Year 1 Introductory Lectures, part of 1VGLA and 1RA

Chris Good, Sara Jabbari, Olga Maleva

September 2024

Contents

Introduction

0.1 Why?

These lectures cover some basic material that will be used across all of your modules in the first twop weeks of term. The lectures are part of both Vectors, Geometry and Linear Algebra (1VGLA) and Real analysis (1RA), so use the first 4 lectures of each of those modules.

0.2 When

- The first two weeks will be delivered by Chris Good, Olga Maleva and Sara Jabarri.
- VGLA will then be taught by Andrew Treglown and Galane Luo.
- RA will then be taught by Yuzhao Wang and Andrew Morris.

The lectures will be at the following times:

- Monday, 0900, VGLA, University Centre, Avon Room, Chris Good
- Tuesday, 1700, VGLA, School of Education, Vaughan Jeffreys Theatre, Chris Good
- Tuesday, 1700, VGLA, Watson Building, LRC, Sara Jabbari
- Thursday, 1400, RA, School of Education, Vaughan Jeffreys Theatre, Chris Good
- Thursday, 1400, RA, Gisbert Kapp Building, NG16, Olga Maleva
- Friday, 1300, RA, School of Education, Vaughan Jeffreys Theatre, Chris Good
- Friday, 1300, RA, Watson Building, LTA, Olga Maleva

0.3 What?

We are going to talk about

- Language, Terminology and Notation: how you do and write maths
- Proof: in particular Mathematical Induction
- Inequalities and the Modulus
- Naive Set Theory
- Functions

1 Mathematical Language, Terminology and Notation

There are two aspects to mathematics: **doing maths**, working through proofs, methods and examples from lectures, books and papers so that you understand them, modelling, doing calculations, solving problems, thinking hard about stuff, and; **presenting maths**, writing up your work. Doing maths can be messy, frustrating, difficult, challenging, fun, exciting, satisfying (just like anything that is worth doing). Writing up your maths can be tedious but it is one of the key skills you learn as mathematician because it allows you and others to check whether your working is correct.

(Note that when I use the term 'mathematics,' I am including pure maths, applied maths, statistics, optimization, etc.)

1.1 Chris's Golden Rules for Doing Maths

Doing maths is the fun part, but it is not all plain sailing and it at times there will be things that you really have to struggle with. This is OK, it is not supposed to be easy, it is supposed to be rewarding. Here are my top tips

1) Make sure you know your definitions. If you don't know the definition of a prime number, how can you begin to understand whether there are infinitely many of them?
2) Make sure you understand the notation so that you can talk about what you are looking at.

3) Work through relevant proofs and techniques from lectures and books, writing them out line by line. Try to understand them, rather than just memorize them. Look for similarities with other arguments, often they are almost identical. Break proofs and arguments down, try to work out the key idea and the bits that are just routine manipulation. Be confident that you can recreate the routine calculations, so you only need to remember the key ideas.
4) Work in rough first.
5) Expect to get stuck and expect to go wrong. Don't be frightened of this, there is no shame in being wrong, it is a normal part of learning how to do new things. The key point is to think about why you went wrong and try something else. Keep going.
6) Put the problem away and sleep on it ... let you brain work on things in the background.
7) Play around with the problem. Try toy examples. Experiment with ideas. Can you solve a simpler version of the problem? Can you solve extreme cases?
8) Work with others and try to solve things together.
9) Look things up in your notes, in books or online, ask ChatGPT. Can you find the question or something similar? Be careful though to make sure what you find is correct (when writing these notes, I used ChatGPT for additional ideas and it confidently told me that $3! > 2^3$!)
10) Use any feedback you have to help you next time.

Remember, the only way to learn maths is to do maths.

1.2 Notation

1.2.1 Using Notation

We use a lot of notation in maths (and therefore in physics, computer science, engineering, etc), using symbols to represent words **so that we can express complex ideas clearly, concisely and accurately**.

In your degree programme, maths should be written in properly punctuated English sentences. Any notation is just a substitute for words. For example: ‘Let $y = x^2$.’ is a sentence with nouns x and y, an adjective ^{2} and a verb $=$.

Once we are used to it, the careful use of notation helps us to understand what we are looking at. For example, because we are all familiar with the notation, it is much easier to read and understand

- ‘If $2x^2 + 5x + 1 = 0$, then $x = \dfrac{-5 \pm \sqrt{5^2 - 4.2.1}}{2.2}$.’

than it is to read and understand

- ‘If 2 times the unknown squared plus 5 times the unknown plus 1 equals 0, then the unknown equals minus 5 plus or minus the square root of 5 squared minus four times 2 times 1 all divided by 2 times 2.’

So notation is there to help us understand what we are looking at. When we use notation, it good practice to try to keep it standard and coherent. For example, we often use: x, y and z for variables; f, g and h for functions; i, j, k, n and m for whole numbers; θ, ϕ and ψ for angles; $\underline{u}$ and $\underline{v}$, $\underline{w}$ or $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ for vectors; A, B, C for sets or matrices, and so on.

When you are choosing notation, **try to choose notation that helps your reader**; doing maths is hard enough without making it harder with poor notation. For example, if we have two sets A and B and we want to choose elements from each of them, it would be confusing to call the element we choose from A by b and the element we choose from B by a. Again, it is much easier to remember and understand the statement

- ‘If $ax^2 + bx + c = 0$, then $x = \dfrac{-b \pm \sqrt{b^2 - 4ac}}{2a}$,’

than it is to understand

- ‘If $ba^2 + xa + y = 0$, then $a = 0.5(-x \pm \Xi^{1/2})/b$, where $\Xi = -4by + x^2$,’

even though both statements say the same thing, because the notation is confusing, unfamiliar and unnecessarily complicated.

How much notation you use is a matter of taste. At the extreme end, a set theorist or logician might write

- ‘$\forall x \forall y[\forall z(z \in x \Leftrightarrow z \in y) \Rightarrow x = y]$.’

(we read this as ’for all x and for all y, if, for all z, z is an element of x if and only if z is an element of y, then x equals y) to mean

- ‘Two sets are equal if they have the same elements.’

An example that you will see in your analysis module is the definition of continuity. The following are all read out allowed in exactly the same way:

1) Let f be a function from the reals to the reals and let a be a real number. We say that f is **continuous at the point** a if, for all ϵ greater than 0, there exists δ greater than 0 such that, for all x in $\mathbb{R}$, the distance from f of x to f of a is less than ϵ whenever the distance from x to a is less than delta. (Note we might say ‘mod f x minus f a’ instead of ‘distance from f of x to f of a’.)

2) Let $f : \mathbb{R} \to \mathbb{R}$ and let $a \in \mathbb{R}$. We say that f is **continuous at the point** a if, for all $\epsilon > 0$, there exists $\delta > 0$ such that, for all $x \in \mathbb{R}$,

$$|f(x) - f(a)| < \epsilon \qquad \text{whenever} \qquad |x - a| < \delta.$$

3) Let $f : \mathbb{R} \to \mathbb{R}$ and let $a \in \mathbb{R}$. We say that f is **continuous at the point** a if,

$$(\forall \epsilon > 0)(\exists \delta > 0)(\forall x \in \mathbb{R})\ (|f(x) - f(a)| < \epsilon \quad \Leftarrow \quad |x - a| < \delta)\,.$$

However, to my mind 2) is the clearest one to look at and remember.

You will be introduced to new notation throughout your modules. **Remember that notation is just short hand for words, so when you see new notation, always try to find out how to say it.** It is much, much easier to learn what something says if you know how to pronounce the words in the first place! An example of this is the (classical) Greek alphabet, which is used a lot in maths (we have already used ϵ and δ). It is easier to talk about the Riemann ζ function if you know that ζ is the Greek letter 'zeta' (however you pronounce 'zeta'), rather than the Riemann 'weird squiggle' function.

Sometimes the same notation is used to mean different things, for example we might use A to represent a matrix or a set. In this case, things should be clear from the context (or be made clear by the author). A very common example of this is the use of $\mathbb{N}$ to represent the natural numbers. Some people include 0 as a natural number, so that they use $\mathbb{N}$ to represent the set $\{0, 1, 2, 3, \ldots\}$. Others use $\mathbb{N}$ to represent the set of positive integers $\{1, 2, 3, \ldots\}$ and $\mathbb{N}_0$ to represent the non-negative integers $\{0, 1, 2, 3, \ldots\}$.

On the other hand, some notation can be read using different words that mean exactly the same thing, for example we might read $\forall$ as 'for all', 'for every' or 'for any' and we might say '$d\ \ f$ by $d\ \ x$' or 'the derivative of f with respect to x' when we see $\frac{df}{dx}$.

Below are two lists of some key notation that might be unfamiliar to you. We will introduce more new notation in the sections that follow as well.

1.2.2 The Greek Alphabet

Notation	Read		Notation	Read		Notation	Read
A, α	alpha		ι	iota		P, ρ	rho
B, β	beta		K, κ	kappa		$\Sigma, \sigma, \varsigma$	sigma
Γ, γ	gamma		Λ, λ	lambda		T, τ	tau
Δ, δ	delta		M, μ	mu		Υ, υ	upsilon
E, ϵ	epsilon		N, ν	nu		Φ, ϕ, φ	phi
Z, ζ	zeta		Ξ, ξ	xi		X, χ	chi
H, η	eta		O, o	omicron		Ψ, ψ	psi
Θ, θ	theta		Π, π	pi		Ω, ω	omega

1.2.3 Some Other Common Symbols

Notation	Read		Notation	Read
$\mathbb{N}, \mathbb{N}_0$	the natural numbers		$\forall$	for all, for any, for every
$\mathbb{Z}$	the integers		$\exists$	there exist
$\mathbb{Q}$	the rationals		$\exists!$	there exists a unique
$\mathbb{R}$	the reals		$\nexists$	there does not exist
$\mathbb{C}$	the complex numbers		$\neg$	not
$\sum$	the sum of		$\vee$	or
$\prod$	the product of		$\wedge$	and
$\Rightarrow$	implies		$\Leftarrow$	is implied by
iff, $\Leftrightarrow$	if and only if		$:=$	is defined to be
$\therefore$	therefore		$\because$	because

The symbol $\Rightarrow$ is used to join two statements when **statement 1** implies **statement 2**. So **statement 1** $\Rightarrow$ **statement 2** reads '**statement 1** implies **statement 2**' or 'if **statement 1**, then **statement 2**'. But notice that $A \Rightarrow (B \Rightarrow C)$ says something different from $(A \Rightarrow B) \Rightarrow C$ (can you think of an example?).

We usually write the word iff when writing mathematical text, instead of using the symbol $\Leftrightarrow$.

You need to be a little bit careful with some of these symbols. For example, $=$ means 'equals' and $\Rightarrow$ means 'implies'. Consider the following example. In Solution 1, $=$ has been used to mean something like 'which is equivalent to saying that.' You certainly cant read out this out loud simply by saying 'equals' wherever you see $=$. Solution 1 is just wrong. In Solution 2, $\Rightarrow$ has been used to mean something like 'which in turn implies that.' This is not great but lots of people do this. My advice is not to, but it is a personal choice. No one can really argue with Solution 3.

Example 1.1.
Let a , b and c be real numbers. Suppose that $b^2-4ac=0$. Show that $ax^2+bx+c=0$ has only one solution.

Solution 1:

$$ax^2+bx+c=0 \qquad = \qquad x=\frac{-b\pm\sqrt{b^2-4ac}}{2a} \qquad = \qquad x=\frac{-b}{2a}.$$

Solution 2:

$$ax^2+bx+c=0 \qquad \Longrightarrow \qquad x=\frac{-b\pm\sqrt{b^2-4ac}}{2a} \qquad \Longrightarrow \qquad x=\frac{-b}{2a}.$$

Solution 3:
If $ax^2+bx+c=0$, then $x=\dfrac{-b\pm\sqrt{b^2-4ac}}{2a}$. Since $b^2-4ac=0$, it follows that there is only one solution, $x=\dfrac{-b}{2a}$.

1.2.4 Set Notation

We make the naive convention that a set is any collection of objects that are called the elements of the set. Typically the element of a set will be mathematical in nature, so we might have sets of numbers, sets of points in the plane, sets of functions, even sets of sets.

- If A is a set and a is an element of A we write $a\in A$ and read this as ‘ a is an element of A ’ or ‘ a is in A ’
- We can specify a set by listing its elements, for example,

  $$A=\{2,5,9,16\} \qquad \text{or} \qquad P=\{2,3,5,7,11,\dots\},$$

  where there is enough information to infer that P is meant to be the set of all prime numbers.

- We can also specify a set using set builder notation, for example,

$$P = \{x \in \mathbb{N} : x \text{ is prime}\},$$

which reads 'P is the set of all x in $\mathbb{N}$ such that x is prime.'

- The **union** of two sets A and B is the set

$$A \cup B := \{x : x \in A \text{ or } x \in B\},$$

that is 'A union B is defined to be the set of all x such that x is in A or x is in B.'

- The **intersection** of two sets A and B is the set

$$A \cap B := \{x : x \in A \text{ and } x \in B\},$$

that is 'A intersect B is the set of all x such that x is in A and x is in B.'

1.2.5 Intervals in the Real Line

Certain subsets, called **intervals**, of the real line, $\mathbb{R}$, are important enough to have their own special notation.

Let a and b be real numbers such that $-\infty < a < b < \infty$.

Terminology	Notation	Definition
open interval	(a, b)	$(a, b) := \{x \in \mathbb{R} : a < x < b\}$
closed interval	$[a, b]$	$[a, b] := \{x \in \mathbb{R} : a \leq x \leq b\}$
half-open intervals	$(a, b]$ $[a, b)$	$(a, b] := \{x \in \mathbb{R} : a < x \leq b\}$, $[a, b) := \{x \in \mathbb{R} : a < x \leq b\}$,
open half-lines	(a, ∞) $(-\infty, a)$	$(a, \infty) := \{x \in \mathbb{R} : x > a\}$ $(-\infty, a) := \{x \in \mathbb{R} : x < a\}$
closed half-lines	$[a, \infty)$ $(-\infty, a]$	$[a, \infty) := \{x \in \mathbb{R} : x \geq a\}$ $(-\infty, a] := \{x \in \mathbb{R} : x \leq a\}$

- We read, for example, (a, b) as 'the open interval a to b' and $(-\infty, a]$ as 'the closed interval minus infinity to a.'

- In this notation, the real line $\mathbb{R}$ could be written as the interval $(-\infty, \infty)$

- The intervals (a, b), $[a, b]$, $(a, b]$ and $[a, b)$ are said to be **bounded intervals**. The points a and b are called the **endpoints** of the interval.

- The intervals (a, ∞), $(-\infty, a)$, $[a, \infty)$, $(-\infty, a]$, and $(-\infty, \infty)$ are said to be **unbounded intervals**.

The symbols ∞ and $-\infty$ read 'infinity' and 'minus infinity' respectively. These are not real numbers and we cannot treat them as if they are. However, they are useful notation. One can think of them as the endpoints of the real line $\mathbb{R}$ and we define the **extended real line** $\overline{\mathbb{R}} : \mathbb{R} \cup \{-\infty, \infty\}$, extending the usual order on $\mathbb{R}$ by stipulating that $-\infty < x < \infty$ for all $x \in \mathbb{R}$.

1.3 Chris's Golden Rules for Writing Maths

1) Write in properly punctuated sentences, in English. What you have written has to make sense as a piece of English, even if the maths is not right.

2) You should be able to read out exactly what you have written without adding any additional words.

3) Make things as easy as possible for your reader (including future you). Maths is hard enough without making it harder still by not explaining things clearly.

4) Connect your statements together so that they follow logically. You should have lots of words and phrases like 'therefore,' 'hence,' 'because,' and 'it follows that' in what you write.

5) Don't write in columns, don't connect different parts of the page with squiggles and arrows.

6) Try to write in the way that your best lecturers write. What would you like to see

as an explanation?

7) If your sentences are starting to get long and complicated, break them down in to simpler ones.

8) If you have to say ‘You know what I mean’ then it is not a proper explanation! Don’t make your reader try to guess what you mean. Write it out clearly!

For more on this see Kevin Houston’s book ‘How to think like a mathematician’ (www.kevinhouston.net) and the extract on writing maths that you can see here: https://www.kevinhouston.net/pdf/htwm.pdf

2 Proof

The power of mathematics lies in its ability to deduce logically guaranteed conclusions from stated assumptions through rigorous argument. Nowhere is this more apparent than in the idea of proof. The deductive argument of a proof allows us to be absolutely certain that the conclusions are true given the assumptions. We can be logically certain of the statement of a theorem. By contrast, no matter how powerful and effective, the empirical or inductive arguments of science only establish reasonable expectations. Mathematics has theorems; science has theories.

Proof is not the only important idea in mathematics; physical and statistical modelling, optimization, data analysis, and computation are just as fundamental to the role mathematics plays. But rigorous logical argument is what allows you and others to be certain that your conclusions follow from your assumptions (assuming you didn’t make a mistake, of course).

You will encounter many proofs in your degree and you will learn how to prove a number of powerful facts through a number of different proof techniques.

2.1 Methods of Proof

2.1.1 Mathematical Induction

See the section below. Note that despite its name, mathematical induction is a form of deductive reasoning and, as such it is not a form of inductive reasoning.

2.1.2 Direct Proof

Sometimes called ‘definition chasing,’ in a direct proof you start with your assumptions and directly deduce your conclusions. The following is a simple example.

Theorem 2.1.
If m and n are odd integrers, then mn is also odd.

Proof.
Since m and n are odd integers, there are integers j and k such that $m = 2j + 1$ and $n = 2k+1$. But then $mn = (2j+1)(2k+1) = 2(2jk+j+k)+1$. Since $(2jk+j+k)$ is an integer, we see that mn is odd. □

2.1.3 Proof by Contraposition

The statement ‘if P, then Q’ is logically equivalent to the statement ‘if not Q, then not P.’ It is sometimes easier to prove the contrapositive.

Theorem 2.2.
Let n be an integer. If n^2 is even, then n is even.

Proof.
Suppose that n is not even. Then n is odd. Letting $n = m$ in Theorem 2.1, we see that $nm = n^2$ is odd. Hence if n is not even, n^2 is not even. Thus if n^2 is even, n is even. □

2.1.4 Proof by Contradiction

Sometimes called ‘reductio ad absurdum,’ in a proof by contradiction we assume some statement to be true and then show that this implies a contradiction or absurdity, which implies that the statement must be false.

Theorem 2.3.
The square root of 2 is irrational.

Proof.
Suppose that $\sqrt{2}$ is rational. Then there are integers p and q such that $\sqrt{2} = p/q$. By cancelling common factors, we may assume that p and q have no common factors. Since $q\sqrt{2} = p$, we see that $p^2 = 2q^2$ is even. By Theorem 2.2, it follows that p is even and there is some integer r such that $p = 2r$. But then $2q^2 = p^2 = (2r)^2 = 4r^2$, so that $q^2 = 2r^2$ is even. However, if p and q are both even, then they have a common factor of 2. This contradicts the fact that they have no common factor. Hence, $\sqrt{2}$ is not rational. $\square$

2.1.5 Non-constructive Proof

In a non-constructive proof, we conclude that something must be true by showing that a particular thing must exist, without being able to give an explicit instance of that thing. The most famous (and controversial) example is Cantor's proof that the real numbers are uncountably infinite. Here we present an example due to Dov Jarden from the 1950s.

Theorem 2.4.
An irrational number to an irrational power can be rational.

Proof.
By Theorem 2.3, $\sqrt{2}$ is irrational. If $\sqrt{2}^{\sqrt{2}}$ is rational then we are done. If $\sqrt{2}^{\sqrt{2}}$ is

irrational, then $(\sqrt{2}^{\sqrt{2}})^{\sqrt{2}} = \sqrt{2}^{\sqrt{2}.\sqrt{2}} = \sqrt{2}^2 = 2$ is rational. □

Notice that in this proof we do not know whether the example of an irrational power of an irrational being rational is $\sqrt{2}^{\sqrt{2}}$ or $(\sqrt{2}^{\sqrt{2}})^{\sqrt{2}}$.

2.1.6 Constructive Proof

In a constrictive proof, we explicitly come up with instance showing that something is true. For example, e and $\log 2$ are both irrational (indeed transcendental) and by definition $e^{\log 2} = 2$, which proves Theorem 2.4. However, proving that e is irrational is non=trivial. Here is an alternative constructive proof. First we need a lemma.

Lemma 2.5.
The number $\log_2 9$ is irrational.

Proof.
Suppose that $\log_2 9$ is rational, so that there are non-zero integers p and q such that $\log_2 9 = p/q$. By the definition of $\log_2$, $2^{p/q} = 2^{\log_2 9} = 9$. This implies that $2^p = 9^q$. However, a similar proof to that of Theorem 2.1 shows that 9^q is odd, where as 2^p is even. This contradiction implies that $\log_2 9$ must be irrational □

We now have a constructive proof of Theorem 2.4:

Proof.
From the above, $\sqrt{2}$ and $\log_2 9$ are both irrational, but

$$\sqrt{2}^{\log_2 9} = \sqrt{2}^{\log_2(3^2)} = \sqrt{2}^{2\log_2 3} 2^{\log_2 3} = 3$$

□

2.1.7 Proof by Cases

Also called proof by exhaustion. Sometimes you just have to work through every case.

Theorem 2.6.

If n is an integer, then the last digit of n^2 is not 2, 3, 7 or 8.

Proof.

Let n be an integer. Then the last digit of n^2 is determined by the last digit of n.

- If the last digit of n is 0, then the last digit of n^2 is 0.
- If the last digit of n is 1, then the last digit of n^2 is 1.
- If the last digit of n is 2, then the last digit of n^2 is 4.
- If the last digit of n is 3, then the last digit of n^2 is 9.
- If the last digit of n is 4, then the last digit of n^2 is 6.
- If the last digit of n is 5, then the last digit of n^2 is 5.
- If the last digit of n is 6, then the last digit of n^2 is 6.
- If the last digit of n is 7, then the last digit of n^2 is 9.
- If the last digit of n is 8, then the last digit of n^2 is 4.
- If the last digit of n is 9, then the last digit of n^2 is 1.

The result follows. $\square$

2.2 The Principle of Mathematical Induction

Suppose that we want to know whether something is true for every natural number $n \in \mathbb{N}$. For example we might want to know whether or not the following are true.

1) For every $n \in \mathbb{N}$, n is prime.

2) Every natural number has a unique prime factorization.

3) For every $n \geq 1$, $\sum_{i=1}^{n} i = 1 + 2 + \cdots + n = \frac{1}{2}n(n+1)$.

4) For any $n \in \mathbb{N}$, n straight lines divide the plane in to at most $\frac{n^2+n+2}{2}$ distinct regions.

5) In any group of n or more people, the probability of at least two people sharing the same birthday is greater than $1/2$.

All of these claims can be expressed in the form

- 'For every $n \in \mathbb{N}$, $P(n)$, where $P(n)$ is the statement ...'

In the case of 1) to 5), the corresponding statements $P(n)$ are:

1) $P(n)$ is the statement 'n is prime'.

2) $P(n)$ is the statement 'n has a unique prime factorization.'

3) $P(n)$ is the statement '$\sum_{i=1}^{n} i = 1 + 2 + \cdots + n = \frac{1}{2}n(n+1)$.'

4) $P(n)$ is the statement 'n straight lines divide the plane in to at most $\frac{n^2+n+2}{2}$ distinct regions.'

5) $P(n)$ is the statement 'in a group of n people, the probability that at least two of them have the same birthday is greater than $1/2$.'

For any fixed n, the statement $P(n)$ might be true or might be false, for example some natural numbers are prime and some are not. However, we are interested in whether $P(n)$ **is true for every** $n \in \mathbb{N}$. Of the above, 1) is not true for all n, 2), 3) and 4) are true for every n and 5) is true for every $n \geq 23$.

How do we prove that something holds for every natural number? One approach is to use a domino effect: if we can deduce the second case from the first case, the third case from the second case, the fourth case from the third case and so on, then as long as the first case is true, we will eventually see that every case is true. This is the idea behind mathematical induction.

Let us start with an example (we will recast this proof in a bit using mathematical induction):

Theorem 2.7.
If $n > 0$ is a natural number, then $\sum_{i=1}^{n} i = 1 + 2 + \cdots + n = \frac{1}{2}n(n+1)$

Proof.
It is certainly true that $\sum_{i=1}^{1} i = 1$.

Now suppose that

$$\sum_{i=1}^{k} i = 1 + 2 + \cdots + k = \frac{1}{2}k(k+1).$$

Notice that

$$\sum_{i=1}^{k+1} i = (1 + 2 + \cdots + k) + (k+1).$$

But we are supposing that $1 + \cdots + k = \frac{1}{2}k(k+1)$, so that

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

We have shown that $\sum_{i=1}^{n} i = \frac{1}{2}n(n+1)$ is true when $n = 1$ and **if it is true for** $n = k$**,** then it is also true for $n = k + 1$.

Since it is true for $n = 1$, it is also true for $n = 1+1 = 2$, and therefore for $n = 2+1 = 3$, and for $n = 3 + 1 = 4$, and so on. Hence it $1 + \cdots + n = \frac{1}{2}n(n+1)$ for all n.

$\square$

We can use the Principle of Mathematical Induction to formalise this and other arguments

Induction (The Principle of Mathematical Induction)**.**
Let n be a natural number and let $P(n)$ be a statement that depends on n.

Suppose that

Base Step: $P(1)$ is true, and

Inductive Step: if $P(k)$ is true, then $P(k+1)$ is also true.

Then $P(n)$ is true for all $n \geq 1$.

Let's re-do the previous theorem using the Principle of Mathematical Induction (induction, for short).

Proof of Theorem 2.7. Let $P(n)$ be the statement

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

Base step: Let $n = 1$. Notice that

$$\frac{1}{2}1(1+1) = \frac{1}{2}.2 = 1 = \sum_{i=1}^{1} i,$$

so that $P(1)$ is true.

Inductive Step: Let k be a natural number and assume that $P(k)$ is true, so that

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

(We often call this the inductive hypothesis.) We want to show that $P(k+1)$ is true. Notice that by the inductive hypothesis,

$$\sum_{i=1}^{k+1} i = \Big[1+2+\cdots+k\Big]+(k+1) = \frac{1}{2}k(k+1)+(k+1) = \frac{1}{2}(k^2+3k+2) = \frac{1}{2}(k+1)\left[(k+1)+1\right],$$

so that $P(k+1)$ is indeed true.

Hence by the Principle of Mathematical Induction, $P(n)$ is true for all $n \in \mathbb{N}$. $\square$

It is useful to stick to a standard format for your induction arguments. This allows you to concentrate on the difficult bit, letting the induction take care of itself. We will see this in some more examples.

Theorem 2.8.

For any natural number $n > 0$, $\sum_{i=1}^{n}(4i-1) = 3+7+11+\cdots+(4n-1) = n(2n+1)$.

Proof.

Let $P(n)$ be the statement $\sum_{i=1}^{n}(4i-1) = 3+7+11+\cdots+4n-1 = n(2n+1)$.

Base Step: If $n=1$, then

$$\sum_{i=1}^{n}(4i-1) = \sum_{i=1}^{1}(4i-1) = 4.1-1 = 3 = 1(2.1+1) = n(2n+1),$$

which shows that $P(1)$ is true.

Inductive Step: Suppose that $P(k)$ is true. We want to show that this implies that $P(k+1)$ is true. But

$$\sum_{i=1}^{k+1}(4i-1) = \Big(\sum_{i=1}^{k}(4i-1)\Big) + (4(k+1)-1)$$

$$= \Big[k(2k+1)\Big] + (4(k+1)-1) \qquad \text{by the inductive hypothesis}$$

$$= 2k^2+5k+3 = (k+1)(2(k+1)+1),$$

so that $P(k+1)$ is true.

Hence, by induction, $P(n)$ is true for all $n \in \mathbb{N}$. □

The induction does not have to start at $n=1$.

Theorem 2.9.

For all $n \geq 4$, $n! > 2^n$.

Proof.

Let $P(n)$ be the statement $n! > 2^n$.

Notice that $1! = 1 < 2 = 2^1$, $2! = 2 < 4 = 2^2$, $3! = 6 < 8 = 2^3$, so n must be at least 4 for this to hold.

Base Step: If $n=4$, then $n! = 4! = 24 > 16 = 2^4$, so $P(4)$ is true.

Inductive Step: Suppose that $P(k)$ is true, so that $k! > 2^k$. Then $(k+1)! = k!(k+1) > 2^k(k+1)$, by hypothesis, but $k+1 > 2$, so that $(k+1)! = k!(k+1) > 2^k(k+1) > 2^{k+1}$, and we see that $P(k+1)$ is true.

Hence, by induction, $P(n)$ is true for all $n \in \mathbb{N}$. □

Usually the base step is the easier step, but this is not always true. Recall statement 5) above that $P(n)$ is the statement 'in a group of n people, the probability that at least two of them have the same birthday is greater than $1/2$.' This induction starts at $n = 23$, which is a little tricky to prove (you might come across this in a probability module). The inductive step is trivial though: if there are more people in the group it is even more likely that two will share a birthday.

A slightly more complicated looking version, Complete Induction, is actually equivalent to induction but is sometimes useful.

Induction (Complete Induction)**.**
Let n be a natural number and let $P(n)$ be a statement that depends on n.

Suppose that

Base Step: $P(1)$ is true, and

Inductive Step: if $P(1)$, $P(2)$, ..., $P(k)$ are all true, then $P(k+1)$ is also true.

Then $P(n)$ is true for all $n \geq 1$.

The Fibonacci sequence is the sequence

$$f_0 = 0, \quad f_1 = 1, \quad f_2 = f_0+f_1 = 1, \quad f_3 = f_1+f_2 = 2, \quad f_4 = f_2+f_3 = 3, \quad \ldots, \quad f_{k+1} = f_{k-1}+f_k, \quad \ldots$$

Theorem 2.10.
Let f_n be the n^{th} Fibonacci number. Then

$$f_n = \frac{A^n - B^n}{A - B},$$

where $A = \frac{1+\sqrt{5}}{2}$ and $B = \frac{1-\sqrt{5}}{2}$.

Notice that we could have written

$$f_n = \frac{(1+\sqrt{5})^n - (1-\sqrt{5})^n}{2^n\sqrt{5}}$$

and not used A and B , but writing it in this way is neater and draws out the symmetry, which makes the proof a little easier.

Proof.

Let $P(n)$ be the statement

$$f_n = \frac{A^n - B^n}{A - B},$$

where $A = \frac{1+\sqrt{5}}{2}$ and $B = \frac{1-\sqrt{5}}{2}$.

Base Step: If $n = 0$, then

$$\frac{A^n - B^n}{A - B} = \frac{A^0 - B^0}{A - B} = \frac{1-1}{A - B} = 0 = f_0,$$

so that $P(0)$ holds.

If $n = 1$, then

$$\frac{A^n - B^n}{A - B} = \frac{A^1 - B^1}{A - B} = \frac{A - B}{A - B} = 1 = f_1,$$

so $P(1)$ is also true.

Inductive Step: Assume that $P(0), \dots, P(k)$ are true. We want to prove that $P(k+1)$ is true.

$$f_{k+1} = f_k + f_{k-1} \qquad \text{(by definition)}$$

$$= \frac{A^k - B^k}{A - B} + \frac{A^{k-1} - B^{k-1}}{A - B} \qquad \text{(by the induction hypothesis)}$$

$$= \frac{A^{k-1}(A+1) - B^{k-1}(B+1)}{A - B}$$

(note that we used both $P(k)$ and $P(k-1)$ in applying the induction hypothesis). On the other hand,

$$\frac{A^{k+1} - B^{k+1}}{A - B} = \frac{A^{k-1} \cdot A^2 - B^{k-1} \cdot B^2}{A - B}.$$

By comparing the two formulas, we see that they are equal if and only if

$$A^2 = A + 1 \qquad \text{and} \qquad B^2 = B + 1.$$

However, this is easily checked by recalling that $A = \frac{1+\sqrt{5}}{2}$ and $B = \frac{1-\sqrt{5}}{2}$. Indeed

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

and

$$B^2 = \left(\frac{1-\sqrt{5}}{2}\right)^2 = \frac{1+5-2\sqrt{5}}{4} = \frac{1-\sqrt{5}}{2} + 1 = B + 1,$$

as required.

Hence, by complete induction, $P(n)$ holds for all $n \geq 0$ □

3 Inequalities and the Modulus

3.1 The Modulus

Definition 3.1.

The **modulus** or **absolute value** of $x \in \mathbb{R}$ is defined by

$$|x| := \begin{cases} x & \text{if } x \geq 0, \\ -x & \text{if } x < 0. \end{cases} \tag{1}$$

We read $|x|$ as the 'modulus' or 'absolute value of x' or sometimes simply as 'mod x'. It is often very useful to think of $|x - y|$ as the distance from x to y, so that $|x|$ is the distance of x from the origin. This notation is so useful that we often use it in other, similar, contexts, for example to indicate the size of a set, the length of a vector or the distance between two point in 2 or 3 (or higher) dimensional Euclidean space.

The following theorem captures some basic, yet important, properties of the modulus

Theorem 3.1.

Let $x, y \in \mathbb{R}$.

1) $0 \leq |-x| = |x|$.

2) $-|x| \leq x \leq |x|$.

3) $|x| \leq y$ iff $-y \leq |x| \leq y$.

4) $|xy| = |x||y|$.

5) **(The Triangle Inequality)**

$$|x + y| \leq |x| + |y|.$$

6) **(The Reverse Triangle Inequality)**

$$\Big||x| - |y|\Big| \leq |x - y|.$$

The triangle inequality says 'mod x plus y is less than or equal to mod x plus mod y'.

Because $|.|$ (we read this as 'mod of something', the $.$ stands for a blank we can fill in later) is defined in terms of cases, using case by case analysis is often useful in analysing or proving things about the modulus.

Proof.
The statements are all proved by considering cases.

For example, to prove $|xy| = |x|y|$ there are four cases to consider, depending on the sign of x and y. To keep things tidy we might also deal with the case that x or y is 0 separately. For example, suppose that $x < 0$ and $y \geq 0$. If $y = 0$, then $|xy| = |0| = 0 = |x||y|$. So we can assume that $y > 0$. Note that $xy < 0$. By definition, $|x| = -x$, $|xy| = -xy$ and $|y| = y$. Hence $|xy| = -xy = (-x)y = |x||y|$ as required.

Statements 1), 2), 3) and 4) all follow in a similar way.

To prove the triangle inequality suppose first that $x + y \geq 0$. Then

$$|x + y| = x + y \leq |x| + |y|,$$

since we know that $x \le |x|$ and $y \le |y|$ by 2). On the other hand if $x + y < 0$, then

$$|x + y| = -(x + y) = (-x) + (-y) \le |-x| + |-y| \le |x| + |y|,$$

again because $-x \le |-x| = |x|$ and $-y \le |-y| = |y|$ by 1) and 2).

We can prove the reverse triangle inequality in the same way but a slicker method is the following. By the triangle inequality,

$$|x| + |y - x| \ge |x + y - x| = |y| \qquad \text{and} \qquad |y| + |x - y| \ge |y + x - y| = |x|,$$

so that

$$|y - x| \ge |y| - |x| \qquad \text{and} \qquad |x - y| \ge |x| - |y|$$

Since $|x - y| = |y - x|$, this implies that

$$|x - y| \ge \pm\Big(|x| - |y|\Big),$$

from which the result is immediate. $\square$

Note that a very natural way to start to try to prove the triangle inequality would be to consider whether the sign of x and y separately rather than the sign of $x + y$. This works but end up being very tedious.

The triangle inequality is fundamental to analysis and it is often useful to think about it in the forms

$$|x - y| \le |x| + |y| \text{ and}$$

$$|x - y| \le |x - z| + |y - z|.$$

The first of these follows by writing $x + y = x - (-y)$. The second uses a common trick in analysis: introduce the term you want by adding it, taking it away again, and seeing what happens.

$$|x - y| = |x - z + z - y| = |(x - z) - (y - z)| \le |x - z| + |y - z|.$$

If we now think about x, y and z as the vertices of a triangle and $|a - b|$ as the distance from a to b, it is clear why it is called the triangle inequality: it is shorter to go from x to y then it is to go from x to z and then from z to y.

3.2 Solving Inequalities

Typically the best systematic way to solve inequalities is through case by case, sign analysis.

Example 3.1.

Determine the values of x for which

$$x^3 \geq 2x - 1.$$

Solution: Rearranging, this is equivalent to

$$x^3 - 2x + 1 \geq 0.$$

Spotting that 1 is a root, we can factorise

$$0 \leq x^3 - 2x + 1 = (x-1)(x^2 + x - 1) = (x-1)(x-a)(x-b),$$

where $a = -(1+\sqrt{5})/2$ and $b = (\sqrt{5}-1)/2$ and $a < b < 1$. (We don't need to call these roots a and b, but it makes writing things down a little easier.) The sign of $x^3 - 2x + 1$ therefore depends on the signs of $(x-1)$, $(x-a)$ and $(x-b)$. We can put this into a sign analysis table.

	$x < a$	$x = a$	$a < x < b$	$x = b$	$b < x < 1$	$x = 1$	$1 < x$
$x - a$	$-$	0	$+$	$+$	$+$	$+$	$+$
$x - b$	$-$	$-$	$-$	0	$+$	$+$	$+$
$x - 1$	$-$	$-$	$-$	$-$	$-$	0	$+$
$x^3 - 2x + 1$	$-$	0	$+$	0	$-$	0	$+$

From the table, it is clear that $x^3 - 2x + 1 \geq 0$ if and only if $a \leq x \leq b$ or $x \geq 1$. In other words, the solution set $\{x \in \mathbb{R} : x^3 \geq 2x - 1\} = [a, b] \cup [1, \infty)$.

Example 3.2.

Determine the values of $x \in \mathbb{R}$ for which

$$|x - 1| > x^2 - 2x.$$

Solution: The definition of $|.|$ means that there are two cases to consider, namely $x-1<0$ and $x-1\geq 0$.

Case 1: $x-1<0$, i.e. $x<1$, so that $|x-1|=-(x-1)$. This implies that $|x-1|>x^2-2x$ iff $1-x>x^2-2x$, or $0>x^2-x-1$. This quadratic has roots $x=\dfrac{1\pm\sqrt{5}}{2}$. Hence, when $x<1$, $|x-1|>x^2-2x$ when $x\in\left(\dfrac{1-\sqrt{5}}{2},\dfrac{1+\sqrt{5}}{2}\right)$. We need to be a bit careful here! Note that both $x<1$ and $x\in\left(\dfrac{1-\sqrt{5}}{2},\dfrac{1+\sqrt{5}}{2}\right)$, have to hold. But $\dfrac{1-\sqrt{5}}{2}<1<\dfrac{1+\sqrt{5}}{2}$, so in fact the solution set in this case is

$$\left(\frac{1-\sqrt{5}}{2},1\right).$$

(Note that $x<1$, so $x=1$ is not in this solution set.)

Case 2: $x-1\geq 0$, so that $|x-1|=x-1$. This implies that $|x-1|>x^2-2x$ iff $x-1>x^2-2x$, or $0>x^2-2x+1$. This quadratic has roots $x=\dfrac{3\pm\sqrt{5}}{2}$. Hence, when $x\geq 1$, $|x-1|>x^2-2x$ when $x\in\left(\dfrac{3-\sqrt{5}}{2},\dfrac{3+\sqrt{5}}{2}\right)$. Since $\dfrac{3-\sqrt{5}}{2}<1<\dfrac{3+\sqrt{5}}{2}$, the solution set is

$$\left[1,\frac{3+\sqrt{5}}{2}\right).$$

(Note that here $x\geq 1$, so $x=1$ is in the solution set.)

Putting these together we see that the solution set is

$$\left(\frac{1-\sqrt{5}}{2},1\right)\cup\left[1,\frac{3+\sqrt{5}}{2}\right)=\left(\frac{1-\sqrt{5}}{2},\frac{3+\sqrt{5}}{2}\right).$$

Example 3.3.

Solve the inequality

$$\frac{x^2-3x+2}{x^3+x}\geq 0.$$

Solution Factorising we see that

$$x^2-3x+2=(x-1)(x-2) \qquad \text{and} \qquad x^3+x=x(x^2+1).$$

Since $x^2+1\geq 1>0$ for all x, the sign of $\dfrac{x^2-3x+2}{x^3+x}$ depends on the signs of x, $(x-1)$ and $(x-2)$. The sign analysis table is therefore

	$x<0$	$x=0$	$0<x<1$	$x=1$	$1<x<2$	$x=2$	$2<x$
$x-2$	$-$	$-$	$-$	$-$	$-$	0	$+$
$x-1$	$-$	$-$	$-$	0	$+$	$+$	$+$
x	$-$	0	$+$	$+$	$+$	$+$	$+$
x^2+1	$+$	$+$	$+$	$+$	$+$	$+$	$+$
$\frac{x^2-3x+2}{x^3+x}$	$-$	n.d.	$+$	0	$-$	0	$+$

where 'n.d.' stands for 'not defined.'

From the table we see that $\dfrac{x^2-3x+2}{x^3+x} \geq 0$ iff $x \in (0,1] \cup [2,\infty)$

Sometimes inequalities involve more complicated functions, and some knowledge of the properties of these functions is needed.

Example 3.4.
Find the solution set to the inequality

$$(x-2)\log(x^2-16) > 0.$$

Solution: As a real valued function, the natural logarithm $\log x$ is only defined for $x > 0$. So for $(x-2)\log(x^2-16)$ to be defined,

$$x^2 - 16 > 0.$$

Moreover, $\log x > 0$ iff $x > 1$, so that $\log(x^2-16) > 0$ iff $x^2-16 > 1$, i.e iff $x^2-17 > 0$.

Therefore, we need to analyse the signs of the three polynomials $x-2$, $x^2-16 = (x-4)(x+4)$, $x^2-17 = (x-\sqrt{17})(x+\sqrt{17})$, giving a sign analysis table

		$-\sqrt{17}$		-4		2		4	
$x-4$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	0	$+$
$x+4$	$-$	$-$	$-$	0	$+$	$+$	$+$	$+$	$+$
x^2-16	$+$	$+$	$+$	0	$-$	$-$	$-$	0	$+$
$x-\sqrt{17}$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$	$-$
$x+\sqrt{17}$	$-$	0	$+$	$+$	$+$	$+$	$+$	$+$	$+$
x^2-17	$+$	0	$-$	$-$	$-$	$-$	$-$	$-$	$-$
$\log(x^2-16)$	$+$	0	$-$	n.d.	n.d.	n.d.	n.d.	n.d.	$-$
$x-2$	$-$	$-$	$-$	$-$	$-$	0	$+$	$+$	$+$
$(x-2)\log(x^2-16)$	$-$	0	$+$	n.d.	n.d.	n.d.	n.d.	n.d.	$-$

From which we can see that the inequality is satisfied if and only if x belongs to the set

$$(-\sqrt{17},-4)\cup(\sqrt{17},\infty).$$

4 Naive Set Theory

4.1 Sets, Elements and Subsets

Set theory is an advanced branch of pure mathematics. It seeks to derive all of mathematics by expressing mathematical objects in terms of objects called sets that are defined in terms of a small set of axioms. However, for the purposes of what we want to do here we make the following naive definition.

Definition 4.1.
A **set** is a collection of objects that are called the **elements** of the set.

The elements of a set are typically mathematical in nature. They might be numbers, functions, points in the plane, or even other sets. We usually use capital letters to denote sets and lower case letters to denote the elements of a set.

Notation 4.1.
If A is a set and a is an element of A and b is not an element of A, then we write $a \in A$ and $b \notin A$. We read this as 'a is an element of A' or 'a is in A' or a in A and 'b is not in A' or 'b is not an element of A'.

We can specify a set by:

- defining it directly, for example, by defining P to be the set of prime numbers;
- explicitly listing its element, for example, letting $A := \{1, 2, 3, 6\}$, which says 'A is defined to be the set containing 1, 2,3, and 6';
- implicitly listing its elements, for example, seeing $P = \{2, 3, 5, 7, 11, 13, \dots\}$, which says P is defined to be the set containing 2, 3, 5, 7, 11, 13, and so on, we can reasonably infer that P is the set of primes';
- using **set-builder** notation, so that we define a set by some property, for example

  $$P := \{n \in \mathbb{N} : n \text{ is prime}\},$$

  which says 'P is defined to be the set of n in $\mathbb{N}$ such that n is prime.'

Definition 4.2.
The **empty set**, $\emptyset$, is the set with no elements.

Definition 4.3.
Let A and B be sets. We say that:

- A is a **subset** of B, denoted $A \subseteq B$ or $B \supseteq A$, if every element of A is an element of B;

- A **equals** B, denoted $A = B$, if A and B have exctly the same elements;

- A is a **proper subset** of B, denoted $A \subsetneq B$, if $A \subseteq B$ but $A \neq B$.

If $A \subseteq B$ we sometime say that B is a **superset** of B.

Note that some people use $\subset$ instead of $\subseteq$ and some people use $\subset$ instead of $\subsetneq$.

The definition of set equality means that $A = B$ iff $A \subseteq B$ and $B \subseteq A$. It also means that it does not matter which order the elements of a set are written in or how many times that they are listed, so that

$$\{1,2,3,6\} = \{3,6,2,1\} = \{10/5, 6/3, 3, 6/2, 1, 4/2, 12/2\}.$$

If we take the integers, $\mathbb{Z}$, and the real numbers, $\mathbb{R}$, as given, then we might write

$$\mathbb{N}_0 = \{n \in \mathbb{Z} : n \geq 0\},$$

$$\mathbb{N} = \{n \in \mathbb{Z} : n > 0\},$$

$$\mathbb{Q} = \{p/q : p, q \in \mathbb{Z}, q \neq 0\},$$

$$\mathbb{C} = \{x + iy : x, y \in \mathbb{R}\},$$

and we have

$$\emptyset \subseteq \mathbb{N} \subseteq \mathbb{N}_0 \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}.$$

4.2 The Algebra of Sets

Just as with numbers, we can combine sets to create new ones.

Definition 4.4.
Let X be a set and let A and B be subsets of X.

- The union of A and B is the set

$$A \cup B := \{x \in X : x \in A \text{ or } x \in B\}.$$

- The intersection of A and B is the set

$$A \cap B := \{x \in X : x \in A \text{ and } x \in B\}.$$

- A minus B is the set

$$A - B := \{x \in X : x \in A \text{ but } x \notin B\}.$$

- The complement of A (in X) is the set

$$A^c := \{x \in X : x \notin A\}.$$

Some people write $A \setminus B$ for $A - B$. We sometimes call $A - B$ the 'the complement of A in B'. A^c is also written A'. Clearly $A^c = X - A$.

Notice that we can define $A \cup B$, $A \cap B$ and $A \setminus B$ without reference to a 'universal' set X containing both A and B. However the complement of A is not well-defined without reference to the superset we are considering. For example, if $A = \{1, 2\}$, $X = \{1, 2, 3\}$ and $Y = \{0, 1, 2\}$, then A is a subset of both X and Y and the complement of A in X is $\{3\}$ whereas the complement of A in Y is $\{0\}$.

Theorem 4.1.
Let A, B and C be sets.

1) (Transitivity of $\subseteq$) If $A \subseteq B$ and $B \subseteq C$, then $A \subseteq C$.

2) (Associativity)

$$(A \cup B) \cup C = A \cup (B \cup C), \text{ and}$$

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

3) (Distributivity)

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

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

4) (De Morgan's Laws)

$$A - (B \cup C) = (A - B) \cap (A - C), \text{ and}$$

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

Note that associativity means that we can write $A \cup B \cup C$ or $A \cap B \cap C$ without confusion.

We leave the proof of this as an exercise, but give an example of how to approach things like this here.

Example 4.1.
Let us show that $A - (B \cup C) = (A - B) \cap (A - C)$. To show that two sets are equal we need to show that they have the same elements.

Suppose that $x \in A - (B \cup C)$.

This is equivalent to saying that $x \in A$ and $x \notin B \cap C$.

Now for an element to be in $B \cap C$ it must be in both B and C, so $x \notin B \cap C$ iff either $x \notin B$ or $c \notin C$.

Hence $x \in A - (B \cup C)$ iff $x \in A$ and either $x \notin B$ or $c \notin C$.

This is equivalent to saying that either $x \in A$ but $x \notin B$ or $x \in A$ but $x \notin C$, i.e. either $x \in A - B$ or $x \in A - C$, which is true iff $x \in (A - B) \cup (A - C)$.

Notice that if A and B are subsets of the set X, then De Morgan's Laws imply that

$$(A \cup B)^c = A^c \cap B^c \qquad \text{and} \qquad (A \cap B)^c = A^c \cup B^c.$$

Suppose that $a \in A$ and $b \in B$. Then the set $\{a, b\}$ is the same as the set $\{b, a\}$. We would like a notion of an ordered pair, so that, for example, we can have a first and second coordinate. We are familiar with the the Euclidean or x-y plane $\mathbb{R}^2$. This is an example of a Cartesian product of sets.

Definition 4.5.
Let A and B be sets.

- For $a \in A$ and $b \in B$, we say that (a, b) is an **ordered pair** if whenever $a, a' \in A$, $b, b' \in B$ and $(a, b) = (a', b')$, then $a = a'$ and $b = b'$.

- The **Cartesian product**

$$A \times B := \{(a, b) : a \in A, b \in B\},$$

  the set of all order pairs whose first coordinate is in A and second coordinate is in B.

Note that there is a clash of notation for open intervals and ordered pairs. However, context usually make the meaning clear.

4.3 Arbitrary Unions and Intersections

You are unlikely to need the ideas in this section unless you study some of the more advanced pure mathematics modules in later years.

From what we have seen, we can write $A \cup B \cup C$ and $A \cap B \cap C$ without ambiguity. What if we want to talk about the union or intersection of a larger, possibly infinite, collection of sets. numbers of sets. For example, for each prime p, define N_p to be the set of all natural numbers divisible by p. What is the union of all the sets N_p? What is the intersection?

The following is a more general defintion of union and intersection.

Definition 4.6.
Let $\mathcal{C}$ be a collection of sets.

- We define the union of the sets in $\mathcal{C}$ to be the set

$$\bigcup \mathcal{C} = \bigcup_{C \in \mathcal{C}} C = \{x : x \in C \text{ for some } C \in \mathcal{C}\}$$

- We define the intersection of the sets in $\mathcal{C}$ to be the set

$$\bigcap \mathcal{C} = \bigcap_{C \in \mathcal{C}} C = \{x : x \in C \text{ for all } C \in \mathcal{C}\}$$

For example, if $\mathcal{C} = \{A, B, C\}$, then $A \cap B \cap C = \bigcap \mathcal{C}$ and, with $\mathbb{N}_p$ as above,

$$\bigcup \{N_p : p \text{ is prime}\} = \{n : n \in N_p \text{ for some prime } p\}$$

$$= \{n \in \mathbb{N} : n \text{ is divisible by } p \text{ for some prime } p\}$$

$$= \{n \in \mathbb{N} : n > 1\}.$$

A very similar proof to that of Theorem 4.1 shows the following also holds.

Theorem 4.2.

Let A and $\mathcal{C}$ be a collection of sets.

1) (Distributivity)

$$A \cup \bigcap \mathcal{C} = \bigcap \{(A \cup C) : C \in \mathcal{C}\}$$

$$A \cap \bigcup \mathcal{C} = \bigcup \{(A \cap C) : C \in \mathcal{C}\}.$$

2) (De Morgan's Laws)

$$A - \bigcup \mathcal{C} = \bigcap \{(A - C) : C \in \mathcal{C}\},$$

$$A - \bigcap \mathcal{C} = \bigcup \{(A - C) : C \in \mathcal{C}\}.$$

5 Functions

5.1 The Defintion of a Function

Definition 5.1.

A **function** f from X to Y consist of three things:

1) a set $X = \mathrm{dom}(f)$ called the **domain** of f ;

2) a set $Y = \mathrm{codom}(f)$ called the **co-domain** of f ;

3) a **well-defined rule** that assigns a **unique element** $f(x) \in Y$ to each element $x \in X$.

When talking about functions, we often call the elements of X and Y **points**.

Definition 5.2.
We say that a rule f assigning elements of X to elements of Y is **well-defined** or **unambiguous** provided whenever $x_1 = x_2$, we have $f(x_1) = f(x_2)$.

So we want a function from X to Y to give an output for each element of x . That output should be in Y , it should be unique and its value should be unambiguously defined.

Notation 5.1.
If f is a function from X to Y , we write

$$f : X \to Y,$$

which says 'f is a function from X to Y .'

If $f(x) = y$, then we often write

$$f : x \mapsto y.$$

which says 'f maps x to y .' We say that y is the **image** of x under f or the **value** of f at x .

We call x the **independent variable** and y the **dependent variable**.

The domain of f is the set of input values for f and the co-domain is the set of possible outputs. Notice that not every element of the co-domain needs to be an output of the function. For example, if

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

-1 is not equal to $f(x)$ for any x. So we define another set called the range or image of f.

Definition 5.3.

Let $f : X \to Y$. The **range** or **image** of f is the set

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

More generally if $A \subseteq X$, the **image** of A is the set

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

If $B \subseteq Y$, the **pre-image** of B is the set

$$f^{-1}(B) = \{x \in X : f(x) \in B\}.$$

In particular, if $y \in Y$, then $f^{-1}(\{y\})$ is the set of points in X that are mapped to the point y. We often abuse notation by writing $f^{-1}(y)$ (but be a little careful not to confuse this with the inverse function f^{-1}, if it exists).

Example 5.1.

Do the following represent functions?

1)

$$f : \mathbb{R} \to \mathbb{R},$$

$$x \mapsto x + x^2 + x^3.$$

2)

$$f : \mathbb{R} \to \mathbb{R},$$

$$x \mapsto \sin x.$$

3)

$$f : \mathbb{R} \to [-1, 1],$$

$$x \mapsto \sin x.$$

4)

$$f : \mathbb{R} \to [0, 1],$$

$$x \mapsto \sin x.$$

5)

$$f : [0, 1] \to \mathbb{R},$$

$$x \mapsto \sin x.$$

6)

$$f : \mathbb{R} \to \mathbb{R},$$

$$x \mapsto 1/x.$$

7)

$$f : \{1, 2, 3, 4\} \to \{1, 2, 3, 4\},$$

$$1 \mapsto 2,$$

$$2 \mapsto 2,$$

$$3 \mapsto 2,$$

$$4 \mapsto 2.$$

8)

$$f : \mathbb{Q} \to \mathbb{Z},$$

$$p/q \mapsto p + q.$$

9) $f : [0, \infty) \to \mathbb{R}$ with f defined by the rule $f(x)^2 = x$.

Definition 5.4.
Two functions f and g are said to be **equal** if

1) $\mathrm{dom}(f) = \mathrm{dom}(g)$,

2) $\mathrm{codom}(f) = \mathrm{codom}(g)$, and

3) $f(x) = g(x)$ for all $x \in \mathrm{dom}(f)$.

Example 5.2.
1) Technically $f : \mathbb{R} \to \mathbb{R}; x \mapsto \sin x$ and $f : \mathbb{R} \to [-1, 1]; x \mapsto \sin x$ are different functions (although we usually do not worry about this).

2) The functions $f : \mathbb{R} \to \mathbb{R}; x \mapsto x^2 - 2x - 3$ and $g : \mathbb{R} \to \mathbb{R}; x \mapsto (x + 1)(x - 3)$ are equal.

3) The functions $\zeta : [1, \infty) \to \mathbb{R}; s \mapsto \sum_{n=1} \frac{1}{n^s}$ and $\overline{\zeta} : [1, \infty) \to \mathbb{R}; s \mapsto \prod_{p \text{ is prime}} \frac{1}{1 - p^{-s}}$ are also equal, although it is not obvious.

Definition 5.5.
Let $f : X \to Y$ and $g : Y \to Z$. The **composition** of g with f is the function

$$g \circ f : X \to Z$$

$$x \mapsto g(f(x)).$$

Definition 5.6.

Let $f : X \to Y$. If $A \subseteq X$, the **restriction** of f to A is the function

$$f|_A : A \to Y$$

$$x \mapsto f(x).$$

Definition 5.7.

Let X be a set. The **identity function** on X is the function

$$\mathrm{id}_X : X \to X x \quad \mapsto x.$$

Definition 5.8.

The function $f : X \to Y$ is said to be a **constant** iff there is some $c \in Y$ such that $f(x) = c$ for all $x \in X$.

5.2 Injections, Surjections and Bijections

Suppose that $f : X \to Y$. A reasonable question might be to ask what values in X give a particular value y in Y . We saw above that these values are the points in the pre-image set $f^{-1}(\{y\})$. We can then ask whether there is a function that undoes what f does, and inverse function.

Definition 5.9.

Let $f : X \to Y$. The **inverse** of f , if it exists, is the function $f^{-1} : Y \to X$ defined by the rule $f^{-1}(y) = x$ iff $f(x) = y$.

Note that we use the notation f^{-1} in two different ways: $f^{-1}(B)$ is always defined when B is a subset of Y ; the function f^{-1} need not always be defined. The question then is whether f^{-1} exists as a function from Y to X .

- For f to be a function we need it to be defined for every $x \in X$. If f^{-1} is going to be a function, it needs to be defined for every y in Y . From the definition of f^{-1}, this means that for every $y \in Y$ there has to be an $x \in X$ such that $f(x) = y$.

Definition 5.10.
Let $f : X \to Y$. We say that f is **onto** (or **surjective** or that f is a surjection) iff for every $y \in Y$, there is at least one $x \in X$ such that $f(x) = y$.

The following is obvious.

Theorem 5.1.
Let $f : X \to Y$. Then f is onto iff $\mathrm{ran}(f) = \mathrm{codom}(f)$ iff $f(X) = Y$.

- The other condition we require for f to be a function is that the rule assigning points in Y to points in X should be well-defined and that for each x there should be a unique y in Y such that $f(x) = y$. For this to be true of f^{-1}, we require that if $y_1 = y_2$, then $f^{-1}(y_1) = f^{-1}(y_2)$. From the definition of f^{-1}, if it is defined, and $f(x_1) = y_1$ and $f(x_2) = y_2$, then f^{-1} is well-defined, iff whenever $f(x_1) = f(x_2)$, then $x_1 = x_2$. This is the same as saying that for all $y \in Y$, there is at most one $xinX$ such that $f(x) = y$.

Definition 5.11.
Let $f : X \to Y$. We say that f is **one-to-one** (or **injective** or that f is an injection) iff for every $y \in Y$, there is at ‘most one $x \in X$ such that $f(x) = y$.

Theorem 5.2.
Let $f : X \to Y$. The following are equivalent:

1) f is one-to-one;

2) if $x_1 \neq x_2$, then $f(x_1) \neq f(x_2)$;

3) if $f(x_1) = f(x_2)$, then $x_1 = x_2$.

So for a well-defined inverse of $f : X \to Y$ to exist f must be both one-to-one and onto.

Definition 5.12.
Let $f : X \to Y$. We say that f is bijective (or that f is an bijection) iff it is both an injection and a surjection.

From the definition of f^{-1}, $f^{-1}(y) = x$ if and only if $f(x) = y$. So $f(f^{-1}(y)) = f(x) = y$ and $f^{-1}(f(x)) = f^{-1}(y) = x$. So we have the following theorem.

Theorem 5.3.
Let $f : X \to Y$. The inverse of f, $f^{-1} : Y \to X$ exists iff f is a bijection.

In this case,

$$f^{-1} \circ f(x) = f^{-1}(f(x)) = x \qquad \text{and} \qquad f \circ f^{-1}(y) = f(f^{-1}(y)) = y.$$

Example 5.3.
Let $f : \mathbb{R} - \{2\} \to \mathbb{R} - \{1\}$ be defined by the rule

$$f(x) = \frac{x-1}{x-2}.$$

Prove that f is one-to-one and onto and determine the inverse of f.

**Solution:* To see that f is one-to-one, suppose that $f(x_1) = f(x_2)$, so that $\frac{x_1 - 1}{x_1 - 2} = \frac{x_2 - 1}{x_2 - 2}$. Rearranging we have $(x_1 - 1)(x_2 - 2) = (x_2 - 1)(x_1 - 2)$, which simplifies to show that $x_1 = x_2$. Hence f is one-to-one.

To see that f is onto, let $1 \neq y = f(x) = \frac{x-1}{x-2}$. Rearranging to solve for x in terms of y, we see that $xy - 2y = x - 1$ or $xy - x = 2y - 1$ so that $x = \frac{2y-1}{y-1}$. This implies that for any $y \neq 1$, there is an x such that $f(x) = y$. Hence f is onto.

Since f is a bijection it has an inverse, the formula for which is given by $f^{-1}(y) = \frac{2y-1}{y-1}$. So $f^{-1} : \mathbb{R} - \{1\} \to \mathbb{R} - \{2\}; x \mapsto \frac{2x-1}{x-1}$.