \mathbb{R} \to \mathbb{R}_+, x \mapsto x^2$
f has no inverse because $f(-2) = f(2) = 4$ so any inverse g would have to satisfy $g(4) = 2$ and $g(4) = -2$, which is impossible.
(Note we do have a *right-inverse*, namely $g\colon \mathbb{R}_+ \to \mathbb{R}, x \mapsto \sqrt{x}$ since $f(g(x)) = x \quad \forall x \in \mathbb{R}_+$.)

(iii) $f\colon \mathbb{R}_+ \to \mathbb{R}, x \mapsto x^2$
f has no inverse because if it did we would need $(g(-5))^2 = -5$, which is impossible. Problem: not all elements of $\mathbb{R}$ have preimages.

Definition (Injection)**.** Let $f\colon A \to B$. f is an *injection* if no elements of B have more than one preimage. Equivalently, f is an injection if

$$f(x) = f(x') \implies x = x'.$$

(This is the most convenient formulation for showing that f is an injection.) Sometimes one says f is one-to-one.

Definition (Surjection)**.** f is a *surjection* (or onto map) if very element of B has at least one preimage. Equivalently, it is a surjection if $f(A) = B$.

Definition (Bijection)**.** f is a *bijection* (or one-to-one correspondence) if it is both an injection and a surjection — i.e. every element of B has exactly one preimage.

Definition (Identity function)**.** If A is a set, then the *identity function* on A, written ι_A (or I_A or 1_A or ι_A or ...) is defined by

$$\iota_A\colon A \to A, x \mapsto x.$$

Then in the definition of inverses (i) says $g \circ f = \iota_A$ and (ii) says $f \circ g = \iota_B$. In the first case, we say that g is a *left-inverse* for f, and in the second case that it is a *right-inverse*.

Proposition 3.1. Let $f\colon A \to B$, $A \neq \emptyset$.

(i) f is an injection if and only if f has a left-inverse.

(ii) f is a surjection if and only if f has a right-inverse.

(iii) f is a bijection if and only if f has an inverse.

Proof. (i) Suppose that $g\colon B \to A$ is a left-inverse of f. Let $f(x) = f(y)$. Then $g(f(x)) = g(f(y))$, so $x = y$ (as g is a left-inverse). This shows that f is an injection.

Now suppose that f is an injection. Let $y \in B$. Define $g(y)$ to be x if $f(x) = y$ (which can happen for at most one x, as f is an injection) and otherwise let $g(y)$ be anything. Then $gf(x) = x$ for all $x \in A$.

(ii) Let $g\colon B \to A$ be a right-inverse of f. Let $y \in B$. Then $g(y)$ is a preimage of y since $fg(y) = y$ since g is a right-inverse. So f is a surjection.

Now suppose that f is a surjection and let $y \in B$. Let $g(y)$ be some preimage of y. Then $fg(y) = y$ so g is a right-inverse for f.

(iii) If g is an inverse for f then by (i) and (ii) f is a bijection. If f is a bijection let $g(y)$ be the unique preimage of y. Then g is an (= the) inverse of f.

□

Example. Functions don't have to have nice definitions, e.g.

$$f\colon \mathbb{R} \to \mathbb{R}$$
$$x \mapsto \begin{cases} x+1 & \text{if } x \text{ is an integer} \\ 13 & \text{if } x = \pi \\ \sqrt{x} & \text{if } 10\tfrac{1}{2} \leq x \leq 10\tfrac{3}{4} \\ x^3 + 16x + \sqrt{2} & \text{otherwise} \end{cases}$$

3.8 Cartesian products

Definition (Cartesian product)**.** If A and B are two sets then their *cartesian product* $A \times B$ is the set

$$\{(x, y) : x \in A \text{ and } y \in B\}.$$

An *ordered pair* (x, y) is a bit like a set except that repeats are allowed and order matters. The main point is that $(x, y) = (z, w)$ if and only if $x = z$ and $y = w$.

Example. $\{1, 3\} \times \{2, 4, 6\} = \{(1, 2), (1, 4), (1, 6), (3, 2), (3, 4), (3, 6)\}$

Definition (Graph)**.** If $f\colon A \to B$ then the *graph* Γ of f is the set

$$\{(x, f(x)) : x \in A\}.$$

Γ is a subset of $A \times B$. It has the property that for every $x \in A$ there is exactly one $y \in B$ such that $(x, y) \in \Gamma$. Conversely, given any Γ with this property, one can define $f\colon A \to B$ by $x \mapsto$ the y such that $(x, y) \in \Gamma$. This is often adopted as the formal definition of a function.

3.9 Relations

Definition (Relation)**.** A *relation* on a set A is better described as a potential relationship. R is a relation on A, if, when $x, y \in A$ and you write xRy, then you obtain a sentence that may be true or false.

Example. If $A = \mathbb{N}$ or $\mathbb{Z}$ or $\mathbb{Q}$ or $\mathbb{R}$ then $=, <, \leq, >, \geq$ are all relations.

A relation is

- *reflexive* if xRx for every $x \in A$.
- *symmetric* if xRy always implies yRx.
- *transitive* if xRy, yRz always implies xRz.

Example. (i) $<$ on $\mathbb{R}$
is not reflexive, since for example $17 \not< 17$.
is not symmetric, e.g. $1 < 2$ but $2 \not< 1$.
is transitive.

(ii) on $\mathbb{R}$, define xRy to mean $|x - y| < 1$,
is reflexive and symmetric, but not transitive (e.g. $0R\frac{1}{2}$, $\frac{1}{2}R1$, $0\not R1$.)

3.10 Equivalence relations and partitions

Definition (Equivalence relation)**.** A relation R is an *equivalence relation* if it is reflexive, symmetric and transitive.

Definition (Partition)**.** Let A be a set. Then a *partition* of A is a collection $\{B_\gamma : \gamma \in \Gamma\}$ of subsets of A such that every $x \in A$ belongs to exactly one B_γ. The sets B_γ are called the *cells* of the partition.

Example. Let $A = \mathbb{R}^2$ and for each $t \in \mathbb{R}$ let B_t be the line $\{(x, y) \in \mathbb{R}^2 : x = t\}$. Then $\{B_t : t \in \mathbb{R}\}$ is a partition of $\mathbb{R}^2$.

Definition (Equivalence class)**.** If $\sim$ is an equivalence relation on A and $x \in A$ then the *equivalence class* of x is $\{y \in A : x \sim y\}$.

Proposition 3.2. Let $\sim$ be an equivalence relation on a set A. Then the equivalence classes form a partition of A.

Proof. Write E_x for the equivalence class of x. Let $x \in A$. Then $x \in E_x$ since $x \sim x$, since $\sim$ is reflexive. Now suppose that $E_x \cap E_y \neq \emptyset$. We must show that $E_x = E_y$.

Let $s \in E_x$. We can find some $z \in E_x \cap E_y$. We are given $x \sim s$, $x \sim z$, $y \sim z$ and must prove $y \sim s$. Since $x \sim z$ and $\sim$ is symmetric, $z \sim x$. Since $y \sim z$, $z \sim x$ and $\sim$ is transitive, $y \sim x$. Since $y \sim x$, $x \sim s$ and $\sim$ is transitive, $y \sim s$. This shows $E_x \subset E_y$. Similarly, $E_y \subset E_x$.

So overlapping equivalence classes are equal. □

Remark. Given a partition $\{B_\gamma : \gamma \in \Gamma\}$ of a set A, we can define an equivalence relation $\sim$ by $x \sim y$ iff x and y lie in the same B_γ.

Definition (Quotient set)**.** If $\sim$ is an equivalence relation on a set A, then we write $A/\sim$ for the set of all equivalence classes of $\sim$. This is called the *quotient set*.

Given R on A we can define $X \subset A \times A$ by $\{(x, y) \in A \times A : xRy\}$. Conversely, given $X \subset A \times A$, we can define R by xRy iff $(x, y) \in X$.

3.11 Binary Operations

Definition (Binary operation)**.** A *binary operation* on a set A is a means of combining pairs of elements of A to produce new ones.

More formally, it is a function from $A \times A$ to A.

Example.

- $+$ and $\times$ on the sets $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$
- $-$ on the sets $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$
- $\div$ on the sets $\mathbb{Q} \setminus \{0\}, \mathbb{R} \setminus \{0\}, \mathbb{C} \setminus \{0\}$

Properties of binary operations

Let $*$ be a binary operation on a set A.

- $*$ is *commutative* if $\forall x, y \in A \quad x * y = y * x$
- $*$ is *associative* if $\forall x, y, z \in A \quad x * (y * z) = (x * y) * z$

- An element $e \in A$ is called an *identity* for $*$ if $\forall x \in A \quad e * x = x * e = x$. If e and f are both identities then $e = e * f = f$, so any identity must be *unique*.
- If $*$ has an identity e and $x \in A$, then y is an *inverse* for x if $x * y = y * x = e$.
- If $*$ and $\square$ are two binary operations on A, then $*$ is *distributive over* $\square$ if

$$\begin{aligned} \forall x, y, z \in A \quad & x * (y \square z) = (x * y) \square (x * z) \\ \text{and} \quad & (y \square z) * x = (y * x) \square (z * x). \end{aligned}$$

Example. (i) On $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ or $\mathbb{C}$,
$+$ and $\times$ are commutative and associative. $\times$ is distributive over $+$.

(ii) On $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ or $\mathbb{C}$,
0 is the identity for $+$ and $-x$ is an inverse for x.

(iii) A binary operation with non-unique inverses.

	e	x	y
e	e	x	y
x	x	e	e
y	y	e	y

(iv) If A is $\{B : B \subset X\}$ then $\cup$ and $\cap$ are commutative, associative and each is distributive over the other. $\emptyset$ is an identity for $\cup$ and X is an identity for $\cap$.

(v) If X is a set and $A = \{f \colon X \to X \text{ :f is a bijection}\}$ then $\circ$ (composition) is associative, not commutative (unless X has at most 2 elements). The identity function ι_X is the identity for $\circ$, and f^{-1} (which exists $\forall f$, since f is a bijection) is a unique inverse for f (in the binary-operations sense!).

Definition (Group)**.** A *group* is a set A together with a binary operation $*$ that is associative and has an identity and is such that every element has an inverse. (So the last example was a group.)

Chapter 4

Induction and counting

4.1 Principle of mathematical induction

The principle of mathematical induction is as follows.

PMI 1. Let $P(1), P(2), \ldots$ be a sequence of statements. Suppose that $P(1)$ is true, and for every k, $P(k) \implies P(k+1)$. Then $P(n)$ is true for every n.

It has two other useful formulations.

PMI 2. Let $P(1), P(2), \ldots$ as before. Suppose that $P(1)$ is true, and that for every k, $P(1) \wedge \cdots \wedge P(k) \implies P(k+1)$. Then $P(n)$ is true for every n.

Well-ordering principle. Every non-empty subset of $\mathbb{N}$ has a least element.

PMI 1 $\iff$ PMI 2
It is easy to see that PMI 2 $\implies$ PMI 1. Indeed, if the assumptions of PMI 1 hold then the assumptions of PMI 2 hold. Now suppose that the assumptions of PMI 2 hold. For each n, let $Q(n)$ be the statement $P(1) \wedge \cdots \wedge P(n)$. Then $Q(1)$ is true, and $\forall k \quad Q(k) \implies Q(k+1)$. So, by PMI 1, $Q(n)$ is true for every n. But $Q(n) \implies P(n)$.

PMI 2 $\implies$ WOP
Let $A \subset \mathbb{N}$. Suppose that A does not have a smallest element. We shall prove (by PMI 2) that $A = \emptyset$.

Let $P(n)$ be the statement $n \notin A$. Then $P(1)$ is true, since otherwise 1 would be the smallest element of A. If $P(1), \ldots, P(k)$ are all true, then none of $1, \ldots, k$ belong to A. So $k+1 \notin A$ since otherwise it would be the smallest element, i.e. $P(1) \wedge \cdots \wedge P(k) \implies P(k+1)$. So by PMI 2, $\forall n \quad P(n)$, which says $A = \emptyset$.

WOP $\implies$ PMI 2
Assume $P(1)$ and that $\forall k \quad P(1) \wedge \cdots \wedge P(k) \implies P(k+1)$. We would like to show that $P(n)$ is true for all n. If this is not true, then $A = \{n : P(n) \text{ is false}\} \neq \emptyset$, so A has a least element n, by WOP. But then $P(1), \ldots, P(n-1)$ are all true, so $P(n)$ is true, by hypothesis. This contradiction implies PMI 2.

Another useful fact: Every non-empty set $A \subset \mathbb{N}$ that is *bounded above* (which means that there is some $M \in \mathbb{N}$ that is larger than every element of A) has a largest element.

To see this, let $B = \{M - n : n \in A\}$. Then $B \subset \mathbb{N}$, $B \neq \emptyset$, so B has a least element m. Then $M - m$ is the largest element of A.

4.2 Permutations and combinations

Definition (k-permutation)**.** Let X be a set of size n. Then a *k-permutation* of X is

- informally, a k-tuple of elements with no repeats allowed
- formally, an injection $\phi\colon \{1,\dots,k\} \to X$.

Example. If $X = \{1,2,5,7\}$ then the 2-permutations are $(1,2)$, $(1,5)$, $(1,7)$, $(2,1)$, $(2,5)$, $(2,7)$, $(5,1)$, $(5,2)$, $(5,7)$, $(7,1)$, $(7,2)$, $(7,5)$. The pair $(1,7)$ would be the function $\phi\colon \{1,2\} \to \{1,2,5,7\}, 1 \mapsto 1, 2 \mapsto 7$.

Lemma 4.1. The number of k-permutations of X is $\frac{n!}{(n-k)!}$.

Proof. We will prove this by induction on k.
The number of 1-permutations is clearly n.
Suppose that we have shown that there are $\frac{n!}{(n-k)!}$ k-permutations. Then for each k-permutation $(x_1,\dots,x_k)$ there are precisely $(n-k)$ $(k+1)$-permutations $(y_1,\dots,y_k,y_{k+1})$ with $y_1 = x_1,\dots,y_k = x_k$, since there is one for each $y_{k+1} \in X \setminus \{x_1,\dots,x_k\}$. Therefore, the number of $(k+1)$-permutations is $(n-k)\frac{n!}{(n-k)!} = \frac{n!}{(n-(k+1))!}$. □

Definition (k-combination)**.** A *k-combination* of a set X is a set $A \subset X$ of size k.

Lemma 4.2. The number of k-combinations of an n-element set is $\frac{n!}{k!(n-k)!}$.

Proof. Let us regard k-permutations as injections $\phi\colon \{1,\dots,k\} \to X$.
Let $P(n,k)$ be the set of all k-permutations.
Let $C(n,k)$ be the set of all k-combinations.
Define a function

$$f\colon P(n,k) \to C(n,k) \text{ by}$$
$$f\colon \phi \mapsto \mathrm{Im}(\phi).$$

(I.e. for each k-permutation ϕ let $f(\phi)$ be the set of values it takes.)

Now let $A \in C(n,k)$. Then A has $k!$ preimages, since there are precisely ${}^k\mathrm{P}_k = k!$ injections from $\{1,2,\dots,k\}$ to A (with image A).

Hence, the size of $P(n,k)$ is $k!$ times the size of $C(n,k)$. So the result is proved. □

Lemma 4.3 (Vandermonde's theorem)**.**

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

Proof. (i) Check it boringly.

(ii) $\binom{n+1}{k+1}$ is the number of subsets of $\{1,\dots,n+1\}$ of size $k+1$. Of those $\binom{n}{k+1}$ do not contain the element $(n+1)$ and $\binom{n}{k}$ do include the element $(n+1)$.

□

Notation (k-combinations)**.** ${}^n\mathrm{C}_k$, $\binom{n}{k}$.

Theorem 4.4.

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

Proof. Let us count in two different ways the number of subsets of $\{1, 2, \ldots, n\}$.

First, if we count subsets of size $0, 1, 2, \ldots, n$ in turn, then we get the LHS.

But we can also that it is 2^n, by induction. True when $n = 1$ (the subsets of $\{1\}$ are $\emptyset$ and $\{1\}$). If it is true for $n = k$, then every subset $A \subset \{1, 2, \ldots, k\}$ produces two subsets A, $A \cup \{k+1\}$ of $\{1, 2, \ldots, k+1\}$ so it is true for $k+1$ as well. □

Theorem 4.5 (The binomial theorem)**.**

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

Proof. Induction on n. Obvious if $n = 1$.

Suppose we know it for n. Then

$$\begin{aligned}(x+y)^{n+1} &= (x+y)(x+y)^n \\ &= (x+y)\left(x^n + \binom{n}{1}x^{n-1}y + \cdots + \binom{n}{n-1}xy^{n-1} + y^n\right).\end{aligned}$$

When we expand the right-hand side, the coefficient of $x^{n+1-k}y^k$ is

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

This is because the $x^{n+1-k}y^k$ term comes from $x \cdot \binom{n}{k}x^{n-k}y^k$ and $y \cdot \binom{n}{k-1}x^{n-(k-1)}y^{k-1}$. By Vandermonde's theorem, this is $\binom{n+1}{k}$, which is the correct coefficient for $n+1$. □

This gives another proof that $\binom{n}{0} + \cdots + \binom{n}{n} = 2^n$: just apply the binomial theorem to $(1+1)^n$.

4.3 The Inclusion-Exclusion Formula

Notation (Cardinality)**.** The number of elements in a set A is written $|A|$.

Example.

$$\begin{aligned}|A \cup B| &= |A| + |B| - |A \cap B| \\ |A \cup B \cup C| &= |A| + |B| + |C| - |A \cap B| - |B \cap C| - |A \cap C| + |A \cap B \cap C|\end{aligned}$$

Theorem 4.6 (The Inclusion-Exclusion Formula)**.** Let $A_1, \ldots, A_n$ be a collection of finite sets. Then

$$|A_1 \cup \cdots \cup A_n| = \sum_{k=1}^{n}(-1)^{k+1} \sum_{1 \le i_1 < \cdots < i_k \le n} |A_{i_1} \cap \cdots \cap A_{i_k}|.$$

Proof. Let x be an element of $A_1 \cup \cdots \cup A_n$. We must show that x contributes 1 to the right-hand side.

Let $\Gamma = \{i : x \in A_i\}$. Then $x \in A_{i_1} \cap \cdots \cap A_{i_k}$ iff $\{i_1, \ldots, i_k\} \subset \Gamma$. So the number of $i_1 < \cdots < i_k$ such that $x \in A_{i_1} \cap \cdots \cap A_{i_k}$ is $\binom{m}{k}$, where $m = |\Gamma|$. So the contribution of x to the RHS is

$$\sum_{k=1}^{n} (-1)^{k+1} \binom{m}{k} = \binom{m}{1} - \binom{m}{2} + \binom{m}{3} - \cdots + (-1)^{m+1} \binom{m}{m}.$$

But $0 = (1-1)^m = 1 - \binom{m}{1} + \binom{m}{2} - \cdots - (-1)^{m+1}\binom{m}{m}$ so $\binom{m}{1} - \binom{m}{2} + \cdots + (-1)^{m+1}\binom{m}{m} = 1$, as we wanted. □

Definition (Characteristic function)**.** If X is a given universal set and $A \subset X$ then the *characteristic function* of A, denoted χ_A, is the function

$$\chi_A \colon X \to \{0, 1\}$$
$$x \mapsto \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

It is easy to see that

$$\chi_{A \cap B}(x) = \chi_A(x) \cdot \chi_B(x)$$

or, more concisely,

$$\chi_{A \cap B} = \chi_A \chi_B.$$

Also, $\chi_{A^C} = 1 - \chi_A$.
It follows that

$$\begin{aligned} \chi_{A \cup B} &= \chi_{(A^C \cap B^C)^C} = 1 - \chi_{A^C \cap B^C} = 1 - (1 - \chi_A)(1 - \chi_B) \\ &= \chi_A + \chi_B - \chi_{A \cap B}. \end{aligned}$$

More generally,

$$\begin{aligned} \chi_{A_1 \cup \cdots \cup A_n} &= \chi_{(A_1^C \cap \cdots \cap A_n^C)^C} = 1 - \chi_{A_1^C \cap \cdots \cap A_n^C} \\ &= 1 - (1 - \chi_{A_1}) \cdots (1 - \chi_{A_n}) \\ &= \sum_{k=1}^{n} (-1)^{k+1} \sum_{i_1 < \cdots < i_k} \chi_{A_{i_1}} \cdots \chi_{A_{i_k}} \\ &= \sum_{k=1}^{n} (-1)^{k+1} \sum_{i_1 < \cdots < i_k} \chi_{A_{i_1} \cap \cdots \cap A_{i_k}} \end{aligned}$$

Now, summing over all x the values of the LHS and the RHS we recover the inclusion-exclusion formula.

Chapter 5

Elementary number theory

Definition. A natural number n is *prime* if the only factors of n are 1 and n, and $n \neq 1$. A *factor* of n means a natural number m such that $n = am$ for some natural number a. If m and n are integers, we write $m \mid n$ and say m *divides* n, if $n = am$ for some integer a.

Theorem 5.1. Every natural number $n \geq 2$ can be written as a product of primes.

Proof. Suppose not. Then let n be the smallest number that cannot be written as a product of primes. Then n isn't a prime, so we can write $n = ab$ with $a, b < n$. By the minimality of n, a and b are products of primes, so $n = ab$ is also a product of primes. □

Theorem 5.2 (Euclid)**.** There are infinitely many prime numbers.

Proof. Assume not, and let all the primes be written in a list $p_1, \ldots, p_n$. Now let $N = p_1 \cdots p_n + 1$. Then, by Theorem 3.1, N is a product of primes. However, every p_i divides $N - 1$, so it cannot divide N. Hence, there must be primes other than $p_1, \ldots, p_n$. □

Definition (Highest common factor)**.** Let m and n be integers. Then the *highest common factor* of m and n, written $\mathrm{hcf}(m, n)$ or (m, n), is the largest (positive) integer d such that $d \mid m$ and $d \mid n$.

Example. $(25, 105) = 5$, $(34, 55) = 1$ and $(47, 141) = 47$.

Lemma 5.3. Let m and n be positive integers. Then there exists integers q and r such that $n = qm + r$ and $0 \leq r < m$.

Proof. Let q be the largest integer such that $qm \leq n$. [This exists because $0m \leq n$ and if $q > n$ then $qm > n$. Hence $\{q : qm \leq n\}$ is non-empty and bounded above, so it has a largest element.] Clearly $r \geq 0$. If $r \geq m$, then $n - qm \geq m$ so $(q + 1)m \leq n$, contradicting the maximality of q. □

Lemma 5.4. Let m and n be positive integers and suppose that $n = qm + r$. Then $(m, n) = (r, m)$.

Proof. Suppose $d \mid m$ and $d \mid n$, and write $m = ad$, $n = bd$. Then $r = n - qm = d(a - qb)$, so $d \mid r$. Hence ($d \mid m$ and $d \mid n$) $\implies$ ($d \mid r$ and $d \mid m$). Conversely, if $d \mid r$ and $d \mid m$, then $d \mid qm + r = n$. So ($d \mid r$ and $d \mid m$) $\implies$ ($d \mid m$ and $d \mid n$). Therefore, the highest common factors are the same (since the common factors of m and n are precisely the common factors of r and m). □

5.1 Euclid's algorithm

Lemma 3.4 leads immediately to *Euclid's algorithm*: to find (m, n), write $n = qm + r$ with $0 \leq r < m$ and the answer will be the same as (r, m). But (r, m) is a smaller pair, so the problem has become easier. Now repeat.

Example. What is $(100, 142)$?

$$\begin{aligned} 142 &= 1 \times 100 + 42 \\ 100 &= 2 \times 42 + 16 \\ 42 &= 2 \times 16 + 10 \\ 16 &= 1 \times 10 + 6 \\ 10 &= 1 \times 6 + 4 \\ 6 &= 1 \times 4 + 2 \\ 4 &= 2 \times \mathbf{2} + 0 \end{aligned}$$

Example. What is $(144, 100)$?

$$\begin{aligned} 144 &= 1 \times 100 + 44 \\ 100 &= 2 \times 44 + 12 \\ 44 &= 3 \times 12 + 8 \\ 12 &= 1 \times 8 + 4 \\ 8 &= 2 \times \mathbf{4} + 0 \end{aligned}$$

We can express 4 as an integer combination of $100, 144$.

$$\begin{aligned} 4 &= 12 - 1 \times 8 \\ &= 12 - 1 \times (44 - 3 \times 12) \\ &= 4 \times 12 - 44 \\ &= 4 \times (100 - 2 \times 44) - 44 \\ &= 4 \times 100 - 9 \times 44 \\ &= 4 \times 100 - 9 \times (144 - 100) \\ &= 13 \times 100 - 9 \times 144 \end{aligned}$$

Theorem 5.5 (Bézout's theorem)**.** Let x and y be positive integers. Then there exists integers h and k such that

$$hx + ky = \mathrm{hcf}(x, y).$$

Proof. Let d be the smallest positive integer that can be written in the form $hx + ky$ with h, k integers.

By Lemma 3.3 we can write

$$x = qd + r \quad \text{with } 0 \leq r < d.$$

But then

$$\begin{aligned} r &= x - qd = x - q(hx + ky) \\ &= (1 - qh)x - qky \end{aligned}$$

So r is a smaller integer combination of x and y, which is a contradiction unless $r = 0$. Therefore $x = qd$, so $d \mid x$. Similarly, $d \mid y$.

Now let's suppose $c \mid x$ and $c \mid y$. Then $c \mid hx + ky = d$. This shows that d is the *highest* common factor of x and y. □

5.2 Solving linear equations in integers

Suppose we are given an equation of the form $ax + by = c$, a, b, c integers, and asked to find integer solutions x and y.

Let $d = (a, b)$. Then $d \mid ax + by$ for any pair x, y, so if $d \nmid c$ then there are no solutions. If $d \mid c$, then $ax + by = c \iff \frac{a}{d}x + \frac{b}{d}y = \frac{c}{d}$. So dividing through by d, we can concentrate on the case $(a, b) = 1$.

To solve $ax + by = c$ when $(a, b) = 1$, find h, k such that $ah + bk = 1$ (using Euclid's algorithm) and set $x = ch$ and $y = ck$.

Now we'd like to find *all* solutions to $ax + by = c$, still assuming $(a, b) = 1$. First look at the "homogeneous equation" $ax + by = 0$. Let h, k be integers such that $ha + kb = 1$. Then, if $ax + by = 0$ we have

$$\begin{aligned} & hax + hby = 0 \\ \implies & (1 - kb)x + hby = 0 \\ \implies & x = b(kx - hy) \end{aligned}$$

So $b \mid x$. Writing $x = \lambda b$, we deduce (from $ax + by = 0$) that $y = -\lambda a$. So all solutions have the form $x = \lambda b$, $y = -\lambda a$. Conversely, all such pairs *are* solutions.

Now suppose we have found *some* solution x_0, y_0 to $ax + by = c$. If $ax' + by' = c$ as well, then $a(x_0 - x') + b(y_0 - y') = 0$, so we can find an integer λ such that $x_0 - x' = \lambda b$, $y_0 - y' = -\lambda a$.

Therefore, all solutions of $ax + by = c$ have the form

$$x = x_0 + \lambda b, \quad y = y_0 - \lambda a$$

and all such pairs are solutions.

5.3 The fundamental theorem of arithmetic

Theorem 5.6. Let p be a prime number and let a, b be positive integers such that $p \mid ab$. Then $p \mid a$ or $p \mid b$.

Proof. If $p \nmid a$, then $(a, p) = 1$ (since p's only factors are 1 and p). So, by Theorem 3.5 we can find h and k such that $ha + kp = 1$. Then $hab + kpb = b$. Since $p \mid ab$, $p \mid hab + kpb = b$. So $p \nmid a \implies p \mid b$, which proves the theorem. □

Corollary 5.7. Let p be a prime and let $a_1, \ldots, a_k$ be positive integers. If $p \mid a_1 \cdots a_k$, then $p \mid a_i$ for some i.

Proof. Induction on k. Obvious if $k = 1$.
Suppose we have it for k and suppose $p \mid a_1 \cdots a_k a_{k+1} = (a_1 \cdots a_k)a_{k+1}$. By Theorem 3.6, either $p \mid a_1 \cdots a_k$ or $p \mid a_{k+1}$. In the first case, $p \mid a_i$ for some i, by hypothesis. In the second, $p \mid a_{k+1}$. □

Theorem 5.8 (The fundamental theorem of arithmetic)**.** Every positive integer $n \geq 2$ can be written in exactly one way as a product of the form

$$p_1^{a_1} \cdots p_k^{a_k},$$

where $p_1 < \cdots < p_k$ are primes and $a_1, \ldots, a_k$ are positive integers.

Proof. We have already proved that at least one such expression exists. It remains to show that there is only one.

If this is not always the case, then there must be a smallest counterexample, that is, a least n that can be written as both

$$n = p_1^{a_1} \cdots p_k^{a_r} = q_1^{b_1} \cdots q_r^{b_r}$$

where the two products are different.

Then p_1 cannot equal any q_i since if it did $\frac{n}{p_1}$ would be a smaller number with two distinct factorizations.

But $p_1 \mid n = q_1^{b_1} \cdots q_r^{b_r}$, so, by Corollary 3.7, $p_1 \mid q_i$ for some i. Since q_i is prime, $p_1 = q_1$. Contradiction. □

Example. Consider $\mathbb{Z}(\sqrt{-5})$ with addition and multiplication as follows.

$$(a + b\sqrt{-5}) + (c + d\sqrt{-5}) = (a + c) + (b + d)\sqrt{-5}$$
$$(a + b\sqrt{-5})(c + d\sqrt{-5}) = (ac - 5bd) + (ad + bc)\sqrt{-5}$$

Then 6 has two factorizations.

$$6 = 2 \times 3$$
$$6 = (1 + \sqrt{-5})(1 - \sqrt{-5})$$

Chapter 6

Modular arithmetic

Loosely speaking, modular arithmetic is like the arithmetic of clocks: when you get to a certain number, you go back to the beginning again.

There two ways of thinking about it more precisely.

6.1 First view

Arithmetic mod m.
Two integers x and y are *congruent modulo* m if $y - x$ is a multiple of m. We write $x \equiv y \pmod{m}$.

Lemma 6.1. Congruence (mod m) is an equivalence relation.

Proof. R: Let $x \in \mathbb{Z}$. Then $x - x = 0m$.
S: If $y - x = am$ then $x - y = (-a)m$.
T: If $y - x = am$, $z - y = bm$ then $z - x = (a + b)m$.

□

The objects of study are the equivalence classes.

Example. If $m = 3$ then the equivalence classes are

$$\{\dots, -6, -3, 0, 3, 6, 9, \dots\}$$
$$\{\dots, -5, -2, 1, 4, 7, 10, \dots\}$$
$$\{\dots, -4, -1, 2, 5, 8, 11, \dots\}$$

Lemma 6.2. If $x \equiv z \pmod{m}$ and $y \equiv w \pmod{m}$ then $x + y \equiv z + w \pmod{m}$ and $xy \equiv zw \pmod{m}$.

Proof. We can write $z = x + am$, $w = y + bm$. Then

$$\begin{aligned} z + w &= (x + am) + (y + bm) \\ &= x + y + (a + b)m \\ &\equiv x + y \\ zw &= (x + am)(y + bm) \\ &= xy + (ay + bx + abm)m \\ &\equiv xy. \end{aligned}$$

□

Definition. This allows us to make the following definitions. Let $[x]$ stand for the equivalence class of x. Then

$$[x] + [y] = [x + y]$$
$$[x].[y] = [xy].$$

Example. $[2].[2] = [4] = [1]$.

6.2 Second view

The objects in arithmetic mod m are the numbers

$$0, 1, \ldots, m - 1.$$

Addition and multiplication are defined as follows.
$x +_m y$ means the remainder when you divide $x + y$ by m.
$x \times_m y$ means the remainder when you divide xy by m.

Example. For example, let $m = 13$. Then

$$7 + 8 \equiv 2$$
$$7 \times 8 \equiv 4$$
$$7^8 \equiv 3 \quad \text{since } 7^2 \equiv 49 \equiv -3, 7^4 \equiv 9, 7^8 \equiv 3.$$

6.3 Third view

The basic objects are integers, but you are always allowed to replace x by y if y differs from x by a multiple of m, that is, if $y \equiv x \pmod{m}$.

6.4 Prime moduli

When m is prime, certain facts hold that give modular arithmetic a different flavour. The main one is the existence of *multiplicative inverses*.

Lemma 6.3. Let p be a prime. Then for every $x \not\equiv 0 \pmod{p}$ there exists y such that $xy \equiv 1 \pmod{p}$.

Proof. $(x, p) = 1$, so by Bézout's theorem we can find integers h, k such that $hx + kp = 1$. But then $hx \equiv 1 \pmod{p}$ so let $y = h$. □

Corollary 6.4 (No zero divisors)**.** Let p be prime. Then if $xy \equiv 0 \pmod{m}$ then either $x \equiv 0 \pmod{p}$ or $y \equiv 0 \pmod{p}$.

Proof. (i) This says $p \mid xy \implies p \mid x$ or $p \mid y$ so we've already seen it.

(ii) If $xy \equiv 0$ and $x \not\equiv 0$, then x has an inverse w, by Lemma 4.3, so $w(xy) \equiv w.0 \equiv 0$. But $w(xy) = (wx)y \equiv 1.y = y$, so $y \equiv 0$.

□

Corollary 6.5 (Cancellation law)**.** Let p be prime and $a \not\equiv 0 \pmod{p}$. Then

$$\forall x, y \quad ax \equiv ay \implies x \equiv y \pmod{p}.$$

Proof. Multiply both sides by the inverse of a. □

Theorem 6.6 (Wilson's theorem)**.** Let p be prime. Then

$$(p-1)! \equiv -1 \pmod{p}.$$

Proof. Suppose that $x^2 \equiv 1 \pmod{p}$. Then $(x+1)(x-1) \equiv x^2 - 1 \equiv 0 \pmod{p}$ so, by Corollary 4.4, $x + 1 \equiv 0$ or $x - 1 \equiv 0$, that is, $x \equiv \pm 1$.

The idea is to partition the numbers into *inverse pairs*. That is, we partition them into pairs $\{x, y\}$ such that $xy \equiv 1$. All these pairs have size 2 apart from the cases $y = x$ — but the only such examples are $\{1\}$ and $\{-1\}$. (No overlapping since if $\{x, y\}$ and $\{z, y\}$ are such pairs then $xy \equiv zy \equiv 1$, so $x \equiv z$ by Corollary 4.5.)

So $(p-1)! \equiv -1 \times 1 \times 1^{\frac{p-3}{2}} = -1$. □

Theorem 6.7 (Fermat's little theorem)**.** Let p be prime and let $a \not\equiv 0 \pmod{p}$ then

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

Proof. Consider the sequence $1, a, a^2, a^3, \ldots$.
By the pigeonhole principle, there must be $r < s$ such that $a^r \equiv a^s \pmod{p}$ (as only p possibilities a^r). Dividing by a^r (i.e., multiplying by $(a^{-1})^r$) we deduce that $1 \equiv a^{s-r}$, so $\exists d > 0$ such that $a^d \equiv 1$. Hence, there is a smallest such d. We shall show that $d \mid p-1$, so that $a^{p-1} \equiv (a^d)^l \equiv 1^l \equiv 1$.

To do this, define an equivalence relation on the numbers $1, 2, \ldots, p-1$. Say that $x \sim y$ if $y \equiv a^r x$ for some $r \in \mathbb{Z}$.

Refl.	Let $r = 0$.
Symm.	$y \equiv a^r x \implies x \equiv a^{-r} y$.
Trans.	$y \equiv a^r x, z \equiv a^s y \implies z \equiv a^s(a^r x) \equiv a^{s+r} x$.

If $x \sim y$ then $y \equiv a^r x$ and we can assume (after adding a multiple of d to r) that $r \geq 0$. Then $r = qd + t$ and $0 \leq t < d$. So $y \equiv a^{qd+t} x \equiv (a^d)^q a^t x \equiv a^t x$.

So the equivalence class of x is

$$\{x, ax, a^2 x, \ldots, a^{d-1} x\}.$$

If $r, s < d$ and $a^r x \equiv a^s x$ then $a^{s-r} \equiv 1$ (by cancellation), $s - r < d$. So $s - r = 0$ (by the minimality of d) and $s = r$.

So the equivalence classes have size exactly d, as we wanted. □

Proof. We shall show by induction that a^p is always congruent to $a \pmod p$. Then, if $(a,p) = 1$ (i.e. a isn't a multiple of p) we can apply the cancellation rule and get from $a^p \equiv a$ to $a^{p-1} \equiv 1$.

The statement is true for $a = 0, a = 1$.

Suppose we know that $a^p \equiv a$ for some particular a. Then

$$(a+1)^p = a^p + \binom{p}{1}a^{p-1} + \binom{p}{2}a^{p-2} + \cdots + \binom{p}{p-1}a + 1.$$

It can be shown that $p \mid \binom{p}{k}$ when p is prime and $1 \leq k < p$. So $(a+1)^p \equiv a^p + 1 \equiv a + 1$ by the inductive hypothesis. □

Lemma 6.8. Let $m \geq 2$ be an integer and let a, b be integers. Then if $(a,m) = (b,m) = 1$, then $(ab,m) = 1$.

Proof. Pick h, k s.t. $ha + km = 1$ and u, v s.t. $ub + vm = 1$. Then

$$\begin{aligned} &(ha+km)(ub+vm) = 1 \\ \implies &huab + (kub + hva + kvm)m = 1 \end{aligned}$$

So any common factor of ab and m divides 1. So $(ab,m) = 1$. □

It is also true that every a s.t. $(a,m) = 1$ has an inverse $\pmod m$ — just apply Bézout's theorem (as we did when m was prime).

This shows that the set of integers a between 0 and m such that $(a,m) = 1$ forms a group under multiplication $\pmod m$.

Definition. If $(a,m) = 1$ then we say that a and m are *coprime*.

Definition (Euler's totient function)**.** *Euler's totient function* ϕ is defined as follows: $\phi(m)$ is the number of positive integers $\leq m$ that are coprime to m. Equivalently, it is the size of the group we have just defined.

Example. $\phi(12) = 4$ since the integers ≤ 12 and coprime to 12 are 1, 5, 7 and 11.

Lemma 6.9. Let m be a positive integer with prime factorization $p_1^{a_1} \cdots p_k^{a_k}$. Then

$$\begin{aligned} \phi(m) &= m(1 - \frac{1}{p_1}) \cdots (1 - \frac{1}{p_k}) \\ &= p_1^{a_1 - 1}(p_1 - 1) \cdots p_k^{a_k - 1}(p_k - 1). \end{aligned}$$

Proof. Let A_i be the set of all multiples of p_i less then or equal to m. Then $\phi(m) = m - |\bigcup_{i=1}^k A_i|$. We shall calculate this by inclusion-exclusion. That gives

$$m + \sum_{j=1}^{k} (-1)^j \sum_{1 \leq i_1 < \cdots < i_j \leq k} |A_{i_1} \cap \cdots \cap A_{i_j}|.$$

$A_{i_1} \cap \cdots \cap A_{i_j}$ is the set of all multiples of $p_{i_1}p_{i_2}\cdots p_{i_j}$ and the number of these is $\frac{m}{p_{i_1}p_{i_2}\cdots p_{i_j}}$. So

$$\begin{aligned}\phi(m) &= m\left(1 + \sum_{j=1}^{k}(-1)^j \sum_{1\le i_1<\cdots<i_j\le k} \frac{1}{p_{i_1}p_{i_2}\cdots p_{i_j}}\right)\\ &= m(1-\frac{1}{p_1})(1-\frac{1}{p_2})\cdots(1-\frac{1}{p_k}).\end{aligned}$$

□

Theorem 6.10 (Euler's theorem)**.** Let $m \ge 2$ be a positive integer and let $(a, m) = 1$. Then

$$a^{\phi(m)} \equiv 1 \pmod{m}.$$

(Note that $\phi(p) = p - 1$ if p is prime.)

Proof. As in FLT we can find $r < s$ such that $a^r \equiv a^s$. Because a has an inverse we can deduce that $a^{s-r} \equiv 1$. So we can find a minimal $d > 0$ such that $a^d \equiv 1$.

Now define an equivalence relation on $\mathbb{Z}_m^*$ — the set of integers $\le m$ and coprime to m.

As before, set $x \sim y$ iff $\exists r \in \mathbb{Z}$ such that $y \equiv a^r x$. As before, the equivalence class of any x is $\{a, ax, \ldots, a^{d-1}x\}$. As before, if $0 \le r < s < d$ and $a^r x \equiv a^s x$ then $a^{s-r} \equiv 1$ (since we have cancellation) and $s - r < d$ so $s - r = 0$, $s = r$.

Thus, all equivalence classes have size d. These partition $\mathbb{Z}_m^*$, which has size $\phi(m)$ so $d \mid \phi(m)$. If $\phi(m) = hd$ then

$$a^{\phi(m)} = (a^d)^h \equiv 1^h = 1.$$

□

6.5 Public-Key Cryptography

This is a method for secure encryption of messages. It works as follows.

Let p and q be large primes. Let $m = pq$. Then $\phi(m) = (p-1)(q-1)$. Let r be an integer $< m$ such that $(r, \phi(m)) = 1$.

To encrypt message: First encode them as a sequence of numbers between 0 and m (in a way that is *easy* to translate back). Given one of these numbers, x, say, raise it to the power $r \pmod{m}$. Then $y = x^r$ is the encryption of x.

To decrypt messages: By Euler's theorem, $x^{\phi(m)} \equiv 1 \pmod{m}$. (As long as $(x, m) = 1$ but this is true with enormously high probability.) Hence, if $rs \equiv 1 \pmod{\phi(m)}$ then $x^{rs} = x^{t\phi(m)+1}$ for some t, and is congruent to x: $x^{t\phi(m)+1} \equiv (x^{\phi(m)})^t x \equiv 1^t x \equiv x$.

But Euclid's algorithm gives us a pair h and k such that $hr + k\phi(m) = 1$, so we can take $s = h$ (and reduce $\pmod{m}$ so that $0 < s < m$).

To encrypt, you need to know m and r.

To decrypt, you need to know s, and to work out s from r and m you need to know $\phi(m)$. But if you know pq and $(p-1)(q-1) = pq - p - q + 1$, then you know $p + q$.

Then you can solve the quadratic equation $x^2 - (p+q)x + pq$ and calculate p and q. *So* to decrypt, you have to be able to factorize m. But this seems to be computationally infeasible.

The other calculations can all be done efficiently. For example, to work out x^r, write r in binary as $\sum_{i=0}^{k} \varepsilon_i 2^i$ where each ε_i is 0 or 1. E.g. $13 = 2^3 + 2^2 + 2^0$. Then work out $x, x^2, x^4, x^8, \ldots, x^{2^k} \pmod m$ by repeatedly squaring. Then multiply together (mod m) the ones you need. E.g. if $r = 13$, work out $x \cdot x^{2^2} \cdot x^{2^3}$.

6.6 Chinese remainder theorem

Lemma 6.11. Let $a_1, a_2, \ldots, a_k$ be positive integers such that $(a_i, a_j) = 1$ whenever $i \neq j$. Then the smallest integer n such that

$$a_i \mid n \quad \text{for all } i,$$

called the *lowest common multiple* of $a_1, a_2, \ldots, a_k$ is $a_1 a_2 \cdots a_k$.

Proof. First, we prove the theorem for $k = 2$.

Suppose that $(a, b) = 1$ and $a \mid n$, $b \mid n$. Write $n = ua = vb$. Pick h and k such that $ha + kb = 1$. Then $hau + kbu = u$ so $hvb + kub = u$ as $ua = vb$. So $b \mid u$. But $n = ua$, so $ab \mid n$. So $\operatorname{lcm}(a, b) = ab$.

To get the general result, apply induction, and use the fact that if $(a_i, a_{k+1}) = 1$ for all $i \leq k$ then $(a_1 a_2 \cdots a_k, a_{k+1}) = 1$. □

Theorem 6.12 (Chinese remainder theorem)**.** Let $a_1, a_2, \ldots, a_k$ be positive integers with $(a_i, a_j) = 1$ for all $i \neq j$. For each i, let $0 \leq r_i < a_i$ be some integer. Then there is an integer x such that

$$x \equiv r_i \pmod{a_i} \quad \text{for every } i.$$

Moreover, there is exactly one such x with $0 \leq x < a_1 a_2 \cdots a_k$.

Example.

$$\begin{aligned} x &\equiv 2 \pmod 3 \\ &\equiv 5 \pmod 7 \\ &\equiv 4 \pmod{11} \end{aligned}$$

Note that $x \equiv 26 \pmod{3 \times 7 \times 11}$.

Existence. First we do the case $r_1 = 1$, $r_2 = \cdots = r_k = 0$. We know, by an earlier lemma (says $(a, m) = (b, m) = 1 \implies (ab, m) = 1$), that

$$(a_1, a_2 \cdots a_k) = 1.$$

So we can find u, v such that $a_1 u + a_2 a_3 \cdots a_k v = 1$. Then $a_2 a_3 \cdots a_k v$ is a multiple of a_i when $i \geq 2$ and is $\equiv 1 \pmod{a_1}$. Call this number x_1.

Similarly, for the other i, we can find x_i such that

$$x_i \equiv \begin{cases} 1 \pmod{a_j} & j = i \\ 0 \pmod{a_j} & j \neq i \end{cases}$$

Then $\sum_{i=1}^{k} r_i x_i \equiv r_j \pmod{a_j}$ since $a_j \mid x_i$ when $i \neq j$ and $x_j \equiv 1 \pmod{a_j}$. □

Example.

$$\begin{aligned} x &\equiv 2 \pmod 3 \\ &\equiv 3 \pmod 7 \end{aligned}$$

We have

$$7 \equiv 1 \pmod 3 \text{ and } 7 \equiv 0 \pmod 7$$
$$15 \equiv 0 \mod 3 \text{ and } 15 \equiv 1 \pmod 7$$

and deduce that

$$x \equiv 2 \times 7 + 3 \times 15 = 59 \equiv 17 \pmod{21}.$$

Uniqueness. Suppose that for every i we have $x \equiv r_i \pmod{a_i}$ and $y \equiv r_i \pmod{a_i}$. Then $a_i \mid x - y$ for every i, so, by the lemma, $a_1 a_2 \cdots a_k \mid x - y$. If $0 \leq x, y < a_1 a_2 \cdots a_k$ it follows that $x = y$. □

Chapter 7

Building numbers from scratch

7.1 The Peano axioms

A formal approach to the natural numbers begins with the *Peano axioms*.

We assume that we have a set, called $\mathbb{N}$, together with a function $\sigma\colon \mathbb{N} \to \mathbb{N}$ called the *successor* function, satisfying the following properties. (There is also a privileged element, which we call 1.)

(i) If $\sigma(x) = \sigma(y)$ then $x = y$. (σ is an injection.)

(ii) There is no x such that $1 = \sigma(x)$. (1 is not the successor of anything.)

(iii) Let S be a subset of $\mathbb{N}$ such that $1 \in S$ and $\sigma(x) \in S$ whenever $x \in S$. Then $S = \mathbb{N}$. (Principle of induction.)

2 is, by definition, $\sigma(1)$
3 is, by definition, $\sigma(2)$ etc.

Definition (Addition)**.** This is defined *inductively*. We say

(i) $n + 1 = \sigma(n) \quad \forall n \in \mathbb{N}$

(ii) $n + \sigma(m) = \sigma(n + m) \quad \forall m, n \in \mathbb{N}$

Example. To show $2 + 3 = 5$.

$$\begin{aligned} \sigma(1) + \sigma(\sigma(1)) &= \sigma(\sigma(1) + \sigma(1)) \quad \text{by (ii)} \\ &= \sigma(\sigma(\sigma(1) + 1)) \quad \text{by (ii)} \\ &= \sigma(\sigma(\sigma(\sigma(1)))) \quad \text{by (i)} \\ &= 5 \quad \text{by definition} \end{aligned}$$

Lemma 7.1.

$$n + \sigma(m) = \sigma(n) + m \quad \forall m, n \in \mathbb{N}$$

Proof. First, we do the case $m = 1$, so we want to show

$$n + \sigma(1) = \sigma(n) + 1.$$

Bute $n + \sigma(1) \overset{(ii)}{=} \sigma(n+1) \overset{(i)}{=} \sigma(\sigma(n)) \overset{(i)}{=} \sigma(n) + 1$. Now suppose we have proven the result for m. Then

$$\begin{aligned} n + \sigma(\sigma(m)) &= \sigma(n + \sigma(m)) && \text{by (ii)} \\ &= \sigma(\sigma(n) + m) && \text{by ind. hyp.} \\ &= \sigma(n) + \sigma(m) && \text{by (ii)} \end{aligned}$$

□

Theorem 7.2 (Commutativity)**.** Addition on $\mathbb{N}$ is commutative. That is,

$$\forall m, n \quad n + m = m + n.$$

Proof. First we prove this for $m = 1$, i.e. that $\forall n \quad n + 1 = 1 + n$. This we prove by induction *on* n. It's true when $n = 1$. If we know it for n, then

$$\begin{aligned} \sigma(n) + 1 &= n + \sigma(1) && \text{by the lemma} \\ &= \sigma(n+1) && \text{by (ii)} \\ &= \sigma(1+n) && \text{by ind. hyp.} \\ &= 1 + \sigma(n) && \text{by (ii)} \end{aligned}$$

Now suppose we know the result for m (and for all n). Then

$$\begin{aligned} n + \sigma(m) &= \sigma(n+m) && \text{by (ii)} \\ &= \sigma(m+n) && \text{by ind. hyp.} \\ &= m + \sigma(n) && \text{by (ii)} \\ &= \sigma(m) + n && \text{by the lemma} \end{aligned}$$

□

Similar arguments can be used to show that addition is associative and also the *cancellation law*:

$$\begin{aligned} & x + y = x + z \\ \implies & y = z. \end{aligned}$$

One can define multiplication,

$$\begin{aligned} m.1 &= m \\ m.\sigma(n) &= m.n + m \end{aligned}$$

or m^n,

$$\begin{aligned} m^1 &= m \\ m^{\sigma(n)} &= m.m^n \end{aligned}$$

or $n!$,

$$\begin{aligned} 1! &= 1 \\ \sigma(n)! &= \sigma(n).n! \end{aligned}$$

One can also prove the usual rules: multiplication is commutative, associative, distributive over $+$, 1 is an identity, and cancellation ($xy = xz \implies y = z$). This is left as an exercise.

One can also define an *ordering* $<$. We say $m < n$ if $\exists k \in \mathbb{N}$ s.t. $m+k = n$. By inductive arguments one can prove that for any m, n exactly one of the following statements holds

$$m < n, m = n, n < m.$$

E.g. if $m < n$ and $n < m$ then

$$\begin{aligned} & \exists k, l \text{ s.t. } n = m + k, m = n + l \\ \implies & n = n + k + l \quad \text{(associativity)} \\ \implies & n + 1 = n + k + l + 1 \\ \implies & 1 = k + l + 1 \quad \text{(cancellation)} \\ \implies & 1 = \sigma(k + 1) \end{aligned}$$

Contradiction.

7.2 Construction of $\mathbb{Z}$

The formal way to define integers in general is as equivalence classes of ordered pairs (m, n) of natural numbers. (One thinks of (m, n) as $m - n$.) We say $(m, n) \sim (r, s)$ iff $m + s = n + r$. Write $[m, n]$ for the equivalence class of (m, n).

Example.

$$\begin{aligned} [2, 5] &= \text{(which we secretly know is } -3) \\ &= \{(1, 4), (2, 5), (3, 6), (4, 7), \dots\}. \end{aligned}$$

Then $[m_1, n_1] + [m_2, n_2]$ is defined to be $[m_1 + m_2, n_1 + n_2]$, and $[m_1, n_1].[m_2, n_2]$ is defined to be $[m_1 m_2 + n_1 n_2, m_1 n_2 + m_2 n_1]$. One must check that these operations are well-defined. Suppose that $(m_1, n_1) \sim (r_1, s_1)$ and $(m_2, n_2) \sim (r_2, s_2)$. Then $(m_1 + m_2, n_1 + n_2) \sim (r_1 + r_2, s_1 + s_2)$ since

$$\begin{aligned} m_1 + m_2 + s_1 + s_2 &= (m_1 + s_1) + (m_2 + s_2) \\ &= (n_1 + r_1) + (n_2 + r_2) \\ &= n_1 + n_2 + r_1 + r_2 \end{aligned}$$

Similarly for multiplication. (We are proving that if we take any two elements out of the same equivalence classes, the elements we get must be in the same equivalence class.)

Now we take $-n$ as notation for $[1, 1 + n]$ and 0 as notation for $[1, 1]$. Then 0 is an additive identity and $-n$ is an inverse for n. We define $m - n$ to mean $m + (-n)$. (In general, $-[m, n]$ means $[n, m]$.)

7.3 Construction of $\mathbb{Q}$

$\mathbb{Q}$ is built up as a set of equivalence classes of ordered pairs of integers. If (m, n) and (r, s) are such pairs with $n \neq 0$, $s \neq 0$ then define $(m, n) \sim (r, s)$ iff $ms = nr$. Write $[m, n]$ for the equivalence class.

Define $[m_1, n_1][m_2, n_2] = [m_1 m_2, n_1 n_2]$ and $[m_1, n_1] + [m_2, n_2] = [m_1 n_2 + m_2 n_1, n_1 n_2]$. Again we must check that the operations are well-defined. Suppose that $(m_1, n_1) \sim (r_1, s_1)$ and $(m_2, n_2) \sim (r_2, s_2)$.

$$\begin{aligned} & m_1 s_1 = r_1 n_1, \quad m_2 s_2 = r_2 n_2 \\ \implies & s_1 s_2 (m_1 n_2 + m_2 n_1) = n_1 n_2 (r_1 s_2 + r_2 s_1) \end{aligned}$$

Then $(m_1 n_2 + m_2 n_1, n_1 n_2) \sim (m_1 n_2 s_1 s_2 + m_2 n_1 s_1 s_2, n_1 n_2 s_1 s_2) = (n_1 r_1 m_2 s_2 + n_2 r_2 n_1 s_1, n_1 n_2 s_1 s_2) \sim (r_1 s_2 + r_2 s_1, s_1 s_2)$.

By this kind of argument one can check that $\mathbb{Q}$ is an example of an algebraic structure called a *field*.

Definition (Field)**.** A *field* is a set A together with two binary operations, $\square$ and $*$, with the following properties:

$\square$ and $*$ are commutative and associative, and they both have (different) identity elements. Every $x \in A$ has an inverse for $\square$, and every $x \in A$ apart from the $\square$-identity has an inverse for $*$. $*$ is distributive over $\square$.

"Useful notation" write $+$ for $\square$, $\cdot$ or $\times$ for $*$, 0 for $\square$-identity, 1 for $*$-identity.

Lemma 7.3. Let $\mathbb{F}$ be a field and let $x \in \mathbb{F}$. Then $0 \cdot x = 0$.

Proof.

$$(0 + 0)x = 0x + 0x \quad \text{(distributivity)}$$

But

$$0 + 0 = 0 \quad \text{(0 an additive identity)}$$

so

$$0x = 0x + 0x.$$

Let u be the additive inverse of $0x$. Then

$$\begin{aligned} 0 &= u + 0x \\ &= u + (0x + 0x) \\ &= (u + 0x) + 0x \quad (+ \text{ is associative}) \\ &= 0 + 0x \\ &= 0x \quad (0 \text{ is an additive identity}) \end{aligned}$$

□

Example. Examples of fields include $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{F}_p$ (integers (mod p) for p prime), $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$.

7.4 Ordered fields

Definition (Ordered field)**.** An *ordered field* is a field $\mathbb{F}$ together with a transitive relation $<$ with the following properties:

(i) For every pair $x, y \in \mathbb{F}$ exactly one of the following is true

$$x < y, x = y, y < x$$

(We say $<$ is a *total ordering*. Transitivity is part of this definition.)

(ii) If $x, y, z \in \mathbb{F}$ and $y < z$ then $x + y < x + z$.

(iii) If $x, y, z \in \mathbb{F}$ and $0 < x, y < z$ then $xy < xz$.

Example. $\mathbb{Q}$ and $\mathbb{R}$ are ordered fields with their usual orderings. To define the ordering on $\mathbb{Q}$ formally, say $\frac{p}{q} < \frac{r}{s}$ iff $ps < qr$ assuming that $q, s > 0$, as we can.

Lemma 7.4. In any ordered field, $0 < 1$.

Proof. It can be shown that $(-1)^2 = 1$. (See examples sheet.) By (i) exactly one is true of $1 < 0, 1 = 0, 0 < 1$. We know $1 \neq 0$. If $1 < 0$ then $0 < -1$ by (ii) with $x = -1$.

$$\begin{aligned} &\implies (-1) \cdot 0 < (-1)^2 \quad \text{by (iii)} \\ &\implies 0 < 1, \quad \text{contradicting (i)} \end{aligned}$$

□

Notation. We write $x > y$ to mean $y < z$ and $x \leq y$ to mean $x < y$ or $x = y$ and $x \geq y$ for $y \leq x$.

Proposition 7.5. Every ordered field contains a copy of $\mathbb{Q}$.

A formal statement of this would be as follows. Let $\mathbb{F}$ be an ordered field. Then there is an injection $\phi \colon \mathbb{Q} \to \mathbb{F}$ such that for all $x, y \in Q$

$$\begin{aligned} \phi(x + y) &= \phi(x) + \phi(y) \\ \phi(xy) &= \phi(x)\phi(y) \\ x < y &\implies \phi(x) < \phi(y) \end{aligned}$$

Sketch of a proof. In $\mathbb{F}$ let's underline elements.

We write $\underline{0}$ and $\underline{1}$ for the additive and multiplicative identities in $\mathbb{F}$. Write $\underline{n}$ for $\underline{1} + \underline{1} + \cdots + \underline{1}$. (More formally, $\underline{n+1} = \underline{n} + \underline{1}$.)

We must show that $\underline{0}, \underline{1}, \underline{2}, \ldots$ are all distinct. But this is easy by induction. $\underline{0} < \underline{1}$. If $\underline{0} < \underline{n}$ then

$$\underline{1} < \underline{n} + \underline{1} \implies \underline{0} < \underline{n+1}.$$

More generally, if $m < n$ then $n = m + k$ for some $k \in \mathbb{N}$.

$$\implies \underline{n} = \underline{m} + \underline{k}^{[1]}$$

so

$$\underline{0} < \underline{k} \implies \underline{m} < \underline{m} + \underline{k} = \underline{n}.$$

[1] This needs to be proved by induction.

To find a copy of $\mathbb{Z}$ in $\mathbb{F}$, let $\underline{\text{-m}}$ be $-\underline{\text{m}}$, that is, the additive inverse of $\underline{\text{m}}$. (One needs to check things like that $\underline{\text{-m-n}} = \underline{\text{mn}} = \underline{\text{mn}}$.)

To embed $\mathbb{Q}$, we define

$$\frac{p}{q} \quad \text{to be} \quad \underline{\text{p}}(\underline{\text{q}})^{-1}.$$

If $\frac{p}{q} = \frac{r}{s}$ then

$$\begin{aligned} ps = qr &\implies \underline{\text{ps}} = \underline{\text{qr}}^{2} \\ &\implies \underline{\text{p}}(\underline{\text{q}})^{-1} = \underline{\text{r}}(\underline{\text{s}})^{-1} \\ &\implies \frac{p}{q} = \frac{r}{s}. \end{aligned}$$

If $\frac{p}{q} < \frac{r}{s}$ with $q, s > 0$ then

$$ps < qr \implies \underline{\text{ps}} < \underline{\text{qr}} \implies \underline{\text{ps}} < \underline{\text{qr}}$$

and also $\underline{\text{q}}, \underline{\text{s}} > 0$, so $(\underline{\text{q}})^{-1}, (\underline{\text{s}})^{-1} > 0$ (see examples sheet).

$$\begin{aligned} &\implies \underline{\text{p}}(\underline{\text{q}})^{-1} < \underline{\text{r}}(\underline{\text{s}})^{-1} \quad \text{by axiom (iii)} \\ &\implies \frac{p}{q} < \frac{r}{s} \end{aligned}$$

Etc. etc.

□

Let $\mathbb{F}$ be an ordered field. Let P be $\{x \in \mathbb{F} : x > 0\}$. Then P has the following properties.

(i) For every $x \in \mathbb{F}$ exactly one of the following is true:

$$x \in P, x = 0, -x \in P.$$

(ii) If $x, y \in P$ then $x + y \in P$.

(iii) If $x, y \in P$ then $xy \in P$.

These three facts follow easily from the corresponding axioms for an ordered field.

This implication can be *reversed*. Given $P \subset \mathbb{F}$ satisfying (i), (ii), (iii), say that

$$x < y \quad \text{iff} \quad y - x \in P.$$

E.g. to deduce property (iii) of an ordered field, let $x > 0$, $y < z$. Then

$$\begin{aligned} &x \in P, z - y \in P \\ \implies &x(z - y) \in P \\ \implies &xz - xy \in P \\ \implies &xy < xz. \end{aligned}$$

[2] Because we would have checked this for integers.

7.5 The least upper-bound axiom

Let $\mathbb{F}$ be an ordered field. $\mathbb{F}$ is said to satisfy the least upper bound axiom if for every non-empty subset $X \subset \mathbb{F}$ that is bounded above (i.e. there is some M such that $x \leq M$ for every $x \in X$), there exists a least upper bound, that is, an element $s \in \mathbb{F}$ such that

(i) $x \leq s$ for every $x \in X$
(s is an upper bound)

(ii) for every $t < s$ there is some $x \in X$ such that $x > t$
(s is the *least* upper bound)

The element s is called the *least upper bound*, or *supremum*, of X.

Example. Let X be the open intervall $(0, 1)$. Then $\sup X = 1$.

(i) If $x \in X$ then $x < 1$.

(ii) If $0 \leq t < 1$ then $\frac{t+1}{2} \in X$ and $\frac{t+1}{2} > t$. If $t < 0$ then $\frac{1}{2} \in X$ and $\frac{1}{2} > t$.

An ordered field that satisfies the least upper bound axiom is called *complete*. It can be shown

(i) that complete ordered fields exist

(ii) that any two are isomorphic (i.e., essentially the same)

We call "the" complete ordered field $\mathbb{R}$.

Theorem 7.6. $\sqrt{2}$ exists. I.e., if $\mathbb{F}$ is a complete ordered field then there exists x such that $x > 0$ and $x^2 = 2$.

Proof. Let $X = \{x \in \mathbb{F} : x^2 < 2\}$.
Let $s = \sup X$. We shall show that $s > 0$ and $s^2 = 2$. (Note that if $x > 2$ then $x^2 > 4 > 2$ so X is bounded above by 2, s exists.)
Suppose that $s^2 < 2$ and write $s^2 = 2 - \delta$ for some $\delta > 0$. We have just shown that $s \leq 2$. Also, since $1^2 < 2$, $s \geq 1$. Hence, $(s + \frac{\delta}{8})^2 = s^2 + \frac{s\delta}{4} + \frac{\delta^2}{64}$. Since $s \geq 1$, $\delta \leq 1$, so

$$s^2 + \frac{s\delta}{4} + \frac{\delta^2}{64} < 2 - \delta + \frac{\delta}{2} + \frac{\delta}{64} \leq 2.$$

So $s + \frac{\delta}{8} \in X$ and s was not an upper bound. Contradiction.
If $s^2 > 2$ then write $s^2 = 2 + \delta$ with $\delta > 0$. Let $0 \leq \eta \leq \frac{\delta}{4}$. Then

$$\begin{aligned}(s - \eta)^2 &= s^2 - 2\eta s + \eta^2 \\ &\geq 2 + \delta - \frac{\delta s}{2} + \eta^2 \\ &\geq 2 + \eta^2 \\ &\geq 2.\end{aligned}$$

Hence, $s - \frac{\delta}{4}$ is an upper bound for X. So s was not the least one. Contradiction.
So if s^2 is not > 2 or < 2 it must be 2. □

7.6 The Archimedean property

Theorem 7.7 (Version 1)**.** $\mathbb{N}$ is unbounded in $\mathbb{R}$. (I.e. if S is a subset of a complete ordered field such that $1 \in S$ and $x+1 \in S$ whenever $x \in S$ then S does not have an upper bound.

Proof. $\mathbb{N}$ is non-empty. So if it is bounded then it has a least upper bound x. Since x is least, there must be some $n \in \mathbb{N}$ such that $n > x-1$. But then $n+1 > x$. Contradiction. □

Theorem 7.8 (Version 2)**.** For every $x > 0$ (in $\mathbb{R}$) there is a positive integer n such that $\frac{1}{n} < x$. (Equivalently, if $0 \le x < \frac{1}{n}$ for every $n \in \mathbb{N}$ then $x = 0$.)

Proof. If $x > 0$ and $\frac{1}{n} > x$ for every $n \in \mathbb{N}$ then $\frac{1}{x}$ exists and $\frac{1}{x} > n$ for every $n \in \mathbb{N}$, contradicting version 1. □

Theorem 7.9 (Version 3)**.** Between any two real numbers there is a rational number.

Proof. Let $a < b$. If $a \le 0 < b$ then we look at $\frac{b}{2}$ and b. If $a < 0 \le b$ we can look at a and $\frac{a}{2}$ instead. And if $a < b < 0$ then we can find $\frac{p}{q} \in (-b, -a)$ and take $\frac{-p}{q}$.

So we can assume that $0 < a < b$.

By version 2, there is some $n \in \mathbb{N}$ such that $\frac{1}{n} < b-a$. By version 1, there exists some $m \in \mathbb{N}$ such that $m > an$. Take the least such m. Then $m > an \implies \frac{m}{n} > a$. If $\frac{m}{n} > b$ then $m > bn \implies m-1 > bn-1 \implies \frac{m-1}{n} > b - \frac{1}{n} > a$ since $\frac{1}{n} < b-a$. So m was not least. Contradiction. □

7.7 Sequences

$1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \cdots$	series
$1, 1\frac{1}{2}, 1\frac{3}{4}, 1\frac{7}{8}, 1\frac{15}{16}, \ldots$	sequence (of partial sums)

Definition (Convergence)**.** Let $x_1, x_2, \ldots$ be a sequence of real numbers. We say that x_n *converges to* a or *tends to* a, and write $x_n \to a$, if

$$\forall \varepsilon > 0 \quad \exists N \quad \forall n \ge N \quad |x_n - a| < \varepsilon.$$

7.8 The monotone-sequence axiom

A sequence $x_1, x_2, \ldots$ is called *monotone* if $x_1 \le x_2 \le \cdots$ or $x_1 \ge x_2 \ge \cdots$. It is *increasing* if $x_1 \le x_2 \le \cdots$ and *strictly increasing* if $x_1 < x_2 < \cdots$. (Similarly for *decreasing*.)

Proposition 7.10. Every increasing sequence that is bounded above converges to some limit.

Example.

$$1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \frac{1}{25} + \cdots = \frac{\pi}{6}$$

Proof. Let $x_1 \leq x_2 \leq \cdots$ and let $x_n \leq M$ for every n. Let $X = \{x_1, x_2, \dots\}$ and let $a = \sup X$. (a exists since $X \neq \emptyset$ and is bounded above by M.)

We show that $x_n \to a$. Let $\varepsilon > 0$. Since $a = \sup X$, $a - \varepsilon$ is not an upper bound for X, so $\exists N$ s.t. $x_N > a - \varepsilon$. But since $x_1 \leq x_2 \leq \cdots$ this means that $x_N > a - \varepsilon$ for every $n \geq N$. As a is an upper bound, $x_n \leq a$ for all n. Therefore,

$$\begin{aligned} \forall n \geq N \qquad & a - \varepsilon < x_n \leq a \\ \implies & |x_n - a| < \varepsilon \end{aligned}$$

□

7.9 Decimal expansions

Theorem 7.11. For every positive real number x there is a sequence of integers $a_0, a_1, a_2, \dots$ such that $a_0 \geq 0$ and $0 \leq a_n \leq 9$ for $n > 0$ and such that $s_n \to x$, where

$$s_n = a_0 + \frac{a_1}{10} + \frac{a_2}{100} + \cdots + \frac{a_n}{10^n}.$$

We write $x = a_0.a_1a_2a_3\dots$.

Proof. The set of integers $\leq x$ contains 0. Also, there is some $N > x$ (by Archimedean principle) so it is bounded above in $\mathbb{N}$. Hence, it has a largest element. Let this be a_0. Then $0 \leq x - a_0 < 1$ or else $a_0 + 1$ would be a larger one.

Now let a_1 be the largest integer such that $\frac{a_1}{10} \leq x - a_0$. Then $0 \leq a_1 \leq 9$ and hence $0 \leq x - a_0 - \frac{a_1}{10} < \frac{1}{10}$. Let a_2 be the largest integer such that $\frac{a_2}{100} \leq x - a_0 - \frac{a_1}{10}$. Then $0 \leq a_2 \leq 9$ and $x - a_0 - \frac{a_1}{10} - \frac{a_2}{100} = x - s_2$ lies between 0 and $\frac{1}{100}$.

Continuing this way, we have

$$s_n = a_0 + \frac{a_1}{10} + \frac{a_2}{100} + \cdots + \frac{a_n}{10^n}$$

with $0 \leq x - s_n < \frac{1}{10^n}$, so $s_n \to x$. (For this last assertion, we need that $\forall \varepsilon > 0 \quad \exists n$ s.t. $\frac{1}{10^n} < \varepsilon$. Use Archimedean principle again!) □

Proposition 7.12. If $a_0 + \frac{a_1}{10} + \frac{a_2}{100} + \cdots = b_0 + \frac{b_1}{10} + \frac{b_2}{100} + \cdots$ are two decimal expansions of x, then either $a_n = b_n$ for every n or they are equal up to some $k-1$, then $a_k = b_k + 1$ and $a_j = 0, b_j = 9 \quad \forall j > k$ (or same with $b_k = a_k + 1$).

Proof. If the sequences aren't equal, then let k be minimal such that $a_k \neq b_k$ and without loss of generality $a_k > b_k$. Then

$$\begin{aligned} & \frac{a_k}{10^k} + \frac{a_{k+1}}{10^{k+1}} + \cdots = \frac{b_k}{10^k} + \frac{b_{k+1}}{10^{k+1}} + \cdots \\ \implies & \frac{a_k - b_k}{10^k} = \frac{b_{k+1} - a_{k+1}}{10^{k+1}} + \frac{b_{k+2} - a_{k+2}}{10^{k+2}} + \cdots \\ \implies & \frac{a_k - b_k}{10^k} \leq \frac{9}{10^{k+1}} + \frac{9}{10^{k+2}} + \cdots = \frac{1}{10^k} \end{aligned}$$

Equality occurs only if $b_j - a_j = 9 \quad \forall j > k$ which means $b_j = 9$, $a_j = 0$.

Since $\frac{a_k - b_k}{10^k} \geq \frac{1}{10^k}$ this must happen and a_k must be $b_k + 1$. □

Define e to be the number $\frac{1}{0!} + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \cdots$. Why does this exist?

Let $e_N = \frac{1}{0!} + \frac{1}{1!} + \cdots + \frac{1}{N!}$. Then $e_0 < e_1 < e_2 < \cdots$ so the sequence is *increasing*. Also, $e_N \leq 1 + 1 + \frac{1}{2} + \cdots + \frac{1}{2^{N-1}} < 3$ so it is *bounded above*.

So the sequence $e_0, e_1, e_2, \ldots$ has a limit which we call e.

Theorem 7.13. e is irrational.

Proof. If $e = \frac{p}{q}$ with $p, q \in \mathbb{N}$ then qe is an integer, so $q!e$ is an integer. But

$$\begin{aligned} q!e &= \frac{q!}{0!} + \frac{q!}{1!} + \frac{q!}{2!} + \cdots + \frac{q!}{q!} + \frac{q!}{(q+1)!} + \frac{q!}{(q+2)!} + \cdots \\ &= M + \frac{1}{q+1} + \frac{1}{(q+1)(q+2)} + \frac{1}{(q+1)(q+2)(q+3)} + \cdots \end{aligned}$$

for some integer M.

But

$$0 < \frac{1}{q+1} + \frac{1}{(q+1)(q+2)} + \cdots < \frac{1}{q+1} + \frac{1}{(q+1)^2} + \frac{1}{(q+1)^3} + \cdots = \frac{1}{q}.$$

So the integer $q!e$ lies strictly between the integer M and $M + \frac{1}{q}$. Contradiction. □

7.10 Algebraic and transcendental numbers

Definition. A real number is *algebraic of degree d* if there is a polynomial of degree d with integer coefficients (or one could say rational coefficients, it is equivalent) with that number as a root.

A number is *transcendental* if it is not algebraic.

Example. $\frac{1+\sqrt{5}}{2}$ is algebraic of degree 2 as it is a root of $x^2 - x + 1 = 0$.

Theorem 7.14 (Liouville)**.** Let θ be an irrational number that is algebraic of degree d. Then there exists some number $c > 0$ (which can depend on θ) such that

$$|\theta - \frac{p}{q}| > \frac{c}{q^d}$$

for every pair $p, q \in \mathbb{N}$.

Proof. Let $r_1, r_2, \ldots, r_k$ be the rational roots of a polynomial P of degree d that has θ as a root. Since θ is irrational, it does not equal any r_i. Let $c_1 > 0$ be the minimum of $|\theta - r_i|$. (If there are no r_i, let $c_1 = 1$.) Now let $\alpha = \frac{p}{q}$, $\alpha \notin \{r_1, \ldots, r_k\}$. We wish to show that θ is not too close to α. Note that $P(\theta) = 0$ and $P(\alpha)$ is a multiple of $\frac{1}{q^d}$. Since $\alpha \notin \{r_1, \ldots, r_k\}$, $P(\alpha) \neq 0$ so $|P(\alpha)| \geq \frac{1}{q^d}$ so $|P(\alpha) - P(\theta)| \geq \frac{1}{q^d}$. We shall exploit this.

Suppose

$$P(x) = a_d x^d + a_{d-1} x^{d-1} + \cdots + a_1 x + a_0.$$

Since

$$\alpha^k - \theta^k = (\alpha - \theta)(\alpha^{k-1} + \alpha^{k-2}\theta + \alpha^{k-3}\theta^2 + \cdots + \theta^{k-1})$$

we have

$$P(\alpha) - P(\theta) = (\alpha - \theta)\big[a_d(\alpha^{d-1} + \alpha^{d-2}\theta + \cdots + \theta^{d-1}) + a_{d-1}(\alpha^{d-2} + \cdots + \theta^{d-2}) + \cdots + a_2(\alpha + \delta) + a_1\big].$$

Let us suppose that $|\alpha - \theta| \leq 1$. In that case $|\alpha| \leq |\theta| + 1$, so the bit in the square brackets is certainly at most

$$c_\theta = |a_d|d(|\theta| + 1)^{d-1} + |a_{d-1}|(d-1)(|\theta| + 1)^{d-2} + \cdots + |a_2|2(|\theta| + 1) + |a_1|.$$

So for such α,

$$|\alpha - \theta| \geq \frac{|P(\alpha) - P(\theta)|}{c_\theta} \geq \frac{1}{c_\theta q^d}.$$

If α is one of the r_i then $|\alpha - \theta| \geq c_1 \geq \frac{c_1}{q^d}$.

If $|\alpha - \theta| \geq 1$ then $|\alpha - \theta| \geq \frac{1}{q^d}$.

Let $c = \min\{1, c_1, \frac{1}{c_\theta}\}$. Then $|\theta - \frac{p}{q}| \geq \frac{c}{q^d}$ for all $\frac{p}{q}$. □

Corollary 7.15. The number $\theta = \frac{1}{10^{1!}} + \frac{1}{10^{2!}} + \frac{1}{10^{3!}} + \cdots$ is transcendental.

Proof. Let $q = 10^{n!}$ and write θ as $\frac{p}{q} + \frac{1}{10^{(n+1)!}} + \frac{1}{10^{(n+2)!}} + \cdots$ for some suitable integer p. Then $|\theta - \frac{p}{q}| \leq \frac{2}{10^{(n+1)!}} = \frac{2}{q^{n+1}}$.

Now suppose that $|\theta - \frac{p}{q}| \geq \frac{c}{q^d}$ for all p, q. This is a contradiction if $n \geq d$ and $\frac{2}{q} < c$, both of which we can get by taking n large enough. □

7.11 Countability and Uncountability

Definition (Cardinality)**.** A set X is said to have *size* or *cardinality* n if there is a bijection between X and $\{1, 2, \ldots, n\}$. We write $|X|$. X is *finite* if there exists some $n \in \mathbb{N}$ such that $|X| = n$ (or $X = \emptyset$), otherwise *infinite*.

Lemma 7.16. There is no injection from $\{1, 2, \ldots, n+1\}$ to $\{1, 2, \ldots, n\}$ for any n.

Proof. Induction on n. Clearly there is no injection from $\{1, 2\}$ to $\{1\}$.

Suppose we have now a function ϕ from $\{1, 2, \ldots, n+1\}$ to $\{1, 2, \ldots, n\}$ and suppose it is an injection, and that n is minimal with this property. Define a function

$$\psi\colon \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n-1\}$$

as follows

$$\psi(m) = \begin{cases} \phi(m) & \text{if } \phi(m) \neq n \\ \phi(n+1) & \text{if } \phi(m) = n \end{cases}$$

If $\psi(m) = \psi(m')$ then $\phi(m) = \phi(m')$ so $m = m'$. So ψ is an injection, contradicting the minimality of n. So the lemma is proved. □

Corollary 7.17. If $m \neq n$ then $|X|$ cannot be m and n simultaneously.

Proof. Without loss of generality, $m < n$. If $|X| = n$ and $|X| = m$ then we have an injection $\phi\colon \{1, \dots, n\} \to X$ and an injection $\psi\colon X \to \{1, \dots, m\}$.

Composing these we get an injection from $\{1, \dots, n\}$ to $\{1, \dots, m\}$.

Since $n > m$, $n \geq m + 1$, so we can restrict attention to $\{1, \dots, m + 1\}$ and contradict the lemma. □

Proposition 7.18. $\mathbb{N}$ is infinite.

Proof. If $\mathbb{N}$ is finite then there is a bijection $\phi\colon \mathbb{N} \to \{1, \dots, n\}$ for some n. Restricting ϕ to $\{1, \dots, n + 1\}$ contradicts the lemma. □

Proposition 7.19. A set X is infinite if and only if there is an injection $\phi\colon \mathbb{N} \to X$.

Proof. If such a ϕ exists and $\psi\colon X \to \{1, \dots, n\}$ is a bijection then $\psi \circ \phi$ is an injection from $\mathbb{N}$ to $\{1, \dots, n\}$. Contradiction.

Conversely, suppose that X is infinite. Build an injection $\phi\colon \mathbb{N} \to X$ inductively as follows.

Step 1 $X \neq \emptyset$ so we can find $x \in X$. Let $\phi(1) = x$.

Step $k + 1$ Now suppose we have found $\phi(1), \dots, \phi(k) \in X$ and that they are all distinct. Then ϕ can be thought of as a bijection from $\{1, \dots, k\}$ to $\{\phi(1), \dots, \phi(k)\} \neq X$ so we can find $x \notin \{\phi(1), \dots, \phi(k)\}$. Call this $\phi(k + 1)$.

□

Definition (Countable)**.** A set X is called *countable* if there is an injection $\phi\colon X \to \mathbb{N}$.

Proposition 7.20. X is countable if and only if there is a surjection from $\mathbb{N}$ to X.

Proof. Suppose $\phi\colon X \to \mathbb{N}$ is an injection. Then ϕ has a left-inverse $\psi\colon \mathbb{N} \to X$. Then ϕ is a right-inverse for ψ, so ψ is a surjection. Similarly the other way round. □

Proposition 7.21. If X is infinite then X is countable iff there is a bijection between X and $\mathbb{N}$.

Proof. Let $\phi\colon X \to \mathbb{N}$ be an injection and let $Y = \phi(X)$ be its image. Then we can regard ϕ as a bijection between X and Y. We can write $Y = \{n_1, n_2, \dots\}$ with $n_1 < n_2 < \cdots$. (n_k is the minimal element of $Y \setminus \{n_1, \dots, n_{k-1}\}$.) Then define $\psi\colon Y \to \mathbb{N}$ by

$$\psi\colon n_j \mapsto j.$$

Then ψ is a bijection. So $\psi \circ \phi$ is bijection from X to $\mathbb{N}$. □

Theorem 7.22. $\mathbb{N} \times \mathbb{N}$ is countable.

Proof. List its elements as $(1,1)$, $(2,1)$, $(1,2)$, $(3,1)$, $(2,2)$, $(1,3)$, $(4,1)$, $(3,2)$ etc. and call these $\phi(1)$, $\phi(2)$, ... Then ϕ is a bijection from $\mathbb{N}$ to $\mathbb{N}^2$. □

Proof. Define $\phi\colon \mathbb{N}^2 \to \mathbb{N}$ by

$$\phi(m,n) = 2^{m-1}(2n-1).$$

This is a bijection. (Use fundamental theorem of arithmetic.) □

If X is infinite, it is countable iff

$$\left.\begin{array}{l} \exists \text{ injection } f\colon X \to \mathbb{N} \\ \exists \text{ surjection } f\colon \mathbb{N} \to X \\ \exists \text{ bijection } f\colon \mathbb{N} \to X \end{array}\right\} \text{ all equivalent}$$

Corollary 7.23. If $A_1, A_2, \dots$ are countable, then so is $\bigcup_{i=0}^{\infty} A_i$.

Proof. Let $A_1, A_2, \dots$ be countable sets. Write $A_r = \{a_{r1}, a_{r2}, a_{r3}, \dots\}$ (terminates if A_r is finite). Define a function $f\colon \mathbb{N} \times \mathbb{N} \to \bigcup_{n=1}^{\infty} A_n$ by

$$f(n,m) = \begin{cases} a_{nm} & \text{if it exists} \\ \text{anything} & \text{if there is no } a_{nm} \end{cases}$$

□

Example. $\mathbb{Q}$ is countable.

Proof. Let $A_n = \{\frac{0}{n}, \frac{1}{n}, \frac{-1}{n}, \frac{2}{n}, \frac{-2}{n}, \dots\}$. Each A_n is countable and $\mathbb{Q} = \bigcup_{n=1}^{\infty} A_n$. □

Proof. Let $A_n = \{\frac{p}{q} : p, q \in \mathbb{Z}, q \neq 0, |p| + |q| \leq n\}$. Each A_n if finite, so countable. $\mathbb{Q} = \bigcup_{n=1}^{\infty} A_n$. □

Example. Let $\mathbb{A}$ be the set of all *algebraic* numbers. $\mathbb{A}$ is countable.

Proof. Let A_n be the set of all roots of polynomials of the form

$$a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

with $-n \leq a_j \leq n$ for each j. There are $(2n+1)^{n+1}$ such polynomials and each has $\leq n$ roots, so A_n is finite. But $\mathbb{A} = \bigcup_{n=1}^{\infty} A_n$ so $\mathbb{A}$ is countable. □

Theorem 7.24 (Cantor)**.** $\mathbb{R}$ is uncountable.

Proof. Since there is a surjection from $\mathbb{R}$ to $[0,1)$, it is enough to show that $[0,1)$ is uncountable.

Let $\phi\colon \mathbb{N} \to [0,1)$ and write

$$\begin{aligned} \phi(1) &= 0.a_{11}a12a_{13}\dots \\ \phi(2) &= 0.a_{21}a22a_{23}\dots \\ \phi(3) &= 0.a_{31}a32a_{33}\dots \end{aligned}$$

Let $x = 0.b_1b_2b_3\dots$ where

$$b_n = \begin{cases} 1 & \text{if } a_{nn} \neq 1 \\ 4 & \text{if } a_{nn} = 1 \end{cases}.$$

Then the n^{th} decimal place of x is not the same as the n^{th} decimal place of $\phi(n)$.

Since there are no 0's or 9's in the expansion of x, $x \neq \phi(n) \quad \forall n \in \mathbb{N}$, so ϕ is not a surjection. □

Proof. Let $\phi\colon \mathbb{N} \to \mathbb{R}$ be a function.

Pick $a_1 < b_1$ such that $\phi(1) \notin [a_1, b_1]$. Now choose $a_2 < b_2$ such that $[a_2, b_2] \subset [a_1, b_1]$ and $\phi(2) \notin [a_2, b_2]$. Then choose $[a_3, b_3] \subset [a_2, b_2]$ such that $a_3 < b_3$ and $\phi(3) \notin [a_3, b_3]$ and so on ....

The sequence $a_1, a_2, a_3, \ldots$ is increasing an bounded above by b_1. So it has limit x. It is an easy exercise to show that $a_n \leq x$ for every n and $x \leq b_n$ for every n. So $x \in \bigcap_{n=1}^{\infty} [a_n, b_n]$. Since $\forall n \quad \phi(n) \notin [a_n, b_n]$, x does not equal any $\phi(n)$. So ϕ is not a surjection. □

Corollary 7.25. There are transcendental numbers.

Proof. $\mathbb{A}$ is countable and $\mathbb{R}$ is not. $\mathbb{A} \neq \mathbb{R}$. So $\mathbb{R} \setminus \mathbb{A} \neq \emptyset$. □

Theorem 7.26. Let X be any set and let $\wp(X)$ be the set of all subsets of X (the *power set* of X). Then there is no surjection from X to $\wp(X)$.

Proof. Let $\phi\colon X \to \wp(X)$. Let $Y = \{y \in X : y \notin \phi(y)\}$. Suppose there were some y such that $\phi(y) = Y$.

- If $y \in Y$ then $y \in \phi(y)$, so $y \notin Y$.
- If $y \notin Y$ then $y \notin \phi(y)$, so $y \in Y$.

So there cannot be a y with $\phi(y) = Y$. So ϕ is not a surjection. □

Corollary 7.27. $\wp(\mathbb{N})$ is uncountable.

Proof that $\mathbb{R}$ is uncountable. Let $\phi\colon \wp(\mathbb{N}) \to \mathbb{R}$ be defined by $\phi(A) = \sum_{n \in A} 10^{-n}$. If $A \neq B$ then $\phi(A)$ and $\phi(B)$ have different 9-free decimal expansions, so $\phi(A) \neq \phi(B)$. So ϕ is an injection. If $\mathbb{R}$ were countable, we could compose ϕ with an injection $\psi\colon \mathbb{R} \to \mathbb{N}$ and deduce that $\wp(\mathbb{N})$ was countable. Contradiction. □