Numbers and Sets *

Zhiyuan Bai

Compiled on September 18, 2022

This document serves as a set of revision materials for the Cambridge Mathematical Tripos Part IA course *Numbers and Sets* in Michaelmas 2019. However, despite its primary focus, readers should note that it is NOT a verbatim recall of the lectures, since the author might have made further amendments in the content. Therefore, there should always be provisions for errors and typos while this material is being used.

Contents

*Based on the lectures under the same name taught by Prof. I. B. Leader in Michaelmas 2019.

0 Introduction: Proofs and why we want them

Vaguely speaking, a proof is a kind of logical argument (or a series of it) which establishes a conclusion. This is not the most rigorous way to define it, but it would be enough for now. The big question, however, is why we need them. Why can't we just assume something that looks nice to be correct? Why can't we declare something to be right if it holds merely for the scope that we can reach? Of course, there is an answer. If not, mathematics would not have existed.

0.1 Why we need proofs

There are a few reasons why proofs are important and essential. The obvious one is that we need to be sure about the truthfulness of something. There are tonnes of example in mathematcis where a theorem holds for a shockingly large selection of numbers but not for all. We could, of course, give some silly examples like the claim "all positive integers are less than N" where you can substitute N for some very large integer, but there is a more interesting example:

Example 0.1 (Polya's Conjecture)**.** For any positive integer $N > 2$, at least half of the positive integers less than or equal to N has an odd number of prime factors (counting multiplicity).

Disproof. The smallest counterexample occurs at $N = 906150257$. □

We do have computers now that can test some of our conjectures up to some very large numbers, but for mathematicians, most of those computational powers are useless, [1] since they can do nothing about how the general picture is like.

0.2 Why we want proofs

We can need something that we don't want, like nuclear weapons. However, this is not the case for proofs. We need them, yes. But we still want them, desperately. Some say that the study of language is proofreading, then the study of mathematics would be proof-reading. What we really want is not only the however elegant results that we, or someone else, have proven. We should be more attracted to the (clever) idea behind. Why? Because if we come through

[1] There are some proofs that harness the power of computers, like the four-color theorem. But still – it is the proof we need, not any sort of verification.

another statement of similar kind, it would be in our advantage to solve them using the same technique. After all, mathematics is a creative art (of problem-solving). Quoting theorems won't get you anywhere in maths, what you really need is to empower yourself with the tricks in the proof.

0.3 Correct and incorrect proofs

First let us see how a correct proof is done:

Claim. *For any integer n, $n^3 - n$ is a multiple of 3.*

Note that when you state some variable, you need to state as well where its home is at. The statement would be false if n is not restricted to integers. Now here comes the proof.

Proof. $n^3 - n = (n-1)n(n+1)$. Since one of any three consecutive integers is a multiple of 3, one of $n-1, n, n+1$ is. So $n^3 - n$ is a multiple of 3. □

A proof almost always comes with a smart idea. The "smart" here doesn't necessarily mean that it is really smart, but a very important point that constitutes the basic of (part of) the proof. In this case, the smart idea is to try and factorize the expression to something that we can easily handle, which makes the problem quite obivous.
Now here come a non-proof, or incorrect proof:

Claim. *Let n be a integer. If n^2 is even, so is n.*

Non-proof. If n is even, then $n = 2k$ for some integer k, so $n^2 = 4k^2 = 2(2k^2)$. Therefore n^2 is even. □

Is it factually wrong? No. The logic in the proof itself is impeccable. So both the statement and the proof logic is correct, what is wrong? Quite obviously, it has proved the wrong thing. We want something like $A \implies B$, but what has been proven is $B \implies A$. So even both the statement and the proof logic are alright, it might still not be a proof. (Needless to say, if either of these two things is not, it would not be a proof either.) So back on the track.

Proof. If n^2 even but n is not, then $n = 2k + 1$ for some integer k, so $n^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1$ which is odd. This is a contradiction. □

This is a new way of proof. We do not take $A \implies B$ directly, but we use $\neg B \implies \neg A$. The equivalence of these two is called *reducio ad absurdum*, and the trick to use this in the proof is called a *proof by contradiction*.

0.4 "iff"

We want to prove the following claim:

Claim. *The solutions to $x^2 - 5x + 6$ are $x = 2$ and $x = 3$.*

Note that this claim is actually two claims: that $2, 3$ are indeed solutions to the equation, and that the equation does not have any other solutions. Therefore, to rephrase the question, we want to prove that $x^2 - 5x + 6 = 0$ *if and only if* $x = 2$ or $x = 3$. Here, *if and only if* means, well, what it literally means, that the two statements are equivalent to each other. They are interchangable and any one of them implies the other. So our proof must consist of both the forward and backward implication. Otherwise, it would be incomplete. One way of doing it is to seperately prove both sides of the claim, but we can also use a chain of "iff"s:

Proof. $x^2 - 5x + 6 = 0 \iff (x-2)(x-3) = 0 \iff (x = 2 \vee x = 3)$ □

However, if such a proof is to be used, one must make sure that the adjascent statements are indeed equivalent to each other.

0.5 Implicit use of assumption

The final proof error that we state here would be that sometimes we use an inappropriate assumption in the proof which make it invalid. Consider the following claim:

Claim. 1 *is the smallest positive real number.*

Nonsense. Let r be the smallest positive real number. If $r < 1$, then $r^2 < r$, contradiction. If $r > 1$, then $\sqrt{r} < r$, contradiction. Therefore $r = 1$ □

Why is this nonsense a nonsense? Because the proof (and the claim) both used a wrong assumption that there exists a smallest positive real number. But there isn't. Actually, the above proof and the fact that $1/2 < 1$ provides a proof of this fact.
So in proofs, we cannot assume, and need to avoid assuming implicitly, anything that might not be correct. If you get to assume something, prove it first.
Now that we are done with proofs, here comes the genuine stuff.

1 The Peano Axioms

We know what the natural numbers are, or do we? In school, we were often told a rather vague idea of the notion of natural numbers. We know that they are $1, 2, 3, \ldots$, but what actually are they? As every mathematical discipline, we do not know what is it until we define it using axioms. In the case of natural numbers, this is done by the Peano axioms:

Definition 1.1. Natural numbers is a triple $(\mathbb{N}, 1, +1)$, where $\mathbb{N}$ is a set, $1 \in \mathbb{N}$ its element, and $+1 : \mathbb{N} \to \mathbb{N}$ an operation on $\mathbb{N}$. It satisfies the following axioms:
1. For any $n \in \mathbb{N}$, $n + 1 \neq 1$.
2. For $n, m \in \mathbb{N}$, if $n \neq m$, then $n + 1 \neq m + 1$.
3. Let P be a proposition on $\mathbb{N}$. If $P(1)$ is true and $P(n) \implies P(n+1)$ for all $n \in \mathbb{N}$, then $P(n)$ is true for all $n \in \mathbb{N}$.

Definition 1.2 (Addition)**.** We define the operation $+k$ inductively by $n + (k + 1) = (n + k) + 1$.

Proposition 1.1. *The operation $+k$ is defined for all natural number k.*

Proof. Induction. □

Similarly, using induction, we can define multiplication, exponentiation and order ($a < b \iff \exists k \in \mathbb{N}, a + k = b$) in the obvious way.

Proposition 1.2. *1.* $(a + b) + c = a + (b + c)$.
2. $a + b = b + a$.
3. $(ab)c = a(bc)$.
4. $ab = ba$.
5. $a(b + c) = ab + ac$.
6. $a < b \wedge b < c \implies a < c$.
7. $\neg(a < a)$.

Proof. Induction. Induction. Induction. Induction. Induction. Induction. Induction. □

There is a more useful form of induction. Induction says that if we have some proposition P such that $P(1)$ is true and $P(n) \implies P(n + 1)$, then $P(n)$ is true. But in fact, we can have a "stronger" induction hypothesis.

Theorem 1.3 (Strong induction)**.** *Suppose that P is a proposition on $\mathbb{N}$. If $P(1)$ is true and for any $n \in \mathbb{N}$,*

$$\forall k \leq n, P(k) \implies P(n + 1)$$

Then $P(n)$ is true for any $n \in \mathbb{N}$.

Proof. Apply our usual induction on the proposition $Q(n)$ meaning $\forall k \leq n, P(k)$ □

Technically, we do not need to check the base case. But it is often safer to check it. If we want to use strong induction on P, we can prove $P(n)$ assuming the case for smaller numbers, as induction says if it would help to assume $P(m)$ for some $m < n$, feel free to do so. There are a few equivalent forms of strong induction which can be pretty useful.

Corollary 1.4. *If some n has $P(n)$ false, then there exists an n with $P(n)$ false but $P(m)$ true for every $m < n$.*

That is, if there is a counterexample, then there is a minimal counterexample.

Corollary 1.5 (Well-ordering Principle)**.** *If some n has $P(n)$ true, then there exists a minimal n with $P(n)$ true.*

2 The Integers and the Rationals

2.1 Integers

The integers $\mathbb{Z}$ consists of all expressions $n, -n$ where n is a natural number, and 0. We can define $+, \times$ etc. in the obvious way. And it is easy and trivial to check all the necessary rules apply. And we define $a < b$ as $a + c = b$ for some natural number c. All previous rules appply except we need c to be positive to have $a < b \implies ac < bc$. Also, for any a, $a + 0 = a$ and that there is a b such that $a + b = 0$. This makes the integers a group.

2.2 Rationals

We can define the rationals $\mathbb{Q}$ as well. It shall consist of all expressions a/b for some integers a, b with $n \neq 0$. And we shall have

$$a/b = c/d \iff ad = bc$$

And we define

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}$$

We have to check that this is well defined, since $\mathbb{Q}$ is constructed based on equivalence classes. For example, we cannot assign an operation sending $a/b \to a^2/b^3$ because it will be ill defined, as $1/2$ and $2/4$ go to different places.

Proposition 2.1. *The addition is well-defined on $\mathbb{Q}$.*

Proof. Trivial. □

Proposition 2.2. *We can define multiplication similarly that satisfies all usual rules and that every nonzero rational number has an inverse under multiplication.*

Proof. Trivial. □

So the rationals excluding 0 is a group.
As for order, $a/b < c/d \iff ad < bc$ for $b, d > 0$. We can also check all the rules we want

3 Elementary Number Theory

Let n be an natural number. The multiples of n are all integers kn where $k \in \mathbb{Z}$. For example, $2n, 5n, n, -4n, 0$ are all multiplies of n.

Definition 3.1. If m is a multiple of n, we say n divides m, or n is a divisor of m, written as $n|m$.

Definition 3.2. We say a natural number $n \geq 2$ is a prime if it has no divisors apart from 1 and n.
Otherwise, we call it to be composite.

Example 3.1. $2, 3, 5, 7, 11, 13, 17, \ldots$ are prime.
$10, 25, 34, 44, 57$ are composite.

Our aim now is to break up a number into primes, for example $63 = 3$, and hopefully it would be unique.

Proposition 3.1. *Every natural number $n \geq 2$ is expressible as a product of primes.*

Proof. Strong induction.
The statement is obvious true for $n = 2$. Given an $n > 2$, if it is prime, then it is done. Otherwise, we can write $n = ab$ where $n > a, b > 1$. By induction hypothesis, we can write a, b as a product of primes.

$$a = p_1 p_2 \cdots p_k, b = q_1 q_2 \cdots q_l$$

where p_i, q_i are primes. So $n = p_1 p_2 \cdots p_k q_1 q_2 \cdots q_l$, therefore the statement is true for n. And it's done. □

Remark. 1. Although we start at 2, sometimes you can regard 1 as a product of no primes.
2. There is no nice pattern of primes. There is no algebraic formula for the n^{th} prime.

Theorem 3.2. *There are infinitely many primes.*

Proof. Suppose for the same of contradiction that there is only finitely many, let

$$p_1, p_2, \ldots, p_n$$

be the primes. So consider $k = p_1 p_2 \cdots p_n + 1$, it does not have any prime factors as none of the primes divide k. This is a contradiction to the fact that k has a prime factorisation. □

Theorem 3.3. *The prime factorization of any positive integer $n \geq 2$ is unique.*

The uniqueness is taken up to re-ordering. Why is it true? Why can't $41 \times 101 = 67 \times 73$? We need $p|ab \implies p|a \vee p|b$ where p is a prime number. We need p to be prime since $6|8 \times 3$ but $6 \nmid 8$ and $6 \nmid 3$. This has to be hard to prove since it is about prime dividing things instead of things dividing primes, as in the definition.

3.1 Highest Common Factors

Definition 3.3. For integers a, b we say the positive integer c is the HCF (Highest Common Factor) of a, b if
1. $c|a$ and $c|b$. (c is a common factor of a, b)
2. For any positive integers d such that $d|a$ and $d|b$, then $d|c$. (every common factor of a, b divides c).

Example 3.2. The HCF of 18 and 12 is 6.

We want to show that an HCF always exists

Proposition 3.4 (Division Algorithm)**.** *For natural numbers n, k, we can write $n = qk + r$, for some integers q, r with $0 \leq r < k$.*

Proof. Induction on n. $n = 1$ is trivial.
Given $n > 1$, we have $n - 1 = qk + r$ for some integers q, r with $0 \leq r < k$.
If $r < k-1$, then we have $n = qk+(r+1)$. If $r = k-1$, then $n = (q+1)k+0$. □

We can find HCF by Euclid's Algorithm.

Definition 3.4. The HCF of a and b where $a \geq b$.
Write $q_1 b + r_1$ (where $0 \leq r_1 < b$).
Then write $b = q_2 r_1 + r_2$ (where $0 \leq r_2 < r_1$).
Then write $r_1 = q_3 r_2 + r_3$ (where $0 \leq r_3 < r_2$).
Continue this process until some remainder $r_{n-1} = q_{n+1} r_n + r_{n+1}$ where $r_{n+1} = 0$. (If necessary, take $a = r_{-1}, b = r_0$) That is, some remainder goes to 0.

Example 3.3. $372, 162$
$372 = 2 \times 162 + 48$.
$162 = 3 \times 48 + 18$.

$48 = 2 \times 18 + 12$.
$18 = 1 \times 12 + 6$.
$12 = 2 \times 6 + 0$.
So the HCF of $372, 162$ is 6.

Theorem 3.5. *The Euclid's Algorithm works.*
That is, the HCF exists and can always be found by Euclid's Algorithm.

Proof. Let r_n be the output. Firstly, it is a common factor of a, b, since $r_n|r_{n-1}, r_n|r_{n-2}| \cdots, r_n|b, r_n|a$ inductively.
Then, for any other common factor d of a, b, we know that $d|r_1$, inductively $d|r_i$ for every i, then $d|r_n$.
Note that the algorithm always terminate, since the sequence r_i is strictly decreasing. In fact, there are at most b steps before it terminates. □

Example 3.4. $82, 57$
$82 = 57 \times 1 + 25$.
$57 = 25 \times 2 + 7$.
$25 = 7 \times 3 + 4$.
$4 = 3 \times 1 + 1$.
$3 = 3 \times 1 + 0$.
So the HCF of $82, 57$ is 1.

If two positive integers have HCF 1, we say that they are coprime.
Can we write $1 = 82x + 57y$ for some integers $x, y \in \mathbb{Z}$? So we have $1 = 4 - 3 = 4 - (7 - 4) = 2 \times 4 - 7 = 2 \times (25 - 3 \times 7) - 7 = 2 \times 25 - 7 \times 7 = 2 \times 25 - 7 \times (57 - 2 \times 25) = -7 \times 57 + 16 \times 25 = -7 \times 57 + 16 \times (82 - 57) = -23 \times 57 + 16 \times 82$. In fact, this algorithm provides a proof that it always works that for any coprime p, q we have $x, y \in \mathbb{Z}$ with $1 = xp + yq$.

Theorem 3.6. $\forall a, b \in \mathbb{N}, \exists x, y \in \mathbb{Z}, \mathrm{HCF}(a, b) = xa + yb$.

Proof. Run the Euclid's Algorithm on a, b to r_n. We have r_n written as an integral combination of r_{n-1}, r_{n-2}. Then substitute for r_{n-1} to obtain r_n as an integral combination of r_{n-2}, r_{n-3}. Inductively, we can write $\mathrm{HCF}(a, b) = r_n$ as an integral combination of a, b. □

Remark. Euclid is telling us that such x, y exists and how to find them in practice.

There is a second proof to this statement.

Alternative Proof. Consider the set $\{ax + by : x, y \in \mathbb{Z}\}$, and let h be the least positive integer in this set.
We claim that h is the HCF of a and b. If $d|a$ and $d|b$ then $d|ax + yb \implies d|h$.
Now h must be a common factor of a, b. Suppose that $h \nmid a$, then $a = qh + r$ where $q \in \mathbb{Z}, 0 < r < h$. However, r would be an integral combination of a, b, but this contradicts the minimality of h.
So $h|a$ and similarly $h|b$, so h is an HCF of a, b. □

Remark. The alternative proof is abstract, nice, and more concise, but it is non-constructive. It does not tell us how to find such an integral combination.

One of the applications of this fact is to solve (linear) Diophantine Equations. Fix $a, b \in \mathbb{N}$, when can we solve $ax = b, x \in \mathbb{Z}$? Obviously we have an solution if and only if $a|b$ and the solution is b/a. But how about 2 variables? Suppose $a, b, c \in \mathbb{N}$, when can we solve $ax+by = c$ in $\mathbb{Z}$? We cannot solve $116x+212y = 13$ due to parity problems, but we can solve $82x + 57y = 13$ by multiplying 13 to the integral combination of 1 by $82, 57$

Corollary 3.7 (Bezout Theorem)**.** *$ax + by = c$ is solvable in $\mathbb{Z}$ if and only if* $\mathrm{HCF}(a, b)|c$.

Proof. h the HCF of a, b.
If there is a solution, we have $c = ax + by$ for some $x, y \in Z$ so $h|ax + by = c$.
Conversely, if $h|c \implies c = qh, q \in \mathbb{N}$, since we can write $h = ax + by$ for some $x, y \in Z$, $c = qh = a(qx) + b(qy)$. □

We are now ready to prove

Proposition 3.8. *Let p be prime, $a, b \in \mathbb{N}$, then $p|ab \implies p|a \vee p|b$.*

Proof. Given that $p|ab$, suppose that $p \nmid a, p \nmid b$. So the HCF of p, a is 1, therefore $\exists x, y \in \mathbb{Z}, px + ay = 1$. Note that here we have produced a positive statement here. So $pbx + aby = b$, but since $p|ab, p|pbx + aby = b$, which is a contradiction. □

This establishes the statement we have claimed in the preceding section.

Remark. Immediately, if $p|a_1a_2 \dots a_n$, then $p|a_i$ for some i.

3.2 Fundamental Theorem of Arithmetic

When we have established the theory so far, we are ready to show

Theorem 3.9 (Fundamental Theorem of Arithmetic)**.** *Any positive integer $n \geq 2$ can be written as a product of primes uniquely up to reordering.*

Proof. We already know that any positive integer $n \geq 2$ can be written as such a product, so the rest is to prove its uniqueness up to reordering.
Uniqueness can be proved by (strong) induction on n. $n = 2$ is obvious. Given $n > 2$, suppose that

$$p_1p_2 \cdots p_k = n = q_1q_2 \cdots q_l$$

where p_i, q_i are primes, we want to show $k = l$ and after reordering $p_i = q_i, \forall i$.
We have $p_1|q_1q_2 \cdots q_l$, so $p_1|q_i$ for some i. We can reorder such that $q_i \mapsto q_1$. Note that due to primity, $p_1 = q_1$, thus

$$p_2p_3 \cdots p_k = q_2q_3 \cdots q_l$$

By induction hypothesis, $k - 1 = l - 1 \implies k = l$ and we can reorder such that $p_i = q_i, i \geq 2$. So the theorem is proved. □

What ideas are involved? We took the things that cannot be broken up (i.e. primes) and we break everything up into the product of those 'unbreakables' (irreducibles).
However, this may not be the case. Despite the fact that we are used to it, unique factorization really isn't obvious.

Example 3.5. Consider the set $\mathbb{Z}(\sqrt{-3}) = \{a + b\sqrt{-3} : a, b \in \mathbb{Z}\}$.
We can do addition and multiplication in it in the obvious way, and the set is closed under both. So we can define 'divides' and 'factor of' etc. One can show that everything can be broken into irreducibles. But $4 = 2 \times 2 = (1+\sqrt{-3})(1-\sqrt{-3})$, and $2, 1 \pm \sqrt{-3}$ are all irreducibles.
Therefore in this funny set, the fundamental theorem of arithmetic doesn't work.

There are a few applications of unique factorization. First of all, we want to look into factors. Consider $n = 2^3 \cdot 3^7 \cdot 5 \cdot 11$, we can spot factors of the form $2^a \cdot 3^b \cdot 5^c \cdot 11^d, 0 \le a \le 3, 0 \le b \le 7, 0 \le c \le 1, 0 \le d \le 1$. We do not have others because of unique factorization. In general, the factors of $n = p_1^{a_1} \cdots p_k^{a_k}$ where p_i are distinct primes are of the form $p_1^{b_1} \cdots p_k^{b_k}$ where $0 \le b_i \le a_i$.
Also, we can find HCF easily. For example, the common factors of $2^3 \cdot 3^2 \cdot 5 \cdot 11$, $2^2 \cdot 3^6 \cdot 5 \cdot 11$ are of the form $2^a \cdot 3^b \cdot 11^c$ where $0 \le a \le 2, 0 \le b \le 2, 0 \le c \le 1$, so the HCF is $2^2 \times 3^2 \times 11$. In general, the HCF of $p_1^{a_1} \cdots p_k^{a_k}$ and $p_1^{b_1} \cdots p_k^{b_k}$ is $p_1^{\min\{a_1,b_1\}} \cdots p_k^{\min\{a_k,b_k\}}$.
We can find LCM as well. The LCM of the numbers in the last example would be $2^3 \times 3^6 \times 5 \times 7 \times 11$ In general, the LCM of $p_1^{a_1} \cdots p_k^{a_k}$ and $p_1^{b_1} \cdots p_k^{b_k}$ is $p_1^{\max\{a_1,b_1\}} \cdots p_k^{\max\{a_k,b_k\}}$.
Note that $\mathrm{HCF(a, b)} \times \mathrm{LCM(a, b)} = a \times b$.

3.3 Modular Arithmetic

Definition 3.5. Let $n \ge 2$ be a positive integer, then $\mathbb{Z}_n$ consists of integers with two of them regarded as the same if their difference is a multiple of n.

For example, in $\mathbb{Z}_7$, 2 and 16 are the same.
If a and b are the same in $\mathbb{Z}_n$, we write $a \equiv b \pmod{n}$.
In this world, we only care about the remainder when we divide it by n. In $\mathbb{Z}_n$, $0, 1, \ldots, n-1$ are distinct and every $k \in \mathbb{Z}_n$ is one of them by division algorithm. So we can view $\mathbb{Z}_n$ as an n-clock.
We can do addition and multiplication in $\mathbb{Z}_n$. Note that parity does not make sense in $\mathbb{Z}_{2n-1}$. So we would have to check that they are well-defined.

Proposition 3.10. *Suppose* $a \equiv a' \pmod{n}, b \equiv b' \pmod{n}$, *then* $a+b \equiv a'+b' \pmod{n}, ab \equiv a'b' \pmod{n}$.

Proof. Trivial. □

All the usual laws of arithmetic applies as $\mathbb{Z}_n$ inherited them from $\mathbb{Z}$.
Something we have done so far can already be expressed in terms of modular arithmetic. For example, $ab \equiv 0 \pmod{p} \implies a \equiv 0 \pmod{p} \vee b \equiv 0 \pmod{p}$ if p is a prime number. Or equivalently, there is no zero divisor in $\mathbb{Z}_p$.
The structure of $\mathbb{Z}_n$ under addition is boring enough (just a cyclic group), but how about it under multiplication?

Definition 3.6. In $\mathbb{Z}_n$, we say a is the inverse of b if $ab \equiv 1 \pmod{n}$.

Example 3.6. 1. In $\mathbb{Z}_{10}$, $3 \times 7 \equiv 1 \pmod{10}$ so 7 is the inverse of 3.
2. (non-example) There is no inverse of 4 in $\mathbb{Z}_{10}$ since $4b$ is even for all $b \in \mathbb{Z}$ so we can never have $4b \equiv 1 \pmod{10}$.

If a has an inverse b, we write $b = a^{-1}$ given that the n in $\mathbb{Z}_n$ is understood.

Remark. 1. If inverse exists, it is unique (in $\mathbb{Z}_n$). Indeed, suppose $ab \equiv ac \equiv 1 \pmod{n}$, but then $b \equiv bab \equiv bac \equiv c \pmod{n}$.
2. If $ab \equiv ac \pmod{n}$ and a has an inverse, then $b \equiv c \pmod{n}$ by multiplying both sides by a^{-1}. However, if a does not have an inverse, you cannot really cancel it. For example $4 \times 5 \equiv 4 \times 0 \pmod{10}$, but $5 \not\equiv 0 \pmod{10}$.

$\mathbb{Z}_n$ is quite nice if n is prime.

Proposition 3.11. $\forall a \not\equiv 0 \pmod{p}, \exists b \in \mathbb{Z}_p, ab \equiv 1 \pmod{p}$.

Proof. We know that $(a, p) = 1$, so there are some $x, y \in \mathbb{Z}, ax + py = 1 \implies ax \equiv 1 \pmod{p}$. We can take $b = x$. □

Alternative proof. IN $\mathbb{Z}_p$, consider $0a, 1a, 2a, \ldots, (p-1)a$. Our task is to show that one of these equals 1. Note that no two of them are equal, since $ia \equiv ja \pmod{p} \implies (i-j)a \equiv 0 \pmod{p} \implies i \equiv j \pmod{p}$.
Therefore $\{ka : k \in \{0, 1, \ldots, p-1\}\} = \{0, 1, \ldots, p-1\}$. so there is some b such that $ba \equiv 1 \pmod{p}$. □

How about the case in $\mathbb{Z}_n$ when n is composite?

Proposition 3.12. $\forall a \in \mathbb{Z}$, *there is some* b *such that* $ab \equiv 1 \pmod{n}$ *if and only if* $(a, n) = 1$.

Proof. If a is invertible, then there is some $b, y \in \mathbb{Z}$ such that $ab + ny = 1$, which means that $(a, n) = 1$.
Conversely, if $(a, n) = 1$, then there is some $x, y \in \mathbb{Z}, ax + ny = 1 \implies ax \equiv 1 \pmod{n}$, so we can take $b = x$. □

Definition 3.7. The Euler ϕ function is defined by

$$\phi(n) = |\{0 < a < n : (a, n) = 1\}|$$

Equivalently, $\phi(n)$ is the number of invertible elements in $\mathbb{Z}_n$.

Example 3.7. 1. $\phi(p) = p - 1$ for any prime p.
2. $\phi(p^2) = p^2 - p = p(p-1)$ for any prime p.
3. $\phi(pq) = pq - p - q + 1 = (p-1)(q-1) = \phi(p)\phi(q)$ for any distinct primes p, q.
4. $\phi(p^n) = p^n - p^{n-1} = p^{n-1}(p-1)$ for any prime p and positive integer n.

We now introduce the order of an element by a few observations: In $\mathbb{Z}_7$, $2^1 \equiv 2, 2^2 \equiv 4, 2^3 \equiv 1$ and things go around again. In $\mathbb{Z}_{11}$, $2^1 \equiv 2, 2^2 \equiv 4, 2^3 \equiv 8, 2^4 \equiv 5, 2^5 \equiv 10, \ldots, 2^{10} \equiv 1 \pmod{1}1$ and things repeat.

Theorem 3.13 (Fermat's Little Theorem (FℓT))**.** *Let* p *be a prime, then in* $\mathbb{Z}_p$, *every* $a \not\equiv 0$ *has* $a^{p-1} \equiv 1$.

Proof. Note that $\{ka : k \in \{1, \ldots, p-1\}\} = \{1, \ldots, p-1\}$.
So we have

$$\prod_{k=1}^{p-1} ak \equiv \prod_{k=1}^{p-1} k \pmod{p} \implies (a^{p-1} - 1)(p-1)! \equiv 0 \pmod{p}$$

Now $(p-1)! \not\equiv 0 \pmod{p}$ since all the terms in the product is invertible. Therefore the theorem. □

For composite n, we have a similar proposition.

Theorem 3.14 (Fermat-Euler Theorem)**.** *Let $n \geq 2$ be a positive integer, then in $\mathbb{Z}_n$, every invertible a has $a^{\phi(n)} \equiv 1$.*

Proof. Note that $\{ka : 0 < k < n, (k,n) = 1\} = \{k : 0 < k < n, (k,n) = 1\}$. Since a is invertible, we have $ia \equiv ja \pmod{n} \implies i \equiv j \pmod{n}$
So we have

$$\prod_{0<k<n,(k,n)=1} ak \equiv \prod_{0<k<n,(k,n)=1} k \pmod{n}$$
$$\implies (a^{\phi(n)} - 1) \prod_{0<k<n,(k,n)=1} k \equiv 0 \pmod{n}$$

Now $\prod_{0<k<n,(k,n)=1} k \neq 0 \pmod{n}$ since all the terms in the product is invertible. Therefore $a^{\phi(n)} \equiv 1 \pmod{n}$. □

Lemma 3.15. *Let p be prime, then in $\mathbb{Z}_p$, the solutions to $x^2 \equiv 1 \pmod{p}$ are $x = \pm 1$.*

Note that it is not true when p is not prime. For example, $1^2 \equiv 3^2 \equiv 5^2 \equiv 7^2 \equiv 1$.

Proof. $x^2 \equiv 1 \pmod{p} \iff (x-1)(x+1) \equiv 0 \pmod{p} \iff x \equiv \pm 1 \pmod{p}$ since p is prime. □

Remark. Any nonzero polynomial of degree d in $\mathbb{Z}_p$ has at most d solutions.

Note that in the proof of FℓT, there is the expression $(p-1)!$. It looks like as if it is some interesting thing. Note that $(3-1)! \equiv -1 \pmod{3}, (5-1)! \equiv -1 \pmod{5}, (7-1)! \equiv -1 \pmod{7}$. So it is natural to state the following, which is in fact true:

Theorem 3.16 (Wilson's Theorem)**.** *Let p be a prime, then $(p-1)! \equiv -1 \pmod{p}$*

Proof. When $p = 2$ it is trivial, so we assume henceforth that $p > 2$.
Note that all of $1, 2, \ldots, p-1$ are invertible. So we can pair up the elements with their inverses, and they have a product 1. We have the pairs as long as we do not have $x^2 \equiv 1 \pmod{p} \implies x \equiv 1 \vee x \equiv p-1 \pmod{p}$. Hence $(p-1)! \equiv 1(p-1) \equiv -1 \pmod{p}$. □

When is -1 a square modulo p? Is there $x \in \mathbb{Z}$ with $x^2 \equiv -1 \pmod{p}$?

Example 3.8. When $p = 5$, we notice that $2^2 \equiv -1 \pmod{5}$ When $p = 7$, by trying, we know that such x does not exist. When $p = 13$, $5^2 = 25 \equiv -1 \pmod{13}$. when $p = 19$, by trying, we know that such x does not exist.

By looking at the pattern, we can try to prove the following theorem

Theorem 3.17. *For a prime number $p > 2$, then $x^2 \equiv -1 \pmod{p}$ is solvable if and only if $p \equiv 1 \pmod{4}$.*

Proof. If $p = 4k + 3$ for some $k \in \mathbb{N}$ but $x^2 \equiv -1 \pmod{p}$ for some $x \in \mathbb{Z}$. But we also have $-1 \equiv (x^2)^{2k+1} = x^{4k+2} = x^{p-1} \equiv 1 \pmod{p}$, which is a contradiction.
Conversely, if $p = 4k + 1$, Wilson's theorem tells us that $(4k)! \equiv -1 \pmod{p}$. Note that $4k - r \equiv -r - 1 \pmod{p}$, therefore $((2k)!)^2 = ((2k)!)^2(-1)^{2k} \equiv (4k)! \equiv -1 \pmod{p}$. □

Now we go back solving linear congruences.

Example 3.9. If we want to solve $7x \equiv 4 \pmod{30}$. Firstly we can find $7 \cdot 13 \equiv 1 \pmod{30}$, we knew we can do this since $(7, 30) = 1$. So $7x \equiv 4 \pmod{30} \iff 13 \cdot 7x \equiv 13 \cdot 4 \pmod{30} \iff x \equiv 22 \pmod{30}$.

Example 3.10. Solve $10x \equiv 12 \pmod{34}$. Note that in this case $(10, 34) \neq 1$ so we cannot do the same thing again. However, we can throw it back to $\mathbb{Z}$, so $10x \equiv 12 \pmod{34} \iff \exists y \in \mathbb{Z}, 10x = 12 + 34y \iff \exists y \in \mathbb{Z}, 5x = 6 + 17y \iff 5x \equiv 6 \pmod{17}$. So from here on we can do the same thing again. $5x \equiv 6 \pmod{17} \iff x \equiv 7 \times 6 \equiv 8 \pmod{17}$.

Now we want to try simultaneous ones.

Example 3.11.

$$\begin{cases} x \equiv 3 \pmod{17} \\ x \equiv 5 \pmod{19} \end{cases}$$

Can we solve it? We would expect the answer being yes, since modulo $17, 19$ should not "intervene each other" since $(17, 19) = 1$. That is not a proof, but that is our intuition.

Example 3.12.

$$\begin{cases} x \equiv 5 \pmod{30} \\ x \equiv 8 \pmod{34} \end{cases}$$

It then becomes immediate that this is not solvable due to parity problem. What has gone wrong looks like that two modulos are related as $(30, 34) = 2$.

Theorem 3.18 (Chinese Remainder Theorem)**.** *The system of linear equations*

$$\begin{cases} x \equiv a \pmod{u} \\ x \equiv b \pmod{v} \end{cases}$$

is solvable if $(u, v) = 1$*. Moreover, the solution is unique modulo* uv*.*

Proof. Existence: Since they are coprime, we have $su + tv = 1$ for some $s, t \in \mathbb{Z}$, then $x = bsu + atv$ solves the system.
Uniqueness: If x, x' are solutions, then $x - x' \equiv 0 \pmod{u}$ and $x - x' \pmod{v}$, hence $x - x' \equiv 0 \pmod{uv}$. Conversely it is obvious that if $x \equiv x' \pmod{uv}$, then if x is a solution so is x'. □

The exact same thing would work if you have more than 2 equations in the system modulo pairwisely coprime numbers. This can be proved by CRT and induction.

3.4 The RSA Cryptosystem

The RSA code is an example of an application of the Fermat-Euler theorem. Consider the following scenario: we want to send an encoded message to the receiver who is supposed to have a way to decode it.
It seems "obvious" that knowing how to decode equals knowing how to encode. However, there is a way such that even if you know how to encode, it is still "very hard" to decode. That is, the process of finding an inverse function will takes a long period of time for the message to expire.
We encode in the following way: Pick two large primes p, q, say a hundred digits each and take their product. A message is just a sequence of digits, so we can decompose it into not-so-long blocks such that the number represented by each block is less than both of the primes. Say one of these blocks is x, and we take an exponent $e \geq 2$ such that $(e, \phi(pq)) = 1$, and the encoding message would be the smallest positive integer with $x^e \equiv m \pmod{pq}$ So to decode, we can find some d with $de \equiv 1 \pmod{\phi(pq)}$, so $m^d \equiv x \pmod{pq}$ by Fermat-Euler. This is the way to decode the message, so we just need to find such d, which is easy and fast by Euclid given that we do know $\phi(pq)$.
Now we only need to know pq and e to encode, and we need to know d (or e and $\phi(pq) = (p-1)(q-1)$, since we can run Euclid) and pq to decode. Observe here that if the decoder only know all eise information but $\phi(pq)$ explicitly, he will have to factorize pq. Thus, we can publish pq and e, then anyone else can send a message to us, however, since only us know d, no one else can decode the message quickly since they will have to run a slow algorithm to try factorizing pq.

4 The Reals

4.1 The Need of the Reals

We already have $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$, then why are we introducing the reals? Why don't we stop at $\mathbb{Q}$?

Proposition 4.1. *There is no rational x with $x^2 = 2$.*

we can assume x is positive since $(-x)^2 = x^2$.

Proof. Suppose we have some rational x with $x^2 = 2$, then suppose $x = a/b$ for some positive integers a, b. So $2b^2 = a^2$, but the power of 2 in the prime factorization of a^2 is even but that in $2b^2$ is odd. This is a contradiction. □

Remark. Same argument shows that if $k \in \mathbb{N}$ is x^2 for some $x \in \mathbb{N}$, then k is a perfect square.

Alternative proof. Again suppose that $x = a/b$ where $a, b \in \mathbb{N}$ has $x^2 = 2$. So every $cx + d$ where $c, d \in \mathbb{Z}$ is of the form e/b for some $e \in \mathbb{Z}$. So $cx + d > 0 \implies cx + d \geq 1/d$, but then $0 < x - 1 < 1$, so $0 < (x-1)^n < 1/b$ for n large, which is a contradiction since all $(x-1)^n$ is of the form $cx + d$ for some integer c, d using $x^2 = 2$. □

How, in $\mathbb{Q}$, could we say $\mathbb{Q}$ has a gap? Consider the set of rationals whose squares are less than 2. Now all of these rationals are less than 2, so 2 is an upper bound. $1.5, 1.42, 1.415, \ldots$ are all upper bounds of this set, but there is no least upper bound. So that establishes the notion of there being a "gap", which is the thing we try to eliminate in our dream model of $\mathbb{R}$.

Definition 4.1. The real numbers $\mathbb{R}$ is an ordered field with the least upper bound property.
Put it into axiom, the reals consists of a set $\mathbb{R}$, the binary operations $+, \times$, elements $1 \neq 0$, and an order $<$ where
1. $(\mathbb{R}, +, 0)$ is an abelian group.
2. $(\mathbb{R}, \times, 1)$ is an abelian group.
3. $\times$ is distributive over $+$, so $\forall a, b, c \in \mathbb{R}, a \times (b + c) = a \times b + a \times c$.
4. For any $a, b \in \mathbb{R}$, exactly one of $a < b, a = b, a > b$ is true. Also $a < b \wedge b < c \implies a < c$.
5. $\forall a, b, c \in \mathbb{R}$, $a < b \implies a + c < b + c$, $c > 0 \implies (a < b \implies ac < bc)$.
6. (Least-upper-bound property) For any nonempty $S \subset \mathbb{R}$ such that it is bounded from above (i.e. $\exists M \in \mathbb{R}, \forall x \in S, x < M$), there is a least upper bound of S.

Remark. 1. We can already conclude $0 < 1$ from axioms 1 to 5. Indeed, if not, then $1 < 0$, so $0 < -1$, so $0 = 0(-1) < (-1)(-1) = 1$, contradiction.
2. We can embed $\mathbb{Q}$ into $\mathbb{R}$, but the Least-upper-bound property is false in $\mathbb{Q}$.
3. We do need the "nonempty" and "bounded above" conditions. We need the former to ensure that there is indeed some element in S to talk about order, and we need the latter to ensure that there is at least one upper bound.
4. We could constuct $\mathbb{R}$ from $\mathbb{Q}$ and we can check that axioms 1 to 6 hold.

We write $\sup S$ to denote the least upper bound of S.

Example 4.1. 1. $S = \{x \in \mathbb{R} : 0 \leq x \leq 1\}$ be the closed interval $[0, 1]$. 2 is an upper buond for S since $\forall x \in S, x \leq 2$, but $3/4$ is not since $1 \in S$ but $1 > 3/4$. The least upper bound is 1, since 1 is an upper bound for S and for every upper bound s of S must be at least 1 since $1 \in S$.
2. $S = \{x \in \mathbb{R} : 0 < x < 1\}$ be the open interval $(0, 1)$. Again, 2 is an upper bound but $3/4$ is not since $S \ni 5/6 > 3/4$. The least upper bound is 1 since 1 is an upper bound and for any upper bound s, if $s < 1$ (note that $s < 0$), we must have $S \ni (1 + s)/2 > s$ which is a contradiction.
3. $S = \{1 - 1/n : n \in \mathbb{N}\}$, then $\sup S = 1$, since 1 is clearly an upper bound and if there is an upper bound $s \in S$ such that $0 < s < 1$, then we take a natural number $n > 1/(1 - s)$, then $1 - 1/n > 1 - (1 - s) = s$.

Note that in example 3, we have assumed the following proposition which we shall prove now.

Proposition 4.2 (Axiom of Archimedes)**.** $\forall x \in \mathbb{R}, \exists n \in \mathbb{N}, n > x$.

Proof. Suppose not, the $\mathbb{N}$ is bounded above, so let $c = \sup \mathbb{N}$, then $c - 1$ is not an upper bound of $\mathbb{N}$, so $\exists n \in \mathbb{N}, n > c-1 \implies \mathbb{N} \ni n+1 > c$, contradiction. □

Corollary 4.3. *Let $t \in \mathbb{R}_{>0}$, then there is some $n \in \mathbb{N}$ such that $1/n < t$.*

Proof. Take $n > 1/t$. □

So the reals do not contain infinity or infinitesimal.

Remark. 1. $\sup X$ might not be in X. E.g. $\sup(0,1) = 1 \notin (0,1)$
2. The least-upper-bound property gives as well the existence of greatest lower bound for bounded-below subsets of $\mathbb{R}$. Indeed, suppose $S \neq \varnothing$ is bounded below, then $-S = \{-s : s \in \mathbb{S}\}$ is bounded above, so there is some $x = \sup -S$. Immediately $-x$ would be the greatest lower bound of S. We denote it by $\inf S$.

Theorem 4.4. *There is $x \in \mathbb{R}$ such that $x^2 = 2$.*

Proof. Consider the set $S = \{r \in \mathbb{R} : r^2 < 2\}$ which is nonempty ($1 \in S$) and bounded above ($\forall s \in S, s < 2$). So there is some $c = \sup S$, also $1 \leq c < 2$.
Consider c^2, if $c^2 < 2$, then for any $0 < t < 1$, we have $(c+t)^2 = c^2 + 2ct + t^2 \leq c^2 + 5t$, so we choose any $0 < t < (2 - c^2)/5$, then $(c+t)^2 < 2$, so $c + t \in S$ but $c + t > c$, contradicting the assumption that c is an upper bound.
If $c^2 > 2$, then for any $0 < t < 1$, we have $(c-t)^2 = c^2 - 2ct + t^2 \geq c^2 - 4t$, so we can choose t such that $0 < t < (c^2 - 2)/4$, therefore $(c-t)^2 > 2$, thus $c - t$ is an upper bound of S as well, contradicting the fact that c is the least upper bound.
Therefore $c^2 = 2$. □

Similarly, for any $x > 0, n \in \mathbb{N}$, $\sqrt[n]{x}$ exists.

Definition 4.2. A real x that is not rational is called irrational.

Example 4.2. $\sqrt{2}, \sqrt{3}, \sqrt{5}, 3 + 5\sqrt{2}$ are irrational.

Proposition 4.5. *For any $a, b \in \mathbb{R}, a < b, \exists q \in \mathbb{Q}, a < q < b$.*

Proof. Suppose $a < b$ and WLOG $a, b \geq 0$. Choose $n \in \mathbb{N}$ with $1/n < |b - a| = b - a$. So we can find $k \in \mathbb{N}$, $k/n \leq a$ and $(k+1)/n > a$. Now if $(k+1)/n \geq b$, then $1/n < |b-a| \leq (k+1)/n - k/n = 1/n$, contradiction. So $a < (k+1)/n < b$. □

Corollary 4.6. *For any $a, b \in \mathbb{R}, a < b, \exists i \in \mathbb{R} \setminus \mathbb{Q}, a < i < b$.*

Proof. $\exists q \in \mathbb{Q}, a/\sqrt{2} < q < b/\sqrt{2}$, so $i = q\sqrt{2}$ works. □

4.2 Sequences and Their Limits

What does $1 + 1/2 + 1/4 + 1/8 + \ldots = 2$ mean? Why does $0.33333\ldots = 1/3$? When we come to think about it, we mean that, for example in the first case, $1, 1 + 1/2, 1 + 1/2 + 1/4, \ldots$ "$\to$" 2. But what do we mean by that? We do not mean that the sequence will be eventually x, but can be "arbitrarily close" to that.

Definition 4.3. The absolute value function $|x|$ is defined by

$$\forall x \in \mathbb{R}, |x| = \begin{cases} x, \text{ if } x \geq 0 \\ -x, \text{ otherwise} \end{cases}$$

Regarding absolute value, we also have the triangle inequality

Proposition 4.7. $|x - y| \leq |x - z| + |z - y|$

Proof. Trivial. □

Definition 4.4. For a sequence (x_n) of reals and $x \in \mathbb{R}$, we say $x_n \to x$ (as $n \to \infty$) or

$$\lim_{n\to\infty} x_n = x$$

if $\forall \epsilon > 0, \exists N \in \mathbb{N}, \forall n > N, |x_n - x| < \epsilon$.

So it means x_n will "eventually" go ϵ-close to x.

Example 4.3. 1. Consider the sequence $1/2, 1/2+1/4, 1/2+1/4+1/8, \ldots$, so $x_n = 1 - 1/2^n$ inductively. $\forall \epsilon > 0$, we can choose $N \in \mathbb{N}$ such that $1/N < \epsilon$, for any $n \geq N$, $|1 - x_n| = |1/2^n| \leq |1/n| \leq 1/N < \epsilon$, so $x_n \to 1$.
2. Any constant sequence converges to that constant. Indeed, for any $\epsilon > 0$, choose $N = 1$, so for every $n > N$, the sequence is c hence is ϵ-close to the constant.
3. Consider $x_n = (-1)^n$. This sequence does not converge to a limit. Suppose it does, then suppose the limit is c, then we can choose $\epsilon = 1$, then there is some $N \in \mathbb{N}$, such that $\forall k > N, |x_n - c| < \epsilon = 1$, but we can choose n such that $2n > N$, but $2 = |x_{2n} - x_{2n+1}| \leq |x_{2n} - c| + |x_{2n+1} - c| < 2\epsilon = 2$, contradiction.
3. Sequence needs not have a closed form, we can take

$$x_n = \begin{cases} 1/n, \text{ if } n \text{ is odd} \\ 0, \text{ otherwise} \end{cases}$$

Then it is trivial that $x_n \to 0$. $\forall \epsilon > 0$, we choose N such that $N > 1/\epsilon$, then $\forall n \geq N, |x_n - 0| \leq 1/N < \epsilon$.

Remark. 1. If $x_n \to c$ for some c, we say the sequence (x_n) or $(x_n)_{n=1}^{\infty}$ is convergent. If it is not the case, then we say it is divergent. Note that the example $(-1)^n$ shows that a divergent sequence needs not to go to infinity.
2. Limits, if exist, are unique. If $x_n \to c$ and $x_n \to d$, then $c = d$. Indeed, if $c \neq d$, then $|c - d| > 0$, so we choose $\epsilon = |c - d|/2$, so $\exists N_1 \in \mathbb{N}, \forall n \geq N, |x_n - c| < \epsilon, \exists N_2 \in \mathbb{N}, \forall n \geq N_2, |x_n - d| < \epsilon$, so let $N = \max\{N_1, N_2\}$, so $|c - d| = 2\epsilon > |x_N - c| + |x_N - d| \geq |c - d|$, contradiction.

A sequence given in the form $x_1, x_1 + x_2, x_1 + x_2 + x_3, \ldots$ is called a series. We can write

$$\sum_{n=1}^{\infty} x_n$$

for the series, and the k^{th} term of it is

$$\sum_{n=1}^{k} x_n$$

This is called a partial sum of the series. If a series is convergent, one can also write the infinite sum as its limit, so

$$\sum_{n=1}^{\infty} \frac{1}{2^n} = 1$$

We should not write and cannot write something like "let c be the limit of (x_n) and blah blah blah" before ensuring that the sequence does converge since such a number may not exist.
Limits do behave nicely.

Example 4.4. 1. If $x_n \leq d$ always, and $x_n \to c$, then $c \leq d$. Indeed, if $c > d$, then we can take $\epsilon = c - d$, so $|c - x_n| = c - x_n = (c - d) + (d - x_n) \geq c - d = \epsilon$, contradiction.
2. If $x_n < d$ always, and $x_n \to c$, then we need not have $c < d$, since we can take for example $x_n = -1/n \to 0$ and $d = 0$.

Proposition 4.8. *If $x_n \to x$, $y_n \to y$, then $z_n = x_n + y_n \to x + y$.*

Proof. $\forall \epsilon > 0$, choose N_1 with $\forall n > N_1, |x_n - x| < \epsilon/2$ and N_2 with $\forall n > N_2, |y_n - y| < \epsilon/2$, then choose $N = \max\{N_1, N_2\}$, then $\forall n > N, |z_n - (x+y)| \leq |x_n - x| + |y_n - y| < 2\epsilon/2 = \epsilon$, so $z_n \to x + y$. □

We say a sequence $x_1, x_2, \ldots$ is increasing if $x_n \leq x_{n+1}$ for every $n \in \mathbb{N}$. It is called bounded above if $\{x_n : n \in \mathbb{N}\}$ is bounded above.

Theorem 4.9. *An incresing sequence that is bounded above is convergent.*

Note that this is false in $\mathbb{Q}$.

Proof. Suppose (x_n) is such a sequence. We claim that it converges to $c = \sup\{x_n : n \in \mathbb{N}\}$. For any $\epsilon > 0$, $c - \epsilon$ cannot be an upper bound, so there is some $N \in \mathbb{N}$ such that $c \geq x_N > c - \epsilon$, then for any $n > N$, $c \geq x_n \geq x_N > c - \epsilon$, so $|c - x_n| < \epsilon$. □

Remark. 1. It is equivalent to say that a decreasing sequence (i.e. $x_n \geq x_{n+1}$ for all $n \in \mathbb{N}$) bounded below is convergent. So as a corollary, a bounded monotone sequence is convergent.
2. In series, the equivalent form would be that if the partial sums of the series is bounded above and all terms are nonnegative, then it converges by corollary just stated.

There are a few applications of the theorem (and its corollaries).

Proposition 4.10. *$\sum_{n=1}^{\infty} n^{-1}$ diverges, and $\sum_{n=1}^{\infty} n^{-2}$ converges.*

Note that as sequences, neither of the sums has a (nice) closed form. Our main idea would be doing comparisons with other series whose sums we are more familiar with.

Proof. For the first part of the proposition, choose any partial sum

$$P = \sum_{n=1}^{k} n^{-1}$$

then consider the least r such that $2^r > k$, so

$$\sum_{n=1}^{2^{r+1}} \frac{1}{n} \geq P + \sum_{n=2^r}^{2^{r+1}} \frac{1}{n} \geq P + \frac{2^r}{2^{r+1}} = P + \frac{1}{2}$$

So it is unbounded, hence does not converge. For the second part,

$$\sum_{n=2^r}^{2^{r+1}-1} \frac{1}{n^2} \leq \frac{2^r}{2^{2r}} = \frac{1}{2^r} \implies \sum_{n=1}^{\infty} \frac{1}{n^2} \leq \sum_{r=0}^{\infty} \frac{1}{2^r} = 2$$

So it is bounded, hence it converges. □

The last sum actually tends to $\pi^2/6$, the proof of this will be covered somewhere else.
Secondly, decimal expansion. When we are thinking of $0.a_1a_2a_3\ldots$, how do we know the limit $0.a_1, 0.a_1a_2, \ldots$ always exists? Obviously, we want to analyze the infinite sum

$$\sum_{n=1}^{\infty} \frac{a_n}{10^n}$$

This does converge, since

$$\sum_{n=1}^{\infty} \frac{a_n}{10^n} \leq \sum_{n=1}^{\infty} \frac{10}{10^n} < \sum_{k=0}^{\infty} \frac{1}{10^k} < 2$$

Conversely, given any real number $0 < x < 1$, we first choose $a_1 \in \{0, 1, \ldots, 9\}$ such that $a_1/10 \leq x < (a_1 + 1)/10$, and when we have chosen a_k, we choose $a_{k+1} \in \{0, 1, \ldots, 9\}$ by choosing it to be such that

$$\sum_{n=1}^{k} \frac{a_n}{10^n} + \frac{a_{k+1}}{10^{k+1}} \leq x < \sum_{n=1}^{k} \frac{a_n}{10^n} + \frac{a_{k+1}+1}{10^{k+1}}$$

such a_{k+1} always exists due to the definitions of $a_1, a_2, \ldots, a_k$.

Remark. 1. The decimal (or binary or other bases) expansion of any rational number is periodic. Conversely, a real number whose decimal (or binary or other bases) expansion is periodic is rational.
2. Decimal expansion may not be unique (in a way) since $0.500\ldots = 0.499\ldots$, but this is kind of the only way that this would happen.

Definition 4.5. We define e to be $1 + 1/1! + 1/2! + 1/3! + 1/4! + \cdots$

Proposition 4.11. *e is well defined.*

Proof. The partial sum is bounded since

$$\sum_{i=0}^{\infty} \frac{1}{i!} \leq 1 + \sum_{k=0}^{\infty} \frac{1}{2^k} = 3$$

Also each term is positive, therefore the sequence of partial sums is increasing, hence it converges. □

4.3 Transcendental Numbers

Definition 4.6. A real number a is called algebraic if it is the root of some nonzero polynomial with integer coefficients.

Example 4.5. 1. Every rational number is algebraic.
2. $\sqrt[q]{p} + r$ is algebraic for any $p, q, r \in \mathbb{Q}$ (given that it is well-defined).

However, there can be real numbers that is not algebraic. To start with, we shall show that e is irrational.

Proposition 4.12. *e is irrational.*

Proof. Suppose for the sake of contradiction that $e = p/q$ such that $p \in \mathbb{Z}, q \in \mathbb{N}$. Obviously $q > 1$. So

$$\frac{p}{q} = 1 + \frac{1}{1!} + \frac{1}{2!} + \cdots \implies p(q-1)! = q!\left(\sum_{n=0}^{q} \frac{1}{n!}\right) + \sum_{n=q+1}^{\infty} \frac{q!}{n!}$$

Note that the first term in the right hand side is integral, but for the second term,

$$0 < \sum_{n=q+1}^{\infty} \frac{q!}{n!} = \frac{1}{q+1} + \frac{1}{(q+1)(q+2)} + \cdots \leq \sum_{k=1}^{\infty} \frac{1}{(q+1)^k} = \frac{1}{q} < 1$$

Thus the right hand side is not an integer, but the left hand side is. This is a contradiction. □

Definition 4.7. If a real number x is not algebraic, we say it is transcendental.

e is actually transcendental, but we will not prove it here. However, we will give an example of a transcendental number.

Proposition 4.13. *Take*

$$c = \sum_{k=1}^{\infty} \frac{1}{10^{k!}}$$

then c is transcendental.

It is trivial that c is well defined. We will need the following facts about polynomials.

Proposition 4.14. *1. For any polynomial P, we have some constant $K > 0$ such that $|P(x) - P(y)| \leq K|x - y|$ for each $x, y \in [0, 1]$.*
2. A polynomial of degree d has at most d roots.

Proof. 1. Suppose $P(x) = \sum_{i=0}^{d} a_i x^i$, then for $x, y \in [0, 1]$,

$$|P(x) - P(y)| = \left|\sum_{i=1}^{d} a_i(x^i - y^i)\right| \leq d\sum_{i=1}^{d} |a_i||x - y| = K|x - y|, K = d\sum_{i=1}^{d} |a_i|$$

2. Polynomial division. □

Now we are ready for the proof.

Proof that c is transcendental. Suppose for the sake of contradiction that there is some polynomial

$$P(x) = \sum_{k=0}^{d} a_k x^k$$

such that $a_k \in \mathbb{Z}$ and $P(c) = 0$. Let

$$c_n = \sum_{k=1}^{n} \frac{1}{10^{k!}}$$

So $c_n \to c$. Note that $|c_n - c| \leq 2/10^{(n+1)!}$. Since P has at most d roots, $\exists N \in \mathbb{N}, \forall n > N, P(c_n) \neq 0$. As P has integer coefficients, for $n > N$,

$$\frac{1}{10^{n!d}} \leq |P(c_n)| = |P(c_n) - P(c)| \leq K|c_n - c| \leq \frac{2K}{10^{(n+1)!}}$$

by the preceding proposition. This fails when n is large, contradiction. □

Definition 4.8. We say $x \in \mathbb{R}\backslash\mathbb{Q}$ is a Liouville number if $\forall n \in \mathbb{N}, \exists p \in \mathbb{Z}, q \in \mathbb{N}$,

$$\left| x - \frac{p}{q} \right| < \frac{1}{q^n}$$

That is, x has a very good rational approximations.

Theorem 4.15. *Every Liouville number is transcendental.*

Proof. Similar as above. □

Remark. e is not a Liouville number.

5 The Complex Numbers

Some polynomials have no solution in $\mathbb{R}$, e.g. $x^2 + 1 = 0$. So we want so enlarge the reals to seek a solution.

Definition 5.1. The complex numbers, written as $\mathbb{C}$, consists of $\mathbb{R}^2$ (ordered pairs of real numbers) with operations $+, \cdot$ defined by

$$(a, b) + (c, d) = (a + c, b + d)$$

$$(a, b) \cdot (c, d) = (ac - bd, ad + bc)$$

The collection $\{(a, 0) : a \in \mathbb{R}\}$ can be identified as the reals. One can verify straight away that the sum and product reduced to the case that we are familiar with. We can write $i = (0, 1)$, and we have $i^2 =) = (-1, 0)$, also every complex numbers are in the form $a+bi$ for $a, bin\mathbb{R}$. Indeed, $(a, b) = (a, 0)+(b, 0)\cdot(0, 1) = a + bi$.

Remark. We can check that the complex numbers satisfy the usual laws of arithmetic. In particular, for each $z = a + bi \neq 0$, we know that $zw = 1$ where $w = \bar{z}/(z\bar{z})$ where $\bar{z} = a - bi$. The complex number $\bar{z}$ is called the conjugate of z. Such a structure is called a field.

Examples of fields includes $\mathbb{C}, \mathbb{R}, \mathbb{Q}, \mathbb{Z}_p$ where p is prime. But $\mathbb{Z}$ is not due to the lack of multiplicative inverses.

Theorem 5.1 (Fundamental Theorem of Algebra)**.** *Every nonconstant polynomial with complex coefficient has a root in* $\mathbb{C}$.

We will obviously not prove it here.

6 Sets and Functions

6.1 Naive Set Theory

Definition 6.1 (Naive definition of sets)**.** A set is any [2] collection of (mathematical) objects.

Example 6.1. $\mathbb{N}, \mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{Z}, [0,1], \{1,2,3\}, \ldots$

A set is determined by its members: $(a \in A \iff a \in B) \iff (A = B)$. Hence, the set is not ordered per se: $\{1,3,7\} = \{7,1,3\}, \{1,3,3,7\} = \{1,3,7\}$. From the second example, we also know that there is no multiple membership. Of course, we would want to make new sets from old.

Definition 6.2. A set A is a subset of another set B iff $x \in A \implies x \in B$. We write it as $A \subset B$ or $A \subseteq B$.

So $A = B$ is equivalent to $A \subset B$ and $B \subset A$.

Example 6.2. $\{1,7\} \subset \{1,3,7\} \subset \mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \ldots, [0,1] \subset \mathbb{R}$.

Definition 6.3. Given a set A and a property P on A, we can form the set of all elements in A having property P, $\{x \in A : P(x)\}$

Example 6.3. We can construct the primes by $\{n \in \mathbb{N} : n \text{ is prime}\}$.

Definition 6.4. Given sets A, B, U such that $A, B \subset U$, we can construct their union by

$$A \cup B = \{x \in U : x \in A \vee x \in B\}$$

and intersection by

$$A \cap B = \{x \in U : x \in A \wedge x \in B\}$$

We say A, B are disjoint if $A \cap B = \varnothing$.
We also have the minus action

$$A \setminus B = \{x \in A : x \notin B\}$$

We can view $A \cap B$ as a subset selection, since $A \cap B = \{x \in A : x \in B\}$. Unions and intersections are commutative and associatives. Also, the union is distributive over the intersection: $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$. The intersection is distributive over the union as well: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$. One can easily check them.
We can also have arbitrary unions and intersections.

Definition 6.5. For $A_i \subset U$ where $i \in I$ for some index set I, the union of all A_i's can be defined by

$$\bigcup_{i \in I} A_i = \{x \in U : \exists i \in I, x \in A_i\}$$

the intersection by

$$\bigcap_{i \in I} A_i = \{x \in U : \forall i \in I, x \in A_i\}$$

[2] Not quite, see later.

It coincides with our previous definition in the finite case due to the associativity and commutativity of unions and intersections.

Definition 6.6. In $\mathbb{R}$, let $A_n = [1 - 1/n, 1 + 1/n], n \in \mathbb{N}$, so $\bigcup_{n\in\mathbb{N}} A_n = [0, 2], \bigcap_{n\in N} A_n = \{1\}$, note that there is no limiting operation going on here.
Let let $A_n = (1 - 1/n, 1 + 1/n), n \in \mathbb{N}$, so $\bigcup_{n\in\mathbb{N}} A_n = (0, 2), \bigcap_{n\in N} A_n = \{1\}$

Definition 6.7. Given two objects a, b, you can form the ordered pair (a, b), And that $(a, b) = (c, d)$ iff $a = c, b = d$. For sets A, B, one can form the collection $A \times B = \{(a, b) : a \in A, b \in B\}$, which is called the product, or Cartesian product of A, B.

Example 6.4. The plane $\mathbb{R}^2$ can be viewed as $\mathbb{R} \times \mathbb{R}$.

Similarly, we can construct things like $\mathbb{R}^n$ by recognising the collection of all ordered n-tuples.
Note that if we wish to do it, we could define (a, b) by the set $\{\{a\}, \{a, b\}\}$. One can check that $(a, b) = (c, d) \iff a = c, b = d$ under this notion. [3]

Definition 6.8. For any set A, we can form the power set of A, written as $\mathbb{P}(A)$ or 2^A, [4] which is set of all subsets of A.

Example 6.5. Let $X = \{1, 2\}$, then $2^X = \{\varnothing, \{1\}, \{2\}, \{1, 2\}\}$.

Remark. Warning: Given a set A, we know we can form $\{x \in A : P(x)\}$, but you should not form $\{x : P(x)\}$. Suppose we could form $X = \{x : x \notin x\}$, then do we have $X \in X$? Indeed, this gives a contradiction.
This is called the Russel's Paradox.

Similarly, there is not an universal set U such that $\forall x, x \in V$, because it would induce the Russel's Paradox. We can only guarantee that a given set exists if it is obtained, in some way, from known sets.

6.2 Finite Sets and Their Sizes

Definition 6.9. A set A has size n where $n \in \mathbb{N}_0$ if we can write $A = \{a_1, a_2, \ldots, a_n\}$ such that a_i's are distinct.

Example 6.6. 1. $\{1, 3, 7\}$ has size 3.
2. $\varnothing$ has size 0.

Proposition 6.1. *If a set has size n and size m, then $m = n$.*

Proof. It's trivial but whatever.
Indeed, if A has size n and size m with $n > m$, then if $m = 0$ it is trivial, otherwise, removing an element from A gives a set of sizes $n - 1, m - 1$, so it is done by induction on m. □

Proposition 6.2. *If A has size n, then 2^A has size 2^n.*

It is obvious when $n = 0$, so we assume henceforth it is not.

[3] Quoting the lecturer, "You're completely nuts if you really think of this in that way'
[4] Different from the lecturer, the author of this set of notes prefers the latter notation.

Proof. We can relabel the element in A by $\{1, 2, \ldots, n\}$. So to specify a subset S, we must specify if $1 \in S$, $2 \in S$, and so on, so the size of 2^A is $2 \times 2 \times 2 \times \cdots \times 2$ n times, which is 2^n. □

Alternative proof. Induction on n. □

The alternative proof can be viewed as a more formal version of the first proof.
We are tired of saying "the size of blah blah blah" so we write $|A|$ to denote the size of A.
A set of size n is sometimes called an n-set.

Definition 6.10 (Binomial Coefficients)**.** Let $A = \{1, 2, \ldots, n\}$ for $n \geq 1$, then the binomial coefficient is defined as

$$\binom{n}{k} = \{S \subset A : |S| = k\}$$

is the number of ways to choose a k-set from an n-set.

Example 6.7.

$$\binom{4}{2} = 6$$

by listing.

We always have $\binom{n}{n} = \binom{n}{0} = 1$, $\binom{n}{1} = n$ and $\binom{n}{k} = \binom{n}{n-k}$. Also, by Proposition 6.2, we instantly have

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

In addition,

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

since both size counted the number of k-sets in an n-set, that is, the number of k-sets which include some specified element and the number of k sets which do not include. Hence we have the Pascal's triangle.

Proposition 6.3.

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

Proof. We first count the number of ordered k-sets is $n(n-1)(n-2)\cdots(n-k+1)$. But here we have overcounted each k-sets by $k(k-1)(k-2)\cdots 1 = k!$, therefore the formula. □

Corollary 6.4. *When n is big, $\binom{n}{k} \sim n^k/k!$.*

An appplication of the binomial coefficient is the binomial theorem.

Theorem 6.5 (Binomial Theorem)**.**

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

Proof. When we expand $(a+b)^n = (a+b)(a+b)\cdots(a+b)$ where there are altogether n brackets, we obtain terms of the form $a^{n-k}b^k$ where $0 \leq k \leq n$. But the number of terms that we get in each is the number of ways to choose k (or equivalently, $n-k$) brackets from the n brackets, so it is $\binom{n}{k}$, thus the theorem. □

In particular, $(1+x)^n = 1 + nx + \binom{n}{2}x^2 + \cdots + x^n$, hence for given n, $(1+x)^n \sim 1+nx$ when x is small, and we can get better approximations when we take more term(s).
Now, how do sizes of unions and intersections (for finite sets) relate to each other?

Example 6.8. We have $|A \cup B| = |A| + |B| - |A \cap B|$, and $|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |B \cap C| - |C \cap A| + |A \cap B \cap C|$.

Theorem 6.6 (Inclusion/Exclusion Principle)**.** *For finite sets* $(S_i)_{i=1}^n$, $|S_1 \cup S_2 \cup \cdots \cup S_n|$ *equals*

$$\sum_i |S_i| - \sum_{i<j} |S_{ij}| + \sum_{i<j<k} |S_{ijk}| - \cdots + (-1)^{n+1} \sum_{i_1<i_2<\ldots<i_n} S_{i_1 i_2 \ldots i_n}$$

Where $S_{k_1 k_2 \ldots k_r} = S_{k_1} \cap S_{k_2} \cap \cdots \cap S_{k_r}$

Proof. Consider $x \in |S_1 \cup S_2 \cup \cdots \cup S_n|$, suppose we can choose the maximal k such that there is a k-set A contained in $\{1, 2, \ldots, n\}$ such that $x \in S_a$ for each $a \in A$. Then, the number of times that x is counted is

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

by Binomial Theorem. So each element is counted exactly once, hence the theorem is proved. □

6.3 Functions

Definition 6.11 (Intuitive Definition of Functions)**.** For sets A, B, a function $f : A \to B$ is a "rule" that assigns each element a in A exactly one element, called $f(a)$ in B.

Definition 6.12 (Rigourous Definition of Functions)**.** A function $f : A \to B$ is a set $f \subset A \times B$ such that if $(a,b),(a,c) \in f$ for $a \in A$ and $b, c \in B$, then $b = c$ (equivalently $(a,b) = (a,c)$). In other words, $\forall a \in A, \exists! b \in B, (a,b) \in f$. We write $f(a) = b$.

A is called the domain of f and B is called the range (or codomain) of f. The set $\{f(a) : a \in A\}$ is called the image of f.

Example 6.9. 1. We can have a function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$.
2. (non-example) But $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = 1/x$ is not a function since it does not have value at 0.
3. (non-example) The function $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = \pm\sqrt{x^2}$ is not a function due to multiple value.
4. $A = \{1,2,3,4,5\}, B = \{1,2,3,4\}$ and $f : A \to B$ by $1 \mapsto 1, 2 \mapsto 3, 3 \mapsto$

$4, 4 \mapsto 3, 5 \mapsto 5$ is a function.
5. $A = B = \{1, 2, 3\}$ and $f : A \to B$ by $1 \mapsto 2, 2 \mapsto 1, 3 \mapsto 3$ is a function.
6. $A = B = \{1, 2, 3, 4\}$ and $f : A \to B$ by $1 \mapsto 1, 2 \mapsto 2, 3 \mapsto 4, 4 \mapsto 4$.
7. $A = \{1, 2, 3, 4, 5\}, B = \{1, 2, 3, 4\}$ and $f : A \to B$ by $1 \mapsto 2, 2 \mapsto 1, 3 \mapsto 3, 4 \mapsto 4, 5 \mapsto 4$.

Note that it can be ambiguous if we just say $f(x) = x^2$ since we may not know about the domain (and range). We can only define a function with knowing about the domain and the range. Note also that we do not have to have a "closed form" of a function, which is frankly ridiculous to do so.

Definition 6.13. A function $f : A \to B$ is called injective if $\forall x, y \in A, x \neq y \implies f(x) \neq f(y)$, or equivalently, $\forall x, y \in A, f(x) = f(y) \implies x = y$.
It is called surjective if $\forall b \in B, \exists a \in A, f(a) = b$, or equivalently, B is the image of f.

One can easily classify the above examples by injective/non-injective and surjective/non-surjective.

Definition 6.14. If a function f is both injective and surjective, we say it is bijective, or that it is a bijection.

Note that a bijection pair up elements in A and B. It is also called a 1-1 correspondence. Example 5 above is an example of a bijection. Note that injectivity and surjectivity depend strongly on the domain and range of the function.

Remark. If A, B are finite and $f : A \to B$. If $|A| > |B|$, f cannot be injective. If $|A| < |B|$, f cannot be surjective.
If $|A| = |B|$, then surjectivity, injectivity and bijectivity are equivalent. Thus any $f : A \to A$ cannot biject A to a proper subset $A' \subsetneq A$. But this is not necessarily true for infinite (i.e. not finite) sets, for example, the "adding-1" function is a bijection $\mathbb{N} \to \mathbb{N} \setminus \{1\}$ (assuming $\mathbb{N}$ starts with 1), this also shows that injectivity does not imply surjective. Also the function $\mathbb{N} \to \mathbb{N}$ by $1 \mapsto 1$ and $n \mapsto n - 1$ for $n \geq 2$ is surjective but not injective.

We have more examples of functions.

Example 6.10. 1. For any set A, we have the identity function $A \to A$ which sends each element to itself It is a bijection.
2. Given a set X and a subset $A \subset X$, the function $\chi_A : X \to \{0, 1\}$ by $\chi_A(x) = 1$ if $x \in S$ and $\chi_A(x) = 0$ otherwise is a function. This is called the indicator (or characteristic) function of A.
3. A sequence on a set X is a function $\mathbb{N} \to X$.
4. The addition and multiplication, say on $\mathbb{N}$, are functions $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$.
5. A finite set A has set m if and only if there is a bijection $\{1, 2, 3, \ldots, m\} \to A$.

Definition 6.15. Let $f : A \to B, g : B \to C$, then the composition $g \circ f$ is a function $A \to C$ defined by $(g \circ f)(a) = g(f(a))$.

Note that function compositions are not necessarily commutative (if $A = C$).

Example 6.11. If $A = B = C = \mathbb{R}$ and $f : x \mapsto 2x, g : x \mapsto x + 1$, then $f \circ g : x \mapsto 2x + 2, g \circ f : x \mapsto 2x + 1$, which are indeed different.

However, function compositions are associative.

Proposition 6.7. *If $f : A \to B, g : B \to C, h : C \to D$, then $h \circ (g \circ f) = (h \circ g) \circ f$.*

Proof. $\forall x \in A, (h \circ (g \circ f))(x) = h(g(f(x))) = ((h \circ g) \circ f)(x)$. □

Definition 6.16. $f : A \to B$ is invertible if and only if there is some other function $g : B \to A$ such that $f \circ g = \mathrm{id}_B, g \circ f = \mathrm{id}_A$. We say g is the inverse of f. Note that g is invertible as well with inverse f.

Example 6.12. Consider $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = 2x + 1$, consider $g : \mathbb{R} \to \mathbb{R}$ by $g(x) = (x - 1)/2$, then $\forall x \in \mathbb{R}, (f \circ g)(x) = (g \circ f)(x) = x$, thus f is invertible and g is the inverse of f.

Remark. Note that we have to check both sides. Consider $f_0 : \mathbb{N} \to \mathbb{N}$ by $f_0(x) = x + 1$ and $f_1 : \mathbb{N} \to \mathbb{N}$ by $f_1(x) = x - 1$ for $x > 1$ and $f_1(1) = 1$, then $f_1 \circ f_0 = \mathrm{id}_{\mathbb{N}}$ but $f_0 \circ f_1 \neq \mathrm{id}_{\mathbb{N}}$.

Proposition 6.8. *A function is invertible if and only if it is a bijection.*

Proof. Trivial. □

6.4 Equivalence Relations

Definition 6.17. Let A be a set, then a relation R is a subset of $A \times A$. We say $x, y \in A$ are related, or xRy, if $(x, y) \in R$.

Example 6.13. 1. On $\mathbb{N}$, the relation $aRb \iff a \equiv b \pmod 7$ is a relation.
2. On $\mathbb{N}$, $aRb \iff a|b$ is a relation.
3. On any set, $aRb \iff a \neq b$ is a relation.
4. On $\mathbb{N}$, $aRb \iff a = b \pm 1$.
5. On $\mathbb{N}$, $aRb \iff |a - b| \leq 2$.
6. On $\mathbb{N}$, $aRb \iff (a, b < 5 \vee a, b \geq 5)$.

Definition 6.18. A relation R is reflexive iff $\forall x \in A, xRx$.

The first, second, fifth and sixth examples above are reflexive relations.

Definition 6.19. A relation R is symmetric iff $\forall x, y \in A, xRy \implies yRx$

The first, third, fourth, fifth and sixth examples above are symmetric relations.

Definition 6.20. A relation R is transitive iff $\forall x, y, z \in A, (xRy \wedge yRz) \implies xRz$

The first, second and sixth exampels above are transitive relations.

Definition 6.21. We say R is an equivalence relation if it is reflexive, symmetric and transitive.

The first and sixth relations above are equivalence relations.

Example 6.14. Let X be a set, then consider a partition $\{C_i\}_{i \in I}$ of X. That is, $C_i \neq \varnothing$, $i \neq j \implies C_i \cap C_j = \varnothing$ and $X = \bigcup_{i \in I} C_i$. Then $xRy \iff \exists i \in I, x, y \in C_i$ is a equivalence relation.

Proposition 6.9. *All equivalence relations on X can be written in the form of Example 6.14.*

Definition 6.22. For an equivalence relation R on a set X and $x \in X$, the equivalence class containing x, written as C_x or $[x]$, is $\{y \in X : xRy\} = \{y \in X : yRx\}$.

Example 6.15. In the first example, $[2] = 6 + 7\mathbb{Z} = [16] = [23] = \ldots$. Note that we have 7 equivalence classes in total, namely $i + 7\mathbb{Z}, i \in \{0, 1, 2, 3, 4, 5, 6\}$.

Proof of Proposition 6.9. let X be the set with the equivalence relation R. We shall show that the equivalence classes partitions X. Note that $x \in [x]$, so their union is X, so it remains to show that different equivalence classes are disjoint. If $y \in [x]$ and $y \in [z]$, then for any $w \in [x]$, we know wRx, xRy, so wRy, but yRz, so wRz, therefore $w \in [z]$, hence $[x] \subset [z]$. Similarly $[z] \subset [x] \implies [x] = [z]$. $\square$

Hence equivalence relation is just a partition of the set.

Definition 6.23. The collection $X/R = \{[x] : x \in X\}$ is called the quotient of X with respect to an equivalence relation R.

Example 6.16. 1. $xRy \iff x \equiv y \pmod 7$ is an equivalence relation on $\mathbb{Z}$. It gives the quotient $\mathbb{Z}/R = \{i + 7\mathbb{Z} : i \in \{0, 1, 2, 3, 4, 5, 6\}\}$.
2. $(a, b)R(c, d) \iff ad = bc$ is an equivalence relation on $\mathbb{Z} \times \mathbb{N}$. So we could construct $\mathbb{Q}$ by $\mathbb{Q} = (\mathbb{Z} \times \mathbb{N})/R$. The equivalence classes are like, for example, $[(1, 2)] = \{(1, 2), (2, 4), (4, 8), \ldots\}$. In that way we can recognize $a/b = [(a, b)]$.

Definition 6.24. The quotient map (or projection map) $q : X \to X/R$ is defined by $q(x) = [x]$.

This is well-defined since equivalence classes partitions the set X.

7 Countability

We want to find a way to describe the sizes of infinite sets. For example, $\mathbb{N}$ "looks smaller" than $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$, but is that the case?

Definition 7.1. A set A is called countable if either A is finite or there exists a bijection $A \to \mathbb{N}$.

Equivalently, A is countable iff we can list the element of A as $\{a_1, a_2, a_3, \ldots\}$ (which might terminate if A is finite).

Example 7.1. 1. Every finite set is countable.
2. $\mathbb{N}$ is countable.
3. $\mathbb{Z}$ is countable. Consider the listing $\{0, 1, -1, 2, -2, 3, -3, \ldots\}$. Or, written in formula

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

A natural question to ask, following the last example, is whether all sets are countable. For example, is $\mathbb{Q}$ countable? How about $\mathbb{R}$?

Proposition 7.1. *Let A be a set, then A is countable iff there is an injection $f : A \to \mathbb{N}$*

Proof. If A is countable, it is obvious that there is an injection.
Conversely, it is done if A is finite, so we can henceforth assume that A is infinite. Note that $f(A) = \operatorname{Im} f$ is a subset of $\mathbb{N}$ and f is a bijection $A \to f(A)$. So it is enough to show that $f(A)$ is countable. We can order $f(A)$ like we did in $\mathbb{N}$, then define b_n recursively by $b_1 = \min f(A)$ and $b_{n+1} = \min(f(A) \setminus \{b_i : 1 \leq i \leq n\})$. So the sequence $\{b_n\}_{n\in\mathbb{N}}$ lists $f(A)$, hence $f(A)$ is countable, so A is countable. □

Corollary 7.2. *Every subset of a countable is countable.*

Remark. Consider the set $\{n/(n+1) : n \in \mathbb{N}\} \cup \{1\} \subset \mathbb{R}$, then this set is countable, but we cannot hit everything by writing every element in increasing order.

Theorem 7.3. $\mathbb{N} \times \mathbb{N}$ *is countable.*

Proof. Define the listing $a_1, a_2, \ldots \in \mathbb{N} \times \mathbb{N}$ by $a_1 = (1, 1)$ and if we have $a_n = (p, q)$, then

$$a_{n+1} = \begin{cases} (p-1, q+1), \text{ if } p > 1 \\ (p+q, 1), \text{ otherwise} \end{cases}$$

This lists all points in $\mathbb{N}\times\mathbb{N}$. All $(x, y) \in \mathbb{N}\times\mathbb{N}$ are hit by induction on $x+y$. □

Alternative proof. Consider the map $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ by $(a, b) \mapsto 2^a 3^b$ is an injection. □

More generally, the same shows

Corollary 7.4. *Let $\{A_i\}_{i\in\mathbb{N}}$ be a collection of countable sets, then $\bigcup_{i\in\mathbb{N}} A_i$ is countable.*

Proof. Each A_i is countable, so we can list A_i as $\{a_{i1}, a_{i2}, \ldots\}$ (which might terminate if A_i is finite). Now consider the function $f : \bigcup_{i\in\mathbb{N}} A_i \to \mathbb{N}$ by $x \mapsto 2^i 3^j$ where $x = a_{ij}$ such that i is the least such that $x \in A_i$ and j is the least such that $x = a_{ij}$. This is an injection. □

Example 7.2. 1. $\mathbb{Q}$ is countable. Indeed,

$$\mathbb{Q} = \bigcup_{n\in\mathbb{N}} \frac{1}{n}\mathbb{Z}$$

So we are done by the preceding corollary. Alternatively, there is an obvious injection from $\mathbb{Q}$ to $\mathbb{Z} \times \mathbb{Z}$.
2. The set $\mathbb{A}$ of all algebraic numbers is countable. Indeed, each polynomial has only finitely many roots, so it is enough to show that there are couontably many integer polynomials, then the claim is proved by Corollary 7.4. But again by this corollary it is enough to show that there are only countably many integer polynomials of degree d for any $d \in \mathbb{N}$. However this set injects to $\mathbb{Z}^{d+1}$ by $a_0 + a_1 x + \cdots + a_d x^d \mapsto (a_0, a_1, \ldots, a_d)$, so it is countable.

We have got many many countable sets in our stock, so the question is then whether all sets are countable.

Theorem 7.5. $\mathbb{R}$ *is uncountable (that is, not countable).*

Proof. It suffices to show that $(0,1)$ is uncountable. So given any sequence $r_1, r_2, \ldots$ of $(0,1)$, we shall show that there is some $s \in (0,1)$ that is not of the form r_i. We write each number in $\{r_i : i \in \mathbb{N}\}$ as decimals. That is, $r_i = 0.r_{i1}r_{i2}\ldots$. For each $n \in \mathbb{N}$, choose a digit $s_n \in \{5,6\} \setminus \{r_{nn}\} \neq \varnothing$, [5] then $0.s_1s_2\ldots \in (0,1)$ is not of the form a_i. □

The above proof is called Cantor's Diagonal Argument.

Remark. $\mathbb{A} \cap \mathbb{R}$ is countable but $\mathbb{R}$ is not, so there exists transcendental numbers. In fact, "most" reals are transcendental, as the set $\mathbb{R} \setminus \mathbb{A}$ is uncountable.

Theorem 7.6. $2^{\mathbb{N}}$ *is uncountable.*

Proof. Suppose $2^{\mathbb{N}} = \{s_1, s_2, \ldots\}$. Consider the set $S \subset \mathbb{N}$ by $n \in S \iff n \notin s_n$, that is $\{n \in \mathbb{N} : n \notin s_n\}$, so $S \neq s_i$ for any i, contradiction. □

Remark. This really is exactly the proof idea of $\mathbb{R}$ being uncountable. Alternatively, we can actually inject $(0,1)$ to $2^{\mathbb{N}}$ by writing in base 2.

The proof above can be actually extended to the following:

Theorem 7.7. *Let X be a set, then there is no bijection $X \to 2^X$.*

For example, there is no bijection between $\mathbb{R}$ and $2^{\mathbb{R}}$.

Proof. Let $f : X \to 2^X$ be a function. Consider the set $S = \{x \in X : x \notin f(x)\}$. But $S \notin f(X)$ since for any $x \in X, S \neq f(x)$ (as $x \in S \iff x \notin f(x)$), so f is not surjective. □

Note that the set $S = \{x \in X : x \notin f(x)\}$ is very similar to Russel's Paradox. Now we want to explore more countability arguments.

Example 7.3. Any collection $\{A_i\}_{i \in I}$ of pairwise disjoint intervals is countable. The first proof that we can do is to inject the set into the rational numbers by choosing one from each interval. The second proof is by observing that the numbers of intervals in the family with length at least $1/n$ is countable, then we can write our family as a countable union of countable sets, hence is countable.

There are generally two sorts of arguments to show a set is uncountable, either copying the diagonal argument or inject an uncountable set to it. To show it is countable, either we can list it (which however often get fiddly) or we can inject it into a countable set. Another way is to use the fact that a countable union of countable sets is countable.
Now consider A, B nonempty. Intuitively, we think of A bijecting with B as saying that the size of A is the size of B. Similarly, we think of A injecting into B as A has at most as large as B, and A surjecting to B as A has at least as large as B. For these to make sense, we certainly want that A injects into B if and only if B surjects into A. Indeed, given $f : A \to B$ injective, then consider $a \in A$, then $g : B \to A$ by

$$b \mapsto \begin{cases} f^{-1}(b), \text{ if } b \in f(A) \\ a, \text{ otherwise} \end{cases}$$

[5] There is nothing special about $5, 6$, we just need to get rid of the case of $9999\ldots$

Then g is surjective. Conversely, given $g : B \to A$ surjective, consider $f : A \to B$ with $a \mapsto a'$ by choosing an $a' \in g^{-1}(\{a\})$. [6] Also it is intuitive that if A injects into B and B injects into A, then A bijects with B.

Theorem 7.8 (Schröder–Bernstein)**.** *Let $f : A \to B$ and $g : B \to A$ be injective, then there is a bijection $h : A \to B$*

Proof. For $a \in A$, we write $g^{-1}(a)$ for the unique (if exists) point in B such that $g(g^{-1}(a)) = a$. Define similarly $f^{-1}(b)$ for $b \in B$. The "ancestors" of $a \in A$ consists of

$$g^{-1}(a), f^{-1}(g^{-1}(a)), g^{-1}(f^{-1}(g^{-1}(a))), f^{-1}(g^{-1}(f^{-1}(g^{-1}(a)))), \dots$$

which may or may not terminate. Similarly for $b \in B$. Let A_0 be the set of $a \in A$ such that the ancestor sequence of a that terminates in even time, that is it has even length (so it includes those $a \notin g(B)$ since 0 is even). And A_1 be the set of $a \in A$ such that the ancestor sequence of a that terminates in odd time, A_∞ be those whose ancestor sequence does not terminate. Similarly construct B_0, B_1, B_∞. So A_0, A_1, A_∞ partitions A and B_0, B_1, B_∞ partitions B. By definition, $f|_{A_0}$ is a bijection $A_0 \to B_1$ (as every $b \in B_1$ is in the image of $f|_{A_0}$). And simiarly, $g|_{B_0}$ is a bijection $B_0 \to A_1$. As for infinity cases, $f|_{A_\infty}$ is a bijection $A_\infty \to B_\infty$, so the function h can be defined by

$$h(a) = \begin{cases} f(a), \text{ if } a \notin A_1 \\ g^{-1}(a), \text{ if } a \in A_1 \end{cases}$$

Then $h : A \to B$ is well-defined and bijective. □

Example 7.4. $[0, 1]$ and $[0, 1] \cup [2, 3]$ biject. Indeed, $x \mapsto x$ is an injection from $[0, 1] \to [0, 1] \cup [2, 3]$. Also $x \mapsto x/3$ injects $[0, 1] \cup [2, 3]$ to $[0, 1]$

Now, is it true that for any sets A, B, either A injects into B or vice versa? The answer is yes, but very hard and way beyond the scope of this course.
Now given $\mathbb{N}$, we can construct a strictly increasing sequence (in terms of sizes) $\mathbb{N}, 2^{\mathbb{N}}, 2^{2^{\mathbb{N}}}, \dots$, but does every set injects into one of them? The answer is obviously no, since we can take $X = \mathbb{N} \cup 2^{\mathbb{N}} \cup 2^{2^{\mathbb{N}}} \cup \cdots$. Now this is definitely not the biggest set either, since we can now take again $X, 2^X, 2^{2^X}, \dots$, which is again beaten by the union X' of all of them. We can then construct a sequence $X, X', X'', \dots$, which is again beaten by $X \cup X' \cup X'' \cup \cdots$. And we can do the same thing again and again and again and this never ends... (but the course does here).

8 Bonus lecture: Primitive Roots

We work in the multiplicative group $\mathbb{Z}_p^\times$ or $\mathbb{Z}_p^\star$ consisting of $(\mathbb{Z}_p \setminus \{0\}, \times, 1)$.

Definition 8.1. The order of $x \in \mathbb{Z}_p^\star$ is the least $n \in \mathbb{N}$ with $x^n \equiv 1 \pmod p$.

It is trivial that an order indeed exists by FℓT.

Example 8.1. The order of 2 in $\mathbb{Z}_7^\star$ is 3, and the order of 3 there is 6.

[6] Does it depend on the Axiom of Choice?

Definition 8.2. A $x \in \mathbb{Z}_p^\star$ is said to be a primitive root, or generator, if x has order $p-1$.

Indeed, if such an x exists, then every element in $\mathbb{Z}_p^\star$ can be written in the form x^n for some $n \in \mathbb{N}$.

Theorem 8.1. *For any p, there exists a generator.*

That is, $\mathbb{Z}_p^\star \cong C_{p-1}$.

Proposition 8.2. *If x has order n, then if $x^d \equiv 1 \pmod{p}$, then $n|d$.*

Proof. Trivial by minimality of order. □

Proposition 8.3. *Let x have order a and y have order b where a, b are coprime, then xy has order ab.*

Proof. Obviously $(xy)^{ab} \equiv 1 \pmod{p}$. Conversely, if $(xy)^d \equiv 1 \pmod{p}$, then $x^d y^d \equiv 1 \pmod{p}$, so $x^{bd} \equiv 1 \pmod{p}$ hence $a|bd$, but a, b are coprime, so $a|d$. Similarly $b|d$, thus $ab|d$ since again a, b are coprime. □

Also the condition of a, b being coprime is necessary. Indeed, by the same idea, we obtain

Proposition 8.4. *Let x have order a and y have order b, then there is some element have order l, the LCM of a, b.*

Proof. Let

$$a = p_1^{a_1} \cdots p_k^{a_k} q_1^{c_1} \cdots q_j^{c_j}, b = p_1^{b_1} \cdots p_k^{b_k} q_1^{d_1} \cdots q_j^{d_j}$$

where p_i, q_i are distinct primes where $a_i \geq b_i, c_i \leq d_i$. Then $x^{q_1^{c_1} \cdots q_j^{c_j}}$ has order $p_1^{a_1} \cdots p_k^{a_k}$, and $y^{p_1^{b_1} \cdots p_k^{b_k}}$ has order $q_1^{d_1} \cdots q_j^{d_j}$. So $x^{q_1^{c_1} \cdots q_j^{c_j}} y^{p_1^{b_1} \cdots p_k^{b_k}}$ has order $p_1^{a_1} \cdots p_k^{a_k} q_1^{d_1} \cdots q_j^{d_j} = l$ by above. □

Corollary 8.5. *If d is the biggest order in $\mathbb{Z}_p^\star$, then any order divides d.*

Note that all the above works in any finite abelian group. We now bring in number theory.

Proposition 8.6. *Let f be a polynomial in $\mathbb{Z}_p$ or degree $k > 0$, the f has at most k roots in $\mathbb{Z}_p$.*

Proof. Division and induction. □

Remark. This is NOT true in $\mathbb{Z}_n$ in general since we used division (which requires it to be a field). For example, $x^2 \equiv 1$ has 4 roots modulo 8.

Proof of Theorem 8.1. Take d to be the biggest order in $\mathbb{Z}_p^\star$. We want $d = p-1$. Note that for any $x \in \mathbb{Z}_p^\star$, $x^d \equiv 1 \pmod{p}$ by the preceding corollary. But the polynomial $x^d - 1$ can have at most d roots. Then we must have $d \geq p-1$, but $d|p-1 \implies d \leq p-1$, hence $d = p-1$. □

There are some consequences by this theorem. We let g be the generator hereafter. For example, we can prove Fermat by takeing $x \equiv g^a \pmod{p}$, then $x^{p-1} \equiv g^{a(p-1)} \equiv 1 \pmod{p}$.
From now on we set p odd. So we have $\mathbb{Z}_p^\star$ contains exactly $(p-1)/2$ roots by again looking at the exponent. Also, a non-square times a non-square gives a square. Now we look at -1. Then we must have $g^{(p-1)/2} \equiv -1 \pmod{p}$, so -1 is a square if and only if $(p-1)/2$ is even, which happens if and only if $p \equiv 1 \pmod{4}$.
For Wilson's Theorem, we can simply write $(p-1)! \equiv g^{1+2+3+\cdots+(p-1)} \equiv g^{p(p-1)/2} \equiv (-1)^p \equiv -1 \pmod{p}$.
For $p = 3k+1$, the cubes are exactly $g^3, g^6, \ldots, g^{3k}$, so there are exactly $k = (p-1)/3$ cubes. Otherwise, everything is a cube (as shown in example sheets).

Remark. The proof does not tell us how we can find the generator as it only shows existence. Even today, it is not well-understood which element would be a generator. For example, when is 2 a generator? It is true for $\mathbb{Z}_5$ but not $\mathbb{Z}_7$. Nobody knows for which primes 2 is a primitive roots, it is not even known if 2 is a primitive root modulo infinitely many primes p (Artin's Conjecture).