Imperial College London

Introduction to University Mathematics

MATH40001/MATH40009

Dr. T. Bertrand
Prof. K. Buzzard
Dr. M.-A. Lawn
Department of Mathematics
Imperial College London

2023 – 2024

Contents

Chapter 1

Sets and Logic

1.1 Propositions and logic

Definition 1.1.1: Proposition

A *logical proposition*, usually just called a *proposition*, is a true/false statement.

Here are some examples of logical propositions:

- $2+2=4$.
- $2+2=5$.
- For all real numbers x and y, $x+y=y+x$.
- The statement of Pythagoras' theorem.
- The statement of the Riemann Hypothesis.

Each of these statements is either true, or false. In fact, the first, third and fourth statements are true, the second one is false, and nobody knows whether the last one is true or false.

Here are some examples of mathematical objects which are *not* propositions:

- $2+2$
- $+$
- π
- The proof of Pythagoras' theorem.

None of these are true/false statements. The first and third of them are numbers, the second one is a function (which takes two numbers a and b in and outputs their sum $a+b$), and the last one is a proof, which is a sequence of logical arguments. Although it is not part of the course, here is a blog post which explains more about where all these parts of mathematics live and how they fit together.

1.2 Doing mathematics with propositions

You all know that if you want to do mathematics with a general real number, you can just call it x. Then you can talk about things like $x - 3$ or $x^2 + 2x + 1$ and so on. You can also do mathematics with matrices; you can let A and B be general two by two matrices, and then talk about things like $A + B$.

During your degree here, you will learn that you can also do mathematics with some far more general things. You will do mathematics with numbers, matrices, groups, elements of groups, rings, elements of rings, vector spaces, elements of vector spaces and so on. In this section, we will do mathematics with propositions.

If you're doing mathematics with real numbers like x and y, you can do things like considering $x + y$ (a number) or x^2/y (a number, as long as y isn't zero), or $x = y$ (a proposition), or $x < y$ (a proposition).

If you're doing mathematics with propositions, there are different things you can do. In this section we will see the most important things we can do with propositions.

1.2.1 And ($\wedge$)

Good variable names for propositions are things like P or Q or R.

If P and Q are propositions, we can make a new proposition called $P \wedge Q$ (Lean). The symbol $\wedge$ is pronounced "and". The definition of $P \wedge Q$ is that $P \wedge Q$ is true if, and only if, both P and Q are true.

Unlike number variables like x, proposition variables like P can only take two values – `true` and `false`. So we can explain precisely what $P \wedge Q$ is by writing down the so-called *truth table* for $\wedge$:

P	Q	$P \wedge Q$
false	false	false
false	true	false
true	false	false
true	true	true

For example, $(2 + 2 = 4) \wedge (2 + 2 = 5)$ is false, because $2 + 2 = 5$ is false.

1.2.2 Or ($\vee$)

A companion to and is the "or" function. If P and Q are propositions, then $P \vee Q$ (Lean) is the proposition which is true when either P, or Q, *or both*, are true. Be careful! The word "or" is sometimes used differently in English. For example, in the sentence "you get down from that tree, or I will call the police", if P is the proposition "you get down from that tree" and Q is the proposition "I will call the police", then the person who said the sentence *probably* meant that P and Q will not both be true – it will be one or the other. The mathematical $P \vee Q$ is happy for both P and Q to be true. Here's the truth table:

P	Q	$P \vee Q$
false	false	false
false	true	true
true	false	true
true	true	true

For example, $(2 + 2 = 4) \vee (2 + 2 = 5)$ is true, because $2 + 2 = 4$ is true.

1.2.3 Not ($\neg$)

If P is a proposition, then $\neg P$ (Lean) is the "opposite one". It is the proposition which is true if P is false, and false if P is true. Here is the truth table:

P	$\neg P$
false	true
true	false

As an example, if P is the statement of the Riemann Hypothesis, then $P \vee \neg P$ is true, because P is either true or false, exactly one of P and $\neg P$ will be true.

1.2.4 Implies ($\implies$)

In some sense, this is the most subtle of the functions associated with propositions. If P and Q are propositions, then $P \implies Q$ is also a proposition, defined in English as "if P is true, then Q is true". Let's think a little about what this means.

The easiest way to think about this definition is to ask how it can go wrong. When is $P \implies Q$ false? The only way it can fail, is that P is true, but Q is not true. This means that the truth table must look like this:

P	Q	$P \implies Q$
false	false	true
false	true	true
true	false	false
true	true	true

Does this agree with our intuition for how $\implies$ behaves?

If x, y and a are real numbers, and if we know $x = y$, then multiplying both sides by a we deduce that $ax = ay$. So we can write that for all real numbers x, y and a, we have

$$x = y \implies ax = ay.$$

Whatever values x and y and a take, $x = y \implies ax = ay$ will be true.

Let's try some example values for x, y and a, to see what we can deduce.

Trying $x = y = 1$ and $a = 2$, we deduce that $1 = 1 \implies 2 = 2$.

Trying $x = 1$, $y = 2$, $a = 3$, we deduce that $1 = 2 \implies 3 = 6$.

Trying $x = 1$, $y = 2$, $a = 0$ we deduce that $1 = 2 \implies 0 = 0$.

Because we are happy with $x = y \implies ax = ay$, we must then also be happy that $\texttt{true} \implies \texttt{true}$, $\texttt{false} \implies \texttt{false}$ and $\texttt{false} \implies \texttt{true}$ are all true.

On the other hand, we can never prove a false result from a true result (assuming we don't make any mistakes!). So we should also be happy with the idea that $\texttt{true} \implies \texttt{false}$ should be false.

One final note – this symbol is sometimes used backwards. When we say $Q \impliedby P$ we just mean $P \implies Q$.

1.2.5 Logical equivalence ($\iff$)

Two propositions P and Q are said to be *logically equivalent* if $P \implies Q$ and $Q \implies P$. We use the notation $P \iff Q$ (Lean) for this. In words – "if P is true, then Q is true, and if Q is true, then P is true". Let's investigate this idea using truth tables, remembering that $P \iff Q$ is defined to be $(P \implies Q) \wedge (Q \implies P)$.

P	Q	$P \implies Q$	$Q \implies P$	$P \iff Q$
false	false	true	true	true
false	true	true	false	false
true	false	false	true	false
true	true	true	true	true

What is going on here? The third column comes from the definition of $\implies$, and the fourth column does too. The fifth column comes from the definition of $\wedge$, applied to the third and fourth column.

We deduce that $P \iff Q$ is true when both P and Q are true, and also when both P and Q are false, but it is false otherwise. In fact, $P \iff Q$ looks rather like $P = Q$!

Mathematicians believe in something called *propositional extensionality*, which is the idea that two logical propositions P and Q are logically equivalent if and only if they are equal. This way of thinking about the world means that there are only two propositions, namely `true` and `false`. In particular, Fermat's Last Theorem is true, and $2 + 2 = 4$ is true, so Fermat's Last Theorem equals $2 + 2 = 4$.

1.3 Relations between these concepts

Here we talk about relations between $\wedge$, $\vee$, $\neg$, $\implies$ and $\iff$. There are only two possible "values" for a proposition – namely `true` and `false`. On the other hand we have introduced five ways of being able to make new propositions from old ones, namely $\wedge$, $\vee$, $\neg$, $\implies$ and $\iff$. So it is not surprising that there are some relations between these terms. Let's start with the easiest one.

1.3.1 The law of the excluded middle

See also Wikipedia.

One way of stating the law of the excluded middle is as follows.

Lemma 1.3.1

If P is a proposition, then $\neg(\neg P) \iff P$ (Lean).

In words – the proposition $\neg(\neg P)$ is logically equivalent to P. How can we check this?

Proof. We try all the possibilities.

P	$\neg P$	$\neg(\neg P)$	$\neg(\neg P) \iff P$
`true`	`false`	`true`	`true`
`false`	`true`	`false`	`true`

We work out the final column by looking at the first and third column, and seeing if they are equal or not. We see that the final column, $\neg(\neg P) \iff P$, is always true. Hence $\neg(\neg P) \iff P$ is always true. ∎

Exercises: (1) check that $\neg P \iff (P \implies \texttt{false})$. (2) Check that $P \vee (\neg P)$ is true.

1.3.2 De Morgan's laws

Let P and Q be propositions. Augustus de Morgan noted the following two theorems:

Lemma 1.3.2

1. $\neg(P \vee Q) \iff (\neg P) \wedge (\neg Q)$ (Lean);
2. $\neg(P \wedge Q) \iff (\neg P) \vee (\neg Q)$ (Lean)

Because P and Q can only take the values `true` and `false`, we can easily prove these results by just checking all cases. Here we prove the first part; we leave the second to you.

Proof. Part 1 goes like this:

P	Q	$P \vee Q$	$\neg(P \vee Q)$	$\neg P$	$\neg Q$	$(\neg P) \wedge (\neg Q)$	$(\neg(P \vee Q)) \iff ((\neg P) \wedge (\neg Q))$
false	false	false	true	true	true	true	true
false	true	true	false	true	false	false	true
true	false	true	false	false	true	false	true
true	true	true	false	false	false	false	true

As you can see, the fourth column $\neg(P \vee Q)$ and the seventh column $\neg P \wedge \neg Q$ are the same, and hence the eighth column is true in every case.

As for part 2, I'm sure you can fill in the blanks yourself:

P	Q	$P \wedge Q$	$\neg(P \wedge Q)$	$\neg P$	$\neg Q$	$(\neg P) \vee (\neg Q)$	$(\neg(P \wedge Q)) \iff ((\neg P) \vee (\neg Q))$
false	false						
false	true						
true	false						
true	true						

■

As you can see, there is a method for proving these things. In school, there is a method for solving pretty much any question they throw at you. At university, life is usually not so easy – you sometimes have to make your own methods. We will begin to see examples of this later on in the course.

1.3.3 Transitivity of implications, and the contrapositive

In this subsection we talk about two basic properties of $\implies$, both of them of fundamental importance.

The first result is the following. If P, Q and R are propositions, and we know $P \implies Q$ and $Q \implies R$, we can deduce that $P \implies R$. Thinking about our definition of $P \implies Q$ as "if P is true, then Q is true", we can argue as follows.

Say $P \implies Q$ and $Q \implies R$ are both true. We want to prove that $P \implies R$ is true. OK then, let's assume that P is true. Because $P \implies Q$, we deduce that Q is true. Now because $Q \implies R$, we can deduce that R is true. This is what we had to prove, so we're done.

Alternatively, one can of course check that $((P \implies Q) \wedge (Q \implies R)) \implies (P \implies R)$ by checking all eight cases.

This result is called *transitivity of* $\implies$ (Lean). Another well-known transitive operation is $\leq$ on real numbers: if $x \leq y$ and $y \leq z$ then $x \leq z$ (Lean).

The other result we want to talk about in this section is something called the *contrapositive* of an implication. The contrapositive of the implication $P \implies Q$ is the implication $\neg Q \implies \neg P$. It turns out that any implication implies its contrapositive! In symbols, here is the claim (Lean):

$$(P \implies Q) \implies (\neg Q \implies \neg P).$$

Let's see if we can figure out why this is true.

Say we know that if P is true, then Q is true (i.e., we know that P implies Q). Let's say we *also* know that Q is in fact false! Then for sure P cannot be true, because if P were true then, because $P \implies Q$ is also true, we would be able to deduce that Q were true.

This means that if we know that $P \implies Q$ is true, and we also know that Q is false, then we can deduce that P is also false. So if $P \implies Q$, then knowing $\neg Q$ is true, we can deduce $\neg P$ is true. This is exactly what the statement above says.

We could instead just write down the "check all cases" proof:

P	Q	$P \implies Q$	$\neg Q$	$\neg P$	$\neg Q \implies \neg P$	$(P \implies Q) \implies (\neg Q \implies \neg P)$
false	false	true	true	true	true	true
false	true	true	false	true	true	true
true	false	false	true	false	false	true
true	true	true	false	false	true	true

Looking at this table, we see that actually more is true: column three ($P \implies Q$) and column six ($\neg Q \implies \neg P$) are exactly the same! So we could actually write down a stronger true statement (Lean):

$$(P \implies Q) \iff (\neg Q \implies \neg P).$$

Exercise: What is the contrapositive of $\neg Q \implies \neg P$? What is this equivalent to?

1.3.4 Distributivity

Distributivity is what mathematicians call "expanding out the brackets" when they want to look clever. More precisely, the well-known fact that

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

is called "distributivity of multiplication over addition". It's true for real numbers, and it's true for matrices as well. On the other hand, addition does not distribute over multiplication: if we switch $\times$ and $+$ in $a \times (b+c) = a \times b + a \times c$, then we get $a + (b \times c) = (a+b) \times (a+c)$ which is in general not true (exercise: *prove* it is not true, by giving a counterexample!). So multiplication distributes over addition, but addition does not distribute over multiplication.

It turns out that $\wedge$ distributes over $\vee$, and $\vee$ distributes over $\wedge$ as well. In other words:

Lemma 1.3.3

Let P, Q and R be propositions. Then

1. $P \wedge (Q \vee R) \iff (P \wedge Q) \vee (P \wedge R)$ (Lean);
2. $P \vee (Q \wedge R) \iff (P \vee Q) \wedge (P \vee R)$ (Lean).

Proof. The boring proof would be to check all eight cases. One can try and argue a bit more concisely – for example to prove the first result, one could argue that a necessary and sufficient condition for the left hand side $P \wedge (Q \vee R)$ to be true is that P is true, and either Q or R is true, so PQR must be $TT?$ or $T?T$, where T denotes `true` and $?$ denotes a value we don't care about. Similarly for the right hand side $(P \wedge Q) \vee (P \wedge R)$ to be true, either $P \wedge Q$ is true or $P \wedge R$ is true, so again PQR must be $TT?$ or $T?T$. The second result is just as easy. ∎

1.3.5 Appendix: we don't need all the symbols

Strictly speaking, we do not need $\implies$ at all in our mathematics of propositions. It is not hard to check that $P \implies Q$ is logically equivalent to $(\neg P) \vee Q$. Indeed, the only way that $P \implies Q$ can be false is when P is `true` and Q is `false`, and the only way that $(\neg P) \vee Q$ can be false is when both $\neg P$ and Q are `false`, which is equivalent to P being `true` and Q being `false`. Hence we can build $\implies$ from $\vee$ and $\neg$.

Similarly, $P \iff Q$ is defined to mean $(P \implies Q) \wedge (Q \implies P)$, so we can build $\iff$ from $\wedge$ and $\vee$ and $\neg$.

Finally, it is a consequence of de Morgan's laws that $P \wedge Q$ is logically equivalent to $\neg(\neg P \vee \neg Q)$, meaning that we can build $\wedge$ from $\neg$ and $\vee$. We conclude that given $\neg$ and $\vee$, we can build all of the other logical functions that we have been talking about in this section. So why do we use all these other functions? The main reason is that if we were to build everything using $\neg$ and $\vee$, then all our formulae would be much bigger, and much harder to understand. We have a good grasp of what $P \implies Q$ means, and $(\neg P) \vee Q$ just looks more complicated.

1.4 An application

1.4.1 The square root of 2 is irrational.

It is not at all obvious that there exists a positive real number whose square is 2. You will learn the techniques needed to prove this later on this year. However, it is easier to prove that there is no *rational* number whose square is 2, and we shall do this here, using some of the techniques introduced earlier.

Lemma 1.4.1

If n is an integer, then n is even if and only if n^2 is even (Lean).

More formally, if we fix an integer n, and let P be the proposition "n is even" and let Q be the proposition "n^2 is even", then the lemma claims that

$$P \iff Q$$

is true.

Proof. The definition of $P \iff Q$ is: $P \implies Q$, and $Q \implies P$. So we need to prove two things: firstly, that if n is even then n^2 is even, and secondly, that if n^2 is even, then n is even.

Proving that if n is even then n^2 is even: this is easy. We let $n = 2t$ with t an integer, and then $n^2 = 4t^2 = 2(2t^2)$ is twice an integer and hence even.

But how to prove that if n^2 is even, then n is even? We could start by writing $n^2 = 2u$ with u an integer. Then $n = \sqrt{2u}$. Now what? We will talk about this in the online videos. Hint: contrapositive. ■

We will now embark on our proof that there is no rational number whose square is two. Along the way, we will run into the concept of "proof by contradiction", which many of you have met before, and which we will explain during the proof.

Theorem 1.4.1

There is no rational number whose square is 2.

Hence, *if* there is a real number whose square is 2, it must be irrational. How do we know there is a real number whose square is 2, by the way?

Proof. We will prove this result by a technique known by mathematicians as *proof by contradiction*. The idea is simple. Let P be the statement that there is a rational number whose square is 2. We want to prove that $\neg P$ is true. But we saw that $\neg P$ is logically equivalent to $P \implies \texttt{false}$, so it will suffice to prove that $P \implies \texttt{false}$. We will do this by *assuming* (incorrectly!) that P is true, and deducing that a false statement is true.

Let me explain this in another way. If we are doing correct mathematics, then we will not be able to prove anything false. If we do correct mathematics *apart from in one step* where we make an assumption that we are not sure about, and we manage to prove something false, then we must have made an incorrect move somewhere. If everything else is correct, then the one step where we made the dubious assumption must be the incorrect step, and hence we can deduce that our assumption is false. Note in particular the importance of considering false propositions when doing mathematics.

So let's assume that there is a rational number whose square is 2. This is our dubious assumption! Let's press on and prove something false.

By changing the sign of the rational number if necessary, we can assume that it is positive (note that 0 doesn't work because $0^2 = 0$ which isn't 2). Write it as a fraction a/b, with a and b positive integers. Of course, if a and b are both even, then we can cancel a factor of two from both, so we may furthermore assume that at least one of a and b is odd. Our assumption $(a/b)^2 = 2$ easily simplifies to $a^2 = 2b^2$.

Summary so far:

- We made a *dubious assumption*, that there was a rational number whose square was 2.
- We have concluded that there are two positive integers a and b, at least one of which is odd, such that $a^2 = 2b^2$.

Now because $a^2 = 2b^2$, we deduce that a^2 is even. By Lemma 1.4.1 we can deduce that $a = 2t$ is even. Substituting in, we deduce $(2t)^2 = 2b^2$ and, expanding out and cancelling 2, that $2t^2 = b^2$. Hence b^2 is even, and hence again by Lemma 1.4.1, b is even. This means that both a and b are even. We also know that at least one of them is odd! We have proved a contradiction, so we must have made a mistake. The only possible mistake is the *dubious assumption*, which must hence be false.

We conclude that there is no rational number whose square is 2. ■

1.5 Sets

A set is a collection of stuff. The things in a set are called its elements. In this course, this is all you need to know about what a set "is". Mathematicians have got a completely precise list of all the ways you can make sets, and the basic axioms which sets satisfy, but we will not worry about these – take the third year course on logic and set theory if you want to know more.

We use these squiggly brackets $\{$ and $\}$ to make sets. For example if we write "Let $X = \{1, 2, 6\}$" then we mean that X is the set with three elements, the numbers 1, 2 and 6.

Sets can only have elements in once – the set $\{1, 1, 1\}$ is equal to the set $\{1\}$. Furthermore, the elements of sets don't come in any particular order: the set $\{1, 2, 6\}$ is equal to the set $\{1, 6, 2\}$. To give a more convoluted example, if $X = \{1, 1, 1, 2, 2, 2, 2, 3, 3, 3\}$ and $Y = \{3, 2, 1\}$, then $X = Y$.

We have special notation for the empty set $\{\}$ – it is written as $\emptyset$ or $\varnothing$. Historical note: this is not a Greek ϕ, it's the Norwegian/Danish letter. We also have special notation for some infinite sets of numbers:

- The *natural numbers* $\mathbb{N}$ are either $\{0, 1, 2, 3, \ldots\}$ or $\{1, 2, 3, \ldots\}$ – there are unfortunately two conventions. They exist in maths because of an axiom of maths, and induction works because of another axiom.
- The *integers* $\mathbb{Z}$ are $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$.
- The *rationals* $\mathbb{Q}$ are ratios a/b of integers with $b \neq 0$.
- The *reals* $\mathbb{R}$ are the limits of convergent sequences of rationals (NB it is hard work to make this precise; you will learn more about this later).
- The *complexes* $\mathbb{C}$ are numbers of the form $x + iy$ with x and y real.

If an element x is in a set X, then we write $x \in X$. This is pronounced "x is an element of X", or "x is in X". It is a proposition. If $\mathbb{N}$ denotes the *natural numbers*, then every mathematician knows that $3 \in \mathbb{N}$ is true, and that $-3 \in \mathbb{N}$ is false. Different mathematicians have different opinions about whether $0 \in \mathbb{N}$. For example in France this is usually believed to be true, and in the UK it is often believed to be false, at least in the maths department. On the other hand, computer scientists seem to be much more consistent about this matter, believing that $0 \in \mathbb{N}$ is true. In this course we will assume $0 \in \mathbb{N}$.

If $a \in X$ is false, so a is not an element of X, then we often write $a \notin X$.

Sets can be finite or infinite. The squiggly bracket notation above works great for finite sets, but for infinite sets it is more problematic. We can write things like $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, 3, \ldots\}$, but it's hard to imagine writing $\mathbb{R} = \{\pi, \sqrt{2}, 73, -29\frac{1}{2}, \ldots\}$ because what does that "..." there even *mean*?

There is a second usage of the squiggly bracket $\{,\}$ notation, to make subsets of sets. For example, the set of real numbers whose square is less than 7 can be written

$$\{x \in \mathbb{R} \mid x^2 < 7\}.$$

If the context makes it clear that we're only interested in real numbers, we sometimes write the shorter

$$\{x \mid x^2 < 7\}.$$

We'll say a bit more about this notation when we talk about subsets.

One very common mistake I see with students, is that they confuse the set $\{4\}$ with the element 4. A set is not *equal* to its elements, it *contains* its elements.

1.6 New notation for use with sets

Like we did mathematics with propositions, we can do mathematics with sets. We can define some new operators and look for relations between them.

1.6.1 Unions

We start with some notation you might be familiar with. If X and Y are sets, then their *union* $X \cup Y$ is the set of all the things which are either in X or in Y. Symbolically, we can write it like this:

$$X \cup Y = \{t \,|\, t \in X \vee t \in Y\}.$$

Notice that $\cup$ looks a bit similar to $\vee$.
For example, if $X = \{1, 2, 3\}$ and $Y = \{3, 4, 5\}$ then $X \cup Y = \{1, 2, 3, 4, 5\}$.

1.6.2 Intersections

The *intersection* $X \cap Y$ of X and Y is the set of all the elements which are in both X and Y. So we can write this:

$$X \cap Y = \{t \,|\, t \in X \wedge t \in Y\}.$$

Notice that $\cap$ looks a bit similar to $\wedge$.
For example, if $X = \{1, 2, 3\}$ and $Y = \{3, 4, 5\}$ then $X \cap Y = \{3\}$.

1.6.3 Set-theoretic difference.

If X and Y are sets, then the set $X \backslash Y$ is "what you get if you remove Y from X." Formally it is defined like this:

$$X \backslash Y = \{x \in X \,|\, x \notin Y\}.$$

For example, if $X = \{1, 2, 3\}$ and $Y = \{3, 4, 5\}$ then $X \backslash Y = \{1, 2\}$.

1.6.4 Subsets and inclusion

If X and Y are sets, and every element of X is an element of Y, then we say that X is a *subset* of Y, and we write $X \subseteq Y$. Note that some people write $X \subset Y$ for this, but I am not a big fan of this notation and I will stick to $X \subseteq Y$. For example, if $X = \{1, 2\}$ and $Y = \{1, 2, 3, 4, 5\}$ then $X \subseteq Y$.

If X is not a subset of Y, we sometimes write $X \not\subseteq Y$. For example $\{1, 2, 3\} \not\subseteq \{3, 4, 5\}$.

We saw the notation $\{x \in \mathbb{R} \,|\, x^2 < 7\}$ earlier. This generalises in the following way. Let X be any set, and for each $x \in X$, let $P(x)$ be a proposition. For example if X is the real numbers then $P(x)$ could be the proposition $x^2 < 7$. Then we can consider the subset of X consisting of the x such that $P(x)$ is true, and we can write it

$$\{x \in X \,|\, P(x)\}.$$

Every subset of X is of this form. For if Y is any subset of X, we can define $P(x)$ for $x \in X$ to be the proposition $x \in Y$, and then $Y = \{x \in X : P(x)\}$ (because they have the same elements). In this way, we can identify the subsets of X with the maps $P : X \to \{\texttt{true}, \texttt{false}\}$.

There are certain subsets of the real numbers of this form which come up so often that they've got their own specialised notation. For example the *open interval* (a, b) is defined to be $\{x \in \mathbb{R} : a < x < b\}$, the *closed interval* $[a, b]$ is defined to be $\{x \in \mathbb{R}, \, , a \leq x \leq b\}$. What do you think $(a, b]$ means? Note to French people: I know you learnt $]a, b]$ at school for this set of reals, but this notation is never used in the UK and it confuses British markers who've never studied in France, so please try and avoid it if you can.

1.6.5 Interlude: a little philosophy

This little subsection is non-examinable, and can be omitted by students uninterested in foundational matters.

You can make sets from anything. For example you can make sets of sets. If $X = \{1, 2, 3\}$ and $Y = \{3, 4, 5\}$ you can define a new set $Z = \{X, Y\}$, and then you could make a new set $W = \{\pi, X, Z\}$ and so on. One of the foundations of mathematics is called "ZFC set theory" and in this way of setting up mathematics, *every mathematical object is a set*. This is a bit ridiculous if you think about it, because it means that π is a set, which is not really how we think about π (what are its elements? If you know something about the foundations of mathematics in set theory, you'll realise that the answer to this question depends on whether π is a Dedekind cut or an equivalence class of Cauchy sequences...), and it means that $+$ is a set, and so on. Not only that, but sets like W above are not really useful mathematical objects; one element of W is a number, one is a set of numbers, and one is a set of sets of numbers.

Another way to set up the foundations of mathematics is to use something called "dependent type theory". In type theory, you can *only* make sets of elements which are "all of the same type". For example, the real numbers are a type in dependent type theory, so you can make sets of real numbers. Sets of real numbers are a type too, so you can make sets of sets of real numbers. But you can't make a set like W above, because π and X and Z all have different types. Turns out it's OK, because you can use types to make any mathematical structure you need (and we don't need W).

In the next section we will talk about subsets of a fixed "universe" set Ω (note: in the videos this set I use the notation α). In practice, every set you run into will be a subset of a sensible "universe". For example the set $X = \{1, 2, 3\}$ is a subset of the real numbers $\mathbb{R}$.

1.6.6 Complements

Let Ω be a set (for example, the real numbers). In this subsection we will only consider subsets of Ω.

If X and Y are subsets of Ω, then it is easy to see that $X \cup Y$ and $X \cap Y$ are also subsets of Ω. But there are other complements of Ω that we can make. For example, the *complement* $\overline{X}$ of X is all the elements of Ω which are *not* in X. Symbolically we can write this:

$$\overline{X} = \{t \in \Omega \mid t \notin X\}.$$

Example: if Ω is the integers, and X is the even numbers, then $\overline{X}$ is the odd numbers – all the integers which are not even.

Example: if $\Omega = \{1, 2, 3, 4, 5\}$, and $X = \{1, 2, 3\}$, then $\overline{X} = \{4, 5\}$.

More worryingly, if Ω is $\{1, 2, 3, 4\}$ is different but $X = \{1, 2, 3\}$ stays the same, then $\overline{X}$ is now $\{4\}$. This means that the notation $\overline{X}$ is rather dubious notation, because $\overline{X}$ *depends on* Ω but Ω is not mentioned in the notation. However, usually this does not matter, because when we are using complements, Ω is almost always a fixed space. For example, Ω can often be a probability space – a space of possibilities, for example $\Omega = \{\text{heads}, \text{tails}\}$. If $X \subseteq \Omega$ is a subset, for example $X = \{\text{heads}\}$ then $\overline{X}$ is the elements of Ω which are not in X, so in this case $\overline{X} = \{\text{tails}\}$. If we were in this probability situation, and if p is the probability that a random sample from Ω lands in X, then the chances that it lands in $\overline{X}$ would be $1 - p$.

1.7 Sets and propositions.

Let Ω be a fixed "universe" (for example, the real numbers). Let's say X and Y are subsets of Ω. If t is an arbitrary element of Ω, then $t \in X$ and $t \in Y$ are two propositions, which may be either true or false. Let's consider whether $t \in X \cap Y$ is true or false:

$t \in X$	$t \in Y$	$t \in X \cap Y$
false	false	false
false	true	false
true	false	false
true	true	true

Does this remind you of anything? It is exactly the truth table for $\wedge$; this is because the formal definition of $X \cap Y$ involves $\wedge$.

Exercise: draw a similar table for $X \cup Y$. What does it correspond to?

Let's now do the same for complements. Say $X \subseteq \Omega$ is a set, and $t \in \Omega$ is an element of Ω.

$t \in X$	$t \in \overline{X}$
false	true
true	false

What does this remind you of? Note that $\overline{X} = \{t \in \Omega \mid \neg(t \in X)\}$.

Recall a fundamental property of sets: two sets are equal *if and only if* they have the same elements. This has a fancy name – "extensionality of sets". Note that a whole bunch of basic concepts in mathematics have fancy names (we have seen symmetry, transitivity, distributivity and now this). Don't worry too much about these words – you will pick them up in time. We will use set extensionality in the proof of this next lemma.

Lemma 1.7.1

Let Ω be a fixed set, and say $X \subseteq \Omega$ and $Y \subseteq \Omega$ are subsets.

1. $\overline{X \cup Y} = \overline{X} \cap \overline{Y}$;
2. $\overline{X \cap Y} = \overline{X} \cup \overline{Y}$.

What does this lemma remind you of?

Proof. First we prove Part (1). To prove that two sets are equal, we need to check that they have the same elements. So let t be an arbitrary element of Ω. We need to check that $t \in \overline{X \cup Y} \iff t \in (\overline{X} \cap \overline{Y})$. Now $t \in X$ is a proposition, so it's either true or false. Similarly $t \in Y$ is a proposition, so it's either true or false. Let's draw a table of the possibilities.

$t \in X$	$t \in Y$	$y \in X \cup Y$	$t \in \overline{X \cup Y}$	$t \in \overline{X}$	$t \in \overline{Y}$	$t \in \overline{X} \cap \overline{Y}$	$t \in \overline{X \cup Y} \iff t \in \overline{X} \cap \overline{Y}$
false	false	false	true	true	true	true	true
false	true	true	false	true	false	false	true
true	false	true	false	false	true	false	true
true	true	true	false	false	false	false	true

The last column is always true, so the sets $\overline{X \cup Y}$ and $\overline{X} \cap \overline{Y}$ have the same elements, which means that they are equal.

Where have we seen pretty much exactly this table before. What is going on? Can we somehow use previous work instead of redoing it? ■

1.7.1 "For all" and "there exists".

We have already seen the idea of a "family" of propositions indexed by a set. For example if x is a real number, then the proposition $x^2 < 7$ should really be thought of as a proposition $P(x)$ because its truth value depends on x. We can think of P as a map $\mathbb{R} \to \{\texttt{true}, \texttt{false}\}$. We can check $P(0)$ is true and $P(1000)$ is false.

But sometimes $P(x)$ is always true. For example if $P(x)$ is the proposition $(x+1)^2 = x^2 + 2x + 1$, then $P(x)$ is always true. We can write this as follows:

$$\forall x \in \mathbb{R}, (x+1)^2 = x^2 + 2x + 1.$$

That upside-down A is pronounced "for all" (and I just wrote `\forall` in this LaTeX document to get it).

Now clearly

$$\forall n \in \mathbb{Z}, n^2 = 9$$

is *not* true. But how do we disprove it? We have to give an explicit counterexample. We can disprove it by setting $n = 2$, and observing that 2^2 isn't 9.

In general, if X is a set, and $P(x)$ is a function which takes an input an element of X and outputs a proposition, then to prove that

$$\forall x \in X, P(x)$$

is true, we have to prove $P(x)$ for *every single element of* X. To prove it is false, all we have to do is to find *at least one* element of X for which $P(x)$ fails. This brings us on to $\exists$, a companion to $\forall$.

We know that there's an integer whose square is 16. In mathematics we can write this as follows:

$$\exists x \in \mathbb{Z}, x^2 = 16.$$

That upside-down E means "there exists". This statement, "$\exists x \in \mathbb{Z}, x^2 = 16$", is a proposition, and it's true. Note that in fact there are *two* integers whose square is 16, but this is OK. $\exists x \in X, P(x)$ mean that there is *at least one* element x of X such that $P(x)$ is true.

Now say X is a set and for each $x \in X$ we have a proposition $P(x)$. Let's consider the proposition "$\exists x \in X, P(x)$". What is the logical negation of this proposition? Well, to prove it, we have to find $x \in X$ such that $P(x)$ is true. So to disprove it, we must show that for all $x \in X$, $P(x)$ is false.

In summary then,

$$\neg(\forall x \in X, P(x)) \iff \exists x \in X, \neg(P(x))$$

and

$$\neg(\exists x \in X, P(x)) \iff \forall x \in X, \neg(P(x)).$$

1.7.2 Infinite unions and intersections

Using $\forall$ and $\exists$, we can beef up our definition of $\cup$ and $\cap$ so that they apply not just to two sets, but to an arbitrary collection of sets.

If we have infinitely many subsets X_0, X_1, X_2, X_3, ... of Ω, we can define their union

$$\bigcup_{n=0}^{\infty} X_n$$

to be

$$\{x \in \Omega \mid \exists n \in \mathbb{N}, x \in X_n\}$$

(here we are assuming $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$). And we can define their intersection

$$\bigcap_{n=0}^{\infty} X_n$$

to be

$$\{x \in \Omega \,|\, \forall n \in \mathbb{N}, x \in X_n\}.$$

This is not just specific to the natural numbers. If I is any "index set" and we have a subset X_i of Ω for each $i \in I$ then we can define $\bigcup_{i\in I} X_i$ and $\bigcap_{i\in I} X_i$. Try writing down their definition formally. Note the following issue: if I is the empty set, then $\bigcap_{i\in I} X_i$ is Ω (perhaps surprisingly), because for every $x \in \Omega$, the proposition $\forall i \in I, x \in X_i$ is always true as I is empty. What? The empty set can be confusing! Let's have a section on it.

1.7.3 The empty set.

This is trickier than you might think.

Let P be a proposition. Let X be the empty set.

Is $\forall x \in X, P$ true or false? Is $\exists x \in X, P$ true or false? Or do these questions depend on whether P is true or false? What do you think?

These questions can be answered using "Buzzard's rule of thumb". Let P be any proposition at all. If $\exists x \in X, P$ were true, then this would imply that there was some $t \in X$ such that P is true and in particular it would imply the existence of some $t \in X$. But X is the empty set, so no such t can exist. Hence, whatever the truth value of P, $\exists x \in \varnothing, P$ is *always false*. Here $\varnothing$ is the standard notation for the empty set.

What about $\forall x \in X, P$? Well, let's show that this must always be true. Indeed, its logical negation is $\exists x \in X, \neg P$, and this is definitely false by argument in the previous paragraph. Hence $\forall x \in X, P$.

I know that some people find this difficult, so here's another way to think about it. Let's say X has three elements. Say that I ask you to prove that $\forall x \in X, f(x) = g(x)$ is true, where here f and g are two functions (you can think of them as functions from $\mathbb{R}$ to $\mathbb{R}$, and think of X as a set of three real numbers). Because X only has three elements, you just have to check three cases. So you have three little problems to solve, and when you've solved them all, you've proved $\forall x \in X, f(x) = g(x)$. OK now say instead that X has 59 elements. Then you have 59 little puzzles to solve and once you've solved them all, you're done and your proof is complete. Now say that X is the empty set, i.e. the set with no elements. Then your job list has 0 things on it, and so you've done all your jobs already, so you're done and the proof is complete and you didn't even have to do any work.

Now say instead that I ask you to prove that $\forall x \in X, f(x) = g(x)$ is false. If X has three elements then you might want to start by checking all three cases and trying to find an example of some $t \in X$ where $f(t) \neq g(t)$. If X has 59 elements then you are more likely to succeed because there are 59 chances of the equality going wrong and you only need to find one to disprove it. If f and g are random functions, then the bigger X is, the easier it will be to disprove $\forall x \in X, f(x) = g(x)$. If X only has one element then it's 50-50: if t is that element, you have to hope $f(t) \neq g(t)$ if it's your job to disprove $\forall x \in X, f(x) = g(x)$. But if X has no elements then all hope is lost, you definitely won't find that t, so there's no way you can disprove $\forall x \in \varnothing, f(x) = g(x)$.

Chapter 2

Functions and Equivalence Relations

2.1 Functions

You have studied functions $f : \mathbb{R} \to \mathbb{R}$ before – things like $f(x) = \sin(x)$ or $f(x) = 1/x$ or whatever.

But there is a more general and abstract theory of functions between general sets. Let X and Y be sets. A function $f : X \to Y$ is a device which, when you give it an element $x \in X$, is guaranteed to return an element $f(x) \in Y$. Notation: we say X is the *domain* of the function f, and Y is the *codomain*.

It is important that we get some things straight about functions before we go any further.

- The value of $f(x)$ cannot change over time. If we evaluate $f(x)$ twice in a calculation, we will get the same answer twice. Random variables, which can change over time, are not functions in this sense (they can be modelled in mathematics, but we don't use functions).
- $f(x)$ is always *exactly one element of* Y. It cannot be "undefined". It cannot be a vague concept such as $\pm 2i$ (which is a shorthand for one of two different numbers). If you want "square root" to be a function you are going to have to be careful to define exactly which square root you are talking about.

One last basic definition before we move on to examples: it does *not* have to be the case that every element of Y is "used" by f. For example, $\sin$ is a function from $\mathbb{R}$ to $\mathbb{R}$, even though 7 is a real number and there is certainly no real number x such that $\sin(x) = 7$. The *subset* of Y which is actually the stuff that f can spit out, is called the *range* of the function by some people, and called the *image* of the function by some other people (another notational nightmare!). We'll call it the range. Symbolically, the range of f is

$$\{y \in Y \mid \exists x \in X, f(x) = y\}.$$

For example, the range of the $\sin$ function is $[-1, 1]$ – these are the real numbers that can be the sine of something.

2.2 Examples and non-examples.

The function $f(x) = x^2 + 3x + \pi$ is a function from $\mathbb{R}$ to $\mathbb{R}$, and it is also a function from $\mathbb{C}$ to $\mathbb{C}$. It is not a function from $\mathbb{Z}$ to $\mathbb{Z}$ though, because $f(0) = \pi$ which is not an integer.

The "function" $f(x) = 1/x$ is *not* a function from $\mathbb{R}$ to $\mathbb{R}$. This is because it is undefined at zero. There are two ways of fixing this. We could set $X = \mathbb{R}\backslash\{0\}$, that is, the reals

with 0 removed. Then $f(x) = 1/x$ is a function from X to $\mathbb{R}$. Alternatively, we could set $Y = \mathbb{R} \cup \{\infty\}$, that is, we can add a new element to the reals and just call it infinity (this is one of the joys of abstract mathematics – we don't have to philosophise about infinity, we can just get a new set, call it infinity, and add it to the reals, and *define* $1/0 = \infty$). Now $f(x) = 1/x$ is a function from $\mathbb{R}$ to Y.

Let X be the set $\{1, 2, 3\}$ and let Y be the set $\{3, 4, 5\}$. We can define a function $f : X \to Y$ by $f(1) = 4$, $f(2) = 3$ and $f(3) = 4$. For every input element $x \in X$ there is a unique output element $f(x) \in Y$, so f is a function with domain X and codomain Y. *Important note*: a function does *not* have to be "defined by a rule" – it can be completely random.

The Conway Base 13 function is a really cool function – that should be a working link to its Wikipedia page. It is a function f from the reals to the reals with the property that if $a < b$ are two real numbers (imagine them as being really close together) and y is any real number, then there will exist some x such that $a < x < b$ and $f(x) = y$. How might you go about defining that function? John Conway sadly died of COVID-19 in early 2020.

Let's consider the squaring function from $\mathbb{R}$ to $\mathbb{R}$, that is, the function sending x to x^2. Mathematicians have got a cool piece of notation which enables them to define this function without ever giving it a name like f: they write it $x \mapsto x^2$. Computer scientists have also got a cool piece of notation for this function, and unfortunately it is completely different: it is $\lambda\ x, x^2$. You will never see this notation in the mathematics literature, but computer scientists use it all the time, and I think that mathematicians should at least be aware of it. These notations are called "anonymous function notation" because they can be used to define functions without naming them.

Let X be any set at all. Then there is a function $X \to X$ called the *identity function*, defined as follows: if $x \in X$ then the identity function sends x to x. In other words, it's the function $x \mapsto x$.

Let X be the empty set, and let Y be any set at all. I claim that there is a function from X to Y, namely the "empty function"! What is the job of a function? A function from X to Y only has to do the following task: if you give it an element of X, it guarantees that it will return an element of Y. So the empty function can just sit there proudly, announcing that it is a function from the empty set to Y, and saying that if you ever give it an element of the empty set, it will guarantee that it will give you an element of Y. It can say this confidently, because there is no element of the empty set that you can ever give it.

2.2.1 Function composition.

Say A, B and C are sets, and we have a function $f : A \to B$ and a function $g : B \to C$. We can use f and g to build a new function from A to C. Given an element $a \in A$, we first apply f to get an element $f(a) \in B$, and then we apply g to get an element $g(f(a)) \in C$. Because g comes before f in this formula, we write this new function from A to C as $g \circ f$. That little circle is pronounced "composed with", and the process of constructing this new function is called *composition of functions*.

It's a bit annoying that we build this composite function by first applying f and then applying g. In a parallel universe, mathematicians decided that function notation was $(x)f$ instead of $f(x)$. In their universe, the composite was $((x)f)g$ and they write it $f \circ g$. But we live in this universe so we're stuck with $g \circ f$ meaning "first apply f, then g".

2.2.2 Equality of functions

If we have two functions $f_1 : X \to Y$ and $f_2 : X \to Y$, then as mathematicians we say that these functions are *equal* if, for all $x \in X$, we have $f_1(x) = f_2(x)$. This seems like

a really obvious definition of equality: two functions are equal if and only if they take the same values everywhere.

A non-examinable observation: this point of view is called "functional extensionality". A computer scientist might consider a function to be an *algorithm* which solves a problem, for example they might consider a sorting algorithm which sorts a finite list of natural numbers into order as a function. For them, two different sorting algorithms might be very different – one of them might be far more efficient than the other, for example, or it might fit into a smaller amount of memory. But to a mathematician, both algorithms return the same answer, so they are the same function; mathematicians usually don't care about the "algorithm" defining a function, or even whether such an algorithm exists!

2.3 Injectivity, surjectivity and bijectivity

The most important true/false statements that we associate to a function $f : X \to Y$ are the concepts of it being injective, surjective, and bijective. Note: at your school, "injective" might have been called "one to one", and "surjective" might have been called "onto". Let's define these concepts.

Definition 2.3.1

- A function $f : X \to Y$ is called *injective* if distinct elements of X get mapped to distinct elements of Y. More formally, f is injective if $\forall a, b \in X, f(a) = f(b) \implies a = b$.
- A function $f : X \to Y$ is called *surjective* if every element of Y gets "hit" by f. More formally, f is surjective if $\forall y \in Y,\ \exists x \in X,\ f(x) = y$.
- A function $f : X \to Y$ is called *bijective* if it is both injective and surjective.

These concepts are *much* easier to understand if you look at some pictures of functions. However, if you are asked to give such definitions in a test or exam, *don't* draw pictures and *don't* write those slightly ambiguous English sentences above – *write the formal definition*. It's short, simple, and gets you full marks immediately.

2.3.1 Examples

Let's take a look at a non-example first. Let's consider the squaring function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$. This function is not injective. Indeed, the distinct real numbers $a = -3$ and $b = +3$ satisfy $a \neq b$ but $f(a) = 9 = f(b)$. This function is not surjective either – because $-1 \in \mathbb{R}$ is not equal to $f(x)$ for any $x \in \mathbb{R}$, as the square of a real number is always non-negative. Hence f is not bijective.

If X is any set, then the identity map $i : X \to X$ defined by $i(x) = x$, is bijective. Let's check this carefully.

Firstly, let's check injectivity. Say $a, b \in X$ and $i(a) = i(b)$. Then by definition of i we deduce that $a = b$, which is what we wanted to prove.

Next, we check surjectivity. Say $y \in X$ is arbitrary. We need to find $x \in X$ such that $i(x) = y$. This is easy – we just define $x = y$, and then $i(x) = i(y) = y$.

Hence i is bijective, because it is injective and surjective.

2.3.2 Bijectivity and composition.

Say X, Y and Z are sets, and $f : X \to Y$ and $g : Y \to Z$ are functions. Recall that in this situation we can define the composition $h = g \circ f$, a function from X to Z.

What are the relations between injectivity/surjectivity of f, injectivity/surjectivity of g and injectivity/surjectivity of h? We explore some of these here.

Theorem 2.3.1

Let $f : X \to Y$ and $g : Y \to Z$ be functions, and say $g \circ f : X \to Z$ be the composition of f and g.

(a) If f and g are both injective, then so is $g \circ f$.

(b) If f and g are both surjective, then so is $g \circ f$.

(c) If f and g are both bijective, then so is $g \circ f$.

Proof. (a) What are we trying to prove here? What is the question? We are trying to prove that a certain function $g \circ f$, with domain X and codomain Z, is injective. What does this *mean*? It means that if p and q are any elements of X, and if $(g \circ f)(p) = (g \circ f)(q)$, then $p = q$. That's the question. Unfolding the definition of the function $g \circ f$, our goal is to prove that if $g(f(p)) = g(f(q))$, then $p = q$. So let's assume that $g(f(p)) = g(f(q))$, and deduce that $p = q$.

What do we know? What haven't we used yet? We know that f is injective and that g is injective. Set $a = f(p)$ and $b = f(q)$. We know that $g(a) = g(b)$. By injectivity of g, we can deduce that $a = b$, that is, that $f(p) = f(q)$. And now, by injectivity of f, we can deduce that $p = q$, which is what we wanted.

(b) Here we are trying to prove that $g \circ f : X \to Z$ is surjective. So let $z \in Z$ be an arbitrary element of the codomain. We want to find some element $x \in X$ such that $(g \circ f)(x) = z$.

What haven't we used yet? We know that f and g are surjective. Now $g : Y \to Z$, and g is surjective, so there must exist some $y \in Y$ with $g(y) = z$. And now f is surjective and $y \in Y$, so there must exist some $x \in X$ with $f(x) = y$. Now what is $(g \circ f)(x)$? It is $g(f(x))$, which equals $g(y)$, which equals z. So we are done!

(c) If f and g are bijective, then by definition of "bijective" they are injective and surjective. So by the first two parts, $g \circ f$ is also injective and surjective. Hence $g \circ f$ is bijective, and we are done.

■

Exercise. If $f : X \to Y$ is a function, then the true/false statement that f is injective formally means $\forall p \in X, \forall q \in X, f(p) = f(q) \implies p = q$. So the true/false statement that f is *not* injective must be the logical negation of that. What is the logical negation of that statement? Can you write the logical negation of $f(p) = f(q) \implies p = q$ in a more useful way, for example in the form $P \wedge Q$? Now do the same exercise for a surjective function.

2.4 Inverses

What is the "inverse" of a function? This is a topic which can be treated quite sloppily in schools. We know about the $\sin$ function from $\mathbb{R}$ to $\mathbb{R}$. And this function has an inverse, right? The $\sin^{-1}$ function, inverse sine. We know $\sin(\pi/6) = 1/2$, and so $\sin^{-1}(1/2) = \pi/6$. Happy so far?

Is it true that if x is any real number, then $\sin^{-1}(\sin(x)) = x$? It certainly looks like it should be true doesn't it. But this equation *cannot be true*, because look at this:

$$\begin{aligned} 2\pi &= \sin^{-1}(\sin(2\pi)) \\ &= \sin^{-1}(0) \\ &= \sin^{-1}(\sin(0)) \\ &= 0 \end{aligned}$$

We have proved that $2\pi = 0$, which is definitely not true. The problem here is that $\sin$ is *not an injective function*. If x and y are distinct real numbers, and $\sin(x) = \sin(y)$ (which can definitely happen for $\sin$ – and this is *exactly* what it means for $\sin$ to be not injective, by the exercise above), then whatever this so-called "function" $\sin^{-1}$ does, it can't send $\sin(x)$ to x *and* $\sin(y)$ to y, because a function cannot eat the same value twice and spit out two different answers.

What does it *mean* for a function to have an inverse? Here is a very strong definition, it's the most powerful kind of inverse that one can hope for.

Definition 2.4.1

Say $f : X \to Y$ is a function. We say that a function $g : Y \to X$ is a *two-sided inverse of* f, if both of the following are true:

1. For all $x \in X$, $g(f(x)) = x$; *and*
2. For all $y \in Y$, $f(g(y)) = y$.

In particular, $\sin^{-1}$ is not a two-sided inverse for $\sin : \mathbb{R} \to \mathbb{R}$, because $\sin^{-1}(\sin(x))$ is not always equal to x, if x is allowed to be any real number.

Example

> Say $X = Y = \mathbb{R}$ and $f : \mathbb{R} \to \mathbb{R}$ is defined as $f(x) = \frac{3x+10}{7}$. What is a two-sided inverse for this function? If $y = \frac{3x+10}{7}$, how do we recover x? Clearly $x = \frac{7y-10}{3}$. So let's define $g : \mathbb{R} \to \mathbb{R}$ by $g(y) = \frac{7y-10}{3}$. It is now an easy school-level exercise to check that $g(f(x)) = x$ for any real number x, and $f(g(y)) = y$ for any real number y. Hence g is a two-sided inverse for f.

Exercise: check the previous example – substitute in and check just to make sure. Can you do it in your head? Remember that you have to do *two* calculations here, because g is a two-sided inverse, which is two conditions.

Example

> Say $X = Y$. Then the identity function $i : X \to X$ defined by $i(x) = x$ has a two-sided inverse, namely the identity function again! Because if $x \in X$ is arbitrary, then $i(i(x)) = x$.

We have seen above that the identity function is a bijection, and also that it has a two-sided inverse. This is not a coincidence. The final theorem of this section is a classification of exactly which functions have two-sided inverses.

Theorem 2.4.1

A function $f : X \to Y$ has a two-sided inverse *if and only if* it is a bijection.

Proof. This theorem is claiming an " $\iff$ ", so we really have to prove two theorems, an $\implies$ and an $\impliedby$. We do these separately.

Proof of $\implies$:

Our assumption is that $f : X \to Y$ has a two-sided inverse. Let's call it $g : Y \to X$. Our goal is to prove that f is a bijection. Our proof now has to bifurcate again! A bijection is an injection and a surjection, so we have to prove that f is injective, and also that f is surjective.

Proof of injectivity of f:

Say $p \in X$ and $q \in X$, and $f(p) = f(q)$ (recall that these are elements of Y). Our goal is to prove that $p = q$. Well, applying the inverse function g, we deduce that $g(f(p)) = g(f(q))$. But by definition of two-sided inverse, $g(f(p)) = p$ and $g(f(q)) = q$. Hence $p = q$.

Proof of surjectivity of f:

Say $y \in Y$. We need to find some $x \in X$ with $f(x) = y$. However are we going to come up with a random element of X like that? Well, we have a function $g : Y \to X$, so let's define $x = g(y)$ and hope that it works, basically because I can't think of any other way of getting an element of X!

And indeed the check is easy. We compute that $f(x) = f(g(y))$, which equals y by definition of a two-sided inverse. So we are finished with the proof of the $\implies$ direction.

Proof of $\impliedby$:

For this part of the proof, we have a bijection $f : X \to Y$, and we want to prove that it has a two-sided inverse $g : Y \to X$. How can we define this inverse function?

Say $y \in Y$. We would like to define an element $g(y) \in X$. Here's how to do it. We know that f is surjective. So, by definition of surjectivity, there exists *at least one* element $x \in X$ such that $f(x) = y$. But can more than one such element exist? Say $x_1 \in X$ and $x_2 \in X$, and both $f(x_1)$ and $f(x_2)$ are equal to Y. Then in particular $f(x_1) = f(x_2)$. By injectivity of f, this means that $x_1 = x_2$! In particular, there is *at most one* element $x \in X$ such that $f(x) = y$. We deduce that there is *exactly one* element $x \in X$ with $f(x) = y$, and we define $g(y)$ to be that element.

That's the definition of g. Now we have to prove that g satisfies the definition of being a two-sided inverse. Let's go.

Firstly, say $y \in Y$ is arbitrary. We know by definition of $g(y)$ that $g(y)$ is an element x such that $f(x) = y$. In other words, we know $f(g(y)) = y$. That's half the problem solved already.

Finally, say $x \in X$ is arbitrary, and let's define $y = f(x)$. Now we know that $g(y)$ is the *only* element of X with $f(g(y)) = y$, and we know that $f(x) = y$, so x must be this element. We deduce that $g(y)$ must equal x, and hence $g(f(x)) = x$. We are done!

■

2.5 Binary relations.

We all know what $2 + 2 = 4$ means and what $2 < 3$ means. And we all know that 2 and 3 and 4 are numbers. In this course I want to make you think a bit more carefully about what $+$ and $<$ and $=$ are. I think that these things are functions. But they're different kinds of functions.

$+$ is a function from $\mathbb{R}^2$ to $\mathbb{R}$. An element of $\mathbb{R}^2$ is a pair (x, y) of real numbers, and when we apply the $+$ function we get a new number which should be called $+(x, y)$ (like $f(x, y)$ when f is a general function), except historically we use a different more standard notation: $x + y$.

I think that $<$ is a function too, but its codomain is different. Like $+$, the function $<$ also accepts as input a pair (x, y), but its output is the *proposition* $x < y$. So $<$ is a function from $\mathbb{R}^2$ to `Prop`, where `Prop` is the set of all propositions. If I feed an explicit element of

$\mathbb{R}^2$ like $(3,3)$ into the $<$ function, I get out an explicit proposition $3<3$ (in this case, a false one) as the output. If you want to identify a proposition with its truth value, then you could think of $<$ as a function from $\mathbb{R}^2$ to $\{\texttt{true}, \texttt{false}\}$.

Similarly $=$ is a function, also with domain $\mathbb{R}^2$ and codomain `Prop`, because $3=3$ is a proposition (in this case, a true one).

Now let X be a general set. We can construct a new set X^2 – this is just a set whose elements are pairs (a,b) with $a \in X$ and $b \in X$.

Definition 2.5.1

A *binary relation* R on a set X is a function $R : X^2 \to \texttt{Prop}$.

In other words, a binary relation on X is, for each pair of elements $x, y \in X$, a true-false statement $R(x,y)$.

Note: one can allow more general binary relations where x is in a set X and y is in a possibly different set Y (so then a binary relation is a map $X \times Y \to \texttt{Prop}$), but we do not consider those more general relations here.

Examples.

1. If X is the positive integers, we could define a binary relation R on X by saying that the positive integers a and b are related if and only if a and b end in the same digit if you write them out in base 10. So $R(574, 14)$ would be true, and $R(475, 14)$ would be false.

2. if X is any set at all, then $=$ is a binary relation on X. If x, y are two elements of X, then $x = y$ is a true/false statement.

3. $<$ and $\leq$ and $>$ and $\geq$ are all binary relations on $\mathbb{R}$.

4. If X is a set of red, yellow, blue and green triangles and squares, then one could define two relations on X: for $a, b \in X$ we could define $C(a,b)$ to be the proposition "a and b have the same colour", and we could define $S(a,b)$ to mean "a and b have the same shape".

5. if X is the set of all subsets of the real numbers, then $\subseteq$ is a binary relation on X; if A and B are two sets of real numbers, then $A \subseteq B$ is a true/false statement, so $\subseteq$ takes two sets in and outputs a true/false statement and it is hence a binary relation on X (note that it is *not* a binary relation on $\mathbb{R}$, a real number is not equal to a set of real numbers).

6. $\wedge$ and $\vee$ and $\implies$ and $\iff$ are all binary relations on `Prop`. If P and Q are two propositions, then so are $P \wedge Q$, $P \implies Q$ etc.

7. If x, y are real numbers, then $x^2 + y^2 = 1$ is a proposition. So $x^2 + y^2 = 1$ is a binary relation on the real numbers.

We can draw the graph of that last binary relation. The graph of $x^2 + y^2 = 1$ is a circle. The graph of the binary relation $=$, i.e., the graph of $x = y$, is a straight line. What do you think the graph of $<$ looks like? In other words, what is the graph of $x < y$?

This is a bit strange, because usually as students you are asked to draw the graph of a function $f : \mathbb{R} \to \mathbb{R}$. But really when you are asked to "draw a graph of f", you are being asked to draw a graph of the binary relation on $\mathbb{R}$ which sends the pair (x,y) to the binary relation $y = f(x)$. Every binary relation on $\mathbb{R}$ has a graph, you just draw the subset of $\mathbb{R}^2$ consisting of points (x,y) where the binary relation is true.

2.6 Common predicates on binary relations.

A "predicate" is just a property – a proposition attached to each element of a set. For example, injectivity is a predicate on functions. We talked about functions earlier on, but the truth of the matter is that there's not much you can say about a general function – you can compose them, and that's about it. Things got more fun when we started talking about functions which satisfied some predicates, such as injective or surjective or bijective functions.

The same is true with relations. There are certain properties which a binary relation might or might not have, and if they are satisfied then we can say more. Here are some predicates on binary relations. Before we start with definitions, let's think about some famous binary relations, and try and isolate some important properties which they have.

The binary relation $\leq$ on the real numbers has the property that $\forall x \in \mathbb{R}, x \leq x$. Remember that $\leq$ is a function which eats two real numbers, and what we're saying here is that if you give it the same real number twice, it always returns a true proposition. Similarly the binary relation $\subseteq$ on sets satisfies $A \subseteq A$ for all sets A. Equality also has that property – if $x \in X$, then certainly $x = x$. We now formalise this concept.

Definition 2.6.1

A binary relation R on a set X is called *reflexive* if $\forall x \in X, R(x, x)$. In other words, if $R(x, x)$ is true for every x then R is reflexive.

Say X is a set of plastic squares and triangles, and for two plastic shapes a and b, $C(a, b)$ is the relation "a is the same colour as b". One thing we see immediately is that if a is the same colour as b, then b is the same colour as a! So $C(a, b) \implies C(b, a)$. This is not true for all binary relations though – for example if $a < b$ is true, then $b < a$ is *never* true, and if $a \leq b$ is true, then $b \leq a$ is only true when $a = b$.

Definition 2.6.2

Let R be a binary relation on a set X.

1. R is called *symmetric* if $\forall a, b \in X$, $R(a, b) \implies R(b, a)$.
2. R is called *antisymmetric* if $\forall a, b \in X$, $R(a, b) \wedge R(b, a) \implies a = b$.

So the binary relation C above is symmetric, and $\leq$ is antisymmetric. Similarly, the binary relation $\subseteq$ on the set of subsets of the real numbers is antisymmetric. Important note: if A and B are two random subsets of the real numbers, then usually $A \subseteq B$ and $B \subseteq A$ are both false! Antisymmetry does *not* say "if A is not related to B, then B is related to A". It says something a little more subtle than that.

Exercise: Is $=$ symmetric? Is it antisymmetric? Is $<$ antisymmetric?

Another important property, which $\implies$ and $\leq$ and $\geq$ and $<$ and $>$ and $=$ all have (but which $\vee$ doesn't have – try it!), is *transitivity*.

Definition 2.6.3

A binary relation R on a set X is called *transitive* if for every $a, b, c \in X$, we have $(R(a, b) \wedge R(b, c)) \implies R(a, c)$.

For example, $\leq$ is transitive on $\mathbb{R}$ because $a \leq b$ and $b \leq c$ implies $a \leq c$. Similarly $\subseteq$ is transitive on the set of subsets of Ω, because if $A \subseteq B$ and $B \subseteq C$ then $A \subseteq C$.

Many binary relations are reflexive and transitive. Even relations which feel quite different, such as "I am the same shape as you" (expressing a similarity) and "I am less than or equal to you" (expressing some kind of order or hierarchy) are reflexive and transitive. Symmetry and antisymmetry are where things start to differ. These concepts pull in different directions. If R is a binary relation which is symmetric *and* antisymmetric, and if $R(a, b)$ is true, then by symmetry $R(b, a)$ is true, and then by antisymmetry we deduce $a = b$, and hence if two objects are related they must be equal, meaning that the relation is not very mathematically rich.

Many binary relations which you see in practice are either symmetric or antisymmetric, and we shall see in the next few sections how these ideas change the way we store the relations in our brains and think about them.

2.7 Partial and total orders.

We do not have much to say here, but these will come up when you start studying number systems like the natural numbers and the real numbers.

Partial orders are a mathematically precise way of expressing some kind of hierarchy in a system. For example, if X is the set of all people, and R is the relation defined by $R(a, b)$ is true if and only if b is an ancestor of a (that is, $b = a$, or b is a parent of a, or b is a parent of a parent of a, or...), then R will be a partial order.

Total orders are even stronger – they are a way of expressing a complete ordering of a system – a way of putting everything into a line, with small things on the left and big things on the right. For example the binary relation $\leq$ on the real numbers is a total order.

Here are the formal definitions.

Definition 2.7.1

Let X be a set, and let R be a binary relation on X.

1. We say that R is a *partial order* if it is reflexive, antisymmetric, and transitive.
2. We say that R is *total* if for all $a, b \in X$, either $R(a, b) \vee R(b, a)$ is true. In other words, for every a and b, either a is related to b, or b is related to a (or both).
3. We say that R is a *total order* if R is a partial order and R is total.

The standard examples of partial orders are $\subseteq$, on the set of subsets of a set, and $\leq$ on the reals (or on any subset of the reals, such as the integers). In fact $\leq$ is a total order on the reals, because if a and b are any real numbers, then either $a \leq b$ or $b \leq a$. However we cannot really "prove" these things about $\leq$ on the real numbers, because we have not even seen a definition of $\leq$, or even a definition of the real numbers! Later on in the course we will see some ideas about how to define the natural numbers, and $\leq$ on the natural numbers. In the Natural Number Game there is a proof that $\leq$ is a total order on the natural numbers.

2.8 Equivalence relations: definition and examples.

The last topic in this chapter is *equivalence relations*. Equivalence relations are a mathematically precise way of expressing the concept that two objects have an attribute or property in common. For example, if X is a set of plastic triangles and squares, and R is

the relation on X such that $R(a, b)$ is true if and only if a and b are the same shape, then R will be an equivalence relation. Similarly if R is the binary relation on positive integers such that $R(a, b)$ is true if and only if a and b end with the same digit, then R will be an equivalence relation. Here is the formal definition.

Definition 2.8.1

A binary relation R on a set X is said to be an *equivalence relation* if it is reflexive, symmetric, and transitive.

This is a strange definition, given what I said above, because the definition says nothing about attributes or properties, it just says that a bunch of logical statements are true.

Examples.

1) Let X be a set of plastic shapes, and let's say these shapes are red, green, blue, or yellow. Let us define a binary relation C on X by saying that if a and b are plastic shapes, then $C(a, b)$ is true if and only if a is the same colour as b.

Any shape a is the same colour as itself, so this means that $C(a, a)$ is true for all shapes a.

If a and b are two shapes, and a is the same colour as b, then b is the same colour as a! So $C(a, b) \implies C(b, a)$, and so C is symmetric.

If a, b and c are three shapes, and a is the same colour as b, and b is the same colour as c, then a, b and c must all be the same colour, so a is the same colour as c. In symbols, $C(a, b) \wedge C(b, c)$ implies $C(a, c)$. This means that C is transitive.

Hence C is an equivalence relation.

2) Let $\mathbb{Z}$ be the set of integers, and let's define a binary relation on $\mathbb{Z}$ by setting $R(m, n)$ be the proposition "$m - n$ is even". Well, zero is even, so for all $m \in \mathbb{Z}$ we have $R(m, m)$, and hence R is reflexive.

Now say $m, n \in \mathbb{Z}$ and $R(m, n)$ is true. Then $m - n$ is even, and this means that $n - m$ is even (if $m - n = 2t$, then $n - m = 2 \times (-t)$), so $R(n, m)$ is true and hence R is symmetric.

Finally, if $R(m, n)$ and $R(n, p)$ are both true, then $m - n$ and $n - p$ are both even, and hence their sum $m - p$ is even, which means $R(m, p)$ is true. Thus R is transitive.

Hence R is an equivalence relation.

3) The binary relation $\leq$ on $\mathbb{R}$ is not an equivalence relation. It is a total order, as we saw in the previous section, but it is not symmetric. You could just say "this is obvious" – but then you might lose a mark. To prove that R is not symmetric we *must* give a counterexample, we cannot just say it is clear. For example, we could say that $1 \leq 2$ is true, but $2 \leq 1$ is false, and hence R is not symmetric.

Exercise: does there exist a subset X of the reals such that the binary relation $\leq$ on X *is* symmetric?

4) Let X be a set of sets, and define a binary relation $\sim$ on X by saying that for elements $A, B \in X$, $A \sim B$ if and only if there exists a bijection from A to B. One can informally think of this concept like this: A is related to B if and only if A and B have the same size. The notion of size is a subtle one when it comes to infinite sets, but let's not worry about that right now.

Is $\sim$ reflexive? In other words, if A is a set, then does there exist a bijection $A \to A$? Yes there does – the identity function dose the job, as we saw in section 2.3.1.

Is $\sim$ symmetric? In other words, if A and B are sets, and there exists a function from A to B which is a bijection, then does there exist a function from B to A which is a

bijection? Yes there does – it turns out that the two-sided inverse works (exercise: check this!).

Finally, is $\sim$ transitive? In other words, if A, B and C are sets, and there's a bijection $f : A \to B$ and a bijection $g : B \to C$, does there exist a bijection $A \to C$? Yes there does – indeed we saw in Theorem 2.3.1(c) that $g \circ f$ will also be a bijection.

Hence $\sim$ is an equivalence relation.

5) This is an important generalisation of the first example.

Let X be a set, and let V be another set, and say $f : X \to V$ is a function. Here is how I want you to model this situation in your minds. We imagine X to be a set of things (for example X could be a set of plastic shapes), and we should imagine that there is an attribute which these things have (for example, these shapes all have a colour). We think of V as the possible values of this attribute (for example V could be the set {red, yellow, green, blue}). And we should think of f as the function which sends an element of X to the value that its attribute has (for example, f could send a plastic shape to its colour).

Now let's define a binary relation R on X like this: if $x, y \in X$ then we say that $R(x, y)$ is true if and only if $f(x) = f(y)$. In the example above, we are saying that two shapes are related if and only if they have the same colour. The relation completely forgets about whether the shape is a square or a triangle, and just concentrates on the colour attribute of the shape; this is all that matters.

I claim that R is an equivalence relation. The proof is just the same as in the first example. If $x \in X$ then certainly $f(x) = f(x)$, so $R(x, x)$ is true. If $x, y \in X$ and $f(x) = f(y)$, then certainly $f(y) = f(x)$, so $R(x, y) \implies R(y, x)$. Finally, if $x, y, z \in X$ with $f(x) = f(y)$ and $f(y) = f(z)$, then $f(x) = f(z)$. In other words, $R(x, y) \wedge R(y, z) \implies R(x, z)$. Hence R is an equivalence relation.

We've seen the example with plastic shapes already. Here's another example.

Let `Prop` denote the set of all propositions (considered as strings of mathematical symbols if you like). Let V be the set {`false`, `true`} of truth-values (a computer scientist might call this set `Bool`). There is a map from `Prop` to V which sends a proposition to its truth-value. Computer scientists might care about the way a proposition is made, or how much memory it takes to store the string, or other questions like that. Computer scientists might like $2+2 = 4$ better than $2000000+2000000 = 4000000$ because it takes up less memory, but mathematicians don't care about these "implementation issues" at all – mathematicians only care about the truth-value of a proposition. Every single calculation we did with propositions we could just do by checking on truth values. $2 + 2 = 4$ is not *equal* to $2 < 3$ – but it is *logically equivalent*, which for us is the correct notion of "the same". So a mathematician would want to define an equivalence relation on propositions – we want to say that two propositions are equivalent if and only if they have the same truth value. In other words, two propositions P and Q should be related, for us, if and only if either they're both true, or both false. This equivalence relation is logical equivalence $\iff$, and that's the reason it plays such an important role in the earlier sections.

2.9 Quotients and equivalence classes.

We have seen that if $f : X \to V$ is a function (to be thought of as sending x to a "value of a property" (e.g. its colour or its truth value)) and we define $R(a, b)$ to be $f(a) = f(b)$, then R is an equivalence relation.

In one of the equivalence relation videos, I observed that example 2 above was also of this form! In that example, X was the integers, and $R(a, b)$ was defined to be the statement that $a - b$ is even, and we proved it was an equivalence relation directly. Here's the trick. We set $V = \{0, 1\}$, and our map $f : X \to V$ sends an integer a to the remainder when a is divided by 2, so all even integers go to 0 and all odd integers go to 1. The

relation on X corresponding to f is then defined as: a is related to b if and only if the remainders when you divide a and b by 2 are equal. And this is easily checked to be our relation R above; the remainders of a and b modulo 2 are equal if and only if $a - b$ has remainder zero when you divide it by 2, i.e., $a - b$ is even.

The last thing we shall do in this chapter is to show that this is not a coincidence. Given *any equivalence relation at all* on *any* set X, it can informally be thought of as capturing some notion of "some attribute being the same". More formally, I am saying that if X is a set and R is an equivalence relation on X, then there exists a set V, called the *quotient* of X by R, and a surjective map $f : X \to V$, such that for all elements $x, y \in X$, $R(x, y) \iff f(x) = f(y)$. Given X and R only, the hard work is defining the elements of V, and this is what we shall do next.

2.9.1 Equivalence classes.

Let's say that I am playing with my collection of plastic triangles and squares, and thinking about my favourite equivalence relation on these shapes – two shapes are equivalent if and only if they have the same colour. I remember defining a new set $V = \{$red, yellow, green, blue$\}$ and defining a map f from my set of shapes to V, sending each shape to its colour, and I remember that I can define the equivalence relation on my shapes by saying that two shapes s and t are related if and only if $f(s) = f(t)$.

I decide to make a machine which computes this equivalence relation. Some time later I have made the machine. You can now take two of my shapes and put them into two slots in the machine, and the machine says "Yes" if they are the same colour, and "No" if they are not.

Now say that my blind friend pays me a visit. My blind friend can feel the shapes, and put them in the machine, but cannot see what colour any of them are. If they only have the machine, can they figure out all of the colours of all the shapes? Of course they can't – the machine only says yes or no, and my friend can't tell the colours of anything. On the other hand, using the machine my friend *can* still *sort the shapes out into piles corresponding to the four different colours*. Can you see how?

They start by choosing a shape s, and then they go through all of my collection, putting each one into the machine with s. Sometimes the machine says "yes", and sometimes the machine says "no". Each "yes" means that the shape they were testing has the same colour as s. My friend collects up all the "yes" shapes into one pile, and then at the end adds s to that pile. My friend doesn't know what colour those shapes all are, but they know that the pile contains all the shapes of one colour – all the shapes that have the same colour as s.

My friend then puts that s pile aside, chooses a new shape t from the remaining unsorted shapes, and starts again. It is not hard to see that after they have done this four times, they will have sorted my shapes into four piles, corresponding to the four colours. Those four piles correspond to the four elements of V. My friend can even make a set which bijects with V – he can just make the set whose elements are each of the four piles.

Not only that, but they can construct the map from the set of shapes to the set of four piles – you just send each shape to the pile which it's in. So my friend has reconstructed f and V, or at least things "the same as f and V" (the set consisting of "red", "blue", "green" and "yellow" is not *literally equal* to the set consisting of "the red pile", "the blue pile", "the green pile", "the yellow pile" – but clearly they are "the same" for the purposes of this argument).

Now let's see if we can pull off the same trick with a general equivalence relation, not assuming that it has come from some map $X \to V$. We use the traditional notation for an equivalence relation, namely $\sim$, which we write in between the elements of x it's relating, so we'll write $x \sim y$ to mean that x is related to y.

Definition 2.9.1

Let X be a set and let $\sim$ be an equivalence relation on X. Let $s \in X$ be an arbitrary element. We define the *equivalence class of* s, written $\mathrm{cl}(s)$, to be the set of elements of X which s is related to. Formally,

$$\mathrm{cl}(s) = \{x \in X : s \sim x\}.$$

This equivalence class $\mathrm{cl}(s)$ is a subset of X. In our plastic shapes example, it will be all the things which are the same colour as s. In the plastic shapes example, the equivalence classes (the piles) form a *partition* of X: every shape is in exactly one of the piles. Let's prove that this happens for a general equivalence relation. We begin with a lemma.

Lemma 2.9.1

Let X be a set and let $\sim$ be an equivalence relation on X. Say s, t are two elements of X. If $s \sim t$ then $\mathrm{cl}(t) \subseteq \mathrm{cl}(s)$.

Proof. Say $x \in \mathrm{cl}(t)$, or, in other words, $t \sim x$. Our goal is to prove $x \in \mathrm{cl}(s)$, or, in other words, $s \sim x$. But we know $s \sim t$ by assumption, and $t \sim x$, so by transitivity of $\sim$ we deduce $s \sim x$. ■

We now beef this lemma up a little.

Lemma 2.9.2

Let X be a set and let $\sim$ be an equivalence relation on X. Say s, t are two elements of X. If $s \sim t$ then $\mathrm{cl}(t) = \mathrm{cl}(s)$.

Proof. By Lemma 2.9.1 we have $\mathrm{cl}(t) \subseteq \mathrm{cl}(s)$. By assumption, $s \sim t$, so by symmetry of $\sim$ we deduce $t \sim s$. Applying the previous lemma again we deduce $\mathrm{cl}(s) \subseteq \mathrm{cl}(t)$. Hence $\mathrm{cl}(t) = \mathrm{cl}(s)$, because we have proved inclusions in both directions. ■

So we just showed that if two elements of X are related, then their corresponding equivalence classes are equal. In the shapes example, we have shown that if two shapes s and t have the same colour, then the set of shapes which are the same colour as s is equal to the set of shapes which are the same colour as t.

Now we show that if two elements are not related, their piles are disjoint (i.e., have no elements in common).

Lemma 2.9.3

Let X be a set and let $\sim$ be an equivalence relation on X. Say s, t are two elements of X. If $\neg(s \sim t)$ then $\mathrm{cl}(t) \cap \mathrm{cl}(s) = \varnothing$.

Proof. We prove the contrapositive instead. In other words, we prove that $\mathrm{cl}(t) \cap \mathrm{cl}(s) \neq \varnothing \implies s \sim t$. So let's choose $x \in X$ such that $x \in \mathrm{cl}(t)$ and $x \in \mathrm{cl}(s)$, and use it to deduce $s \sim t$. By definition of cl we see that $t \sim x$ and $s \sim x$. By symmetry we deduce $x \sim t$ and by transitivity ($s \sim x \sim t$) we deduce $s \sim t$. ■

In the shapes example, we have shown that if s and t have different colours, then there are no shapes which are simultaneously the same colour as s and as t.

We now formally define a partition of a set X.

Definition 2.9.2

A *partition* of a set X is a set A of non-empty subsets of X with the property that each element of X is in *exactly one* of the subsets.

For example, if $X = \{1, 2, 3, 4, 5, 6\}$ then an example of a partition of X would be $A = \{\{1, 3, 4\}, \{2, 6\}, \{5\}\}$.

Theorem 2.9.1

Let X be a set and let $\sim$ be an equivalence relation on X. Then the set V of equivalence classes $\{cl(s) : s \in X\}$ for $\sim$ is a partition of X.

Proof. First note that every equivalence class is non-empty, the equivalence class $cl(y)$ contains y.

So now we just need to show that every $x \in X$ is in exactly one equivalence class. By reflexivity we know $x \sim x$, so we can deduce $x \in cl(x)$. Now say $x \in cl(y)$ for some equivalence class y. We want to show that $cl(y) = cl(x)$. But $x \in cl(y)$ implies $y \sim x$ (by the definition of cl), and hence $cl(y) = cl(x)$ by Lemma 2.9.2. ■

Note that when we checked that each element was in exactly one equivalence class, we are really using the fact that sets "can't have elements in more than once". In the plastic shapes example, if we have 20 red shapes r_1, r_2, r_3, ..., r_{20} then $cl(r_1) = cl(r_2) = \cdots = cl(r_{20})$ so the set $\{cl(r_1), cl(r_2), \ldots, cl(r_{20})\}$ just equals the set $\{cl(r_1)\}$ consisting of the red pile.

Now let X be a set, let $\sim$ be an equivalence relation on X, and let V be the set of equivalence classes. There is a natural map $f : X \to V$, sending x to $cl(x)$ (in the plastic shapes example, V is a set of size four (the four piles) and f sends every shape to the pile it's in). Now let R_f be the equivalence relation associated to f, so $R_f(s, t)$ is true if and only if $f(s) = f(t)$, or in other words if $cl(s) = cl(t)$.

Theorem 2.9.2

The equivalence relations R_f and $\sim$ are equal.

In other words, we have shown that every equivalence relation can be thought of as one coming from "equality of some property of the elements"; we can think of V as the possible values of that property.

Proof. We need to show that for all $a, b \in X$, $R_f(a, b)$ is true if and only if $a \sim b$. Now $R_f(a, b)$ is true if and only if $f(a) = f(b)$, which means $cl(a) = cl(b)$. So our goal is to prove that $cl(a) = cl(b) \iff a \sim b$. One way is lemma 2.9.2. The other way is easier – if $cl(a) = cl(b)$ then because $b \sim b$ by reflexivity we deduce $b \in cl(b) = cl(a)$ and hence $a \sim b$ by definition of $cl(b)$. ■

2.9.2 Summary / overview of what is to come.

It's very rare for maths texts to have summaries. But I (KMB) have been reading computer science papers recently and they all have summaries, and I've seen people complaining online about how maths papers and books never have summaries and this makes them much harder to read. In response to that, here is a summary of the first two chapters, just

to prove that it can be done, plus some other comments about the objects we have been studying so far.

Propositions (or logical propositions, to give them their full name) are true/false statements. There are various operators which create new propositions from old ones, for example $\wedge$ and $\implies$. We proved various facts relating these operators, for example $(P \implies Q) \iff (\neg Q \implies \neg P)$, the fact that an implication is logically equivalent to its contrapositive (and we even used this to prove that there was no rational number whose square was 2). Maths is full of propositions, so it was important to see these first.

Sets are ways to collect a bunch of objects together. Sets and their elements, and Propositions and their proofs, are the four main kinds of mathematical objects studied throughout pure mathematics. Famous sets of numbers include the natural numbers (which may or may not include 0 depending on who you're talking to), the integers, rationals, reals and complexes, but there are plenty of other abstract sets, even including sets of sets. The empty set is one of the hardest sets. The rule of thumb is that if you're trying to prove "for all x in the empty set,..." then this statement will be true no matter what comes after the ..., and conversely if you're trying to prove "there exists x in the empty set, ..." then this statement will be false no matter what comes after the ....

Functions are maps between sets. It is a basic principle in mathematics that if you are studying objects (such as sets, or more exotic structures such as groups, rings, fields, topological spaces,...) then it's important to figure out how to move between these objects. Functions are the way to move between sets. Things get fun when we introduce predicates describing these functions, such as "injective" or "bijective". These things all have careful definitions, which we understand internally using pictures; however our mathematical language is rich enough to be able to write down precise proofs without needing to resort to pictures. One irony is that a mathematician explains the picture in their head by translating the picture down into formalised strings of symbols, which are then read by another mathematician and converted back into a picture. Mathematicians might store pictures in their heads in very different ways – this phenomenon is not particularly well-understood (at least as far as I can make out).

Equivalence relations are a subtle abstract concept which are often taught at university but very rarely taught at school. They are everywhere in mathematics because the idea of two elements of a set being "equivalent" captures the abstract ideas of them being "the same", or sharing a certain attribute (for example being the same colour, or ending in the same digit) without necessarily being equal.

Equivalence relations are one of the most subtle concepts introduced in these chapters, and many students struggle to understand them. The Wikipedia page on equivalence relations is quite nice, but in these notes and the lectures I adopted a completely different approach, running with one concrete example (plastic shapes) quite unrelated to numbers, in the hope that this example enables students to understand things from another point of view.

2.10 Appendix: the universal property of quotients (not examinable).

This appendix is non-examinable. However it might help some people understand things better.

First year undergraduates are normally very happy with the idea of a subset, but far less happy with the idea of the set of equivalence classes of an equivalence relation. These two concepts are "dual" to each other in some abstract way – I won't make this precise. However this does not change the fact that a set of equivalence classes is a set of sets, and hence an intrinsically more complicated thing than a subset.

So let X be a set, let $\sim$ be an equivalence relation on X, let V be the set of equivalence classes, and let $cl : X \to V$ be the function which takes an element of X and sends it to its equivalence class. For example, X could be a collection of plastic red, green, yellow and blue squares and triangles, the equivalence relation could be: $a \sim b \iff a$ is the same colour as b, and then V is a set with four elements – one element is the set (or pile) of all the red shapes, one is the set of all the yellow shapes, etc. The map from X to V sends a shape to the pile it is in.

In this appendix I'm going to explain a machine which gives you a way to define functions from V to other sets. In fact we're going to see how functions from V are related to certain kinds of functions from X. Let's start with the easy way around – let's say we have a function g from V to some other set Y. This gives the following picture:

$$\begin{array}{ccc} X & & \\ \downarrow{\scriptstyle \mathrm{cl}} & & \\ V & \xrightarrow{g} & Y \end{array}$$

and we can just compose those two maps together to get a map $h = g \circ \mathrm{cl} : X \to Y$ like this:

$$\begin{array}{ccc} X & & \\ \downarrow{\scriptstyle \mathrm{cl}} & \searrow^{h} & \\ V & \xrightarrow{g} & Y \end{array}$$

Now note that the function $h : X \to Y$ has the following property. Say $x_1, x_2 \in X$ and $x_1 \sim x_2$ – the two elements are related by the equivalence relation (e.g., they are the same colour). Then $\mathrm{cl}(x_1) = \mathrm{cl}(x_2)$ by Lemma 2.9.2, and so $h(x_1) = h(x_2)$, because $h(x_1) = g(\mathrm{cl}(x_1)) = g(\mathrm{cl}(x_2)) = h(x_2)$. Look at the picture to understand this proof: if x_1 and x_2 are in X, and cl sends them to the same element of V, then h must send them to the same element of Y. But the proof is not the picture, the proof is the sentence before I tell you to look at the picture.

So g gives us h with some nice property. The whole point of this appendix is to explain to you that the converse is also true: this is called the *universal property* of quotients by equivalence relations.

Theorem 2.10.1

Let X, $\sim$, V, cl be as above. Say Y is an arbitrary set and $h : X \to Y$ is an arbitrary function. Assume that for all $x_1, x_2 \in X$, if $x_1 \sim x_2$ then $h(x_1) = h(x_2)$. Then there exists a unique function $g : V \to Y$ such that $h = g \circ \mathrm{cl}$.

Before we prove the theorem, here are some examples of what is going on.

Let X be a set of plastic red, blue, green and yellow squares and triangles, with the equivalence relation $a \sim b$ iff a is the same colour as b. Let Y be the set $\{$`yellow`, `not-yellow`$\}$. Consider the function h from X to Y which sends a plastic shape to `yellow` if it's yellow, and `not-yellow` if it's not yellow. I claim that it satisfies the hypothesis of the theorem. Indeed, if a is the same colour as b then either a and b are both yellow, so they get mapped to `yellow`, or they're both some other colour, in which case they both get mapped to `not-yellow`. The theorem then says that there should be some map g from the four equivalence classes (the yellow pile, the red pile, the blue pile and the green pile) to Y such that $h = g \circ \mathrm{cl}$. It's pretty clear what this map is – it sends the yellow pile to `yellow` and all three other piles to `not-yellow`.

Now let X be the same set and let's keep the same equivalence relation on it. But this time let Y be the set $\{3,4\}$ and consider the map h from X to Y which sends each shape to the number of sides that it has (so the triangles get mapped to 3 and the squares get mapped to 4). This function h does *not* satisfy the conditions of the theorem, because if two shapes are the same colour then they still might have a different number of sides. And indeed, there is no map g from the set of piles to $\{3,4\}$ because if we have a pile of blue squares and triangles, this is *one element of* V and this one element has to go to *one* of $\{3,4\}$, and there is no good choice. Let's say we send it to 3. Then for a blue square in X, h sends it to 4, but cl sends it to the blue pile so it ends up going to 3 after g, so $h \neq g \circ cl$.

Now let $X = \mathbb{Z}$ be the integers, and let's consider the equivalence relation of being congruent mod 4, so the set V of equivalence classes, which is usually written $\mathbb{Z}/4\mathbb{Z}$, has size 4 and equals $\{cl(0), cl(1), cl(2), cl(3)\}$, sometimes written $\{[0]_4, [1]_4, [2]_4, [3]_4\}$. If Y is the set $\mathbb{Z}/2\mathbb{Z}$, and h is the map sending each integer to its equivalence class mod 2 (i.e. h sends even integers to $[0]_2$ and odd integers to $[1]_2$), then for $a, b \in X$, if $a \sim b$ then $a - b$ is a multiple of 4, so it's a multiple of 2, which means that if a and b are equivalent then $h(a) = h(b)$. By our theorem, there should then be some map $\mathbb{Z}/4\mathbb{Z} \to \mathbb{Z}/2\mathbb{Z}$ such that reducing an integer modulo 4 and then applying this map should be the same as reducing it modulo 2. One checks easily that the map sending $[0]_4$ and $[2]_4$ to $[0]_2$ and sending $[1]_4$ and $[3]_4$ to $[1]_2$ does the job. In words, if an integer is 0 mod 4 or 2 mod 4 then it's even, and if it's 1 mod 4 or 3 mod 4 then it's odd.

Next let X be the integers as above, with the same equivalence relation of being congruent mod 4, and let h be the function which sends an integer to its reduction mod 5. Now h does not satisfy the criterion of the theorem, because 4 and 8 are congruent mod 4 and hence related, but they are different modulo 5. Hence one does not expect there to be a map from $\mathbb{Z}/4\mathbb{Z}$ to $\mathbb{Z}/5\mathbb{Z}$ making the triangle commute (i.e., making $h = g \circ cl$). But what is wrong with the map $\mathbb{Z}/4\mathbb{Z} \to \mathbb{Z}/5\mathbb{Z}$ sending 0 to 0, 1 to 1, 2 to 2 and 3 to 3, you may ask? Well sure, this map exists, but the thing is that if you take a number like 9, its remainder mod 4 is 1 but its remainder mod 5 is 4, which isn't 1, so even though we could define g that way, it won't satisfy $h = g \circ cl$.

These examples show that there is a method for defining maps from equivalence classes – define a map from X and check that it sends equivalent elements to the same thing. This is why one has to work hard when defining things like addition on the rationals (something you'll learn about in Part II of this module); if a is an integer and b is a non-zero integer, and $cl(a,b)$ is the equivalence class representing a/b, one can't just define $cl(a,b) + cl(c,d) = cl(ad+bc, bd)$, because how do we know that if we change (a,b) to an equivalent (a',b'), that the formula defining the function will still give the same answer? The formulae for adding $\frac{1}{2}$ to a fraction and adding $\frac{2}{4}$ to a fraction are *different formulas*, in the sense that one has 4's in and the other doesn't – and one has to check that they still give the same answer. Exactly what needs to be checked is explained by the theorem above.

We finally embark upon a proof of Theorem 2.10.1.

Proof. Assume that h is as in the theorem, and $x_1 \sim x_2 \implies h(x_1) = h(x_2)$. First let's prove that g exists. Say $v \in V$. The map $cl : X \to V$ is surjective, because by definition V is the set of equivalence classes, and "equivalence class" *means* "$cl(x)$ for some $x \in X$". So given this v, we can choose some $x \in X$ such that $cl(x) = v$. I want to define $g(v) = h(x) \in Y$. But this definition is dodgy, because what if $v = cl(x)$ for 100 different choices of x which get sent to 100 different elements of Y by h? Well this is *exactly* what our assumption rules out. If $v = cl(x_1) = cl(x_2)$ then $x_2 \in cl(x_2)$ by reflexivity and hence $x_2 \in cl(x_1)$, so $x_1 \sim x_2$ and by our assumption we deduce $h(x_1) = h(x_2)$. This means that our definition of $g(v)$ does not depend on that auxiliary choice of x. Because of this,

we see that if $x \in X$ then $g(\mathrm{cl}(x))$ must equal $h(x)$, because if $v = \mathrm{cl}(x)$ then to define $g(v)$ we can choose some random element of X whose equivalence class is v, and we may as well choose x. Not only have we defined g, but we've proved that $g \circ \mathrm{cl} = h$.

It remains to show that g is unique. This follows immediately from the surjectivity of cl. Indeed, if g' is another function $V \to Y$ with $g' \circ \mathrm{cl} = h$ and if $v \in V$ is arbitrary, then we can choose $x \in X$ with $\mathrm{cl}(x) = v$, and then $g'(v) = g'(\mathrm{cl}(x)) = h(x) = g(\mathrm{cl}(x)) = g(v)$. Because v was arbitrary, this shows that $g' = g$. ■

Chapter 3

Naturals and Integers

”Die Zahlen sind freie Schöpfungen des menschlichen Geistes, sie dienen als ein Mittel, um die Verschiedenheit der Dinge leichter und schärfer aufzufassen” (DEDEKIND, Was sind und was sollen die Zahlen? Braunschweig 1887, S. III).

[Numbers are free creations of the human intellect, they serve as a means of grasping more easily and more sharply the diversity of things.]

3.1 Introduction

Historically, numbers were used to solve real world problems, and formal *definitions* of numbers did not exist in the literature. For example the ancient Greeks might have thought of positive real numbers as “lengths” in the physical universe. This informal approach is fine in practice if you’re using numbers to do things like make measurements of real world objects (for example if you are building a tower, or keeping track of how much gold or how many sheep you have), but it has deficiencies. For example it doesn’t answer the questions of whether $0.999999\ldots = 1$ and whether infinitely small positive numbers can exist. By the 1600s Newton and Leibniz had discovered the elusive quantities dy and dx, which were giving correct answers to questions in physics via the theory of differential equations. Are these “numbers”? Any attempt to resolve such subtle questions about the nature of numbers via physics will inevitably fail, because by the time we have got down to about 10^{-35} metres (Planck length) then spacetime becomes a quantum foam, matter and antimatter start rapidly appearing and disappearing, and things get pretty weird.

The mathematician knows that quantum foam has nothing to do with the real numbers. What is happening is that the real numbers *do not exist in the physical universe* and hence we cannot use the real universe to model them perfectly. Indeed no infinite thing can “exist” in our physical universe, which is large but finite, and contains only around 10^{70} atoms. This is in sharp contrast to what is happening in mathematics, where one of the first things we learn as children is how to count, together with the implicit idea that you can continue to count forever.

The idea of a “mathematical universe” where things like a perfect circle or a 17-dimensional cube can exist untroubled by physics, goes back to Plato. But it was only around 140 years ago that mathematicians started seriously thinking about *axioms* which defined numbers. There are all kinds of number systems in use in mathematics: in this and the next chapter we shall consider the five most important ones:

- The natural numbers (the counting numbers);
- The integers (where we allow negative whole numbers);
- The rationals (where we allow ratios of integers);

- The reals (where we allow non-periodic decimals);
- The complexes (where we allow square roots of negative reals).

We'll also consider the integers modulo n, a world where some positive numbers can equal zero. People interested in the quaternions, octonions, sedonions, p-adic numbers, hyperreals, nimbers and other exotic number systems will have to wait until later in the course; we already have plenty to get through here.

3.2 Overview

We will begin our story with the *natural numbers*. We will watch these numbers be born, and then we will develop the theory until they become the numbers which we are used to using. A big big warning: while the theory is being developed, we absolutely *cannot* use standard facts about numbers which we all know: for most of this section we will be seeing numbers in a strange half-constructed state, and we will need to be *extremely* careful not to assume anything which we didn't already prove, because we need to avoid *circular arguments*,

We will summon these numbers into existence by using three *axioms*; these will be rules which we decree are true in the mathematical universe, rather like how the laws of physics are true in our physical universe. Once we have the natural numbers up and running, we will then *build* the integers, rationals, reals and complexes from them directly; no more axioms will be required.

One important consequence of this is that we will then be able to *prove* statements like $x+y = y+x$ or that $0.9999\cdots = 1$, without *ever* having to appeal to real world intuition. In particular, in this part of the course, we *reject* arguments such as "if I have x apples in one hand and y apples in the other, and I then swap my hands around, obviously the total number of apples I have didn't change, and obviously this works for all things, not just apples, therefore $x + y = y + x$". The formalist view of mathematics is that it is a puzzle game, like a Sudoku or a chess puzzle. You cannot make an illegal move to solve a chess puzzle because this contradicts the axioms of chess (i.e., the rules of chess). We will take the same view when working with numbers: if you can't deduce your argument from the axioms, then your argument is not allowed.

3.3 Natural numbers

The natural numbers are as old as mankind. But it's really with the work of Dedekind and finally Peano in 1889, that they were described axiomatically. Let's take a look at a modern version of Peano's axioms.

3.3.1 Peano axioms

In this section, we will uniquely characterise the natural numbers by axioms. Informally, what we are trying to *model* with these axioms is the infinite set

$$\mathbb{N} = \{0, 1, 2, 3, 4, \ldots\}$$

of *natural numbers*. Some other lecturers will use the convention that $\mathbb{N} = \{1, 2, 3, 4, \ldots\}$ starts at 1, and historically Peano actually started with 1, but for the material we present here it is much more convenient to start with 0, so we shall start with 0.

Infinity gives rise to lots of paradoxes; the problem with the above definition is that it is unclear what that "..." in the equation above really *means*. Let us give a slightly more formal definition of $\mathbb{N}$, in terms of three axioms.

Axiom 3.3.1: Peano axioms, first version

The natural numbers $\mathbb{N}$ are defined via the following three axioms.

- 0 is a natural number.
- If n is a natural number, then its *successor* $S(n)$ is a natural number.
- That's it.

The *successor* of a natural number is the number which comes after it. For example the successor of 37 is 38. "That's it" means "0 and successors are the only ways to make natural numbers". We will be more precise about this later; we do not need the third axiom for the moment.

You might be thinking "why do we talk about $S(n)$? Why don't we just say that $S(n)$ is $n+1$?" Unfortunately this would be a *circular argument*, and those are not allowed in mathematics. We *have not yet defined addition* on the natural numbers, and the definition of addition will *use* the successor function, so the definition of the successor function cannot use addition. If you think that it is surprising to introduce this funny function $S : \mathbb{N} \to \mathbb{N}$ before addition, then think back to the way very small children are taught about numbers: before they learn to add, they learn to count. The successor function is telling you how to count.

3.3.2 Our first numbers.

Forget the third axiom: let's think about the first two. Which numbers can we make from them? The first axiom only gives us 0. But applying the second axiom gives us new numbers $S(0)$ (the number after 0), $S(S(0))$ (the number after the number after 0) and so on. New numbers have just been born: let's give them names.

Definition 3.3.1: One, two, three, four

We define 1 to be $S(0)$, we define 2 to be $S(1)$, we define 3 to be $S(2)$ and we define 4 to be $S(3)$.

Summary: So far we have numbers $0, 1, 2, 3, 4$, a function S, and *nothing else*. No addition, no multiplication, and no $\leq$. We still have a lot of work to do before numbers become the mathematical object which we are all familiar with.

3.3.3 More on "That's it".

Now we have defined 2 and 4, can we prove that $2 + 2 = 4$? Unfortunately we cannot, because we have not yet defined addition! Let's work towards this now. Let's fix a natural number like 37, and discuss how to define the function which adds 37 to a number, i.e., the function $f(x) = 37 + x$. I stress again: addition does *not yet exist* in our theory – we have to *make* it, and this is what we will do now. More generally, for a fixed number a we will have to define the function sending x to $a + x$.

To define addition, we have say more precisely what the final axiom "that's it" means. Let us give a slightly more precise explanation of this axiom.

Axiom 3.3.2: That's it, version 2

If you want to do something (for example define something, or prove something) for

all natural numbers, then you only need to do the following two things:

- Do it for 0.
- Do it for $S(n)$, assuming you've already done it for n.

The point of "That's it" is that it is saying that 0 and S are the *only ways to make natural numbers*, so if you want to do something for all natural numbers, then it suffices to do it in the 0 case and do it in the $S(n)$ case. Let's now define addition.

3.3.4 Definition of addition.

Recall the set-up: a is a fixed natural number, and we want to define the function which sends x to $a + x$. By "That's it", here's what we have to do:

1. Define $a + 0$;
2. If we have already defined $a + n$, we must define $a + S(n)$.

We want addition to agree with our real world intuition, so let's make the following definition:

Definition 3.3.2: Addition

- Define $a + 0$ to be a.
- If $a + n$ is already defined to be y, then define $a + S(n)$ to be $S(y)$. In other words, $a + S(n)$ is defined to be $S(a + n)$.

"That's it" ensures that this is a valid definition. The logic in the successor case is that $a + S(n)$ is a and n and one more, so it's the number after $a + n$.

Let's now prove that $1 + 1 = 2$.

Theorem 3.3.1

$1 + 1 = 2$.

Proof.

$$\begin{aligned} 1 + 1 &= 1 + S(0) \text{ (by definition of } 1) \\ &= S(1 + 0) \text{ (by definition of } a + S(n)) \\ &= S(1) \text{ (by definition of } a + 0) \\ &= 2 \text{ (by definition of } 2). \end{aligned}$$

■

Exercise. Prove that $2 + 2 = 4$.

Now we have defined addition, the equation $S(x) = x + 1$ finally *makes sense*. Let's prove that it is true.

Lemma 3.3.1

If x is a natural number then $S(x) = x + 1$.

Proof.

$$\begin{aligned} x+1 &= x+S(0) \text{ (by definition of 1)} \\ &= S(x+0) \text{ (by definition of +)} \\ &= S(x) \text{ (by definition of +).} \end{aligned}$$

■

3.3.5 Basic properties of addition

We know that $x+0=x$. Indeed this is how we *defined* addition. But what about $0+x$? Well obviously $a+b=b+a$ so $0+x$ is just $x+0$ which is x. Why is this argument not allowed?

The reason is that we are *not allowed to say* "obviously $a+b=b+a$" – we have to *prove* this! It might be "obvious" from some physical intuition which you have, but we are in the mathematical universe here and there is no physics for you to rely on. In fact deducing $0+x=x$ from $a+b=b+a$ is no good – it is another example of a *circular argument*, because $a+b=b+a$ is a tricky puzzle, and it needs $0+x=x$ as an input, so we will have to prove $0+x=x$ first.

Let's go back to "that's it", and write down explicitly what it tells us in the case when we want to *prove* something.

Axiom 3.3.3: The principle of mathematical induction

Say we have infinitely many true-false mathematical statements P_0, P_1, P_2, P_3, ..., and we want to prove them all. Then it suffices to do the following two things:

- Prove that P_0 is true;
- Prove that if P_n is true, then $P_{S(n)}$ is true.

In the second case, when proving $P_{S(n)}$, we often refer to our assumption P_n as "the inductive hypothesis". Let's use this principle to prove a *lemma*, which is a simple theorem.

Lemma 3.3.2

If x is a natural number, then $0+x=x$.

Proof. We use induction on x. Formally we let P_x be the statement that $0+x=x$ and we use the above principle to prove that P_x is true for all x.

To prove P_0 we need to prove that $0+0=0$, but this follows from the definition of addition.

In the successor case, we may assume that $0+n=n$ and we want to now deduce that $0+S(n)=S(n)$. This follows from the following calculation:

$$\begin{aligned} 0+S(n) &= S(0+n) \text{ by definition of addition} \\ &= S(n) \text{ by the inductive hypothesis.} \end{aligned}$$

■

We also know that $a + S(x) = S(a + x)$, but what about $S(a) + x$? Can we prove that this is $S(a + x)$? This is the last remaining piece we need in order to prove commutativity of addition.

Lemma 3.3.3

If a and x are natural numbers, then $S(a) + x = S(a + x)$.

Proof. Exercise! Hint: imagine that a is fixed, let P_x be the statement that $S(a) + x = S(a + x)$, and use induction on x. ■

We're now ready to prove *commutativity of addition on the natural numbers*, which is a fancy way of saying $x + y = y + x$.

Theorem 3.3.2: Commutativity of addition on $\mathbb{N}$

If a and b are natural numbers, than $a + b = b + a$.

Proof. We fix a and use induction on b.

The base case: we need to show $a + 0 = 0 + a$. But $a + 0 = a$ by definition of addition, and $0 + a = a$ by lemma 3.3.2, so we are done.

The inductive step: we can assume $a + n = n + a$ and we want to prove $a + S(n) = S(n) + a$. We do this as follows.

$$\begin{aligned} a + S(n) &= S(a + n) \text{ by definition of } + \\ &= S(n + a) \text{ by the inductive hypothesis} \\ &= S(n) + a \text{ by lemma } 3.3.3. \end{aligned}$$

■

It is remarkable to think that all of modern pure mathematics can be built up in this way. But it can.

We've talked about adding two numbers, but what about adding three numbers? What does $a + b + c$ mean? This is an *ambiguous term*! It could either mean $(a + b) + c$ or $a + (b + c)$. These are "obviously" the same, but that's no good! We need to *prove* it from the axioms. The pompous name that mathematicians give to this property is *associativity* of addition.

Theorem 3.3.3: Associativity of addition

If a, b, c are natural numbers, then $(a + b) + c = a + (b + c)$.

Proof. Exercise! Hint: fix a and b, and do induction on c. ■

Exercise: by means of an explicit example, show that subtraction is not associative (don't worry about the fact that we haven't defined subtraction yet).

3.3.6 Recursion

I have told you one aspect of "that's it" – it encodes the principle of mathematical induction. But in fact the full formal definition of "that's it" is both the principle of induction and the principle of recursion, which we now state in its most general form. This definition is non-examinable.

Axiom 3.3.4: The principle of mathematical recursion

Say we have infinitely many sets X_0, X_1, X_2, ... You can make a function F which sends a natural number n to an element of the set X_n by doing the following two things:

- Define the element $F(0) \in X_0$;
- Define a function p_n from X_n to $X_{S(n)}$, and define $F(S(n))$ to be $p_n(F(n))$.

In many applications, all the X_n are equal, and often all the p_n are equal too. For example, when we defined $F(x) = a + x$ above, we let all the X_n be $\mathbb{N}$, we let all the p_n be $S : \mathbb{N} \to \mathbb{N}$, and we defined $F(0) = a$. Informally, adding x to a number a is defined by "start at a, and then repeatedly take the successor x times". But don't write this in an exam, because x is a *formal term in a formal system*, and "doing something x times" makes no sense in this formal system. In particuar, trying to define $a + x$ as $\underbrace{S(S(S(\ldots S}_{x \text{ times}}(a))) \ldots)$ is *not allowed*, because it is not at all clear what $\ldots$ *means*; it is attempting to attach an intuitive semantic meaning to a formal term in a formal language. But intuition is exactly what is not allowed here: if you want to do something x times, you must use induction or recursion.

3.3.7 Multiplication

Subtraction is a problem on the natural numbers – if the only numbers available to us are $\{0, 1, 2, 3, \ldots\}$ then we cannot define $2 - 5$ correctly. So let's leave subtraction until later on when we have negative integers, and now define multiplication, which will work fine. We use both notations $a \times x$ and ax for multiplication.

Definition 3.3.3: Multiplication

Fix a natural number a. The function $F(x) = a \times x$ is defined by the following two rules:

- $a \times 0 = 0$;
- $a \times S(n) = a \times n + a$.

Like addition, we're defining the $F(x)$ by saying what it does to 0, and also what it does to $S(n)$ assuming we know what it does to n. By "that's it" (or more formally "by recursion"), this is a valid way to define multiplication.

[Note: A fully formal justification that this is a valid definition of the function $F(x) = ax$ would use the principle of recursion above, with all of the sets X_n defined to be $\mathbb{N}$, with $x_0 = 0$ and $p_n(m) = m + a$. In short, to multiply a general number x by a, you "start at 0, and then repeatedly add a, x times".]

Just like addition, we now have a lot of work to do, in order to prove all the basic properties of multiplication which we know and love.

Proposition 3.3.1: Basic facts about multiplication

Let $x, y, z \in \mathbb{N}$. Then

- $x \times 1 = x$;

- $0 \times x = 0$;
- $S(x)y = xy + y$;
- $xy = yx$;
- $1 \times x = x$;
- $x(y + z) = xy + xz$;
- $(x + y)z = xz + yz$;
- $(xy)z = x(yz)$.

The order above is the recommended order for proving the above results. Proving this proposition is the first question on the first IUM Part II problem sheet.

3.3.8 Peano's remaining axioms

Earlier on we proved $1 + 1 = 2$, but can we also prove that $1 + 1 = 3$? This would be easy to prove if $3 = 2$, but $3 = 2$ looks a bit suspicious. However, can we rule out $3 = 2$ *using only the axioms*? It is amusing to note that we were able to prove lots of theorems about addition and multiplication without ever worrying about this issue.

In Peano's original paper, he add two extra axioms to the system in order to make this easy.

Definition 3.3.4: Peano's extra axioms

- If x is a natural number, then $S(x) \neq 0$.
- If x and y are natural numbers and $S(x) = S(y)$, then $x = y$.

The first of these axioms says that if you keep applying S then you will never end up back at 0, and the second axiom says that S is an *injective* function. Put together, these two axioms guarantee that if we keep applying the S function then it will never output a number which we've seen before ("counting doesn't loop"). Let's use Peano's extra axioms to show that $3 \neq 2$.

Theorem 3.3.4

$3 \neq 2$.

Proof. We prove this by contradiction. Let's assume $3 = 2$. Writing $2 = S(1)$ and $3 = S(2)$ (remember that those are the *definitions* of 3 and 2), we deduce $S(2) = S(1)$ and hence by the second extra axiom we deduce $2 = 1$, or in other words $S(1) = S(0)$. Applying this axiom again we deduce that $1 = 0$ or in other words $S(0) = 0$. But this contradicts the first extra axiom (set $x = 0$). ■

However, using the more modern approach explained in these notes, it is possible to *prove* these two extra axioms of Peano – we don't need them after all. Let's use the axiom of recursion ("that's it") to define two new temporary functions. First let's define a function Z from the natural numbers to the set $\{\texttt{true}, \texttt{false}\}$ by recursion. It's called Z because it detects whether a number is zero.

Set $Z(0) = \texttt{true}$, and if we have already defined $Z(n)$ then let's define $Z(S(n))$ by ignoring what the value of $Z(n)$ was and just setting $Z(S(n))$ to be always `false`. The upshot of this definition is that we have a function which sends 0 to `true` and everything else to `false`.

Now it is easy to prove the first of Peano's extra axioms $S(x) \neq 0$ by contradiction: if $S(x) = 0$ for some x, then applying Z to both sides we deduce that `false=true`, which is the contradiction we seek.

To prove the second axiom, let's define a "predecessor" function $P(x)$ which subtracts 1 from a number. There is a problem with 0 though, because $0-1$ is not a natural number, so we'll have to define $P(0)$ to be something else. Fortunately our proof will never involve looking at $P(0)$, so we can just define it to be anything we like. Formally then, let's define a function P from the naturals to the naturals by recursion: we will define $P(0)$ to be 37, and if we have already defined $P(n)$ then let's define $P(S(n))$ by forgetting all about $P(n)$ and just setting $P(S(n)) = n$.

Now $P(0)$ is a bit weird, but $P(S(n)) = n$, and this is exactly what we need to prove Peano's second extra axiom (injectivity of S). Indeed, if x and y are natural numbers and $S(x) = S(y)$, then applying P we deduce $P(S(x)) = P(S(y))$ and hence $x = y$, which is what we require. We conclude that Peano's extra axioms can be proved from the three axioms we assumed at the beginning.

We saw some slightly odd uses of recursion above; here's a slightly odd use of induction.

Theorem 3.3.5

Every natural number x is either 0 or the successor of some other natural number.

Proof. Induction on x. The base case has $x = 0$ which is zero. For the inductive step we set $x = S(n)$ and we know that the result is true for n. We can ignore the inductive hypothesis though, and conclude because $x = S(n)$ is a successor. ■

3.3.9 Consequences of these new results

The results above are the first theorems that we have seen in this chapter of the form "if some equation is true, then some other equation is also true"; results like commutativity and associativity of addition and multiplication were of the form "this equation is always true", and hence logically simpler. Let's prove some other results of the form "if some equation is true, then some other equation is also true", using Peano's last two axioms $S(x) = S(y) \implies x = y$ and $S(x) \neq 0$.

Theorem 3.3.6

Say x, y and n are natural numbers.

a) If $x + n = y + n$ then $x = y$ (the cancellation property for addition);

b) If $x + n = n$, then $x = 0$;

c) If $x + y = 0$, then $x = y = 0$.

Proof. *a*) Fix x and y, and let's do induction on n. More precisely, we let P_n denote the statement "if $x + n = y + n$ then $x = y$" and we prove P_n for all n, by induction.

The base case is easy: if $x + 0 = y + 0$ then $x = y$ because $x + 0 = x$ and $y + 0 = y$ by definition of addition. For the successor case, we assume "if $x + n = y + n$ then

$x = y$" and we have to prove "if $x + S(n) = y + S(n)$ then $x = y$". So let's assume $x + S(n) = y + S(n)$; by definition of addition we deduce $S(x + n) = S(y + n)$. By injectivity of S (Peano's second extra axiom, which we proved above) we deduce $x + n = y + n$. And finally by our inductive hypothesis we conclude $x = z$.

b) Setting $y = 0$ in the previous part gives $x + n = 0 + n \implies x = 0$, and because we know $0 + n = n$ we're done.

c) Let's first prove that if $x + y = 0$ then $y = 0$. Let's do this by instead showing the contrapositive, namely that if $y \neq 0$ then $x + y \neq 0$. We showed in theorem 3.3.5 that y was either 0 or a successor, so if $y \neq 0$ then we must have $y = S(n)$ for some natural number n. In this case $x + y = x + S(n) = S(x + n)$, and we have to prove that this is nonzero, but this follows immediately from Peano's first extra axiom.

We deduce that $y = 0$, and hence $x + 0 = 0$ which implies that $x = 0$ by definition of addition.

■

3.3.10 The ordering on the naturals

We did not need Peano's extra axioms to develop the theory of addition and multiplication, but we will need them to develop the theory of $\leq$ (this is not surprising: if there was some weird loop in the natural numbers then $\leq$ would not satisfy basic properties which we require from it like transitivity).

Here are the definitions.

Definition 3.3.5: The ordering on the naturals

Let x and y be natural numbers. We say that $x \leq y$ if there exists a natural number n such that $y = x + n$.

Let's do some warm-up exercises.

Lemma 3.3.4

Let x be a natural number. Then

a) $0 \leq x$;

b) $x \leq x$;

c) $x \leq S(x)$;

d) If $x \leq 0$ then $x = 0$.

Proof. *a*) What does the question *mean*? By *definition* of $\leq$ we need to prove that there exists a natural number n such that $x = 0 + n$. We can choose n to be whatever we like, so let's choose $n = x$ and now we have to prove $x = 0 + x$. But we already proved this in lemma 3.3.2.

b) Similarly here, the question means that we have to prove there exists a natural number n such that $x = x + n$. Just take $n = 0$, because $x + 0 = x$ by definition of addition.

c) The question means that we have to prove that there exists n such that $S(x) = x+n$. But we know $S(x) = x + 1$ from Lemma 3.3.1, so $n = 1$ will work.

d) Say $x \leq 0$. Then by definition, there is some natural number n such that $x + n = 0$. By Lemma 3.3.6(c) we deduce $x = n = 0$ and in particular $x = 0$.

■

The main theorem that we want is that $\leq$ is a total order on $\mathbb{N}$. Let's recall what this means.

Theorem 3.3.7: $\leq$ is a total order

Say $x, y, z \in \mathbb{N}$.

a) $x \leq x$ (reflexivity);

b) If $x \leq y$ and $y \leq z$ then $x \leq z$ (transitivity);

c) If $x \leq y$ and $y \leq x$ then $x = y$ (antisymmetry);

d) Either $x \leq y$ or $y \leq x$ (totality).

Proof. We've already done the first part; I'll leave the next two to you, and will do the last one.

a) This is Lemma 3.3.4 b).

b) For this one you can *assume* that $x \leq y$ and $y \leq z$, and you need to *prove* that $x \leq z$. Use the definition of $\leq$ and finish the proof as an exercise.

c) Don't use induction – use some of the results in theorem 3.3.6.

d) This one is fiddly! Let's fix x and do induction on y. Formally, the statement P_n we're proving by induction on n is that $x \leq n$ or $n \leq x$.

For the base case we need to prove $x \leq 0$ or $0 \leq x$. But we already proved in Lemma 3.3.4(a) that $0 \leq x$, and this finishes the base case.

For the inductive step, we may assume $x \leq n$ or $n \leq x$, and we want to prove $x \leq S(n)$ or $S(n) \leq x$. Our assumption says that one of two things is true; let's deal with the two cases separately.

The first case is where $x \leq n$. We need to prove $x \leq S(n)$ or $S(n) \leq x$. Let's prove $x \leq S(n)$; this follows from $x \leq n$ (our assumption), $n \leq S(n)$ (lemma 3.3.4(c)), and transitivity (part b of this lemma).

The second case is the trickiest part of the proof. We assume $n \leq x$ and we need to deduce that either $x \leq S(n)$ or $S(n) \leq x$. The reason this is the hardest case is that we still don't know which of these is going to be true; if $x = n$ then the first one will be true, and if x is much bigger than n then it will be the second one, so we have to break things down even further. We know $n \leq x$ so we know there exists y such that $x = n + y$. We also know that y is either 0 or of the form $S(t)$ for some natural t. So let's do a final case split and finish this problem.

If $y = 0$ then $x = n + 0 = n$, so we can prove $x \leq S(n)$; indeed $n \leq S(n)$ was proved already in Lemma 3.3.4(c).

Finally, if $y = S(t)$ then $x = n + S(t) = S(n + t) = S(n) + t$ by definition of $+$ and lemma 3.3.3, and this means that $S(n) \leq x$ by definition of $\leq$.

■

Theorem 3.3.8: Interaction between $\leq$ and $+$, $\times$

Let a, b and x be natural numbers.

a) If $a \leq b$ then $a + x \leq b + x$.

b) If $a \leq b$ then $a \times x \leq b \times x$.

Proof. Here is a sketch. You should check that you can fill in the details, and that I definitely didn't assume anything which we didn't prove yet.

a) Choose n such that $b = a + n$. Now check that $(b + x) = (a + x) + n$.

b) Choose n such that $b = a + n$. Now check that $bx = ax + nx$.

■

3.3.11 Variants of the induction principle

Now we know about $\leq$ we can state and prove some variants of the principle of mathematical induction. The first is an easy one: you don't have to start at zero. What I mean by this is: if you have some fixed "base" natural number k, and true-false statements Q_k, Q_{k+1}, Q_{k+2},..., then to prove them all it suffices to prove the base case Q_k and the inductive step that for all $n \geq k$, Q_n implies Q_{n+1}. How do we prove that this is valid? Simple: just define P_n to be Q_{k+n} and apply the usual induction principle to the P_n.

A more subtle variant is *strong induction*. Again we have statements P_0, P_1, P_2, ..., and we want to prove all of them. Strong induction relies crucially on the concept of $<$, which we haven't talked about yet and which we develop the theory of in the first Part 2 example sheet. So let's define $x < y$ to mean $S(x) \leq y$, and then we can state the principle of strong induction.

Theorem 3.3.9: The principle of strong induction

Say P_0, P_1, P_2, ... are infinitely many true-false statements. To prove that all of them are true, it suffices to prove the following statement:
(*) For every n, you can deduce P_n if you assume P_t for all $t < n$.

Proof. Let's define infinitely many new true-false statements Q_n by $Q_n = P_0 \wedge P_1 \wedge \ldots \wedge P_n$. This definition is no good because we used ..., but we can define Q_n properly by recursion: we define Q_0 to be P_0, and if we've defined Q_n already then we define $Q_{S(n)}$ to be $Q_n \wedge P_{S(n)}$.

Exercise: check that Q_2 is $(P_0 \wedge P_1) \wedge P_2$.

We now prove Q_n by induction on n. To do the base case we use (*) with $n = 0$; this says that you can deduce P_0 from no hypotheses at all, so in particular P_0 must be true, and hence Q_0 is true.

To do the inductive step, we assume Q_n and we need to deduce $Q_{S(n)}$. Now if $Q_n = P_0 \wedge P_1 \wedge \ldots P_n$ is true then P_t is true for all $t < S(n)$, so by (*) applied at $S(n)$ we deduce that $P_{S(n)}$ is true. Hence $Q_{S(n)} = Q_n \wedge P_{S(n)}$ is true.

Hence Q_n is true for all n. But it is easy to check that $Q_n \implies P_n$ for all n (check the 0 and successor cases separately), and hence P_n is true for all n. ■

Using strong induction we can deduce that the naturals are *well-ordered*, meaning that if a set of naturals has one or more elements, then it has a smallest element. Before

we embark on this, you might want to think about the following exercise, which shows you "the point" of being well-ordered.

Exercise: This section is about the non-negative integers. But assume temporarily that we have already created the non-negative real numbers and all the usual properties we know about them are true. Show that the non-negative real numbers are *not* well-ordered, by proving that the set of positive real numbers is a nonempty set with no smallest element.

Proposition 3.3.2: Well-ordering principle

Every non-empty set $A \subseteq \mathbb{N}$ of natural numbers has a least element, i.e., there exists a number $a \in A$, such that for every $x \in A$ we have $a \leq x$.

Proof. We will prove this by contradiction. Let $A \subseteq \mathbb{N}$ be a set, and assume that it has no least element; we'll show that A is empty. If n is a natural number, define P_n to be the true-false statement "$n \notin A$". Let's prove that P_n is true for all n; this is enough, as it shows that A has no elements.

We prove P_n for all n, by strong induction. This means that for a fixed n, we have to check that if P_t is true for all $t < n$ then P_n is also true. But if P_t is true for all $t < n$ then this means (by definition of P_t) that every number less than n is not in A. So if $n \in A$ then n would be the least element of A, a contradiction. We deduce that n can't be in A either, so P_n is true, and we are done. ∎

3.3.12 Division, divisibility and primes

We can't do exact division in the natural numbers (1 and 2 are natural numbers, but $1/2$ isn't). However we can do division with remainder.

Proposition 3.3.3: Quotient-Remainder Theorem

Let $a, b \in \mathbb{N}$ with $b > 0$. Then there exist naturals q (quotient) and r (remainder) with $0 \leq r < b$ such that

$$a = bq + r.$$

Proof. Fix b and use do induction on a. For the base case we need to solve $0 = qb + r$ so we just let $q = r = 0$. For the inductive step can assume $d = q'b + r'$ and we need to find q and r with $r < b$ such that $d + 1 = qb + r$. There are two cases. If $r' + 1 < b$ then just let $q = q'$ and $r = r' + 1$. If however $r + 1 = b$ then let $q = (q' + 1)$ and $r = 0$. In either case it's easy to check that these choices work. ∎

Even though we can't do exact division in naturals, we can at least record when it would work.

Definition 3.3.6

Let $n, m \in \mathbb{N}$. We say that m *divides* n if there exists a number $k \in \mathbb{N}$, such that $n = m \times k$. Notation: $m \mid n$. We also say that n is *divisible* by m.

Note that this is the analogue of the definition of $\leq$, with addition replaced by multiplication.

Definition 3.3.7

A natural number $n \geq 2$ which is divisible only by 1 and itself is called a prime number. An integer greater than 1 which is not prime is called composite.

Note that 0 and 1 are neither prime nor composite; 0 is zero, and 1 is called a *unit*.

By "a finite product of prime numbers" below, we mean "a product of finitely many prime numbers", so for example $84 = 2^2 \times 3 \times 7$.

Proposition 3.3.4: Prime factorization

Every natural number greater than one can be factored as a finite product of prime numbers.

Proof. Let A be the set of natural numbers greater than one which are *not* a product of primes. We want to prove that this set is empty. If it's not empty, then it has a least element by Proposition **??**. This least element n is in A and hence greater than one (by definition of A) so is either prime or composite. If it's prime then it's the product of one prime number! And if it's composite then we can write $n = ab$ with $1 < a, b < n$. Because n is the smallest element of A, we must have that a, b are not in A, but they're greater than 1, so they must be the product of some finite number of primes. Hence $n = ab$ is also a product of finitely many primes, a contradiction. ■

If you think about it, 1 is also a product of prime numbers: it's the product of no prime numbers at all (for the same reason that $2^0 = 1$).

Here is another consequence of these ideas.

Theorem 3.3.10

There are infinitely many primes.

Proof. We will prove this result by contradiction. Assume that there are only finitely many primes $p_1, p_2, \ldots, p_n$, $n \in \mathbb{N}$ and define

$$X = p_1 p_2 \ldots p_n + 1.$$

By the previous proposition, X must have a prime factor q. Then

$$q | p_1 p_2 \ldots p_n + 1.$$

But also q must be one of the p_i because those are all of the primes. Hence

$$q | p_1, p_2, \ldots, p_n.$$

Hence q divides 1, which is a contradiction. There are therefore infinitely many primes. ■

3.3.13 Euclid's algorithm

Say we have two natural numbers a and b. Here is an algorithm which generates a sequence r_0, r_1, r_2, ... and then outputs another natural number. It's called Euclid's algorithm.

Definition 3.3.8: Euclid's Algorithm

STEP 0) Let r_0 be the largest of a and b, and let r_1 be the smallest. Let $n = 0$ (this number will increase as the algorithm runs).
STEP 1) If $r_{n+1} = 0$ then stop and return r_n.
STEP 2) If instead $r_{n+1} > 0$ then write $r_n = qr_{n+1} + r_{n+2}$ with $r_{n+2} < r_{n+1}$.
STEP 3) Increase n by 1 and go back to step 1.

Let's see an example. Let's start with $a = 28$ and $b = 20$. We set $r_0 = 28$ and $r_1 = 20$ as the first two terms in our sequence. And then we start dividing the last-but-one term in the sequence by the last term, and we adjoin the remainder to the sequence.

$$28 = 1 \times 20 + 8$$
$$20 = 2 \times 8 + 4$$
$$8 = 2 \times 4 + 0$$

We reached 0 so we stop, our sequence of r's is $28, 20, 8, 4, 0$, and we return the last nonzero r which is 4.

Let's prove some facts about the number which is output by this algorithm.

Theorem 3.3.11: Results about Euclid's algorithm

a) The output of the algorithm is always positive, unless both inputs a, b are zero (in which case it's zero).

b) If the output is d, then d divides all of the r_i in the sequence.

c) If x is any number which divides a and b, then it also divides the output d.

Proof. *a*) If we get as far as step 2 then $r_1 > 0$, and then all the r_i are positive until one is zero and then we return the one before, which must be positive. So the only way the algorithm can return 0 is if $r_1 = 0$, in which case it returns r_0, the maximum of a and b. So if it returns 0 then they must both have been 0. Conversely, if they are both 0 then it does indeed return 0.

b) Say the algorithm returned $d = r_n$, and $r_{n+1} = 0$. Then d divides both r_n and r_{n+1}, so it divides $r_{n-1} = qr_n + r_{n+1}$. Continuing this way, we deduce that d divides r_1 and r_0, so it divides a and b.

c) Say x divides a and b. Then it divides r_0 and r_1. Hence it divides $r_2 = r_0 - qr_1$, and then $r_3 = r_1 - q'r_2$ and so on. By induction it divides all the r_i and in particular it divides d.

■

We conclude that the output d divides a and b, and furthermore any common divisor x of a and b must also divide d.

If $d \neq 0$ then any divisor of d must be at most d, so this means that d is the greatest of all the common divisors of a and b. If $d = 0$ then this means $a = b = 0$ so d still does have the property that any divisor of a and b divides d, because *everything* divides 0. In this case d is not the *greatest* common divisor because $37 > 0$ and 37 divides 0. However it is still the divisor that all the other divisors divide.

We have just shown this:

Definition 3.3.9: greatest common divisor

We define the *greatest common divisor* of a and b to be the output d of Euclid's algorithm. Notation: $d = gcd(a, b)$.

Theorem 3.3.12: greatest common divisor facts

If at least one of a and b is positive, then d is the greatest of the common divisors of a and b, and it is even a *multiple* of all common divisors of a and b.

Example: $gcd(28, 20) = 4$.

Note that computing the greatest common divisor of two numbers via Euclid's algorithm is *much* more efficient than factoring both sides, the moment the numbers involved are bigger than 200 or so.

Developing the theory of prime numbers and divisors any further is pretty annoying without a theory of subtraction, and developing a theory of subtraction is pretty annoying without having access to negative numbers. So we stop our development of naturals here, and move on to an explanation of the integers.

3.4 The integers

3.4.1 Construction

We have developed a theory of addition, but there are equations involving addition such as $x + 3 = 2$ which we cannot solve in natural numbers. The approach we will take to fix this is to define a *new* number system, the *whole numbers*, commonly referred to as the *integers*:

$$\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}.$$

Of course we cannot use this as a definition because it has $\ldots$ in. So how do we construct the integers formally from the naturals? We will explain a beautiful method using the theory of equivalence relations and equivalence classes. Before we launch into it, let's jump ahead a little and think about how we might make the positive *rational* numbers like $3/7$, because the same issues will show up here and I think they're a bit easier to understand.

Let's define the positive naturals $\mathbb{N}_{>0} = \{1, 2, 3, \ldots\}$ to be $\mathbb{N}$ with 0 removed. Now, every positive rational number q can be expressed as a ratio a/b of positive naturals, and in particular can be built from two positive naturals. So how about we just *define* the positive rationals to be $\mathbb{N}_{>0} \times \mathbb{N}_{>0}$, the set of ordered pairs of positive naturals? The ordered pair $(3, 7)$ would then correspond to the rational number $3/7$.

But there is a problem here: we have *too many* positive rationals with this definition. The ordered pairs $(1, 2)$ and $(2, 4)$ are different pairs (they are different vectors in the plane, if you like), but the rational numbers $1/2$ and $2/4$ are *equal*. There isn't just *one* way to write a positive rational q as a ratio a/b, there are *lots* of ways.

What we want to do then, is to start with $\mathbb{N}_{>0} \times \mathbb{N}_{>0}$ (which is too big) and then "squish some elements together" and somehow make $(1, 2)$ and $(2, 4)$ "equal" even though they are not actually equal. Once we have squished $(1, 2)$ and $(2, 4)$ and $(3, 6)$ and $(n, 2n)$ together, and then squished together all the other different pairs of positive naturals which give rise to the same rational, then we really will have defined the positive rationals.

This squishing together of different elements of a set is formally done by putting an *equivalence relation* on the set, and then considering the *equivalence classes*. The axioms of an equivalence relation are an attempt to mimic and generalise the concept of equality. When you quotient out a set by an equivalence relation by taking the set of equivalence classes, the concept of equivalence in the bigger set *becomes* the concept of equality in the quotient set: if $x \sim y$ are equivalent, then the equivalence classes $cl(x)$ and $cl(y)$ are *equal*.

Note that there is another possibility here: instead of squishing different pairs together, we could decide to *reject* a pair (a, b) if it's not "in lowest terms", i.e. if there is some prime number p dividing both a and b. This might superficially look simpler, but it causes real problems later on. For example the standard formula $\frac{a}{b} \times \frac{c}{d} = \frac{ac}{bd}$ for multiplying fractions would no longer be valid if we took this approach, because even if a/b and c/d are in lowest terms, ac/bd might not be. If you define the product to be "put $\frac{ac}{bd}$ in lowest terms" then proving things like associativity of multiplication becomes a *nightmare*. Instead of having to manipulate ac/bd to put it into lowest terms, our approach is to allow non-lowest-terms fractions but set things up so that they are *equivalent* to the corresponding lowest term fraction, and then use equivalence classes.

We'll come back to the rationals in the next chapter, but let's use the ideas above to make the integers. Just like every positive rational can be written as the ratio of two natural numbers, every integer can be written as the *difference* of two natural numbers. For example, $3 = 5 - 2$ and $-3 = 5 - 8$. And, just like in the case of fractions, there is more than one way to do this (for example $-3 = 5 - 8 = 6 - 9 = 7 - 10 = \cdots$).

So let's start with pairs $\mathbb{N} \times \mathbb{N}$ of naturals, with the idea that the pair (a, b) of naturals is supposed to represent the integer $a - b$. Next we need to identitfy pairs (a, b) and (c, d) of naturals if they represent the same integer, that is, if $a - b = c - d$. But there is a big problem here: this argument is *circular*, because we haven't defined the integers or subtraction yet, so "$a - b = c - d$" *doesn't make sense* at this point. Can we rewrite this equation in such a way that it doesn't mention subtraction? Yes we can: adding $b + d$ to both sides turns the equation into $a + d = b + c$, which makes perfect sense. But is it an equivalence relation?

Proposition 3.4.1

The following binary relation on $\mathbb{N} \times \mathbb{N}$

$$(a, b) \sim (c, d) \text{ if and only if } a + d = b + c$$

is an equivalence relation.

Proof. Reflexivity: We have to show that $(a, b) \sim (a, b)$. By commutativity of addition we have immediately $a + b = b + a$, hence $\sim$ is reflexive.

Symmetry: Here we must show that if $(a, b) \sim (c, d)$, then $(c, d) \sim (a, b)$. By definition, $(a, b) \sim (c, d)$ means the same as $a + d = b + c$, and similarly $(c, d) \sim (a, b)$ if and only if $c + b = d + a$. So we need to show that $a + d = b + c \implies c + b = d + a$. This follows, because equality is symmetric and addition is commutative.

Transitivity: Assume that $(a, b) \sim (c, d)$ and $(c, d) \sim (e, f)$ or in other words $a+d = b+c$ and $c + f = d + e$. We want to show that $a + f = b + e$. Adding $e + f$ to both sides of the first identity, we deduce $(a + d) + (e + f) = (b + c) + (e + f)$. Now rearraging using commutativity and associativity of addition, we deduce

$$(a + f) + (d + e) = (b + e) + (c + f).$$

Using the second identity $c+f = d+e$ we deduce that $(a+f)+(d+e) = (b+e)+(d+e)$, and now using the cancellation law for addition (Theorem 3.3.6(a)) we deduce $a + f = b + e$, which was what we had to prove. ■

Definition 3.4.1

The set of integers $\mathbb{Z}$ is the set of all equivalence classes for $\mathbb{N} \times \mathbb{N}$ under this equivalence relation.

Note that we did not add any new axioms here, we just made definitions.

3.4.2 Relationship with the naturals

We've defined the naturals $\mathbb{N}$ and the integers $\mathbb{Z}$. But can you see that the standard inclusion $\mathbb{N} \subseteq \mathbb{Z}$ is *not actually true*? The natural number 2 is just $S(S(0))$, whereas the integer 2 is a totally different object: it is an infinite equivalence class $\{(2,0),(3,1),(4,2),\ldots\}$. However we really *want* $\mathbb{N} \subseteq \mathbb{Z}$ to be true. What to do about this?

Let's define a *function* i from $\mathbb{N}$ to $\mathbb{Z}$, sending a natural number n to the equivalence class of $(n,0)$.

Lemma 3.4.1

The function i is injective. In other words, if x, y are two natural numbers and $i(x) = i(y)$, then $x = y$.

Proof. Say x, y are natural numbers, and $i(x) = i(y)$. Then by definition $cl((x,0)) = cl((y,0))$, which means that $(x,0) \sim (y,0)$. By definition this means that $x + 0 = 0 + y$. Hence $x = y$. ■

Mathematicians often *identify* $\mathbb{N}$ with its image $i(\mathbb{N})$ in $\mathbb{Z}$ to avoid this problem. This causes some formal subtleties, which are usually ignored. We'll talk about some of these on the second example sheet.

Definition 3.4.2: The integers 0,1,2

We define $0 \in \mathbb{Z}$ to be $i(0)$ where that last 0 is the natural 0, we define $1 \in \mathbb{Z}$ to be $i(1)$, we define $2 \in \mathbb{Z}$ to be $i(2)$.

3.4.3 Addition, subtraction, multiplication

To test that our definitions are behaving the way we expect, we could check that $2-1 = 1$. But wait: we haven't defined subtraction yet! Or addition or multiplication for that matter. Now we have a new number system $\mathbb{Z}$ we *strictly speaking* have to start again and define $+$ and $\times$ on $\mathbb{Z}$, and we also have to define subtraction, and prove all the standard facts about how they are related such as $(x - y) \times z = x \times z - y \times z$. Fortunately this job is easier than the corresponding job for the natural numbers. I'll get you to do some of it on the second Part II example sheet.

Let's talk about subtraction, because we've not seen it before. A cool definition of subtraction could be to *define* $x-y$ to be $x+(-y)$, where that $-y$ is "minus y", the *negation* of y. Note that even though we use the same symbol $-$ for both ideas, subtraction and

negation are strictly speaking two different things: subtraction eats two integers and spits out an integer; negation only eats one. Let's define negation. This involves introducing a new idea.

The integers $\mathbb{Z}$ were defined to be the *quotient* of the auxiliary object $\mathbb{N}\times\mathbb{N}$ by a certain equivalence relation. So the integer 2 is actually the infinite set $\{(2,0),(3,1),(4,2),(5,3)\ldots\}$ (remember that an ordered pair (a,b) of naturals should be thought of as representing the integer $a-b$, even though this doesn't make sense yet). Let's think of the integers as obtained from $\mathbb{N}\times\mathbb{N}$ by "squishing together" equivalent pairs into one object.

Now a function *has* to have the property that if give it the same input twice, you get the same output. So for example if $a=b$ and f is a function then $f(a)$ has to equal $f(b)$ (note that this is not injectivity; injectivity is the implication the other way. This is an axiom of mathematics). The problem with all this "squishing together" idea is that if two things got squished together then in the quotient they are *equal*, so any function had better do the same thing to them – it's can "unsquish them".

The upshot of all this: if we want to define negation $f(x)=-x$ on the integers, then we should

1. First define a version of negation on $\mathbb{N}\times\mathbb{N}$; then
2. Second check that it sends equivalent inputs to equivalent outputs.

The second condition ensures that no "unsquishing" is taking place.

Let's pretend the integers exist for 10 seconds, and ask what the negation of $a-b$ is; it's $b-a$. This motivates the following definition.

Definition 3.4.3: pre-negation

Define a "pre-negation" function f from $\mathbb{N}\times\mathbb{N}$ to $\mathbb{N}\times\mathbb{N}$, sending (a,b) to (b,a).

Now let's prove that pre-negation *preserves the equivalence relation*. Informally I mean that we need to check that "pre-negation doesn't unsquish squished things". Formally I mean this:

Theorem 3.4.1: pre-negation plays well with equivalence

If $(a,b)\sim(c,d)$ then $f(a,b)=f(c,d)$.

Proof. Our hypothesis is that $a+d=b+c$, and the conclusion we want is that $(b,a)\sim(d,c)$ or in other words that $b+c=a+d$. But this obviously follows (you can say "...because equality is symmetric" if you want to sound clever). ■

Definition 3.4.4: negation on $\mathbb{Z}$

efine negation on $\mathbb{Z}$ (i.e. the function sending x to $-x$) to be the function sending the equivalence class of (a,b) to the equivalence class of (b,a).

We have just proved that this is a valid mathematical definition. The potential issue is that you can have different elements (a,b) and (c,d) of $\mathbb{N}\times\mathbb{N}$ which are equivalent but not equal, and then their equivalence classes are equal, so we have to make sure this function maps them to equal things, but it does by the preceding theorem.

Now let's do addition, subtraction and multiplication.

For addition we first do the following calculation, pretending that the integers already have a subtraction: $(a-b)+(c-d)=(a+c)-(b+d)$, so this motivates the following:

Definition 3.4.5: pre-addition

Define pre-addition to be the function which takes two elements (a, b) and (c, d) of $\mathbb{N} \times \mathbb{N}$ and returns the element $(a + c, b + d)$.

Theorem 3.4.2: pre-addition plays well with equivalence

If $(a, b) \sim (a', b')$ and $(c, d) \sim (c', d')$ then $(a + c, b + d) \sim (a' + c', b' + d')$.

Proof. Our hypotheses are $a+b' = b+a'$ and $c+d' = d+c'$. Our goal is $(a+c)+(b'+d') = (b+d)+(a'+c')$. But the left hand side rearranges to $(a+b')+(c+d')$ and the right hand side to $(b + a') + (d + c')$ and these are equal by our assumptions. ■

This means that pre-addition cannot unsquish things which are already squished – if two things are equivalent and you pre-add them, they stay equivalent. This means that the definition of pre-addition decends to $\mathbb{Z}$.

Definition 3.4.6: addition on $\mathbb{Z}$

Define addition on $\mathbb{Z}$ by saying that $cl(a, b) + cl(c, d) = cl(a + c, b + d)$.

We don't need presubtraction; we can just define subtraction given what we already have.

Definition 3.4.7: subtraction on $\mathbb{Z}$

If $x, y \in \mathbb{Z}$ then define $x - y$ to be $x + (-y)$.

Finally, multiplication. First the thought experiment: $(a - b) \times (c - d) = (ac + bd) - (ad + bc)$. Now the pre-definition.

Definition 3.4.8: pre-multiplication

If (a, b) and (c, d) are in $\mathbb{N} \times \mathbb{N}$, define their pre-product to be $(ac + bd, ad + bc)$.

Theorem 3.4.3

If $(a, b) \sim (a', b')$ and $(c, d) \sim (c', d')$ then $(ac + bd, ad + bc) \sim (a'c' + b'd', a'd' + b'c')$.

Proof. This is probably the nastiest calculation in the theory. We assume $a + b' = b + a'$ and $c + d' = d + c'$, and we need to check that $ac + bd + a'd' + b'c' = ad + bc + a'c' + b'd'$. Remember that these are naturals so we can't do subtraction. This makes things a little tricky! Here is a proof.

Consider the following calculation:

$$\begin{aligned}(ac + bd + a'd' + b'c') + (ad' + bc' + bd' + ac') &= a(c + d') + b(d + c') + (a' + b)d' + (b' + a)c' \\ &= a(d + c') + b(c + d') + (b' + a)d' + (a' + b)c' \\ &= (ad + bc + a'c' + b'd') + (ad' + bc' + bd' + ac').\end{aligned}$$

Now cancelling $(ad'+bc'+bd'+ac')$ (which we are allowed to do by lemma 3.3.6(a)) gives us the result. ■

Definition 3.4.9: multiplication

We define $cl(a,b) \times cl(c,d)$ to be $cl(ac+bd, ad+bc)$.

By definition, $0 \in \mathbb{Z}$ is the equivalence class of $(0,0)$, and $1 \in \mathbb{Z}$ is the equivalence class of $(1,0)$. So now you can check the analogoues of all of the facts about addition and multiplication which we showed for the naturals. Here is an example

Theorem 3.4.4: distributivity

If x, y, z are integers, then $x(y+z) = xy + xz$.

Proof. Let's say x is the class of (a,b), y is the class of (c,d) and z is the class of (e,f). Then the left hand side is the class of $(a,b) \times (c+e, d+f)$ which is $(ac+ae+bd+bf, ad+af+bc+be)$, and the right hand side is the class of $(ac+bd, ad+bc) + (ae+bf, af+be)$ which is $(ac+bd+ae+bf, ad+bc+af+be)$ and these are equal because of standard facts from the theory of addition (commutativity and associativity). ■

From this point on, let us use addition, subtraction and multiplication normally, because we have shown how to justify all the usual rules.

3.4.4 Back to GCDs

In the section on naturals we defined the greatest common divisor of two naturals. If the naturals are $r_0 \geq r_1$ then we defined a decreasing sequence $r_1 > r_2 > r_3 > \cdots > r_n > r_{n+1} = 0$ and returned r_n. Here is another important fact about the r_i which can easily be proved by induction, but only when you have subtraction available.

Theorem 3.4.5: Euclid revisited

If $d = gcd(a,b)$ then there are integers λ and μ such that $d = \lambda a + \mu b$.

Proof. In fact we prove more. Recall that d is one of the r_i; we prove that all the r_i can be written in this way. For a and b it's easy: $a = 1a + 0b$ and $b = 0a + 1b$. And if it's true for r_n and r_{n+1}, then becasue $r_{n+2} = r_n - qr_{n+1}$ we deduce the result for r_{n+2} as well. ■

We say that two natural numbers a and b are *coprime* if $gcd(a,b) = 1$. The following corollary is immediate from the theorem.

Corollary 3.4.1

If a and b are coprime, then there exist integers λ and μ such that $\lambda a + \mu b = 1$.

We use this result to prove a crucial fact about prime numbers.

3.4.5 Primes and unique factorization

We have already proved that every positive integer factors into primes. But we have not yet proved that the factorization is *unique*. It can sometimes be hard to explain to people what the issue is here. Let me attempt to do so. Say $p < q < r < s$ are prime numbers. Say that you are *only* allowed to use the definition of a prime number (that its only factors are 1 and itself). How do you deduce from this that $ps \neq qr$? The problem is that ps and qr are both probably *bigger* than p, q, r, s, so the definition of a prime number (which only tells you things about numbers *less* than the prime) isn't much help.

The following crucial lemma gets us out of this mess.

Theorem 3.4.6: key property of prime numbers

If p is prime and if a, b are natural numbers, and if $p \mid ab$, then $p \mid a$ or $p \mid b$.

See problem sheet 2 for a demonstration of how this and other ideas do not generalise away from primes.

Proof. Assume p is prime and $p \mid ab$. Let's assume the $p \nmid a$, and we'll show that $p \mid b$.

What can we say about $gcd(p, a)$? Certainly it divides p, which means that it must be 1 or p, because p is prime. But it can't be p because $p \nmid a$. Hence it's 1. Thus by Corollary 3.4.1 there exist integers λ and μ such that $\lambda a + \mu p = 1$. Multiplying both sides by b we deduce $\lambda ab + \mu pb = b$. But p divides ab and hence p divides the left hand side of this equality, so p divides the right hand side, which is b. ■

Corollary 3.4.2

If p is prime and if $a_1, a_2, \ldots, a_n$ are natural numbers, and if $p \mid a_1 a_2 \ldots a_n$ then p divides one of the a_i.

Proof. Induction on $n \geq 1$. The base case $n = 1$ has $p \mid a_1$ as an assuption, so we're done. For the inductive step assume that if $p \mid a_1 a_2 \ldots a_n$ then p divides one of the a_i, and we have to deduce that if $p \mid a_1 a_2 \ldots a_n a_{n+1} = (a_1 a_2 \ldots a_n) a_{n+1}$ then it divides one of the a_i. By Theorem 3.4.6 we know p divides either $a_1 a_2 \ldots a_n$ or a_{n+1}. If p divides a_{n+1} we're done immediately, and if p divides $a_1 a_2 \ldots a_n$ then we're done by the inductive hypothesis. ■

Theorem 3.4.7: Uniqueness of prime factorization

Every positive integer is *uniquely* the product of prime numbers, up to re-ordering.

Proof. We have already shown that every positive integer can be written as the product of primes in theorem 3.3.4. What is left to show that if

$$n = p_1 p_2 \ldots p_r = q_1 q_2 \ldots q_s$$

with the p_i and the q_j all prime numbers, then $r = s$ and after possibly re-ordering the q_i, we have $p_i = q_i$ for all i.

If there are any counterexamples to this question, then there will be a smallest counterexample by Proposition 3.3.2, so let n be the smallest counterexample.

By switching the p's and q's if necessary, we can assume $r \leq s$. If $r = 0$ then $p_1 p_2 \dots p_r = 1$ (the product of no things is 1), so $1 = q_1 q_2 \dots q_s$ meaning that s must also be 0, because if $s \geq 1$ then $q_1 q_2 \dots q_s > 1$, and n wasn't a counterexample after all. If however $r \geq 1$ then consider the last prime p_r on the left hand side. It divides $p_1 p_2 \dots p_r$, so it divides $q_1 q_2 \dots q_s$. By the previous corollary p_r must divide one of the q_i. By rearranging the q_i we can assume that it's q_s. But the only divisors of q_s are 1 and q_s, because q_s is prime. Hence $p_r = q_s$, meaning that we can cancel and get $p_1 p_2 \dots p_{r-1} = q_1 q_2 \dots q_{s-1}$. But this number is less than n so isn't a counterexample to unique factorization. Hence we must have $r - 1 = s - 1$ and all the remaining q_j are a permutation of the p_i, so n wasn't a counterexample after all. This contradiction finishes the proof. ■

Something else we can do with subtraction is to prove that the quotient and remainder whose existence we proved in Proposition 3.3.3 are unique! In fact we can extend the quotient-remainder result to integers, as long as the integer we're dividing by is positive.

Proposition 3.4.2: quotient-remainder for integers

If a is an integer and b is a positive integer, there exists integers q and r with $0 \leq r < b$ and $a = qb + r$.

Proof. See problem sheet 2. ■

Note that b is positive and the remainder r is non-negative even if a is negative. For example if we divide -37 by 10 using this method, we get $q = -4$ and $r = 3$.

Theorem 3.4.8: uniqueness of quotient and remainder

If a is an integer, b is a positive integer, and $a = qb + r = q'b + r'$ with $0 \leq r, r' < b$ then $q = q'$ and $r = r'$.

Proof. By swapping q, r with q', r' if necessary, we can assume $r \leq r'$. Now subtracting the equations gives $(q - q')b = r' - r$, so $r' - r$ is a multiple of b. But it's ≥ 0 and $< b$, so it can't be $b \times S(x)$ because this is $bx + b \geq b$ and hence too big. So it must be $b \times 0$, meaning $r = r'$. Hence $qb = q'b$ and because $b \neq 0$ we deduce $q = q'$ from the fact that you can cancel multiplication by non-zero naturals (this needs a proof: see the problem sheets!) ■

3.4.6 Modular arithmetic

We will now give a short introduction to congruences of integers. Here again the notion of equivalence relation is of vital importance. Note that the notion of division extends easily to integers: if x, y are integers then we say $x \mid y$ if there exists an integer z with $y = xz$.

Definition 3.4.10

Let $a, b \in \mathbb{Z}$ and $n \in \mathbb{N}, n > 0$. We say that a is congruent to b modulo n if $n \mid (a - b)$. Notation: $a \equiv b \mod n$.

Example

$22 \equiv 4 \mod 9$, since $22 - 4 = 18$, and $9 \mid 18$.

Proposition 3.4.3

Let $a, b, c \in \mathbb{Z}$, . Then $n \in \mathbb{N}, n > 0$.

1. $a \equiv a \mod n$, for all $a \in \mathbb{Z}$.
2. If $a \equiv b \mod n$, then $b \equiv a \mod n$.
3. If $a \equiv b \mod n$ and $b \equiv c \mod n$, then $a \equiv c \mod n$.

Proof. 1. $a - a = 0$, and $n \mid 0$ for all n, which yields the result.

2. $b - a = -(a - b)$. Hence since $n \mid (a - b)$, then $n \mid -(a - b) = b - a$.
3. If $n \mid (a - b)$ and $n \mid (b - c)$, then $n \mid (a - b) + (b - c) = a - c$, which by definition means that $a \equiv c \mod n$.

■

The last proposition exactly means that congruence modulo n is an equivalence relation on $\mathbb{Z}$. We denote the equivalence class of a by

$$[a]_n := \{b \in \mathbb{Z} : b \equiv a \mod n\}$$

and we call it a *congruence class*.

For example, the congruence class $[7]_{10}$ is all the integers congruent to 7 mod 10, so it's $\{\ldots, -23, -13, -3, 7, 17, 27, 37, \ldots\}$.

Here's a useful result.

Proposition 3.4.4

Let $a, b \in \mathbb{Z}$ and n a positive integer. Then $a \equiv b \mod n$ if and only if a and b have the same remainder after division by n.

Proof. Let $a = qn + r$ and $b = q'n + r'$ be the Euclidean division of a and b by n with $0 \leq r, r' < n$. Then we get immediately $a - b = (q - q')n + (r - r')$.

Assume that $a \equiv b \mod n$. by definition $n \mid (q - q')n + (r - r')$, therefore $n \mid (r - r')$. But $0 \leq r, r' < n$, therefore $-n < r - r' < n$ and the only multiple of n in this range is 0. Hence $r = r'$ and a and b have the same remainder after division by n.

Conversely if $r = r'$, then $a - b = (q - q')n$ and $n \mid a - b$ which by definition means exactly $a \equiv b \mod n$. ■

We want now to find out how many congruence classes exist modulo n. Let us look at it when $n = 1, 2, 3$ to try to understand what is going on.

Note that 1 divides every integer. Hence if a and b are any integers then $a \equiv b \mod 1$. Hence there is only one congruence class modulo 1. In other words

$$[0]_1 = [1]_1 = [2]_1 = \cdots = \mathbb{Z}.$$

If $n = 2$, for any two integers a, b such that $a \equiv b \mod 2$, we have that $a = 2k + b$, which means that $[a]_n = \{\ldots, a-4, a-2, a, a+2, a+4 \ldots\}$, hence a and b have necessarily the same parity (they are either both odd or both even). There are therefore two classes of

integers modulo 2: the even numbers $[0]_2 = \{\ldots, -4, -2, 0, 2, 4, \ldots\}$ and the odd numbers $[1]_2 = \{\ldots, -5, -3, -1, 1, 3, 5, \ldots\}$. We have

$$\cdots = [-2]_2 = [0]_2 = [2]_2 = [4]_2 \ldots \text{ and } \cdots = [-1]_2 = [1]_2 = [3]_2 = [5]_2 = \ldots$$

so $[a] = [a']$ if and only if the integers a and a' are congruent mod n.

If $n = 3$, we have $a = 3k + b$ for some integers k and this reduces to the three cases $3k, 3k+1, 3k+2$. Therefore we get three congruence classes modulo 3, i.e. $[0]_3$, $[1]_3$ and $[2]_3$.

We can recognize a pattern here. In fact we can show the following result.

Proposition 3.4.5

There are exactly n congruences classes modulo n.

Proof. We are first going to show that the equivalence classes $[0]_n, [1]_n, ..., [n-1]_n$ are all different. Let $a, b \in \mathbb{Z}$, $0 \leq a, b \leq n-1$. Assume that $[a]_n = [b]_n$. Then $n \mid a - b$ by Definition. But this is impossible since $-n < a - b < n$ and the only multiple of n in this range is 0, meaning $a = b$.

Now we prove that these are the only possible classes. Again given any integer a, consider the Euclidean division of a by n given by $a = qn + r$, $0 \leq r < n$. Obviously $n \mid a - r$ and consequently $[a]_n = [r]_n$. Therefore since $0 \leq r < n$ any integer a is congruent modulo n to exactly one of the integers $0, 1, \ldots n-1$, which by our previous considerations are not congruent to each other. ■

The last proposition means that every integer is congruent to exactly one of the numbers $0, 1, \ldots, n-1$, so there are exactly n equivalence classes. We write

$$\mathbb{Z}_n =: \{[0]_n, [1]_n, ..., [n-1]_n\}$$

for the set of equivalence classes modulo n. These are infinitely many new worlds of numbers, where, like the integers, we can do addition, subtraction and multiplication (but still not division: you'll have to wait for the next chapter to see that).

From now on n will be fixed, and we'll just write $[a]$ for $[a]_n$. We now want to define operations on this new set. In order to do this we need the following

Lemma 3.4.2: pre-addition and pre-multiplication for integers mod n

Suppose that $a, a', b, b' \in \mathbb{Z}$ are integers such that $a \equiv a' \mod n$ and $b \equiv b' \mod n$. then

1. $a + b \equiv a' + b' \mod n$;
2. $ab \equiv a'b' \mod n$.

Proof. Exercise! See problem sheet 2. ■

This lemma means that if we choose integers a and b and calculate $a + b$ or ab, and then you choose two other representatives $a' \in [a]$ and $b' \in [b]$ and calculate $a' + b'$ or $a'b'$ we get the same element of $\mathbb{Z}_n$. This is the same idea we used to define the addition and multiplication on the integers: we needed our operation to be independent of the

elements of the class. We can therefore define the following well-defined addition and multiplication on $\mathbb{Z}_n$:

$$+ : \mathbb{Z}_n \times \mathbb{Z}_n \to \mathbb{Z}_n, ([a]_n, [b]_n) \mapsto [a+b]_n$$
$$\cdot : \mathbb{Z}_n \times \mathbb{Z}_n \to \mathbb{Z}_n, ([a]_n, [b]_n) \mapsto [ab]_n.$$

We can now ask all the usual questions: whether $(x+y)+z = x+(y+z)$ and $x(y+z) = xy+xz$ and so on for this new world of numbers; the answer is that again all these things are true, and not hard to prove, but let us leave this for now because I want to talk about division.

Chapter 4

Rationals and Reals

4.1 The rationals

We have now all four operations (addition, subtraction, multiplication, division), but we are still not able to divide any integers by any other integer. In other words, we cannot solve yet equations of the type $ax = b$ for any integers a and b. We therefore need a new system of numbers. Similarly to our construction of the integers from the natural numbers, what we want to have is something of the form $x = \frac{b}{a}$, and our idea is to define this as a pair of integers (a, b). But we encounter the same type of problem we had before, since $\frac{1}{2} = \frac{2}{4} = \frac{-1}{-2} \ldots$. We notice that $\frac{m}{n} = \frac{m'}{n'}$ if and only if $mn' = m'n$ and the way to solve our problem is similar to what we did before, meaning that we define rational numbers formally as equivalence classes of pairs of integers.

Definition 4.1.1

Let (m, n), $(m', n') \in \mathbb{Z} \times \mathbb{Z} - \{0\}$ and consider the equivalence relation

$$(m, n) \sim (m', n') \text{ if and only if } mn' = m'n,$$

and define $\frac{m}{n} := cl((m, n)) = \{(x, y) \in \mathbb{Z} \times \mathbb{Z} - \{0\} | (x, y) \sim (m, n)\}$.
$\mathbb{Q}$ is the set of all equivalence classes $\frac{m}{n}$.

We leave as an exercise to check that this is indeed an equivalence relation. Of course we also want to have an addition and a multiplication on $\mathbb{Q}$, satisfying ideally

$$\frac{m_1}{n_1} + \frac{m_2}{n_2} = \frac{m_1 n_2 + n_1 m_2}{n_1 n_2}, \quad \frac{m_1}{n_1} \cdot \frac{m_2}{n_2} = \frac{m_1 m_2}{n_1 n_2}$$

We therefore define these two operations accordingly on $\mathbb{Q}$

$$cl(m_1, n_1) +_{\mathbb{Q}} cl(m_2, n_2) := cl(m_1 n_2 + n_1 m_2, n_1 n_2), \tag{4.1}$$
$$cl(m_1, n_1) \cdot_{\mathbb{Q}} cl(m_2, n_2) := cl(m_1 m_2, n_1 n_2). \tag{4.2}$$

Here we use the notation $+_{\mathbb{Q}}$ and $\cdot_{\mathbb{Q}}$ to distinguish addition and multiplication in $\mathbb{Q}$ from the usual addition and multiplication in $\mathbb{Z}$. As for the integers these operations are compatible with the equivalence relation, i.e. they do not depend on the representative of the class and therefore are well defined on $\mathbb{Q}$.

Proposition 4.1.1

Let (m_i, n_i), $(m_i', n_i') \in \mathbb{Z} \times \mathbb{Z}$, $i = 1, 2$. If $(m_1, n_1) \sim (m_1', n_1')$ and $(m_2, n_2) \sim (m_2', n_2')$,

then

$$\begin{aligned} cl(m_1, n_1) +_{\mathbb{Q}} cl(m_2, n_2) &= cl(m_1', n_1') +_{\mathbb{Q}} cl(m_2', n_2'), \\ cl(m_1, n_1) \cdot_{\mathbb{Q}} cl(m_2, n_2) &= cl(m_1' n_1') \cdot_{\mathbb{Q}} cl(m_2', n_2') \end{aligned}$$

Proof. Exercise! ■

In the following we will just denote the two operations by $+$ and $\cdot$ as usual.

Remark. *We want the rational numbers to be an extension of the integers. We therefore identify $\mathbb{Z}$ with an appropriate subset of $\mathbb{Q}$ in the following way as we did to map the natural numbers inside of the integers. Consider the map*

$$i : \mathbb{Z} \to \mathbb{Q}, n \mapsto i(n) = \frac{n}{1}.$$

We leave as an exercise to show that his map is injective and that therefore $\mathbb{Z} \subset \mathbb{Q}$.

We can now define an ordering on $\mathbb{Q}$.

Definition 4.1.2

Let $\frac{a}{b}, \frac{c}{d} \in \mathbb{Q}$. Assuming $b, d > 0$, we say that $\frac{a}{b} <_{\mathbb{Q}} \frac{c}{d}$ if and only if $ad < bc$.

Note that again we use the symbol $<_{\mathbb{Q}}$ to denote our ordering relation in $\mathbb{Q}$ so as to distinguish it from the usual ordering $<$ in $\mathbb{Z}$. Note also that, given an element in Q, it is always possible to choose a representative (a, b) with $b > 0$.

Proposition 4.1.2

$\leq_{\mathbb{Q}}$ defines a total order on $\mathbb{Q}$.

Proof. Exercise! ■

From now on you can just assume all the rules you know about rational numbers.

4.2 Fields and ordered fields

The last stage in our quest of constructing numbers is to develop the set of real numbers $\mathbb{R}$. Intuitively if we represent the rational numbers as a points on the number line, it is obvious that there are gaps on the line.

Example

$\sqrt{2}$ is not rational.

Proof. Assume that there exists p and q in $\mathbb{Z}$, such that $\gcd(p, q) = 1$ (i.e. we assume without loss of generality that the fraction cannot be simplified further) and $\sqrt{2} = \frac{p}{q}$. Then $p^2 = 2q^2$ which means that $2|p^2$ and consequently $2|p$ as we proved on problem sheet 0. But then $4|p^2$ and consequently $2|q^2$. Again this yields $2|q$. This is a contradiction to the fact that $\gcd(p, q) = 1$. ■

This gaps are exactly the real numbers and they can actually be constructed from the rationals in various different ways: 1) as limits of (Cauchy) sequences of rational numbers, 2) with Dedekind cuts, 3) with nested interval..., but these constructions are tedious and

go beyond the scope of this lecture. So for a first treatment of the real numbers we will formulate them as a collection of axioms exactly characterizing them. The axioms for real numbers fall into three groups, the axioms for fields, the order axioms and the completeness axiom.

Axiom 4.2.1: Axiom of a field

A field is a set $\mathbb{F}$ together with a binary operation $+ : \mathbb{F} \times \mathbb{F} \to \mathbb{F}$ called addition and a binary operation $\cdot : \mathbb{F} \times \mathbb{F} \to \mathbb{F}$ called multiplication such that $\forall x, y, z \in \mathbb{F}$ the following properties hold:
(A) Axioms for addition

1. (A1) $x + y = y + x$ (addition is commutative)
2. (A2) $(x + y) + z = x + (y + z)$ (addition is associative)
3. (A3) $\mathbb{F}$ contains an element 0 such that $x + 0 = x$ (neutral element of addition)
4. (A4) There exists an element $-x \in \mathbb{F}$ such that $x + (-x) = 0$. (additive inverse)

(M) Axioms for multiplication

1. (M1) $xy = yx$ (multiplication is commutative)
2. (M2) $(xy)z = x(yz)$ (multiplication is associative)
3. (M3) $\mathbb{F}$ contains an element $1 \neq 0$ such that $1x = x$. (neutral element of multiplication)
4. (M4) For each $x \neq 0$ there is an element x^{-1}, such that $xx^{-1} = 1$ (multiplicative inverse)

(D) The distributive law: $x(y + z) = xy + xz$.

It is a long but straightforward work to prove the following result using the construction of the rationals we have studied before and we leave it to the interested reader.

Proposition 4.2.1

The rational numbers $\mathbb{Q}$ form a field.

So obviously this list of axioms is not enough to characterize $\mathbb{R}$. Actually there are many examples of fields. Some of them even have a finite number of elements.
Example

> Let $\mathbb{F}_2 = \{0, 1\}$ the set of two elements together with the addition and multiplication
>
> $$0 + 0 = 0,\ 0 + 1 = 1 + 0 = 1,\ 1 + 1 = 0, \quad 0 \cdot 0 = 0,\ 0 \cdot 1 = 1 \cdot 0 = 0,\ 1 \cdot 1 = 1.$$
>
> It can be checked easily that $\mathbb{F}_2$ satisfies all the axioms of a field.

Directly from these axioms, we can prove that the usual rules for addition and multiplication hold. What it means is that as long as we are within any field, we can use them without any problem. We are going to prove some of them here and you will practice more of these proofs in the problem sheets.

Proposition 4.2.2: (cancelation law for the addition)

Let $\mathbb{F}$ be a field, $x, y, z \in \mathbb{F}$. If $x + z = y + z$, then $x = y$.

Proof. Suppose that $x + z = y + z$. Let $(-z)$ be an additive inverse to z, which exists by Axiom (A4). Then $(x + z) + (-z) = (y + z) + (-z)$. By associativity of addition (Axiom A2), we have then $x + (z + (-z)) = y + (z + (-z))$. And finally by Axiom (A4) $x + 0 = y + 0$ and by Axiom A3, $x = y$. ■

Proposition 4.2.3

Let $\mathbb{F}$ be a field, $x \in \mathbb{F}$. Then $x \cdot 0 = 0$

Proof. By Axiom (A3), $x \cdot 0 = x \cdot (0 + 0)$. By distributivity (D), $x \cdot (0 + 0) = x \cdot 0 + x \cdot 0$. By Axiom (A3) again, $0 + x \cdot 0 = x \cdot 0 + x \cdot 0$, and by Axiom (A1), $x \cdot 0 + 0 = x \cdot 0 + x \cdot 0$. Hence $0 = x \cdot 0$ by the preceding proposition. ■

Axiom 4.2.2: Ordered field

An ordered field is a field $\mathbb{F}$ together with a total order $\leq$ such that additionally if $x, y, z \in \mathbb{F}$, $x \leq y$ then

1. (O1) $x + z \leq y + z$
2. (O2) and if moreover $z \geq 0$ then $xz \leq yz$

Actually our axioms (O1) and (O2) imply the following strengthened version.

Proposition 4.2.4

1. For all $x, y, z \in \mathbb{F}$, if $x < y$, then $x + z < y + z$,
2. For all $x, y, z \in \mathbb{F}$, if $x < y$ and $z > 0$, then $xz < yz$

And again from these axioms we can derive the usual rules we know for inequalities. We will prove for example the following result

Proposition 4.2.5

Let $\mathbb{F}$ be an ordered field. For all $x \in \mathbb{F}$. Then $x < 0$ if and only if $-x > 0$ and $x > 0$ if and only if $-x < 0$.

Proof. Let $x < 0$. By axiom (O1) we can add the additive inverse of x (which exists by axiom (A4)) to both sides and we get $x + (-x) < 0 + (-x)$. But this implies immediately $0 < -x$ by axiom (A3). The proof second statement is very similar and we leave it as an exercise. ■

$\mathbb{Q}$ is again an example of an ordered field, which means that despite a lot of axioms already, this is still not enough to define the reals. We will need to add something which is called the axiom of completeness in order to define them uniquely.

4.3 Axiom of completeness

Definition 4.3.1

A set non-empty set S is called bounded above (resp. bounded below) if there is a number B such that $x \leq B$ (resp. $x \geq B$) for all $x \in S$.
Any such B is called an upper bound (resp. lower bound) for S.

Remark. *Note that if they exist upper or lower bounds are of course not unique, neither need they to be in the set S. And they do not need to exist either. For example the set $S := \{n \in \mathbb{N} | n \ positive \ and \ even\} \in \mathbb{Q}$ is bounded below by any negative rational number and 0 (none of them except 0 being in the set S) , but has no upper bound.*

Definition 4.3.2

Let S be a non-empty set. We call s a least upper bound or supremum (resp. largest lower bound or infimum) of S if

1. s is an upper bound (resp. lower bound) for S,
2. if $B < s$ (resp. $B > s$) then B is not an upper bound (resp. lower bound) for S.

This new notion will allow us to give the final axiom of the real numbers.

Axiom 4.3.1: Axiom of completeness

(C) A complete ordered field is an ordered field $\mathbb{F}$ such that if a nonempty subset $S \subseteq \mathbb{F}$ has an upper bound, then S has a supremum which lies within $\mathbb{F}$.

This is a very important property which distinguishes $\mathbb{R}$ from the set $\mathbb{Q}$ of rational numbers. For example, the set $\{x | x^2 < 2\} \subset \mathbb{Q}$ is bounded above by many rational numbers (e.g. by $M = 3/2$) but it does not have a supremum in $\mathbb{Q}$!

Definition 4.3.3

The real numbers $\mathbb{R}$ are the only set satisfying axioms 4.2.1, 4.2.2 and 4.3.2

To understand the importance of the axiom of completeness we are going to look several applications. The first (and maybe most important) consequence of this last axiom is called the Archimedean property or sometimes the axiom of Eudoxus. It basically states that the natural numbers are not bounded above in $\mathbb{R}$. This property, which you will use extensively in analysis may seem obvious, but in fact there are other ordered fields than the reals in which it does not hold.

Proposition 4.3.1: Archimedean Property

For all $x, y \in \mathbb{R}$, $x > 0$ there exists an $n \in \mathbb{N}$, such that $nx > y$.

Proof. Exercise! (Problem sheet 7). ■

There are several equivalent forms of the Archimedean property that are useful in different contexts and which we state in the following immediate corollary. Especially, you will encounter the second formulation very often when you study convergence properties of sequences for example.

Corollary 4.3.1

The following statement is equivalent to the Archimedean property.

1. For all $x \in \mathbb{R}$, there exists an $n \in \mathbb{N}$ such that $n > x$.
2. For all $x \in \mathbb{R}$, $x > 0$ there exists an $n \in \mathbb{N}$ such that $0 < \frac{1}{n} < x$.

The following result is another easy immediate consequence.

Lemma 4.3.1

Let $x \in \mathbb{R}$, $x > 0$. Then there exists $n \in \mathbb{N}$ such that $n - 1 \leq x < n$. Moreover, n is unique.

Proof. Consider the set $S = \{m \in \mathbb{N} | m > x\}$. By the Archimedean property we know that S is non empty and by the well ordering property of $\mathbb{N}$ we also know that there is a unique $n \in \mathbb{N}$ such that for every $m \in S, m \geq n$. Since $n \in S$, we have $n > x$. Now consider two possibilities, $n = 1$ or $n > 1$. If $n = 1$ then $n - 1 = 0$ and by assumption $0 < x$, then $0 < x < 1$. If $n > 1$, then $n - 1 \in \mathbb{N}$ but by construction $n - 1$ is not in S. So $n - 1 \leq x$, implying the desired result. ■

Another extremely important consequence of the completeness axiom is the following fact.

Proposition 4.3.2: $\mathbb{Q}$ is dense in $\mathbb{R}$

For all $x, y \in \mathbb{R}$ such that $x < y$, then there exists a number $r \in \mathbb{Q}$ such that $x < r < y$.

Proof. We will first assume that $x, y > 0$. Since $y - x > 0$ by the Archimedean property there exists a natural number n, such that

$$\frac{1}{n} < y - x. \tag{4.3}$$

Now consider the set $\{\frac{1}{n}, \frac{2}{n}, \ldots, \ldots \frac{k}{n}\}$, $k \in \mathbb{N}$ and pick the largest k for which

$$\frac{k}{n} \leq x \leq \frac{k+1}{n}.$$

Obviously this construction is possible by Lemma 4.3.1.

Will show by contradiction that $\frac{k+1}{n} < y$. Assume therefore that $\frac{k+1}{n} \geq y$. Then

$$\frac{1}{n} = \frac{k+1}{n} - \frac{k}{n} \geq y - x.$$

which contradicts equation (4.3). Defining $r := \frac{k+1}{n}$ finishes the proof. ■

This result, in turn, is of crucial important as well in analysis. We will see this in the following example.

Example

Let $A_1 := \{x \in \mathbb{R} | 1 < x < 2\}$. We want to find the infimum and supremum of this set. Clearly one upper bound is 2 and one lower bound is 1. We claim that actually $\sup A_1 = 2$, $\inf A_1 = 1$. We start with the supremum. Assume that there exists another upper bound $M > 1$, and such that $M < 2$. Then by Proposition 4.3.2 there exists a rational number r, such that $M < r < 2$. Consequently $r \in A_1$, but this is a contradiction that M is an upper bound as for this it would have to be bigger than any element in A_1. Therefore $\sup A_1 = 2$. The proof for the infimum is very similar and we leave it to the reader.

Chapter 5

Vectors and Geometry

5.1 The intimate link between Geometry & Algebra

On one hand, you were introduced to (Euclidean) **Geometry** in two and three dimensions at school. You probably learnt that geometry deals with: points, lines, planes, surfaces, solids ... On the other hand, you have been dealing for a few years now with what is called **Algebra**; you learnt that algebra deals with variables and equations, i.e. statements like $x = 5$, $y = 7$, $y = ax + b$ or even $2x + 6y + 8z = 42$. In particular, **linear algebra** is the study of linear equations; while we will introduce some concepts in this chapter, linear algebra will be the topic of a two-term module in year 1.

Yet, you might not have been taught that there is an intimate link between geometry and algebra. To convince yourself, consider the equation $y = ax + b$ (which is a algebraic object) is the graph of a set of points satisfying the equation; it turns out that the locus of those points traces out a straight line (a geometric object). Graphing an equation is taking an *algebraic concept* and making it a *geometric concept*.

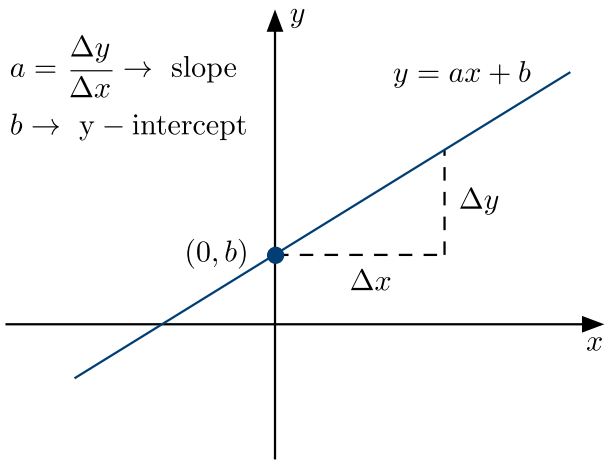


Figure 5.1 Graph of a straight line with equation $y = ax + b$.

In general, the equation of a straight line has the form $y = ax + b$, where a is the slope of the line (i.e. the rate of change of the y-values when changing the x-values) and b is the y-intercept. Thus, the variables in the equation of a line (algebraic concepts) relate to geometric concepts pertaining to the line (slope and intercept)! We will soon see that at the core of the intimate link between algebra and geometry lies the concept of **coordinates**.

5.2 Cartesian product and ordered pairs

Consider two sets A and B and two elements $a \in A$ and $b \in B$, we can construct what is called an **ordered pair** (a, b) with the following property

$$(a, b) = (a', b') \iff (a = a' \wedge b = b') \tag{5.1}$$

Definition 5.2.1: Cartesian product

For two given sets A and B, the set formed by the ordered pairs (a, b), with $a \in A$ and $b \in B$ is called the **Cartesian product** of A and B and is denoted $A \times B$. Thus, we can write:

$$A \times B = \{(a, b) | a \in A, b \in B\} \tag{5.2}$$

When $A = B$, we write the cartesian product $A \times A$ as A^2.

Remark. *Similarly, we can define a triplet* (a, b, c) *verifying the following property*

$$(a, b, c) = (a', b', c') \iff (a = a' \wedge b = b' \wedge c = c') \tag{5.3}$$

as well as the Cartesian product:

$$A \times B \times C = \{(a, b, c) | a \in A, b \in B, c \in C\} \tag{5.4}$$

if $A = B = C$*, we denote this Cartesian product* A^3*.*

More generally, for $n \in \mathbb{N}^*$*, we can define the notion of* n*-tuple* $(a_1, a_2, \cdots, a_n)$ *as well as the ensemble*

$$A_1 \times A_2 \times \cdots \times A_n = \{(a_1, a_2, \cdots, a_n) | a_1 \in A_1, a_2 \in A_2, \cdots a_n \in A_n\} \tag{5.5}$$

if $A_1 = A_2 = \cdots = A_n = A$*, we denote this Cartesian product* A^n*.*

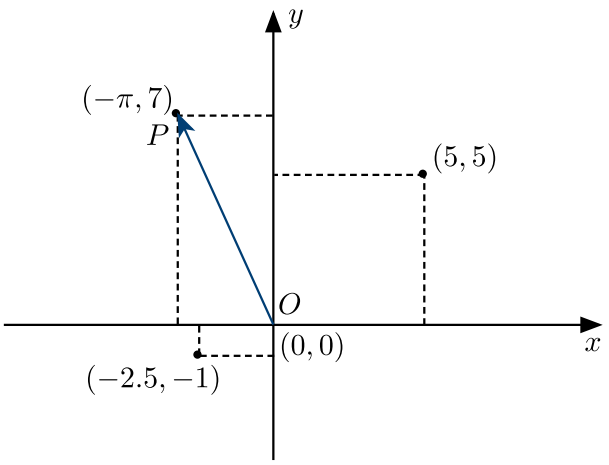


Figure 5.2 Example of Cartesian product, the Cartesian plane.

In the previous chapter, you have axiomatized the **real numbers**; these are numbers densely populating what we call the **real line**, defining the field $\mathbb{R}$. Now consider the Cartesian product $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$. This Cartesian product is defined as:

$$\mathbb{R}^2 = \{(a, b) | a \in \mathbb{R}, b \in \mathbb{R}\} \tag{5.6}$$

and thus represents the whole Cartesian **plane**.

One of the main historical examples of Cartesian products was introduced by René Descartes which assigned to each point in the plane an element of $\mathbb{R}^2$, i.e. an ordered pair (a, b). Geometrically, it is easy to convince yourself that for any ordered pair $(a, b) \neq (b, a)$ (order matters!). If you consider an ordered pair (a, b) representing a point P of the plane, the ordered pair first and second elements are generically called the x-coordinate and y-coordinate of P respectively.

5.3 Vector Algebra

In this section, we will explore a set called $\mathbb{R}^n$ (with $n \in \mathbb{N}$) and introduce the concept of **vectors**.

5.3.1 Definitions and notations

Consider $(-\pi, 7)$ an ordered pair of $\mathbb{R}^2$. This ordered pair can be pictured as a point in the Cartesian plane $\mathbb{R}^2$ or as a *vector* starting at $(0, 0)$ and ending at $(-\pi, 7)$ (see Figure 5.2). If we denote the point $P = (-\pi, 7)$ and the point $O = (0, 0)$ then the vector $\mathbf{OP} = \begin{pmatrix} -\pi \\ 7 \end{pmatrix}$.

It is generally assumed that the points of $\mathbb{R}^2$ are the **same** as arrows starting at the origin $(0, 0)$ and ending at the said point. This allows us to **define** a correspondance between vectors and points in $\mathbb{R}^2$, i.e. one can represent a point in the plane either as the ordered pair (a, b) (a and b are *coordinates of the point*) or the vector $\begin{pmatrix} a \\ b \end{pmatrix}$ (with a and b *components of the vector*). As points and vectors in $\mathbb{R}^n$ are thought to be the same, the terms components and coordinates are commonly used interchangeably.

Example

> If you consider the two dimensional Euclidean space, the ordered pair (a, b) can be seen as the point in $\mathbb{R}^2$ which is a units along the horizontal axis (commonly denoted x-axis) and b units along the vertical axis (commonly denoted y-axis). It can also be seen as the translation which would take the origin $(0, 0)$ to the point (a, b).

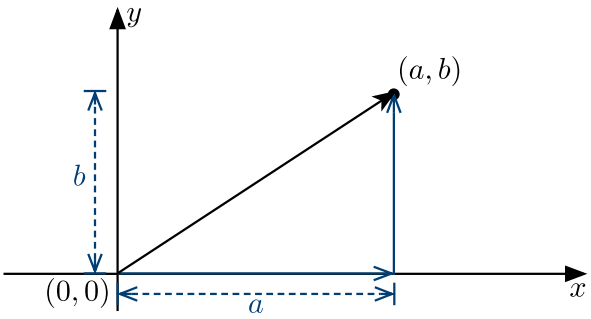


Figure 5.3 Points and vectors.

Definition 5.3.1: n-dimensional vectors

Consider a positive integer n, we call a n-dimensional vector (or n-vector) an element $\mathbf{x} \in \mathbb{R}^n$, i.e. a list of n real numbers $x_1, x_2, x_3, \ldots, x_n$. This list can be arranged either

as a **column vector**, in which case, we write it as follows

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

or as a **row vector** and written as

$$(x_1, x_2, \ldots, x_n)$$

For all $1 \leq i \leq n$, we refer to the x_i as the i^{th} component of vector $\mathbf{x}$.

Remark. *In the rest of these notes, we will denote vectors using bold face symbols such as* $\mathbf{u}$ *or* $\mathbf{OA}$*. In lecture and in problem sessions, this notation is unpractical and we shall use other classical vector notations such as* $\vec{u}$ *or* $\underline{u}$*. Unit vectors will sometimes be denoted with a hat, e.g.* $\hat{\boldsymbol{x}}$ *or* $\hat{\boldsymbol{e}}_1$*. While the student should not be confused by the multiplicity of notations, in this part of the course, it is important to distinguish between scalars and vectors. The reason for this will become apparent in what follows.*

Definition 5.3.2

For a given integer n, we denote the set of all n-dimensional vectors with real-valued components as $\mathbb{R}^n$, which is often referred to as the n-dimensional space. In particular:

- For $n = 2$, we commonly use x and y to denote the components of a two-dimensional vector defining $\mathbb{R}^2 = \{(x, y) : x, y \in \mathbb{R}\}$ as the **xy-plane**.
- For $n = 3$, we commonly use x, y and z to denote the components of a three-dimensional vector defining $\mathbb{R}^3 = \{(x, y, z) : x, y, z \in \mathbb{R}\}$ as the **xyz-space**.

Remark. *You will see in Algebra later in your course at Imperial that* n*-tuples and row vectors are not the same objects. This discussion goes beyond the scope of this introductory module. In linear algebra, one needs to differentiate column and row vectors. Due to the rules of multiplications and additions of vectors and matrices, these are objects that one cannot use interchangeably.*

Definition 5.3.3

We call the **zero vector** (or null vector) the vector in $\mathbb{R}^n$ composed of a list of n zeros $(0, 0, \ldots, 0)$. We denote this vector $\mathbf{0}$. Furthermore, the vectors defined by $(0, \ldots, 0, x_i, 0, \ldots, 0)$ represent in $\mathbb{R}^n$ all the points on the x_i-axis.

In Definition 5.3.3, we have introduced the idea that for $\mathbb{R}^n$ we can define a point of origin $\mathbf{0}$ which lies at the intersection of n axes. These axes are commonly defined using what is called the standard basis for $\mathbb{R}^n$.

Definition 5.3.4: Standard (or canonical) basis for $\mathbb{R}^n$

The n vectors defined by

$$(1, 0, \cdots, 0, 0), (0, 1, \cdots, 0, 0), \quad \cdots \quad , (0, 0, \cdots, 1, 0), (0, 0, \cdots, 0, 1)$$

in $\mathbb{R}^n$ are known as the **standard (or canonical) basis** for $\mathbb{R}^n$. These vectors are commonly denoted $\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \cdots, \hat{\mathbf{e}}_n$. Any vector $\mathbf{u} \in \mathbb{R}^n$ can be written uniquely as a linear combination of $\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \cdots, \hat{\mathbf{e}}_n$, i.e.

$$\forall \mathbf{u} \in \mathbb{R}^n, \exists! \, (a_1, a_2, \cdots, a_n) \in \mathbb{R}^n \mid \mathbf{u} = \sum_{i=1}^{n} a_i \hat{\mathbf{e}}_i \tag{5.7}$$

Further, when $n = 2$, the vectors $(1, 0)$ and $(0, 1)$ form the standard basis for $\mathbb{R}^2$. In this particular case, they are also commonly denoted $\hat{\mathbf{i}}$ and $\hat{\mathbf{j}}$ respectively. Following Definition 5.3.4, any vector $\mathbf{u} = (x, y)$ can be written as a linear combination of $\hat{\mathbf{i}}$ and $\hat{\mathbf{j}}$, i.e. $\mathbf{u} = (x, y) = x\hat{\mathbf{i}} + y\hat{\mathbf{j}}$ and this is the only way to write $\mathbf{u}$ as a sum of scalar multiples of $\hat{\mathbf{i}}$ and $\hat{\mathbf{j}}$. In the case of $\mathbb{R}^3$, the canonical basis is commonly denoted $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$.

5.3.2 An alternative definition of vectors

Progress in mathematics is rarely achieved by working alone. If most of your time as mathematician is spent exchanging with collaborators, you would want to adopt a common language. We have just introduced a definition for the concept of vectors; it is interesting to note though that the answer to a question as simple as "what is a vector?" will depend on your interlocutor (whether he/she is an algebraist, an analyst, a geometer, an applied mathematician ...). Indeed, vectors are usually thought of in two different ways. Let's consider the case of vectors in $\mathbb{R}^2$:

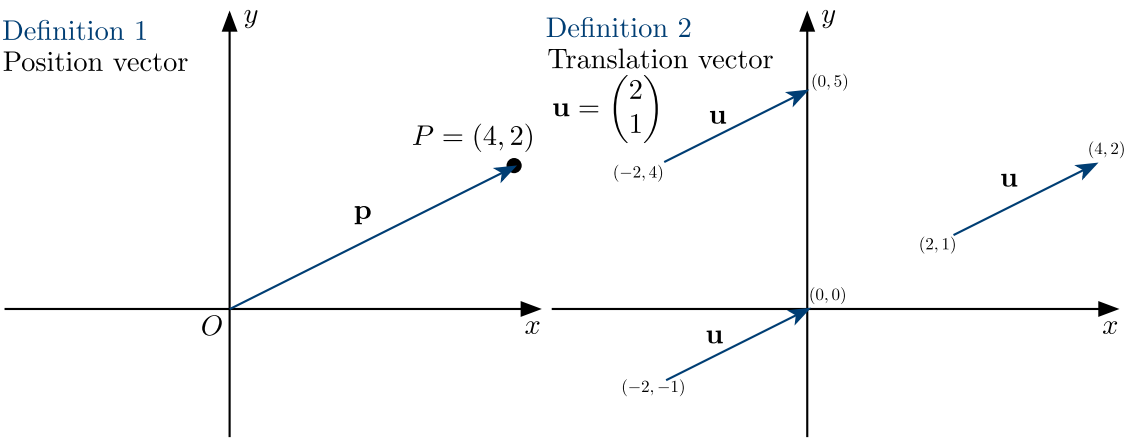


Figure 5.4 Two points of view on the definition of a vector.

- In a first point of view (the one adopted in the previous chapter), a vector $\mathbf{p} = \mathbf{OP} = \begin{pmatrix} a \\ b \end{pmatrix}$ can represent the point P in $\mathbb{R}^2$ which has coordinates a and b (where we denote the origin as the point O). This is called the **position vector of point** P. In practice though, we rarely refer to the **point with position vector** $\mathbf{p}$ but rather simply refer to it as **point** $\mathbf{p}$ when the meaning of otherwise clear. The point $\mathbf{0}$ is usually referred to as the **origin**.

- In a second point of view (quite useful in applied mathematics or differential geometry for instance), a vector can be seen as a movement or a translation from a point to another point. This is illustrated on Figure 5.4. For instance, to get from point $(-2, 4)$ to point $(0, 5)$, we need to translate to the right by 2 and up by 1, we can

define this movement as the **translation vector** $\mathbf{u} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$. The same movement would be required to go from point $(-2, -1)$ to point $(0, 0)$ or from point $(2, 1)$ to point $(4, 2)$; these movements are thus also represented by vector $\mathbf{u}$. From this point of view, the vector $\mathbf{0}$ is understood as *no movement*.

Definition 5.3.5

An alternative definition of a vector is a geometrical object representing a physical quantity which has a **magnitude** and a **direction**. In this definition, the point of origin of the vector is meaningless as the vector represents a translation. This definition of a vector is commonly used in physics and mechanics (including fluid mechanics).

5.3.3 Algebra of vectors

Definition 5.3.6: Vector Addition and Subtraction

Consider two vectors $\mathbf{u} = (u_1, u_2, \cdots, u_n)$ and $\mathbf{v} = (v_1, v_2, \cdots, v_n)$ in $\mathbb{R}^n$, we can add or subtract these two vectors; as you might expect, these operations are done component-wise

$$\mathbf{u} + \mathbf{v} = (u_1 + v_1, u_2 + v_2, \cdots, u_n + v_n) \tag{5.8}$$
$$\mathbf{u} - \mathbf{v} = (u_1 - v_1, u_2 - v_2, \cdots, u_n - v_n) \tag{5.9}$$

Note that two vectors may be added or subtracted if and only if they have the same number of components. For instance, one can not add to a vector in $\mathbb{R}^3$ a vector in $\mathbb{R}^2$.

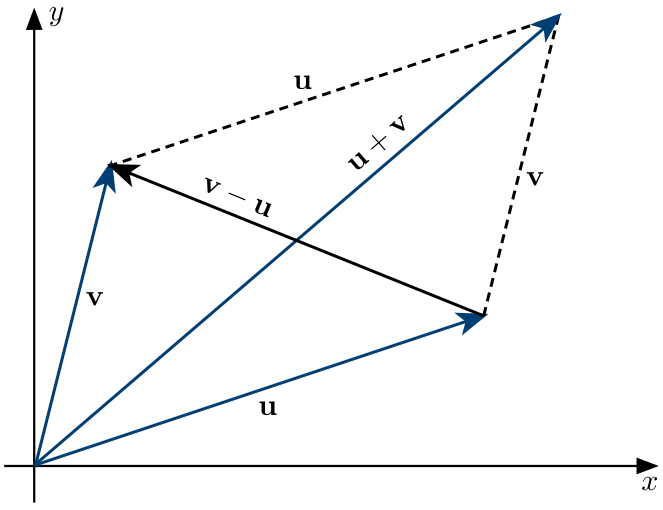


Figure 5.5 Adding and subtracting vectors in $\mathbb{R}^2$.

Example

Consider two particular vectors $\mathbf{u} = (6, 2)$ and $\mathbf{v} = (1, 4)$ in $\mathbb{R}^2$ (see Figure 5.5). On one hand, the vector $\mathbf{u} + \mathbf{v}$ is the translation whose overall effect corresponds to a translation by $\mathbf{u}$ followed by a translation by $\mathbf{v}$. Note that this can also be achieved if

we were to translate first by $\mathbf{v}$ and then by $\mathbf{u}$. formally this means that $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$, the addition of vectors is said to be *commutative*.

On the other hand, It is also interesting to note that the vector $\mathbf{v}-\mathbf{u}$ is the vector that translates the point (with position vector) $\mathbf{u}$ to the point (with position vector) $\mathbf{v}$.

Definition 5.3.7: Multiplication by a scalar

Consider a vector $\mathbf{u}=(u_1,u_2,\cdots,u_n)\in\mathbb{R}^n$ and a real number λ, the **scalar multiplication** is defined as follows

$$\lambda\mathbf{u}=(\lambda u_1,\lambda u_2,\cdots,\lambda u_n) \tag{5.10}$$

- If λ is a positive integer, we can think of the scalar multiple $\lambda\mathbf{u}$ as an overall translation achieved by translating λ times by the vector $\mathbf{u}$.
- As we vary λ in $\mathbb{R}$, the points $\lambda\mathbf{u}$ trace out a straight line passing through the origin and the point $\mathbf{u}$.
- If $\lambda=-1$, we write simply $-\mathbf{u}=(-u_1,-u_2,\cdots,-u_n)$; translating by $-\mathbf{u}$ is the exact inverse of a translation by $\mathbf{u}$.

5.4 Geometry of vectors

The definition, we just introduced, allowed us to manipulate vectors as algebraic objects. As we saw earlier, there is a clear link between algebra and geometry. Thus, vectors also have important geometric properties. Let's start by defining what we just called the magnitude of a vector.

Definition 5.4.1: Magnitude of a vector

Let $\mathbf{u}=(u_1,u_2,\cdots,u_n)\in\mathbb{R}^n$, we define the **magnitude** (or length) of vector $|\mathbf{u}|$ as

$$|\mathbf{u}|=\sqrt{u_1^2+u_2^2+\cdots+u_n^2} \tag{5.11}$$

This is also called the Euclidean norm of vector $\mathbf{u}$.

- The vector $\mathbf{u}$ is said to be a **unit vector** if $|\mathbf{u}|=1$.
- Further, notice that:

$$\forall\mathbf{x}\in\mathbb{R}^n,|\mathbf{x}|\geq 0 \text{ and } |\mathbf{x}|=0 \iff \mathbf{x}=\mathbf{0} \tag{5.12}$$

- Finally, we also have

$$\forall\mathbf{x}\in\mathbb{R}^n,\forall\lambda\in\mathbb{R},|\lambda\mathbf{x}|=|\lambda||\mathbf{x}| \tag{5.13}$$

Remark. *If you look at Figure 5.3, you will quickly convince yourself that this definition of the length of a vector makes sense. Indeed, for vectors in $\mathbb{R}^2$, this is just a reformulation of Pythagoras' theorem!*

Now consider two points U and V in $\mathbb{R}^n$, we saw that these can be thought of as the position vectors $\mathbf{u}=\mathbf{OU}$ and $\mathbf{v}=\mathbf{OV}$ in $\mathbb{R}^n$, the translation from $\mathbf{u}$ to $\mathbf{v}$ is given by the vector $\mathbf{v}-\mathbf{u}$. It is natural then to define the distance between two points of $\mathbb{R}^n$ as follows.

Definition 5.4.2: Distance between two points

Let $\mathbf{u} = (u_1, u_2, \cdots, u_n) \in \mathbb{R}^n$ and $\mathbf{v} = (v_1, v_2, \cdots, v_n) \in \mathbb{R}^n$, we define the **Euclidean distance** between $\mathbf{u}$ and $\mathbf{v}$ (seen as position vectors) as $|\mathbf{v} - \mathbf{u}|$ (or equivalently $|\mathbf{u} - \mathbf{v}|$). Thus, in terms of the vectors components, the distance from $\mathbf{u}$ to $\mathbf{v}$ reads

$$|\mathbf{v} - \mathbf{u}| = \sqrt{\sum_{i=1}^{n} (v_i - u_i)^2} \tag{5.14}$$

Definition 5.4.3: Scalar Product

Given two vectors $\mathbf{u} = (u_1, u_2, \cdots, u_n)$ and $\mathbf{v} = (v_1, v_2, \cdots, v_n)$ in $\mathbb{R}^n$, the **scalar product**, also known as **dot product** or **Euclidean inner product** , denoted $\mathbf{u} \cdot \mathbf{v}$ is defined as the real number

$$\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^{n} u_i v_i \tag{5.15}$$

First, it is interesting to note that knowing the definition of the transpose of a vector/matrix and the rules of vector multiplication (for instance, see Video VIII - Linear Algebra of the pre-arrival material), we can rewrite the scalar product as

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^\intercal \mathbf{v} \tag{5.16}$$

where $\mathbf{u}$ and $\mathbf{v}$ are column vectors.

Note that the **dot product** is also called **Euclidean inner product**; an inner product is usually denoted $\langle \mathbf{u}, \mathbf{v} \rangle$. The dot product is a special case in $\mathbb{R}^n$ of a more general class of operations. You will see in later modules that inner products are generalizations of the idea of the dot product to *almost any* vector space V over a field $\mathbb{F}$; it then associates to two elements of V an element of F, i.e.

$$\begin{aligned} \langle \cdot, \cdot \rangle : \ & V \times V \to \mathbb{F} \\ & \mathbf{u}, \mathbf{v} \mapsto \langle \mathbf{u}, \mathbf{v} \rangle \end{aligned}$$

The following properties are here given for the dot product in $\mathbb{R}^n$ but easily generalize to any inner product.

Proposition 5.4.1

The scalar product (or dot product) is a **symmetric bilinear form**, i.e. for all vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ and for all reals λ and μ, we have

(a) $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$ (symmetry)
(b) $(\lambda \mathbf{u} + \mu \mathbf{w}) \cdot \mathbf{v} = \lambda \mathbf{u} \cdot \mathbf{v} + \mu \mathbf{w} \cdot \mathbf{v}$ (linearity in the first argument)
(c) $\mathbf{u} \cdot (\lambda \mathbf{v} + \mu \mathbf{w}) = \lambda \mathbf{u} \cdot \mathbf{v} + \mu \mathbf{u} \cdot \mathbf{w}$ (linearity in the second argument)

Further, we have the following properties

(d) $\mathbf{u} \cdot \mathbf{u} = |\mathbf{u}|^2 \geq 0 \quad \text{and} \quad \mathbf{u} \cdot \mathbf{u} = 0 \iff \mathbf{u} = \mathbf{0}$

Proof. Let $\mathbf{u} = (u_1, u_2, \cdots, u_n)$, $\mathbf{v} = (v_1, v_2, \cdots, v_n)$ and $\mathbf{w} = (w_1, w_2, \cdots, w_n)$ be vectors in $\mathbb{R}^n$.

- **Symmetry**

$$\begin{aligned}\mathbf{u}\cdot\mathbf{v} &= u_1v_1 + u_2v_2 + \cdots + u_nv_n \\ &= v_1u_1 + v_2u_2 + \cdots + v_nu_n \\ &= \mathbf{v}\cdot\mathbf{u}\end{aligned}$$

- **Linearity in the first argument**

$$\begin{aligned}(\lambda\mathbf{u} + \mu\mathbf{w})\cdot\mathbf{v} &= \sum_{i=1}^{n}(\lambda\mathbf{u} + \mu\mathbf{w})_i v_i \\ &= \sum_{i=1}^{n}(\lambda u_i + \mu w_i)v_i \\ &= \sum_{i=1}^{n}\lambda u_i v_i + \mu w_i v_i \\ &= \lambda\sum_{i=1}^{n}u_i v_i + \mu\sum_{i=1}^{n}w_i v_i \\ &= \lambda\mathbf{u}\cdot\mathbf{v} + \mu\mathbf{w}\cdot\mathbf{v}\end{aligned}$$

- **Linearity in the second argument** — this is done similarly to the proof of the linearity in the first argument.

■

This means that the **length** of a vector can be written in terms of the **scalar product**, namely we have

$$|\mathbf{u}| = \sqrt{\mathbf{u}\cdot\mathbf{u}} \tag{5.17}$$

Now that we have defined a measure of length and distance and related it to an inner product, we are armed to derive some important geometric results.

Proposition 5.4.2: Cauchy-Schwarz Inequality

Let $\mathbf{u}$ and $\mathbf{v}$ be two vectors in $\mathbb{R}^n$, then

$$|\mathbf{u}\cdot\mathbf{v}| \leq |\mathbf{u}||\mathbf{v}| \tag{5.18}$$

with equality when one of $\mathbf{u}$ and $\mathbf{v}$ is a multiple of the other. In this case, we say that the two vectors are linearly dependent.

Proof. This proof is left to the student as an exercise. Please see Part III - Problem Sheet 1. ■

Remark. *Further, the Cauchy-Schwarz Inequality holds for any inner product.*

Proposition 5.4.3: Triangle Inequality

Let $\mathbf{u}$ and $\mathbf{v}$ be two vectors in $\mathbb{R}^n$. Then

$$|\mathbf{u}+\mathbf{v}| \leq |\mathbf{u}|+|\mathbf{v}| \tag{5.19}$$

Proof. Consider two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$. We write

$$|\mathbf{u}+\mathbf{v}|^2 = (\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v}) \leq |\mathbf{u}|^2+|\mathbf{v}|^2+2|\mathbf{u}\cdot\mathbf{v}|$$

By the Cauchy-Schwarz inequality (see Proposition 5.4.2), we know that

$$|\mathbf{u}\cdot\mathbf{v}| \leq |\mathbf{u}||\mathbf{v}|$$

So

$$|\mathbf{u}+\mathbf{v}|^2 \leq |\mathbf{u}|^2+|\mathbf{v}|^2+2|\mathbf{u}||\mathbf{v}| = (|\mathbf{u}|+|\mathbf{v}|)^2 \iff |\mathbf{u}+\mathbf{v}| \leq |\mathbf{u}|+|\mathbf{v}|$$

■

Remark. *To understand why this intuitive result is called triangle inequality, consider vectors in $\mathbb{R}^2$ as in Figure 5.5. If you consider the triangle with vertices $\mathbf{0}$, $\mathbf{u}$ and $\mathbf{u}+\mathbf{v}$, then it is obvious that the length of its sides are given by $|\mathbf{u}|$, $|\mathbf{v}|$ and $|\mathbf{u}+\mathbf{v}|$. $|\mathbf{u}+\mathbf{v}|$ is the distance going straight from $\mathbf{0}$ to point with position vector $\mathbf{u}+\mathbf{v}$; on the other hand, $|\mathbf{u}|+|\mathbf{v}|$ is the combined distance going from $\mathbf{0}$ to $\mathbf{u}+\mathbf{v}$ passing through $\mathbf{u}$. Clearly, this cannot be shorter than go straight from $\mathbf{0}$ to $\mathbf{u}+\mathbf{v}$ and will only be equal if $\mathbf{u}$ is on the straight segment of line between $\mathbf{0}$ and $\mathbf{u}+\mathbf{v}$.*

Following this remark, we realize that we can complete the **triangle inequality** adding that if $\mathbf{v} \neq \mathbf{0}$ then **there is equality** if and only if $\mathbf{u} = \lambda\mathbf{v}$ with a real $\lambda \geq 0$. One would use a similar proof as for the Cauchy-Schwarz inequality to prove this statement.

Definition 5.4.4: Angle between vectors

Consider $\mathbf{u}$ and $\mathbf{v}$ two non-zero vectors in $\mathbb{R}^n$, the **angle between these two vectors** is defined by the following expression

$$\cos^{-1}\left(\frac{\mathbf{u}\cdot\mathbf{v}}{|\mathbf{u}||\mathbf{v}|}\right) \tag{5.20}$$

Note that the Cauchy-Schwarz inequality (see Proposition 5.4.3) ensures that the angle between vectors is well defined as the quantity $|\mathbf{u}\cdot\mathbf{v}|/|\mathbf{u}||\mathbf{v}| \leq 1$. Further, if one takes the value of $\cos^{-1}$ to be in the range $[0,\pi]$, then the above definition measures the smaller angle between the two vectors. You will see in Section 6.1 that this definition for the angle between two vectors is consistent with our usual definition of angles in $\mathbb{R}^2$. Intuitively, the concept of perpendicular vectors follows from this definition, vectors are said to be **perpendicular** (or **orthogonal**) if and only if $\mathbf{u}\cdot\mathbf{v} = 0$.

Example

Let $\mathbf{u} = (1,2,3,1)$ and $\mathbf{v} = (-1,0,2,3)$ vectors in $\mathbb{R}^4$. Find the length of these vectors and the angle between them.

The length of these vectors is given by

$$|\mathbf{u}|^2 = 1^2+2^2+3^2+1^2 \Rightarrow |\mathbf{u}| = \sqrt{15}$$
$$|\mathbf{v}|^2 = (-1)^2+0^2+2^2+3^2 \Rightarrow |\mathbf{v}| = \sqrt{14}$$

Further, the scalar product of $\mathbf{u}$ and $\mathbf{v}$ is given by

$$\mathbf{u} \cdot \mathbf{v} = 1 \times (-1) + 2 \times 0 + 3 \times 2 + 1 \times 3 = 8$$

Finally, the angle between these two vectors is given by

$$\theta = \cos^{-1}\left(\frac{8}{\sqrt{14}\sqrt{15}}\right) = \cos^{-1}\left(\frac{8}{\sqrt{210}}\right) \approx 0.986 \text{ radians.}$$

Proposition 5.4.4

Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in $\mathbb{R}^n$ with $\mathbf{v} \neq \mathbf{0}$. There is a unique real number λ such that $\mathbf{u} - \lambda\mathbf{v}$ is perpendicular to $\mathbf{v}$. This implies that

$$\mathbf{u} = \lambda\mathbf{v} + (\mathbf{u} - \lambda\mathbf{v})$$

with the vector $\lambda\mathbf{v}$ being called the **component of $\mathbf{u}$ in the direction of $\mathbf{v}$** and $\mathbf{u} - \lambda\mathbf{v}$ being called the **component of $\mathbf{u}$ perpendicular to the direction of $\mathbf{v}$**. Further, we obtain

$$\lambda = \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{v}|^2}$$

The vector projection of $\mathbf{u}$ onto $\mathbf{v}$ is denoted

$$\text{proj}_{\mathbf{v}}\mathbf{u} = \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{v}|^2}\mathbf{v}$$

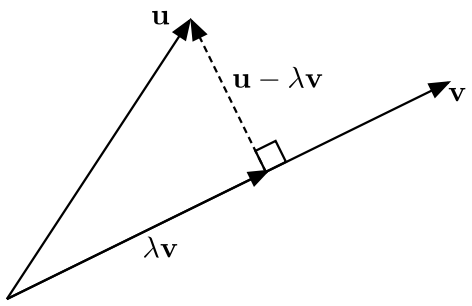


Figure 5.6 Components of a vector $\mathbf{u}$ relative to $\mathbf{v}$.

5.5 Basis and coordinate systems

Let us take a small step back and think about what we have done so far. So far, we have worked in $\mathbb{R}^n$ and in many ways $\mathbb{R}^n$ is special; indeed, we have assumed a certain number of things about $\mathbb{R}^n$ including that: (1) there is a special point that we called the **origin**, (2) there is a set of n independent axes all crossing at the origin, (3) we had a notion of a unit length.

As a mathematician, you will not always work on abstract structures but if you are so inclined you might use mathematics to solve more applied problems and model physical, biological and social systems. In the real world, all the features of $\mathbb{R}^n$ (or even $\mathbb{R}^3$ for that matter) are not obvious. Is there a universal origin? In modelling a physical problem (e.g. the rotation of a planet around a star), we could set it to be the center of the universe (if such a thing exist) but would it actually make sense to do so? In a physics experiment, there may be quite obvious choices for an origin and independent axes. As mathematician, we need to make sure that our definitions and geometric properties do not differ in different coordinates systems provided that we chose our coordinates system appropriately. For instance the distance from a point A to a point B should not depend on where I define the origin to be.

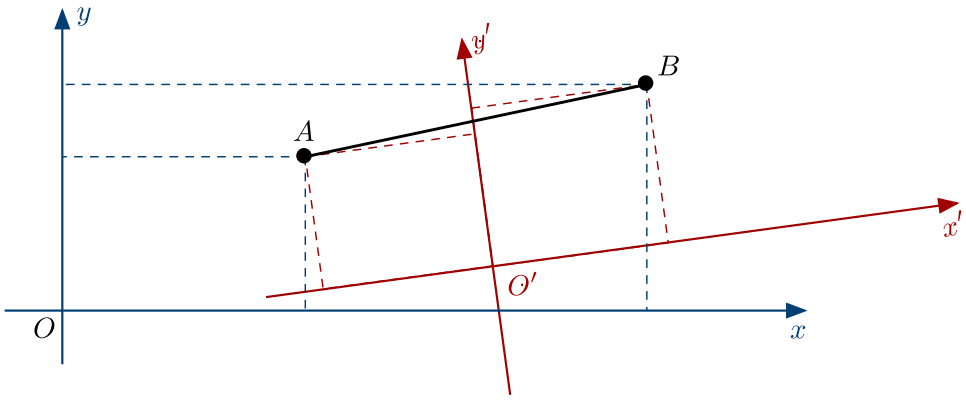


Figure 5.7 Is the distance between points A and B the same in both coordinate systems?

First, we are going to need to define the concept of **basis**. We have seen in Section 5.3.1 that with a basis we can assign unique coordinates to every vector.

Definition 5.5.1: Basis in $\mathbb{R}^n$

Let $\mathbf{u}_1$,$\mathbf{u}_2$, ... , $\mathbf{u}_n$ be n vectors in $\mathbb{R}^n$. We say that $\mathbf{u}_1$,$\mathbf{u}_2$, ... , $\mathbf{u}_n$ form a **basis** of $\mathbb{R}^n$ if for every vector $\mathbf{u} \in \mathbb{R}^n$, there exist a unique set of real numbers λ_1, λ_2, ... , λ_n such that

$$\mathbf{u} = \sum_{i=1}^{n} \lambda_i \mathbf{u}_i$$

In this case, λ_1, λ_2, ... , λ_n are called the coordinates of $\mathbf{u}$ with respect to the basis $\mathbf{u}_1$,$\mathbf{u}_2$, ... , $\mathbf{u}_n$.

Remark. *Note that the addition and multiplication by a scalar rules work in other basis than the standard basis.*

In Section 5.3.1, we have defined what we called the **standard basis** or **canonical basis** of $\mathbb{R}^n$ and denoted the vectors forming this basis $\hat{\mathbf{e}}_1$, $\hat{\mathbf{e}}_2$, ... , $\hat{\mathbf{e}}_n$. By definition of these vectors, we were able to write that

$$\forall u \in \mathbb{R}^n, \mathbf{u} = (u_1, u_2, \cdots, u_n) = u_1\hat{\mathbf{e}}_1 + u_2\hat{\mathbf{e}}_2 + \cdots + u_n\hat{\mathbf{e}}_n. \tag{5.21}$$

Until now, when we referred to the coordinates of $\mathbf{u}$, we meant the real numbers u_1, u_2, ... , u_n; when, in reality, we should have be talking about the coordinates of $\mathbf{u}$ with respect to the canonical basis.

Further, consider the statement that in $\mathbb{R}^2$ the points (x, y) satisfying $x^2 + y^2 = 1$ is a circle. Another mathematician might disagree with you and with reason. This statement is only true if you have chosen your coordinate system such that the two vectors forming what is called the basis of your plane are perpendicular to each other and of the same size! More generally, the locus of the points in $\mathbb{R}^2$ satisfying $x^2 + y^2 = 1$ would be an ellipse.

Let's try to understand why this is. We have just seen that defining a basis of $\mathbb{R}^n$ allows us to define unique coordinates for each vector. In the previous section, we introduced the concept of length and angle between vectors; calculating these quantities requires knowing the coordinates of the vectors. The question is: are the formulas we gave for the length of a vector or the angle between vectors true in any arbitrary basis? The answer is no. At this point, we need to introduce what are called **orthonormal bases**. Note that both the length of a vector and the angle between vectors was defined through the scalar product and in particular, via the fact that we can write the scalar product in the canonical basis of $\mathbb{R}^n$ as

$$\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^{n} u_i v_i \tag{5.22}$$

If we are going to pick another basis than the canonical basis to do geometry, we need to make sure (at the very least) that this definition of the scalar product holds true in our new shiny basis!

Example

Let $\mathbf{w}_1 = (1, 0, -1)$, $\mathbf{w}_2 = (2, 1, -1)$ and $\mathbf{w}_3 = (-2, 1, 4)$ be three vectors in $\mathbb{R}^3$ (as expressed in the canonical basis). These three vectors form a basis of $\mathbb{R}^3$ (you might want to check this), so we can express vectors $\mathbf{u}$ and $\mathbf{v}$ with respect to this basis as

$$\begin{aligned}\mathbf{u} &= u_1\mathbf{w}_1 + u_2\mathbf{w}_2 + u_3\mathbf{w}_3 \\ \mathbf{v} &= v_1\mathbf{w}_1 + v_2\mathbf{w}_2 + v_3\mathbf{w}_3\end{aligned}$$

then, we can write that the scalar product of these two vectors in this new basis is given by

$$\begin{aligned}\mathbf{u} \cdot \mathbf{v} &= (u_1\mathbf{w}_1 + u_2\mathbf{w}_2 + u_3\mathbf{w}_3) \cdot (v_1\mathbf{w}_1 + v_2\mathbf{w}_2 + v_3\mathbf{w}_3) \\ &= \sum_{i=1}^{3}\sum_{j=1}^{3} u_i v_i \mathbf{w}_i \cdot \mathbf{w}_j \\ &= 2u_1v_1 + 3(u_1v_2 + u_2v_1) + 6(u_2v_2 - u_1v_3 - u_3v_1) - 7(u_2v_3 + u_3v_2) + 21u_3v_3 \\ &\neq u_1v_1 + u_2v_2 + u_3v_3\end{aligned}$$

Thus, this basis does not respect our definition of the scalar product!

Proposition 5.5.1

Let $\mathbf{w}_1$, $\mathbf{w}_2$, ... , $\mathbf{w}_n$ be a basis of $\mathbb{R}^n$. Then the expression

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n \tag{5.23}$$

holds for any two vectors $\mathbf{u}$ and $\mathbf{v}$ with coordinates u_i and v_i with respect to the basis $\mathbf{w}_1$, $\mathbf{w}_2$, ... , $\mathbf{w}_n$ **if and only if**

$$\forall i \in [1, n],\ \mathbf{w}_i \cdot \mathbf{w}_i = 1 \quad \text{and} \quad \mathbf{w}_i \cdot \mathbf{w}_j = 0,\ \text{when } i \neq j \tag{5.24}$$

Proof. First, if we suppose that the scalar product is calculated as in Equation 5.23 in terms of coordinates in the basis $\mathbf{w}_1$, $\mathbf{w}_2$, ... , $\mathbf{w}_n$. Then by definition of this basis, we require that any vector $\mathbf{w}_i$ is written as a vector full of zeros except for component i which is equal to 1. So it follows immediately from Equation 5.23 that

$$\forall i \in [1, n],\ \mathbf{w}_i \cdot \mathbf{w}_i = 1 \quad \text{and} \quad \mathbf{w}_i \cdot \mathbf{w}_j = 0,\ \text{when } i \neq j$$

Conversely, if we suppose that $\mathbf{w}_i \cdot \mathbf{w}_i = 1$ for any i and $\mathbf{w}_i \cdot \mathbf{w}_j = 0$ when $i \neq j$, we can expand the following scalar product and write

$$\begin{aligned}
\mathbf{u} \cdot \mathbf{v} &= (u_1\mathbf{w}_1 + u_2\mathbf{w}_2 + \cdot + u_n\mathbf{w}_n) \cdot (v_1\mathbf{w}_1 + v_2\mathbf{w}_2 + \cdot + v_n\mathbf{w}_n) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} u_i v_j \mathbf{w}_i \cdot \mathbf{w}_j \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} u_i v_j \delta_{ij} \\
&= u_1v_1 + u_2v_2 + \cdot + u_nv_n
\end{aligned}$$

where we used δ_{ij} the Kronecker symbol defined as

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

■

Definition 5.5.2: Orthonormal basis

Let $\mathbf{w}_1$, $\mathbf{w}_2$, ... , $\mathbf{w}_n$ be a basis of $\mathbb{R}^n$. This basis is said to be an **orthonormal basis** if

$$\mathbf{w}_i \cdot \mathbf{w}_j = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \tag{5.25}$$

Note that

- A set of n othonormal vectors in $\mathbb{R}^n$ forms a basis. Indeed, orthogonality is sufficient to guarantee their linear independence.
- If $\mathbf{w}_1$, $\mathbf{w}_2$, ... , $\mathbf{w}_n$ is an othonormal basis of $\mathbb{R}^n$ and

$$\mathbf{u} = \alpha_1\mathbf{w}_1 + \alpha_2\mathbf{w}_2 + \cdots + \alpha_n\mathbf{w}_n$$

then note that we have

$$\forall i \in [1, n],\ \alpha_i = \mathbf{u} \cdot \mathbf{w}_i.$$

Remark. *This section raised deep questions in Algebra and Geometry, we shall close this discussion here for this module and leave it to the first-year Linear Algebra and Groups module to cover these in greater details. In particular, bases will be defined in your Linear Algebra and Groups module as sets that are linearly independent and spanning. While you will understand those terms formally in due time, you can develop an intuition for it in the context of what the present material. If the set of vectors of $\mathbb{R}^n$ we chose did not span $\mathbb{R}^n$ it would mean that we would not be able to assign coordinates to some points. If the set of vectors we chose was not linearly independent, it would mean that you could write one of those vectors as a linear combination of the others and in turn, it means that some points would have multiple sets of coordinates associated to them.*

Chapter 6

Geometry of Space

6.1 Euclidean geometry

In this section, we will focus on the special cases where $n = 2$ and $n = 3$ and we will introduce concepts from Euclidean geometry. This will allow us to obtain a geometrical intuition for the ideas we introduced above.

When dealing with geometry in space, one encounters two types of problems:

- problems in **affine geometry** which are mainly concerned with alignment properties, parallel lines, centroids, medians ... Indeed, it is said that affine geometry is what is left of Euclidean geometry when you forget the **metric notions** of distance and angle. Affine geometry can be developed equivalently from a set of axioms in **synthetic geometry** or directly from the results of **linear algebra**. In affine geometry, one extensively uses general cartesian coordinates. For instance, to solve a problem related to a triangle defined by points (A, B, C), one usually work in the frame defined by $\{A, \mathbf{AB}, \mathbf{AC}\}$, where A represents the origin and $\{\mathbf{AB}, \mathbf{AC}\}$ the basis.

- problems in **Euclidean geometry** which are concerned mainly with problems related to scalar product, angles, orthogonality, norms or distances. In Euclidean geometry, the expression for the scalar product or of the distance is only simple in orthonormal bases; thus, it is recommended to use coordinate systems with an orthonormal basis.

Euclid's axioms (you can read about them here!) provide us with a way to construct geometrically objects in the plane based on those postulates alone numerous results in geometry were derived. However, those axioms do not provide us with a way to parametrise these geometrical objects, this need led to the success of linear algebra. Until "recently", Euclidean geometry was thought to describe the physical space surrounding us but the 19th and 20th century have seen the need to develop non-Euclidean geometry, for instance when general relativity was developed for instance.

6.1.1 Geometry of vectors in the plane (n = 2)

Basis in the plane

In this section, we will cover some new concepts as well as concepts that you may have seen at school. We will always make an effort to provide a precise theoretical framework to derive these results in elementary geometry.

A frame in the Euclidean plane $\mathcal{P}$ is a triplet $\{P, \mathbf{u}, \mathbf{v}\}$ where P is a point of the plane (which serves as **origin of the frame**) and $\mathbf{u}$ and $\mathbf{v}$ form a **basis**. Consider the two vectors $\hat{\mathbf{i}} = (1, 0)$ and $\hat{\mathbf{j}} = (0, 1)$. In this section, we will consider the Euclidean plane $\mathcal{P}$ in

which we have defined an orthonormal frame $\mathcal{F} = \{O, \hat{\mathbf{i}}, \hat{\mathbf{j}}\}$ (i.e. $\{\hat{\mathbf{i}}, \hat{\mathbf{j}}\}$ form an orthonormal basis of the plane).

Definition 6.1.1: Collinear vectors

Two vectors $\mathbf{u}$ and $\mathbf{v}$ are said to be **collinear** if there exists $\lambda \in \mathbb{R}$ such that $\mathbf{u} = \lambda\mathbf{v}$ or $\mu \in \mathbb{R}$ such that $\mathbf{v} = \mu\mathbf{u}$.

Remark.
- *If two vectors $\mathbf{u}$ and $\mathbf{v}$ are non zero, it is equivalent to write $\mathbf{v} = \lambda\mathbf{u}$ or $\mathbf{u} = \frac{1}{\lambda}\mathbf{v}$.*
- *Otherwise, to write that $\mathbf{u}$ and $\mathbf{v}$ are collinear we can not only consider one of these equalities. Indeed, without loss of generality, consider $\mathbf{u} = \mathbf{0}$ and $\mathbf{v} \neq \mathbf{0}$. There exists λ such that $\mathbf{u} = \lambda\mathbf{v}$ (in particular, $\lambda = 0$), but we cannot find a real λ such that $\mathbf{v} = \lambda\mathbf{u}$!*

Proposition 6.1.1

In $\mathbb{R}^2$, two vectors that are **not collinear form a basis**. If these vectors are **orthogonal to one another**, then they form an **orthogonal basis**. If these are also **unit vectors**, then they form an orthonormal basis of the plane.

Proposition 6.1.2

Let $\hat{\mathbf{u}}$ be a unit vector such that $\hat{\mathbf{u}} = \alpha\hat{\mathbf{i}} + \beta\hat{\mathbf{j}}$, then $\hat{\mathbf{v}} = -\beta\hat{\mathbf{i}} + \alpha\hat{\mathbf{j}}$ is a unit vector orthogonal to $\hat{\mathbf{u}}$. Thus, the vectors that are orthogonal to $\hat{\mathbf{u}}$ are given by $\lambda\hat{\mathbf{v}}$, with $\lambda \in \mathbb{R}$.

As a consequence, there exist exactly two orthonormal basis of $\mathbb{R}^2$ with first vector $\hat{\mathbf{u}}$: either $\{\hat{\mathbf{u}}, \hat{\mathbf{v}}\}$ or $\{\hat{\mathbf{u}}, -\hat{\mathbf{v}}\}$.

Proof. It is quite trivial to realize that $\hat{\mathbf{v}}$ is a unit vector and that $\lambda\hat{\mathbf{v}}$ is orthogonal to $\hat{\mathbf{u}}$ for any $\lambda \in \mathbb{R}$.

On the other hand, let $\mathbf{w} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}}$ a vector that is orthogonal to $\hat{\mathbf{u}}$. We thus have from the previous section that $\mathbf{w} \cdot \hat{\mathbf{u}} = \alpha x + \beta y = 0$. As $\hat{\mathbf{u}}$ is non zero, we can for instance assume that $\alpha \neq 0$. Thus, we have $x = -(\beta/\alpha)y$, which gives $\mathbf{w} = (y/\alpha)\hat{\mathbf{v}}$. The case where $\beta \neq 0$ is dealt with similarly. ■

So far, we have talked about vectors in general terms and have not specified much what our coordinate system was. In what follows, we will see that both in the two-dimensional plane and in the three-dimensional space, a certain number of systems of coordinates prove particularly useful and their use depends on the geometry and the symmetries of the problem at hand.

Cartesian Coordinates

Definition 6.1.2: Cartesian Coordinates

We call Cartesian coordinates of the point M of the Euclidean plane $\mathcal{P}$ in the orthonomal frame $\mathcal{F} = \{O, \hat{\mathbf{i}}, \hat{\mathbf{j}}\}$ the real numbers x and y such that $\mathbf{u} = \mathbf{OM} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}}$.

We denote this point $M(x, y)$.

Let $M_1(x_1, y_1)$ and $M_2(x_2, y_2)$, the components of the vector $\mathbf{M_1M_2}$ are given by $(x_2 - x_1, y_2 - y_1)$ in the basis $\{\hat{\mathbf{i}}, \hat{\mathbf{j}}\}$. Further, consider λ_1 and λ_2 two real numbers such that $\lambda_1 + \lambda_2 = 1$, then the centroid $G = \lambda_1 M_1 + \lambda_2 M_2$ of the points M_1 and M_2 with weights λ_1 and λ_2 has for coordinates $(\lambda_1 x_1 + \lambda_2 x_2, \lambda_1 y_1 + \lambda_2 y_2)$.

These are direct consequences of the vectorial relations

$$\mathbf{M_1M_2} = \mathbf{OM_2} - \mathbf{OM_1} \quad \text{and} \quad \mathbf{OG} = \lambda_1 \mathbf{OM_1} + \lambda_2 \mathbf{OM_2} \tag{6.1}$$

Remark. *The orthonormal basis in the Cartesian coordinates system are commonly denoted* $\hat{\mathbf{i}}$ *and* $\hat{\mathbf{j}}$ *or respectively* $\hat{\boldsymbol{x}}$ *and* $\hat{\boldsymbol{y}}$.

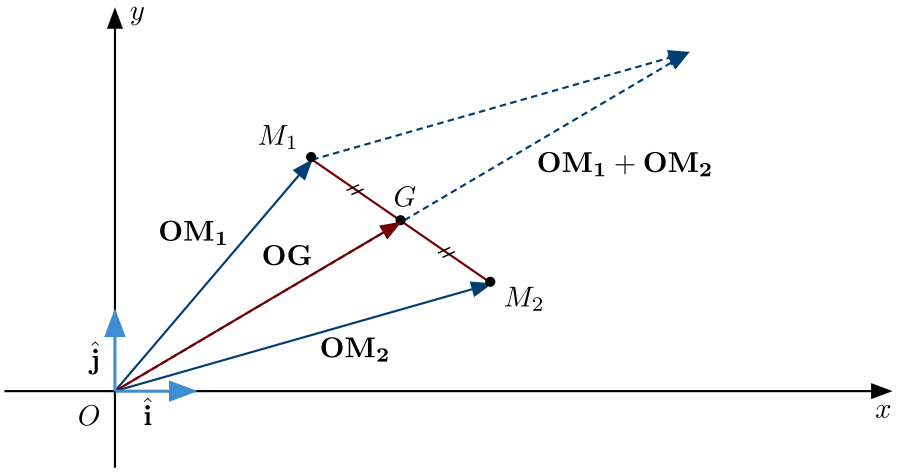


Figure 6.1 Example of position vectors for points $M_1(x_1, y_1)$ and $M_2(x_2, y_2)$ as well as their barycentre $G(\lambda_1 x_1 + \lambda_2 x_2, \lambda_1 y_1 + \lambda_2 y_2)$ (in this example, $\lambda_1 = \lambda_2 = 1/2$).

Definition 6.1.3: Cartesian Equation

An equation $f(x, y) = 0$ is a **Cartesian equation** of a part of the plane $\mathcal{A}$ if we have the following equivalence

$$M(x, y) \in \mathcal{A} \iff f(x, y) = 0 \tag{6.2}$$

Remark. *Under this definition,* $\lambda f(x, y) = 0$ *is also a Cartesian equation for* $\mathcal{A}$ *for any* $\lambda \in \mathbb{R}^*$. *The reciprocal is not true as* $(ax + by + c)^2 = 0$ *and* $ax + by + c = 0$ *are two equations that are non proportional but represent the same part of the plane.*

Example

The level sets of a function $f : \mathbb{R}^2 \to \mathbb{R}$ are the set of equations $f(x, y) = \lambda$ with $\lambda \in \mathbb{R}$.

Complex plane (or Argand plane)

The complex plane is a geometric representation of the ensemble of complex numbers $\mathbb{C}$. It can be thought of as a modified Cartesian plane where the real part of a complex number is represented by a coordinate on the x-axis (or horizontal axis) and the imaginary part is represented by a coordinate on the y-axis (or vertical axis). This representation stems from the fact that the geometry of complex numbers in many ways parallels what we have established for vectors.

Definition 6.1.4

We call **image** of the complex number $z = \alpha + i\beta$ (with $(\alpha, \beta) \in \mathbb{R}^2$) the point of coordinates (α, β) in the orthonormal frame $\mathcal{F} = \{O, \hat{\mathbf{i}}, \hat{\mathbf{j}}\}$. Conversely, the affix of a vector $\alpha\hat{\mathbf{i}} + \beta\hat{\mathbf{j}}$ (with $(\alpha, \beta) \in \mathbb{R}^2$) is the complex number $z = \alpha + i\beta$.

This provides a way to identify the Euclidean plane $\mathcal{P}$ with the set of complex numbers $\mathbb{C}$.

Example

Consider two points A and B with affixes a and b respectively.

- the affix of the vector $\mathbf{AB}$ is $b - a$;
- Consider $(\lambda, \mu) \in \mathbb{R}^2$, the affix of the centroid of A and B with respective weights λ and μ is $\lambda a + \mu b$.

Proposition 6.1.3

- The norm of a vector whose affix is $z \in \mathbb{C}$ is the modulus $|z|$.
- The distance between two points whose affixes are z_1 and z_2 is $|z_2 - z_1|$.

Proof. The first point in the proposition above is a consequence of the following relation $\forall z \in \mathbb{C}, |z| = \sqrt{a^2 + b^2}$ if $z = a + ib$ with $(a, b) \in \mathbb{R}^2$. The second point can be directly deduced from this. ■

Example

- The ensemble $\mathcal{U}$ of complex numbers with modulus 1 has for image the circle of radius 1 centered on O.
- More generally, the circle with center A and radius r is given by the ensemble of points whose affixes z verify the following relation $|z - a| = r$, with a the affix of A.
- Similarly, the open and closed disk with center a and radius r are respectively given by the points z such that $|z - a| < r$ and $|z - a| \leq r$.

Polar coordinates

Depending on symmetries, the Cartesian coordinate system might not be the best system to model a given physical problem under study. If θ is a real number, we denote:

$$\hat{\mathbf{r}} = \cos\theta\hat{\mathbf{i}} + \sin\theta\hat{\mathbf{j}} \quad \text{and} \quad \hat{\boldsymbol{\theta}} = -\sin\theta\hat{\mathbf{i}} + \cos\theta\hat{\mathbf{j}} \tag{6.3}$$

It is trivial to see that the vectors $\hat{\mathbf{r}}$ and $\hat{\boldsymbol{\theta}}$ are orthogonal unit vectors.

Definition 6.1.5: Polar coordinates

The frame $\mathcal{F} = \{O, \hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}\}$ is called the **polar frame** (or polar system of coordinates). The point O is called the **pole** and the straight line $(O, \hat{\mathbf{i}})$ (i.e. the line passing through O and directed by $\hat{\mathbf{i}}$) the **polar axis**. The distance from the pole to a given point M is called **radial distance** or simply **radius** and the angle between the polar axis and

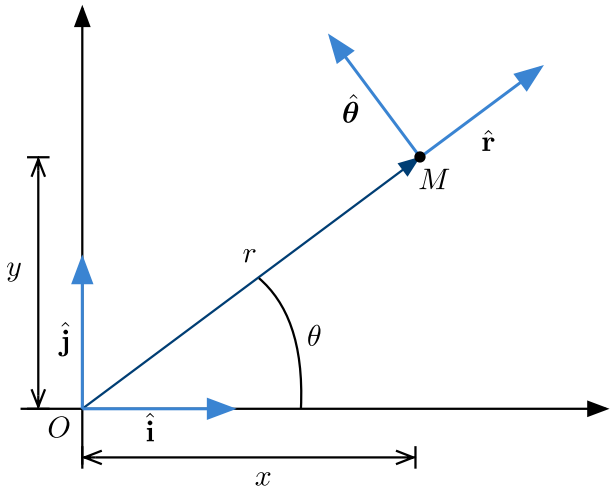


Figure 6.2 Definition of polar coordinates

the vector $\mathbf{OM}$ is called the **polar angle** or **azimuth**.

Remark. *Under this definition, to any ordered pair $(r, \theta) \in \mathbb{R}^2$, one can associate the point M such that $\mathbf{OM} = r\ \hat{\mathbf{r}}$. We just defined a map from $\mathbb{R}^2$ to the Euclidean plane, we need now to show that this map is surjective.*

Proposition 6.1.4

For any point M of the Euclidean plane $\mathcal{P}$, there exists an ordered pair $(r, \theta) \in \mathbb{R}^2$ such that:

$$\mathbf{OM} = \mathbf{r} = r\ \hat{\mathbf{r}} = r(\cos\theta\hat{\mathbf{i}} + \sin\theta\hat{\mathbf{j}}) \tag{6.4}$$

Such an ordered pair is called **a system of polar coordinates** of the point M with respect to the frame $\mathcal{F}$.

Proof.

- If $M = O$, then any ordered pair $(0, \theta)$ works.
- If $M \neq O$, we denote z the affix of point M, the modulus and the argument of z are giving us such an ordered pair. This is due to the fact that the following relations are equivalent $z = re^{i\theta}$ and $\mathbf{OM} = r\ \hat{\mathbf{r}}$.

■

Passing from Cartesian to Polar coordinates

- First, let us see how to obtain the cartesian coordinates of a point/vector whose polar coordinates are known. One can go back to the definition of a point in polar coordinates:

$$\mathbf{OM} = r\ \hat{\mathbf{r}} = r(\cos\theta\hat{\mathbf{i}} + \sin\theta\hat{\mathbf{j}}) = r\cos\theta\hat{\mathbf{i}} + r\sin\theta\hat{\mathbf{j}} \tag{6.5}$$

  We can thus identify the terms and see that:

$$x = r\cos\theta \quad \text{and} \quad y = r\sin\theta \tag{6.6}$$

- Conversely, if we are given the Cartesian coordinates of a point M of the plane (assuming $M \neq O$), we can start by calculating r

$$r = \pm\sqrt{x^2 + y^2} \tag{6.7}$$

then we define θ modulo 2π by

$$\cos\theta = \frac{x}{r} \quad \text{and} \quad \sin\theta = \frac{y}{r} \tag{6.8}$$

Remark. *In particular, for a point M whose coordinates are not of the form $(x, 0)$ with $x \leq 0$, we can get an analytical expression of $r \geq 0$ and of θ in the range $(-\pi, \pi)$ (it is the principal argument of the complex number $x + iy$):*

$$r = \sqrt{x^2 + y^2} \quad \text{and} \quad \theta = 2\arctan\left(\frac{y}{x + \sqrt{x^2 + y^2}}\right) \tag{6.9}$$

Indeed, if we know that $x = r\cos\theta$ and $y = r\sin\theta$ with $r > 0$ then

$$x + \sqrt{x^2 + y^2} = r(1 + \cos\theta) = 2r\cos^2\frac{\theta}{2} \quad \text{and} \quad y = 2r\sin\frac{\theta}{2}\cos\frac{\theta}{2} \tag{6.10}$$

which gives

$$\frac{y}{x + \sqrt{x^2 + y^2}} = \tan\frac{\theta}{2} \tag{6.11}$$

and thus proves the result.

Definition 6.1.6: Polar Equation

An equation $f(r, \theta) = 0$ is a **polar equation** of a part of the plane $\mathcal{A}$ if: a point M is a point of $\mathcal{A}$ if and only if one of its systems of polar coordinates verifies $f(r, \theta) = 0$.

Remark. *A fundamental difference with Cartesians equations is that a point in the plane admits more than one set of polar coordinates and to define a polar equation, we only require that one of these verifies the equation.*

Further, if $f(r, \theta) = 0$ is a polar equation representing a part of the plane $\mathcal{A}$, then $\lambda f(r, \theta) = 0$ is also a polar equation for $\mathcal{A}$. However, we can find equations that are not proportional and representing the same part of the plane $\mathcal{A}$. For instance, the equations $r = \cos\theta + 1$ and $r = \cos\theta - 1$ both represent the same curve: a cardioid. This is due to the fact that the points with polar coordinates (r, θ) and $(-r, \theta + \pi)$ are the same point.

Angles between vectors and the scalar product

In the previous Chapter, we have seen that the scalar product of two vectors is given by:

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 \tag{6.12}$$

with vectors $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$ as expressed in an orthonormal basis. Further, we have seen that the expression for the scalar product in function of the components of the vectors is the same in all orthonormal bases.

Let $\mathbf{u}_1$ and $\mathbf{u}_2$ vectors of $\mathbb{R}^2$, we denote their respective affixes z_1 and z_2 (elements of $\mathbb{C}$). Their scalar product is then defined as

$$\mathbf{u}_1 \cdot \mathbf{u}_2 = \mathrm{Re}(\bar{z}_1 z_2). \tag{6.13}$$

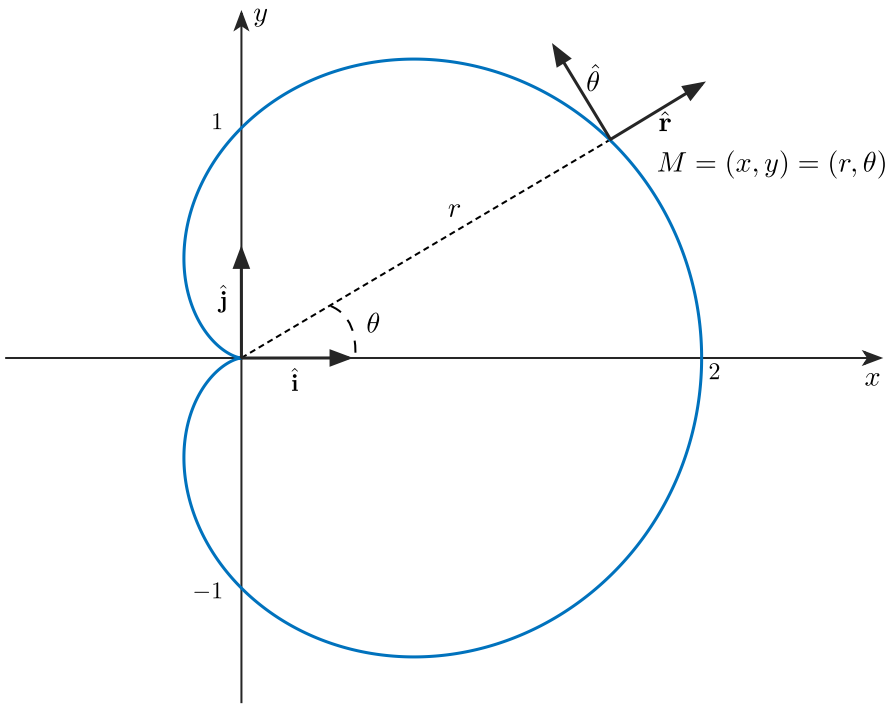


Figure 6.3 The Cardioid.

Proposition 6.1.5: Cauchy-Schwarz inequality

Let $\mathbf{u}_1$ and $\mathbf{u}_2$ be two vectors of the plane, we have

$$|\mathbf{u}_1 \cdot \mathbf{u}_2| \leq |\mathbf{u}_1||\mathbf{u}_2| \tag{6.14}$$

Proof. Let $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$ be the affixes associated to vectors $\mathbf{u}_1$ and $\mathbf{u}_2$, respectively. We have:

$$\mathbf{u}_1 \cdot \mathbf{u}_2 = \mathrm{Re}(\bar{z}_1 z_2)$$

so we obtain

$$|\mathbf{u}_1 \cdot \mathbf{u}_2| \leq |\bar{z}_1 z_2| = |z_1||z_2| = |\mathbf{u}_1||\mathbf{u}_2|$$

■

Definition 6.1.7: Angle between two vectors

Let $\mathbf{u}$ and $\mathbf{v}$ be two vectors such that $\mathbf{u} \neq 0$ and $\mathbf{v} \neq 0$, then

$$\exists!\ \theta \in [0, \pi] \ :\ \mathbf{u} \cdot \mathbf{v} = |\mathbf{u}||\mathbf{v}| \cos\theta$$

This real number is called a **measure of the angle** between vectors $\mathbf{u}$ and $\mathbf{v}$.

Further, let $\mathbf{u}$ and $\mathbf{v}$ be unit vectors then the measure of the angle between the straight lines $\mathcal{D}_u$ and $\mathcal{D}_v$ (respectively, directed by vectors $\mathbf{u}$ and $\mathbf{v}$) is the unique real number $\theta \in [0, \pi/2]$ such that

$$|\mathbf{u} \cdot \mathbf{v}| = \cos\theta$$

Remark. *We had introduced the concept of angle between vectors in Chapter 5, we see now that this definition is consistent with our usual definition of angles between vectors in the plane.*

Example

- In particular, two non-zero vectors are orthogonal to each other if and only if a measure of their angle is $\pi/2$.
- Let A, B and C be three points distinct of each other. If we denote θ a measure of the angle between vectors $\mathbf{AB}$ and $\mathbf{AC}$ then:

$$d(B,C)^2 = d(A,B)^2 + d(A,C)^2 - 2d(A,B)d(A,C)\cos\theta \tag{6.15}$$

  Indeed, we know that $\mathbf{BC} = \mathbf{AC} - \mathbf{AB}$ which gives

$$\mathbf{BC}\cdot\mathbf{BC} = \mathbf{AC}\cdot\mathbf{AC} + \mathbf{AB}\cdot\mathbf{AB} - 2\mathbf{AC}\cdot\mathbf{AB} \tag{6.16}$$

- In particular, we recover **Pythagoras' Theorem**: the triangle (ABC) is a right triangle in A if and only if $d(B,C)^2 = d(A,B)^2 + d(A,C)^2$.

Determinants and angles: Orientations of the plane

Imagine drawing the real line as a horizontal line. When we set that the x-axis is oriented by the vector $\hat{\mathbf{i}}$, we have set what is "left" and is "right". Now imagine laying down another axis, the y-axis orthogonal to the real line. You now draw a vertical line. Let $\hat{\mathbf{u}} = \alpha\hat{\mathbf{i}} + \beta\hat{\mathbf{j}}$ be a unit vector of $\mathbb{R}^2$. We have seen in Proposition 6.1.2 that there are exactly two vectors $\hat{\mathbf{v}}$ such that $\{\hat{\mathbf{u}}, \hat{\mathbf{v}}\}$ is an orthonormal basis of the plane: those are $\hat{\mathbf{v}} = \beta\hat{\mathbf{i}} + \alpha\hat{\mathbf{j}}$ and its opposite. So in particular, if $\hat{\mathbf{u}} = \hat{\mathbf{i}}$, deciding whether we define $\hat{\mathbf{j}} = (0,1)$ or $\hat{\mathbf{j}} = (0,-1)$ (both unit vectors orthogonal to $\hat{\mathbf{i}}$) is deciding what we call "up" and what we call "down". We are thus chosing an orientation of the plane.

Definition 6.1.8: Right-handed orthonormal basis

In the plane $\mathcal{P}$ oriented by the orthonormal basis $\{\hat{\mathbf{i}}, \hat{\mathbf{j}}\}$, an orthonormal basis $\{\hat{\mathbf{u}}, \hat{\mathbf{v}}\}$ is said to be **right-handed** (or positively oriented) if the vectors $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$ are given by

$$\hat{\mathbf{u}} = \alpha\hat{\mathbf{i}} + \beta\hat{\mathbf{j}} \quad \text{and} \quad \hat{\mathbf{v}} = -\beta\hat{\mathbf{i}} + \alpha\hat{\mathbf{j}} \quad \text{with} \quad (\alpha,\beta) \in \mathbb{R}^2 \text{ and } \alpha^2 + \beta^2 = 1$$

If $\hat{\mathbf{u}}$ is a unit vector, we call the **unit vector positively orthogonal** to $\hat{\mathbf{u}}$ the unique vector $\hat{\mathbf{v}}$ such that $\{\hat{\mathbf{u}}, \hat{\mathbf{v}}\}$ is a right-handed orthonormal basis.

Remark. *An orthonormal basis $\{\hat{\mathbf{u}}, \hat{\mathbf{v}}\}$ is said to be* **left-handed** *(or negatively oriented) if the vectors $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$ are given by*

$$\hat{\mathbf{u}} = \alpha\hat{\mathbf{i}} + \beta\hat{\mathbf{j}} \quad \text{and} \quad \hat{\mathbf{v}} = \beta\hat{\mathbf{i}} - \alpha\hat{\mathbf{j}} \quad \text{with} \quad (\alpha,\beta) \in \mathbb{R}^2 \text{ and } \alpha^2 + \beta^2 = 1$$

Definition 6.1.9: Determinant of two vectors

Let $\mathbf{u}$ and $\mathbf{v}$ be non-zero vectors and θ a measure of the signed angle between $\mathbf{u}$ and $\mathbf{v}$. The determinant of $\mathbf{u}$ and $\mathbf{v}$ is given by

$$\det(\mathbf{u},\mathbf{v}) = |\mathbf{u}||\mathbf{v}|\sin\theta$$

If $\mathbf{u} = \mathbf{0}$ or $\mathbf{v} = \mathbf{0}$ then we define $\det(\mathbf{u},\mathbf{v}) = 0$.

Proposition 6.1.6: Collinear vectors

Two vectors $\mathbf{u}$ and $\mathbf{v}$ are collinear if and only if $\det(\mathbf{u}, \mathbf{v}) = 0$.

Proof. If one of the two vectors is zero, then they are collinear and their determinant is zero. Otherwise, we denote θ a measure of the angle between vectors $\mathbf{u}$ and $\mathbf{v}$. The determinant is zero if and only if $\sin\theta = 0$, i.e. $\theta \equiv 0[\pi]$, thus if and only if they are collinear by definition. ■

Corollary 6.1.1

Three points A, B and C are aligned if and only if $\det(\mathbf{AB}, \mathbf{AC}) = 0$.

Example

- If $\mathbf{u}$ and $\mathbf{v}$ are two orthogonal vectors, we have $\det(\mathbf{u}, \mathbf{v}) = \pm|\mathbf{u}||\mathbf{v}|$.
- More precisely, if the vector $\mathbf{v}$ is *positively* orthogonal to $\mathbf{u}$, then $\det(\mathbf{u}, \mathbf{v}) = |\mathbf{u}||\mathbf{v}|$
- In particular, if $(\mathbf{u}, \mathbf{v})$ is a right-handed orthonormal basis

$$\det(\mathbf{u}, \mathbf{v}) = 1 \quad \text{and} \quad \det(\mathbf{v}, \mathbf{u}) = -1$$

Corollary 6.1.2

Let $\mathbf{u}_1 = x_1\hat{\mathbf{i}} + y_1\hat{\mathbf{j}}$ and $\mathbf{u}_2 = x_2\hat{\mathbf{i}} + y_2\hat{\mathbf{j}}$, then

$$\det(\mathbf{u}_1, \mathbf{u}_2) = \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} = x_1y_2 - x_2y_1$$

Remark. *We can easily realize that*

$$\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} = x_1y_2 - x_2y_1 = \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}$$

Thus, the vector $\mathbf{u}_1 = (x_1, y_1)$ and $\mathbf{u}_2 = (x_2, y_2)$ are collinear if and only if the vectors $\mathbf{v}_1 = (x_1, x_2)$ and $\mathbf{v}_2 = (y_1, y_2)$ are collinear.

Further, the collinearity of the vectors $\mathbf{u}_1 = (x_1, y_1)$ and $\mathbf{u}_2 = (x_2, y_2)$ is equivalent to the proportionality of the ordered pairs (x_1, y_1) and (x_2, y_2), thus

$$\exists\lambda \in \mathbb{R}^* : (x_1, y_1) = \lambda(x_2, y_2) \iff \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = 0$$

Proposition 6.1.7: Properties of the determinant

The determinant is an **antisymmetric bilinear** form, i.e. for all vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$

and for all λ and μ in $\mathbb{R}$, we have

(a) $\det(\mathbf{u},\mathbf{v}) = -\det(\mathbf{v},\mathbf{u})$ (antisymmetry)
(b) $\det(\lambda\mathbf{u}+\mu\mathbf{w},\mathbf{v}) = \lambda\det(\mathbf{u},\mathbf{v}) + \mu\det(\mathbf{w},\mathbf{v})$ (linearity in the first argument)
(c) $\det(\mathbf{u},\lambda\mathbf{v}+\mu\mathbf{w}) = \lambda\det(\mathbf{u},\mathbf{v}) + \mu\det(\mathbf{u},\mathbf{w})$ (linearity in the second argument)

Remark. *Let $\mathbf{u}$ and $\mathbf{v}$ be two vectors and x_1, x_2, y_1 and y_2 in $\mathbb{R}$, we have*

$$\begin{aligned}\det(x_1\mathbf{u}+y_1\mathbf{v}, x_2\mathbf{u}+y_2\mathbf{v}) &= x_1x_2\det(\mathbf{u},\mathbf{u}) + (x_1y_2 - x_2y_1)\det(\mathbf{u},\mathbf{v}) + y_1y_2\det(\mathbf{v},\mathbf{v})\\ &= (x_1y_2 - x_2y_1)\det(\mathbf{u},\mathbf{v}).\end{aligned}$$

In particular, we deduce from this that if $(\mathbf{u},\mathbf{v})$ is a right-handed orthonormal basis, we have

$$\det(x_1\mathbf{u}+y_1\mathbf{v}, x_2\mathbf{u}+y_2\mathbf{v}) = x_1y_2 - x_2y_1$$

The expression for the determinant as a function of vector components is thus the same in all right-handed orthonormal bases.

Proposition 6.1.8

Let $\theta \in \mathbb{R}$ be a measure of the signed angle between unit vectors $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$, then

$$\cos\theta = \hat{\mathbf{u}}\cdot\hat{\mathbf{v}} \quad \text{and} \quad \sin\theta = \det(\hat{\mathbf{u}},\hat{\mathbf{v}})$$

We also have:

$$\widehat{(\hat{\mathbf{u}},\hat{\mathbf{v}})} = -\widehat{(\hat{\mathbf{v}},\hat{\mathbf{u}})}$$

Example

- For all vectors $\mathbf{u}$ and $\mathbf{v}$, we have

$$(\mathbf{u}\cdot\mathbf{v})^2 + \det(\mathbf{u},\mathbf{v})^2 = |\mathbf{u}|^2|\mathbf{v}|^2 \quad (*) \tag{6.17}$$

  This is trivial from the expression of the dot product and determinant in terms of affixes.

- We also deduce that if $\mathbf{u}$ and $\mathbf{v}$ are orthogonal, then $|\det(\mathbf{u},\mathbf{v})| = |\mathbf{u}||\mathbf{v}|$.

- The area of a parallelogram $(ABCD)$ is given by $|\det(\mathbf{AB},\mathbf{AD})|$.

- If we write relation $(*)$ in a right-handed (positively oriented) orthonormal basis of the plane, we obtain what is called Lagrange's relation:

$$(a^2+b^2)(c^2+d^2) = (ad-bc)^2 + (ac+bd)^2.$$

  In particular, this formula is useful to prove that if two integers are sums of squares then their product is as well.

6.1.2 Parametrization and properties of geometric objects in the plane

Parametric representation of a line

We work in the Euclidean plane $\mathcal{P}$ in which we have defined the orthonormal frame $\mathcal{F} = \{O,\hat{\mathbf{i}},\hat{\mathbf{j}}\}$. Let $\mathcal{L}$ be the line going through a point P with position vector $\mathbf{p} = (x_0, y_0)$ and

directed by non-zero vector $\mathbf{u} = \alpha\hat{\mathbf{i}} + \beta\hat{\mathbf{j}}$. By extension of Definition 5.3.7, one quickly realizes that this line $\mathcal{L}$ in $\mathbb{R}^2$ is the set of points $R(x, y)$ such that:

$$\begin{cases} x = x_0 + \lambda\alpha \\ y = y_0 + \lambda\beta \end{cases} \quad \text{with} \quad \lambda \in \mathbb{R}$$

We can write this parametrization vectorially and obtain

$$\mathbf{r}(\lambda) = \mathbf{p} + \lambda\mathbf{u},\ \lambda \in \mathbb{R} \tag{6.18}$$

By doing so, we just defined $\lambda \mapsto \mathbf{r}(\lambda) = \mathbf{p} + \lambda\mathbf{u}$ to be a surjection of $\mathbb{R}$ onto $\mathcal{L}$. Here, we even have a bijection because for all points of $\mathcal{L}$, there is a unique $\lambda \in \mathbb{R}$. This equation is said to be in **parametric form** and the real λ is the **parameter**.

Cartesian equation of a line

Proposition 6.1.9

- Any line $\mathcal{L}$ of the plane has at least one cartesian equation of the kind

$$ax + by + c = 0 \quad \text{with} \quad (a, b) \neq (0, 0)$$

- Two such equations represent the same line if they are proportional to each other.
- Conversely, any equation of the form:

$$ax + by + c = 0 \quad \text{with} \quad (a, b) \neq (0, 0)$$

represents a straight line parallel to the vector $(-b, a)$ and thus orthogonal to the vector with components (a, b).

Proof. Let $A = (x_0, y_0)$ be a point of the line $\mathcal{L}$ and $\mathbf{u} = (\alpha, \beta)$ with $(\alpha, \beta) \neq (0, 0)$, a vector orienting the line $\mathcal{L}$.

- A point $M = (x, y)$ is part of $\mathcal{L}$ if and only if $\det(\mathbf{AM}, \mathbf{u}) = 0$. Thus, the straight line is represented by the following equation

$$\beta(x - x_0) - \alpha(y - y_0) = 0 \quad (*)$$

with $(\beta, -\alpha) \neq (0, 0)$.

- Consider $ax + by + c = 0$ to be the cartesian equation of line $\mathcal{L}$. If $B = (x_1, y_1)$ is a point of $\mathcal{L}$ with $B \neq A$, then we have:

$$a(x_1 - x_0) + b(y_1 - y_0) = 0$$

Thus, the vector with components (a, b) is orthogonal to the vector $\mathbf{AB}$ and so it is orthogonal to $\mathcal{L}$, which proves the existence of $\lambda \in \mathbb{R}$ such that $(a, b) = \lambda(-\beta, \alpha)$.

In particular, we notice that we can write the equation $ax + by + c = 0$ as $a(x - x_0) + b(y - y_0) = 0$. This last equation is proportional to equation $(*)$.

- Finally, we can find a point $A = (x_0, y_0)$ verifying the equation: for instance, if $a \neq 0$, we take $y_0 = 0$ and $x_0 = -c/a$. The equation is then equivalent to $a(x - x_0) + b(y - y_0) = 0$, i.e. $\det(\mathbf{AM}, \mathbf{u})$ with $\mathbf{u} = (-b, a)$ or $\mathbf{AM} \cdot \mathbf{v} = 0$ with $\mathbf{v} = (a, b)$. This equation thus represents the line $(A, \mathbf{u})$ which is orthogonal to $\mathbf{v}$.

■

Example

- If $b \neq 0$, then the equation $ax+by+c=0$ can be equivalently written $y=mx+p$; however, this second equation does not allow to represent vertical lines (which are of the kind $x=a$, with $a \in \mathbb{R}$).
- Consider a line $\mathcal{L}$ going through points $M_1=(x_1,y_1)$ and $M_2=(x_2,y_2)$, a point $M=(x,y)$ is part of the line $\mathcal{L}$ if and only if the vectors $\mathbf{MM_1}$ and $\mathbf{MM_2}$ are collinear, i.e.
$$\begin{vmatrix} x_1-x & x_2-x \\ y_1-y & y_2-y \end{vmatrix} = 0 \tag{6.19}$$
- If a and b are real numbers such that $a \neq 0$ and $b \neq 0$, then the line going through the points $A=(a,0)$ and $B=(0,b)$ is represented by the following equation
$$\frac{x}{a}+\frac{y}{b}=1$$
Indeed, this is the equation of a line and it is easy to check that it goes through the points A and B.

Remark. *Note that to check whether a vector $\mathbf{u}=(\alpha,\beta)$ is parallel to a line $\mathcal{L}$ of equation $ax+by+c=0$, one needs to check whether:*

$$a\alpha+b\beta=0$$

and ***NOT*** *$a\alpha+b\beta+c=0$!*

Proposition 6.1.10

Consider two lines $\mathcal{L}_1$ and $\mathcal{L}_2$ with respective equations $a_1x+b_1y+c_1=0$ and $a_2x+b_2y+c_2=0$, then

- The lines $\mathcal{L}_1$ and $\mathcal{L}_2$ are parallel to each other if and only if (a_1,b_1) and (a_2,b_2) are proportional, i.e. if and only if
$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = 0$$
- Otherwise, there exists a unique point of intersection.

Proof. On one hand, the vectors (a_1,b_1) and (a_2,b_2) are orthogonal to $\mathcal{L}_1$ and $\mathcal{L}_2$ respectively and they are collinear if and only if the lines are parallel. If there exists an intersection point A, then both of these lines are equal to the line going through point A and parallel to the vector $(-b_1,a_1)$.

On the other hand, if the lines are not parallel, determining their intersection is solving a 2-by-2 linear system of equations for which the determinant is non zero and thus admits a unique solution (see below). ■

Example

Two lines that are orthogonal to the same non zero vector $\mathbf{u}$ are parallel.

Remark. *Two lines represented by the following equations $a_1x+b_1y+c_1=0$ and $a_2x+b_2y+c_2=0$ are:*

1. *parallel if and only if* $(-b_1, a_1)$ *and* $(-b_2, a_2)$ *are proportional, i.e. if and only if* (a_1, b_1) *and* (a_2, b_2) *are proportional;*

2. *equal to one another if and only if* (a_1, b_1, c_1) *and* (a_2, b_2, c_2) *are proportional.*

Polar equation of a line

Proposition 6.1.11

A line that goes through the pole admits at least one polar equation of the kind:

$$\theta = \theta_0 \quad \text{with} \quad \theta_0 \in \mathbb{R}$$

Conversely, such an equation is the equation of a line going through the pole.

Proof. This proof is left as an exercise to the reader. ■

Proposition 6.1.12

A line that does not pass through the pole admits one polar equation of the kind

$$r = \frac{1}{\alpha \cos\theta + \beta \sin\theta} \quad \text{with} \quad (\alpha, \beta) \neq (0, 0)$$

Proof. A line $\mathcal{L}$ which does not go through the pole O admits a cartesian equation of the kind

$$ax + by = c \quad \text{with} \quad (a, b) \neq (0, 0)$$

A point with polar coordinates (r, θ) is part of $\mathcal{L}$ if and only if $r(a\cos\theta + b\sin\theta) = c$, which gives the expected result with $\alpha = a/c$ and $\beta = b/c$.

Conversely, the equation $r = \frac{1}{\alpha\cos\theta + \beta\sin\theta}$ is a polar equation of the line with cartesian equation $\alpha x + \beta y = 1$. ■

Lines and orthogonality

Consider $\mathbf{u}$ a non zero vector, A a point of the plane and k a real, the set of points M verifying the following relation:

$$\mathbf{u} \cdot \mathbf{AM} = k \quad (**)$$

is a line orthogonal to $\mathbf{u}$. Indeed, we can consider without loss of generality that $\mathbf{u}$ is a unit vector, then equation $(**)$ is equivalent to

$$\mathbf{u} \cdot \mathbf{AM} = \mathbf{u} \cdot (k\mathbf{u})$$

thus, obtaining

$$\mathbf{u} \cdot (\mathbf{AM} - k\mathbf{u}) = 0$$

which proves that the set of points M verifying $(**)$ is the line going through the point $A + k\mathbf{u}$ and orthogonal to $\mathbf{u}$.

Distance to a line

Working in Cartesian coordinates, if you consider a line $\mathcal{L}$ going through point $A(a,b)$ and oriented by vector $\mathbf{u}=(\alpha,\beta)$. We can compute the distance from a given point $M(x,y)$ to the line $\mathcal{L}$ using determinants. Indeed, if we denote H the orthogonal projection of point M on the line $\mathcal{L}$, we have that the distance from M to $\mathcal{L}$ is the length of the vector $\mathbf{HM}$. Further, we know that by bilinearity of the determinant, we have:

$$\det(\mathbf{AM},\mathbf{u})=\det(\mathbf{AH},\mathbf{u})+\det(\mathbf{HM},\mathbf{u})$$

As $\mathbf{AH}$ and $\mathbf{u}$ are collinear by definition of H, we then have:

$$\det(\mathbf{AM},\mathbf{u})=\det(\mathbf{HM},\mathbf{u})$$

and as $\mathbf{HM}$ and $\mathbf{u}$ are orthogonal vectors, we know that

$$|\det(\mathbf{HM},\mathbf{u})|=|\mathbf{HM}||\mathbf{u}|$$

which gives us

$$\boxed{d(M,\mathcal{L})=\frac{|\det(\mathbf{AM},\mathbf{u})|}{|\mathbf{u}|}=\frac{|\beta(x-a)-\alpha(y-b)|}{\sqrt{\alpha^2+\beta^2}}}$$

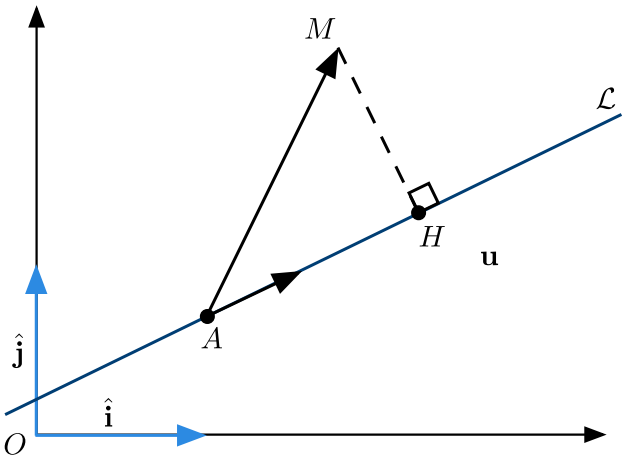


Figure 6.4 Distance from point M to line $\mathcal{L}$.

Further, consider now that we are provided with a Cartesian equation for a line $\mathcal{L}$, say $ax+by+c=0$. Then, the distance between point $M(x,y)$ and the line $\mathcal{L}$ is given by

$$\boxed{d(M,\mathcal{L})=\frac{|ax+by+c|}{\sqrt{a^2+b^2}}}$$

Proof. This result is a direct consequence of the previous result on lines defined by a parametrization in terms of a point and a vector orienting the line. ■

Further link between geometry and linear algebra

When you work in linear algebra, you work with linear equations and in particular, systems of linear equations. Consider the following system of equations

$$\begin{cases} ax+by=e \\ cx+dy=f \end{cases}$$

We saw earlier that the equation $ax + by = e$ is the Cartesian equation of the straight line perpendicular to vector $\mathbf{u} = (a, b)$. Similarly, the second equation can be seen as the Cartesian equation of a line perpendicular to the vector $\mathbf{v} = (c, d)$.

A solution of this set of equations (x_0, y_0) by definition satisfies both equations simultaneously. Geometrically, we can then see solving this system of equations as finding the potential intersection of these lines. Therefore, there are three possible outcomes:

- The vectors are not collinear, then the lines intersect in a single point; this means that the system of equation has a single solution;
- The vectors are collinear, then the lines are parallel; this means that the system of equation has either zero solution or an infinity of solutions.

We have seen that the condition for the vectors $\mathbf{u}$ and $\mathbf{v}$ to be collinear was given in terms of the determinant of these vectors.

Proposition 6.1.13

We call

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

the **determinant of the system of equations**. If this determinant is **non zero**, then the system of equations admits a unique solution given by

$$x = \frac{\begin{vmatrix} e & b \\ f & d \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}} \quad \text{and} \quad y = \frac{\begin{vmatrix} a & e \\ c & f \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}$$

Proof.

- **Unicity:** consider that the system of equations admits two solutions (x_1, y_1) and (x_2, y_2), then $(x, y) = (x_1 - x_2, y_1 - y_2)$ satisfies the following system of equations

$$\begin{cases} ax + by = 0 \\ cx + dy = 0 \end{cases}$$

  In the plane, this means that the vectors (a, b) and (c, d) are orthogonal to (x, y); but these two vectors are not collinear as their determinant

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \neq 0.$$

  Thus, $(x, y) = (0, 0)$ by Corollary 6.1.1.

- **Existence:** it suffices to check that the expression provided above are solutions to the system of equations.

■

Remark.

- *In the next section, we will derive results in three dimensions. Can you see how to translate these ideas to systems of three equations with three unknowns? What is then the geometric intuition?*
- *These ideas will be generalized to higher dimensional spaces in your Linear Algebra & Groups module.*

6.1.3 Geometry of vectors in space (n = 3)

In the previous section, we worked extensively in $\mathbb{R}^2$; in this section, we will consider the xyz-space represented by $\mathbb{R}^3$.

6.1.4 System of coordinates in the Euclidean space

Cartesian system of coordinates

We consider the Euclidean three-dimensional space $\mathcal{S}$ and the frame $\mathcal{F} = \{O, \hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\}$. Any point M of $\mathcal{S}$ can be represented by what are called its **Cartesian coordinates** (x, y, z) in the orthonormal frame $\mathcal{F}$ (see Figure 6.5). This allows us to identify the space $\mathcal{S}$ to $\mathbb{R}^3$, by associating to any point M the triplet $(x, y, z) \in \mathbb{R}^3$ which are the coordinates of the point in the frame $\{O, \hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\}$.

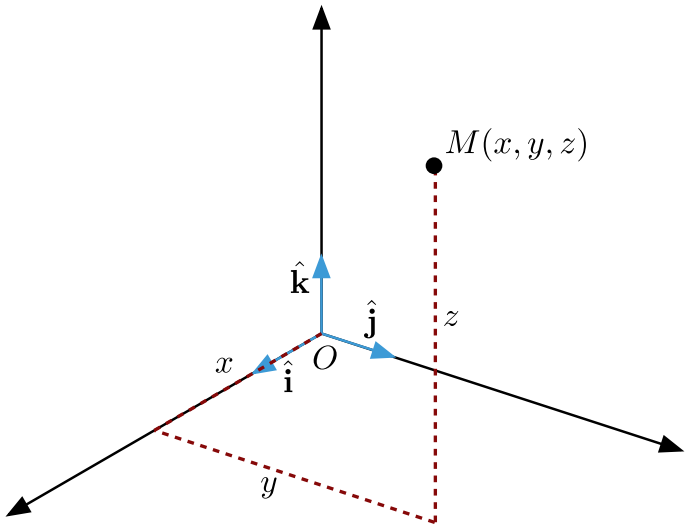


Figure 6.5 Cartesian coordinates in three dimensions

Example

An equation in x, y and z will represent a **surface** in $\mathbb{R}^3$.

- The equation $z = 3$ represents the set $\{(x, y, z) : z = 3\}$, this is the set of all points with z-coordinate 3. It is a plane parallel to the xy-plane and three units above it.
- The equation $y = 5$ represents the set $\{(x, y, z) : y = 5\}$, this is the set of all points with y-coordinate 5. It is a plane parallel to the xz-plane and five units above it.
- Describe the surface in $\mathbb{R}^3$ represented by equation $y = x$.

Remark. *When an equation is given, you will need to understand from context whether the equation describes a curve in $\mathbb{R}^2$ or a surface in $\mathbb{R}^3$. Indeed, the equation $y = 5$ in $\mathbb{R}^2$ could be interpreted as the line parallel to the x-axis and five units above.*

Cylindrical system of coordinates

It is relatively easy to extend to three dimensions, the concept of polar coordinates introduced in the plane. Given a real number θ, we denote

$$\hat{\mathbf{r}} = \cos\theta\hat{\mathbf{i}} + \sin\theta\hat{\mathbf{j}} \quad \text{and} \quad \hat{\boldsymbol{\theta}} = -\sin\theta\hat{\mathbf{i}} + \cos\theta\hat{\mathbf{j}} \tag{6.20}$$

If M is a point with cartesian coordinates (x, y, z) in the orthonormal frame $\mathcal{F} = \{O, \hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\}$, we denote P its orthogonal projection on the xy-plane (with orthonormal frame $\{O, \hat{\mathbf{i}}, \hat{\mathbf{j}}\}$). If we take (r, θ) a system of polar coordinates of P, then the position vector of point M can be written as:

$$\mathbf{r} = \mathbf{OM} = \mathbf{OP} + z\hat{\mathbf{k}} = r\,\hat{\mathbf{r}} + z\hat{\mathbf{k}}$$

Definition 6.1.10: Cylindrical coordinates

Let M be a point of the Euclidean space $\mathcal{S}$ with Cartesian coordinates (x, y, z), we call a **system of cylindrical coordinates** of M with respect to frame $\mathcal{F}$ a triplet $(r, \theta, z) \in \mathbb{R}^3$ such that

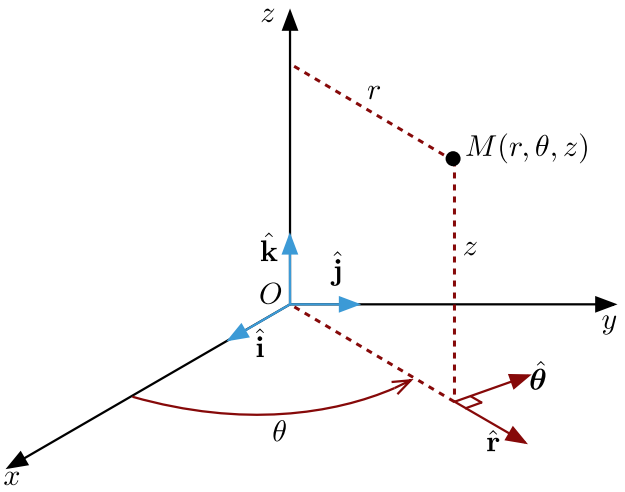
$$\mathbf{r} = \mathbf{OM} = r\,\hat{\mathbf{r}} + z\hat{\mathbf{k}}$$

with the following relations

$$\begin{cases} r = \sqrt{x^2 + y^2} \\ \theta = \arctan(y/x) \end{cases}$$

and unit vectors defined as

$$\begin{cases} \hat{\mathbf{r}} = \cos\theta\hat{\mathbf{i}} + \sin\theta\hat{\mathbf{j}} \\ \hat{\boldsymbol{\theta}} = -\sin\theta\hat{\mathbf{i}} + \cos\theta\hat{\mathbf{j}} \end{cases}$$



Remark. *The orthonormal basis $\{\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}, \hat{\mathbf{k}}\}$ is obtained by rotating the canonical orthonormal basis $\{\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\}$ by an angle θ about the axis $(O, \hat{\mathbf{k}})$.*

Spherical system of coordinates

Finally, one last system of coordinates which proves useful in three dimensions depending on the symmetries of the problem you are trying to solve is the system of spherical coordinates. Let M be a point with cylindrical coordinates (ρ, ϕ, z). We know that in this case we can write:

$$\mathbf{OM} = \rho\hat{\mathbf{r}} + z\hat{\mathbf{k}}$$

Thus, we can define $r = |\mathbf{OM}| = \sqrt{\rho^2 + z^2}$. There exists a real number θ such that $z = r\cos\theta$ and $\rho = r\sin\theta$. Finally, the Cartesian coordinates of M satisfy the following relations:

$$\begin{cases} x = \rho\cos\phi = r\cos\phi\sin\theta \\ y = \rho\sin\phi = r\sin\phi\sin\theta \\ z = r\cos\theta \end{cases}$$

Definition 6.1.11: Spherical coordinates

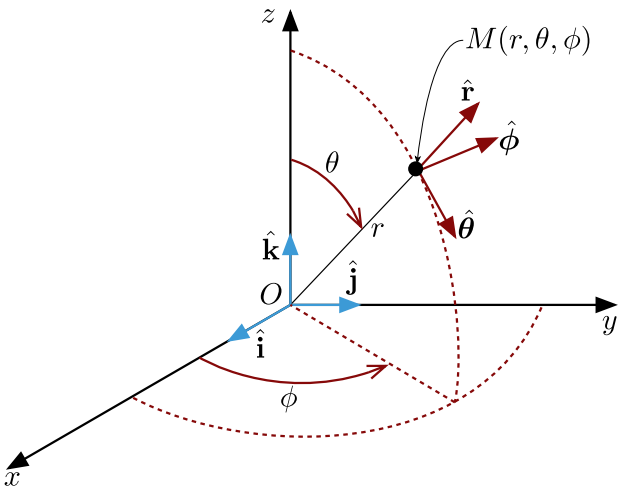
Let M be a point of the Euclidean space $\mathcal{S}$ with Cartesian coordinates (x, y, z), we call a **system of spherical coordinates** of M with respect to frame $\mathcal{F}$ a triplet $(r, \theta, \phi) \in \mathbb{R}^3$ such that $r \geq 0$, $\theta \in [0, \pi]$ and

$$\mathbf{r} = \mathbf{OM} = r\,\hat{\mathbf{r}}$$

with the following relations

$$\begin{cases} r = \sqrt{x^2 + y^2 + z^2} \\ \theta = \arccos(z/r) \\ \phi = \arctan(y/x) \end{cases}$$

and unit vectors defined as



$$\begin{cases} \hat{\mathbf{r}} = \sin\theta\cos\phi\hat{\mathbf{i}} + \sin\theta\sin\phi\hat{\mathbf{j}} + \cos\theta\hat{\mathbf{k}} \\ \hat{\boldsymbol{\theta}} = \cos\theta\cos\phi\hat{\mathbf{i}} + \cos\theta\sin\phi\hat{\mathbf{j}} - \sin\theta\hat{\mathbf{k}} \\ \hat{\boldsymbol{\phi}} = -\sin\phi\hat{\mathbf{i}} + \cos\phi\hat{\mathbf{j}} \end{cases}$$

In this coordinate system, the position of a point is described by the **radial distance** from that point to the fixed origin r, a **polar/inclination angle** θ measured from a fixed zenith (z-axis) and the **azimuth angle** ϕ of its orthogonal projection on the plane that passes through the origin and is orthogonal to the zenith direction (xy-plane).

Remark. *Note that there is some ambiguity in the definition of the spherical coordinates:*

- *Physicists and mathematicians tend to use different symbols for the angles; they use inverse naming convention for θ the polar angle and ϕ the azimuthal angle.*
- *One might express this system of coordinates in terms of the elevation angle measured from the fixed xy-plane rather than the polar angle as measured from the fixed z-axis (zenith). This is done in particular in physical geography.*

You should not be confused by this and be able to re-derive the proper expression independently of the naming convention.

6.1.5 Vector product

Lemma 6.1.1

Let $\mathbf{u}_1 = x_1\hat{\mathbf{i}} + y_1\hat{\mathbf{j}} + z_1\hat{\mathbf{k}}$ and $\mathbf{u}_2 = x_2\hat{\mathbf{i}} + y_2\hat{\mathbf{j}} + z_2\hat{\mathbf{k}}$ be two vectors of $\mathbb{R}^3$. These vectors are collinear if and only if we have

$$\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} = \begin{vmatrix} x_1 & x_2 \\ z_1 & z_2 \end{vmatrix} = \begin{vmatrix} y_1 & y_2 \\ z_1 & z_2 \end{vmatrix} = 0$$

Proof. This is left as a proof to the student, see Part III - Problem Sheet 3. ■

Proposition 6.1.14

Consider $\mathbf{u}_1$ and $\mathbf{u}_2$ two non collinear vectors.

- There exists a vector $\mathbf{v}$ orthogonal to both $\mathbf{u}_1$ and $\mathbf{u}_2$.
- The vectors orthogonal to $\mathbf{v}$ are the linear combinations of the vectors $\mathbf{u}_1$ and $\mathbf{u}_2$.

Definition of the vector product

Along with the dot product, we can define in $\mathbb{R}^3$ (not in general in $\mathbb{R}^n$!) another operation between vectors called the vector product.

Definition 6.1.12: Vector Product

Let $\mathbf{u}_1 = x_1\hat{\mathbf{i}} + y_1\hat{\mathbf{j}} + z_1\hat{\mathbf{k}}$ and $\mathbf{u}_2 = x_2\hat{\mathbf{i}} + y_2\hat{\mathbf{j}} + z_2\hat{\mathbf{k}}$ be two vectors of $\mathbb{R}^3$. We call **vector product** (or **cross product**) of $\mathbf{u}_1$ and $\mathbf{u}_2$ the vector whose components in the orthonormal basis $\{\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\}$ are

$$\left(\begin{vmatrix} y_1 & y_2 \\ z_1 & z_2 \end{vmatrix}, -\begin{vmatrix} x_1 & x_2 \\ z_1 & z_2 \end{vmatrix}, \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix}\right)$$

We denote this vector product $\mathbf{u}_1 \times \mathbf{u}_2$. We also write

$$\mathbf{u}_1 \times \mathbf{u}_2 = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{vmatrix} = (y_1z_2 - y_2z_1)\hat{\mathbf{i}} + (z_1x_2 - z_2x_1)\hat{\mathbf{j}} + (x_1y_2 - x_2y_1)\hat{\mathbf{k}}$$

The vector product is an operation that takes as an input two vectors in $\mathbb{R}^3$ and outputs a vector in $\mathbb{R}^3$ (unlike $\mathbf{u}_1 \cdot \mathbf{u}_2$ which is a scalar!).

- The vector product $\mathbf{u}_1 \times \mathbf{u}_2$ is zero if and only if $\mathbf{u}_1$ and $\mathbf{u}_2$ are collinear.
- The vector product $\mathbf{u}_1 \times \mathbf{u}_2$ is orthogonal to both $\mathbf{u}_1$ and $\mathbf{u}_2$.
- When $\mathbf{u}_1$ and $\mathbf{u}_2$ are not collinear, the vector product is proportional to the unique vector orthogonal to $\mathbf{u}_1$ and $\mathbf{u}_2$.

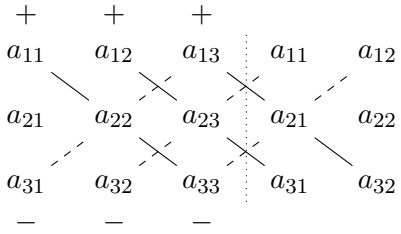
Remark. *We have seen earlier that the determinant of a 2×2 matrix was given by*

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

Computing the expression for the components of the vector product involves a formal 3-by3 determinant; we can make use of the **rule of Sarrus** *to compute determinants of 3×3 matrices. Let A be a 3 by 3 matrix given by*

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Then, the rule of Sarrus provides a visual way to compute the determinant of A, such as described below



This in turn gives the following expression:

$$\det A = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12}$$

Careful, this method of calculation is not valid for $n \geq 4$. It is always good to remember that the determinant is linear in the rows (or columns) of a matrix and that if the matrix has two equals rows (or columns) then its determinant is zero. Further, swapping two rows or columns of a determinant changes its sign!

It is quite important to note that $\hat{\mathbf{i}} \times \hat{\mathbf{j}} = \hat{\mathbf{k}}$, $\hat{\mathbf{j}} \times \hat{\mathbf{k}} = \hat{\mathbf{i}}$ and $\hat{\mathbf{k}} \times \hat{\mathbf{i}} = \hat{\mathbf{j}}$ while we easily realize that $\hat{\mathbf{j}} \times \hat{\mathbf{i}} = -\hat{\mathbf{k}}$, $\hat{\mathbf{i}} \times \hat{\mathbf{k}} = -\hat{\mathbf{j}}$ and $\hat{\mathbf{k}} \times \hat{\mathbf{j}} = -\hat{\mathbf{i}}$.

Example

> Find the vector product of $(1, 2, 3)$ and $(1, 0, 2)$.
>
> By definition, we have
>
> $$(1,2,3)\times(1,0,2) = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{vmatrix} = (2\times 2 - 0\times 3)\hat{\mathbf{i}} + (1\times 3 - 2\times 1)\hat{\mathbf{j}} + (1\times 0 - 1\times 2)\hat{\mathbf{k}} = 4\hat{\mathbf{i}} + \hat{\mathbf{j}} - 2\hat{\mathbf{k}}$$

Remark. *Unlike the dot product, the cross product is not a concept that can be generalized to n dimensions. Actually, it can be shown that the cross product only exists in dimensions 3 and 7. For a proof that the cross product can not exist in $\mathbb{R}^4$, see Part III - Problem Sheet 3 (Note that this is not examinable!)*

Proposition 6.1.15: Bilinear antisymmetric form

The vector product is an **antisymmetric bilinear form**, i.e. for all $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ vectors in $\mathbb{R}^3$ and real numbers λ and μ, we have

(a) $\mathbf{u} \times \mathbf{v} = -\mathbf{v} \times \mathbf{u}$ (antisymmetry)

(b) $(\lambda\mathbf{u} + \mu\mathbf{w}) \times \mathbf{v} = \lambda\mathbf{u} \times \mathbf{v} + \mu\mathbf{w} \times \mathbf{v}$ (linearity in the first argument)

(c) $\mathbf{u} \times (\lambda\mathbf{v} + \mu\mathbf{w}) = \lambda\mathbf{u} \times \mathbf{v} + \mu\mathbf{u} \times \mathbf{w}$ (linearity in the second argument)

Proof. This is a direct consequence of the bilinearity and antisymmetry of the 2×2 determinant. ■

Geometric interpretation of the vector product

Proposition 6.1.16: Properties of the vector product

Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in $\mathbb{R}^3$, we have

$$|\mathbf{u}\times\mathbf{v}|^2=|\mathbf{u}|^2|\mathbf{v}|^2-(\mathbf{u}\cdot\mathbf{v})^2$$

Proof. This proof is left as an exercise to the reader, see Part III - Problem Sheet 3. ■

Corollary 6.1.3

Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in $\mathbb{R}^3$, we have

$$|\mathbf{u}\times\mathbf{v}|=|\mathbf{u}||\mathbf{v}|\sin\theta$$

with $\theta\in[0,\pi]$ a measure of the smaller angle between vector $\mathbf{u}$ and $\mathbf{v}$.

Proof. We have shown that $\mathbf{u}\cdot\mathbf{v}=|\mathbf{u}||\mathbf{v}|\cos\theta$ so

$$|\mathbf{u}\times\mathbf{v}|^2=|\mathbf{u}|^2|\mathbf{v}|^2-(\mathbf{u}\cdot\mathbf{v})^2=|\mathbf{u}|^2|\mathbf{v}|^2(1-\cos^2\theta)=|\mathbf{u}|^2|\mathbf{v}|^2\sin^2\theta$$

The results follows as $0\leq\sin\theta$ for $\theta\in[0,\pi]$. ■

Corollary 6.1.4

Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in $\mathbb{R}^3$, then $|\mathbf{u}\times\mathbf{v}|$ equals the area of the parallelogram with vertices $\mathbf{0}$, $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{u}+\mathbf{v}$.

Corollary 6.1.5

Let $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$ be orthogonal unit vectors in $\mathbb{R}^3$, then if we denote $\hat{\mathbf{w}}=\hat{\mathbf{u}}\times\hat{\mathbf{v}}$, $\{\hat{\mathbf{u}},\hat{\mathbf{v}},\hat{\mathbf{w}}\}$ is an orthonormal basis of the space. Conversely, if $\{\hat{\mathbf{u}},\hat{\mathbf{v}},\hat{\mathbf{w}}\}$ is an orthonormal basis of the space, then we have that $\hat{\mathbf{w}}=\hat{\mathbf{u}}\times\hat{\mathbf{v}}$ or $\hat{\mathbf{w}}=-\hat{\mathbf{u}}\times\hat{\mathbf{v}}$

Orientation and Right-handedness

Let $\mathbf{u}_1=x_1\hat{\mathbf{i}}+y_1\hat{\mathbf{j}}+z_1\hat{\mathbf{k}}$ and $\mathbf{u}_2=x_2\hat{\mathbf{i}}+y_2\hat{\mathbf{j}}+z_2\hat{\mathbf{k}}$ be two vectors in $\mathbb{R}^3$. The vector product of these two vectors

$$\mathbf{w}=\mathbf{u}_1\times\mathbf{u}_2=\begin{vmatrix}\hat{\mathbf{i}}&\hat{\mathbf{j}}&\hat{\mathbf{k}}\\x_1&y_1&z_1\\x_2&y_2&z_2\end{vmatrix}=\hat{\mathbf{i}}(y_1z_2-y_2z_1)+\hat{\mathbf{j}}(z_1x_2-z_2x_1)+\hat{\mathbf{k}}(x_1y_2-x_2y_1)$$

Now, if we were to consider the orthonormal basis $\{\hat{\mathbf{i}},\hat{\mathbf{j}},-\hat{\mathbf{k}}\}$, the vectors are then written

$$\begin{cases}\mathbf{u}_1=x_1\hat{\mathbf{i}}+y_1\hat{\mathbf{j}}-z_1(-\hat{\mathbf{k}})\\\mathbf{u}_2=x_2\hat{\mathbf{i}}+y_2\hat{\mathbf{j}}-z_2(-\hat{\mathbf{k}})\end{cases}$$

In this orthonormal basis, the vector product has the following components

$$\mathbf{w}'=\mathbf{u}_1\times\mathbf{u}_2=\begin{vmatrix}\hat{\mathbf{i}}&\hat{\mathbf{j}}&-\hat{\mathbf{k}}\\x_1&y_1&-z_1\\x_2&y_2&-z_2\end{vmatrix}=\hat{\mathbf{i}}(-y_1z_2+y_2z_1)+\hat{\mathbf{j}}(-z_1x_2+z_2x_1)+(-\hat{\mathbf{k}})(x_1y_2-x_2y_1)$$

and is thus equal to $-\mathbf{w}$. The expression of the vector product in an orthonormal basis thus depends on the basis! This is quite unsatisfactory, how can we solve this? Everything we have seen about the vector product points to the fact that there are **two orientations** in the three dimensional space.

Definition 6.1.13: Right-handedness

Let $\{\hat{\mathbf{u}}, \hat{\mathbf{v}}, \hat{\mathbf{w}}\}$ an orthonormal basis. We say that the basis is **right-handed** if $\hat{\mathbf{w}} = \hat{\mathbf{u}} \times \hat{\mathbf{v}}$. Conversely, if $\hat{\mathbf{w}} = -\hat{\mathbf{u}} \times \hat{\mathbf{v}}$, we say that the basis if **left-handed**.

Remark. *The orientation of bases is determined using what is called the right-hand rule (which is where the name right-handed comes from).*

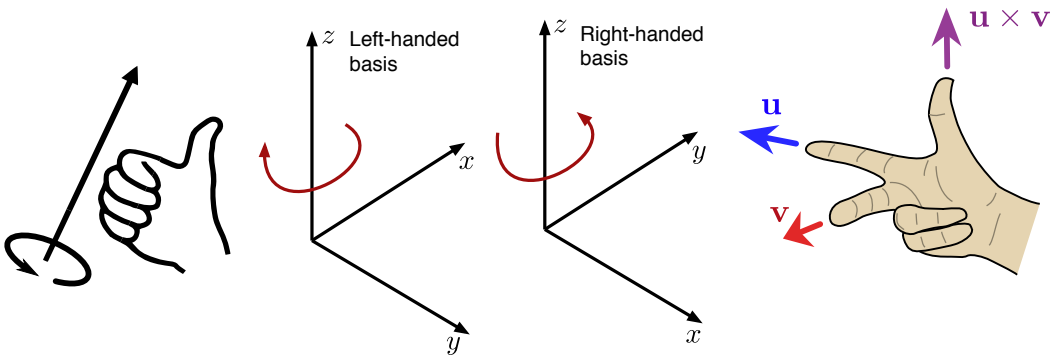


Figure 6.6 Right-hand rule (adapted from wikipedia).

The right-hand rule helps:

- *determine the orientation of a basis;*
- *determine the direction of the vector product of two vectors in a right-handed basis.*

Proposition 6.1.17

If $\{\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2, \hat{\mathbf{u}}_3\}$ is a right-handed orthonormal basis, then $\{\hat{\mathbf{u}}_2, \hat{\mathbf{u}}_3, \hat{\mathbf{u}}_1\}$ and $\{\hat{\mathbf{u}}_3, \hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2\}$ also are (order of vectors in the basis matters!). We have:

$$\hat{\mathbf{u}}_1 \times \hat{\mathbf{u}}_2 = \hat{\mathbf{u}}_3 \quad \text{and} \quad \hat{\mathbf{u}}_2 \times \hat{\mathbf{u}}_3 = \hat{\mathbf{u}}_1 \quad \text{and} \quad \hat{\mathbf{u}}_3 \times \hat{\mathbf{u}}_1 = \hat{\mathbf{u}}_2$$

Proof.

- It is obvious that they are orthonormal basis.
- Let us verify that $\{\hat{\mathbf{u}}_2, \hat{\mathbf{u}}_3, \hat{\mathbf{u}}_1\}$ is right-handed, i.e. that $\hat{\mathbf{u}}_2 \times \hat{\mathbf{u}}_3 = \hat{\mathbf{u}}_1$. We denote (x_i, y_i, z_i) the components of $\hat{\mathbf{u}}_i$ in the right-handed orthonormal basis $\{\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\}$. We have:

  $$x_3 = y_1 z_2 - y_2 z_1, \; y_3 = z_1 x_2 - z_2 x_1 \text{ and } z_3 = x_1 y_2 - x_2 y_1$$

  As $\hat{\mathbf{u}}_1 \neq \mathbf{0}$, we can assume that $z_1 \neq 0$. The third component of the vector $\hat{\mathbf{u}}_2 \times \hat{\mathbf{u}}_3$ is

  $$\begin{aligned} x_2(z_1 x_2 - z_2 x_1) - y_2(y_1 z_2 - y_2 z_1) &= z_1(x_2^2 + y_2^2) - z_2(x_1 x_2 + y_1 y_2) \\ &= z_1(1 - z_2^2) - z_2(x_1 x_2 + y_1 y_2) \text{ as} |\hat{\mathbf{u}}_2| = 1 \\ &= z_1 - z_2(x_1 x_2 + y_1 y_2 + z_1 z_2) \\ &= z_1 \text{ as } \hat{\mathbf{u}}_1 \cdot \hat{\mathbf{u}}_2 = 0 \end{aligned}$$

  As the two vectors $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2 \times \hat{\mathbf{u}}_3$ are collinear and z_1 is non zero, we deduce that they are equal.
- We can proceed similarly for the other two basis.

■

Proposition 6.1.18

Let $\{\hat{\mathbf{u}}, \hat{\mathbf{v}}, \hat{\mathbf{w}}\}$ be a right-handed orthonormal basis and

$$\mathbf{u}_1 = x_1\hat{\mathbf{u}} + y_1\hat{\mathbf{v}} + z_1\hat{\mathbf{w}} \quad \text{and} \quad \mathbf{u}_1 = x_2\hat{\mathbf{u}} + y_2\hat{\mathbf{v}} + z_2\hat{\mathbf{w}}$$

then we have

$$\mathbf{u}_1 \times \mathbf{u}_2 = \left(\begin{vmatrix} y_1 & y_2 \\ z_1 & z_2 \end{vmatrix}, - \begin{vmatrix} x_1 & x_2 \\ z_1 & z_2 \end{vmatrix}, \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} \right)$$

The expression of the vector product as a function of the vector components is thus the same in all right-handed orthonormal basis.

Operation with the vector product

There are important operations combining the dot product and the vector product which we introduce here.

Definition 6.1.14: Scalar Triple Product

Given three vectors $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{R}^3$ we define the **scalar triple product** as

$$[\mathbf{u}, \mathbf{v}, \mathbf{w}] = \mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})$$

If $\mathbf{u} = (u_1, u_2, u_3)$, $\mathbf{v} = (v_1, v_2, v_3)$ and $\mathbf{w} = (w_1, w_2, w_3)$ then

$$[\mathbf{u}, \mathbf{v}, \mathbf{w}] = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$$

Further, as swapping two rows of a determinant changes its sign but cyclic permutations do not, we have:

$$[\mathbf{u}, \mathbf{v}, \mathbf{w}] = [\mathbf{v}, \mathbf{w}, \mathbf{u}] = [\mathbf{w}, \mathbf{u}, \mathbf{v}] = -[\mathbf{u}, \mathbf{w}, \mathbf{v}] = -[\mathbf{w}, \mathbf{v}, \mathbf{u}] = -[\mathbf{v}, \mathbf{u}, \mathbf{w}]$$

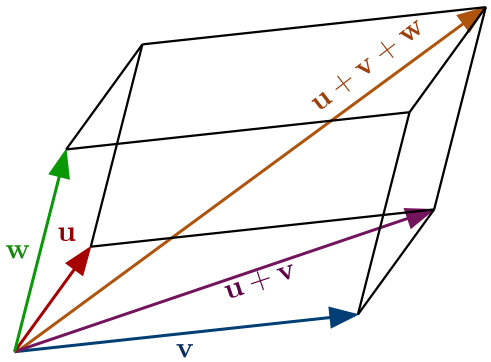


Figure 6.7 Parallelepiped

Example

Consider the parallelepiped (a generalization of the parallelogram in three dimensions) parametrized by $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$. The eight vertices of this solid can be parametrized as $\alpha\mathbf{u} + \beta\mathbf{v} + \gamma\mathbf{w}$ with $\alpha, \beta, \gamma \in \{0, 1\}$. Consider that $\mathbf{u}$ and $\mathbf{v}$ determine the base of the parallelepiped, then we know that this parallelogram has for area $|\mathbf{u} \times \mathbf{v}|$. If we denote θ the angle between $\mathbf{w}$ and a normal to the plane containing $\mathbf{u}$ and $\mathbf{v}$, then the volume of the parallelepiped is given by

$$\text{area of base} \times \text{height} = |\mathbf{u} \times \mathbf{v}||\mathbf{w}|\cos\theta = |(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w}| = |[\mathbf{u}, \mathbf{v}, \mathbf{w}]|$$

We can also form a vector triple product!

Proposition 6.1.19: Vector Triple Product

Given three vectors $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{R}^3$ we define the **vector triple product** as

$$\mathbf{u} \times (\mathbf{v} \times \mathbf{w})$$

For any three vectors $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{R}^3$, then

$$\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) = (\mathbf{u} \cdot \mathbf{w})\mathbf{v} - (\mathbf{u} \cdot \mathbf{v})\mathbf{w}$$

Proof. Both the LHS and RHS of the above expression are linear in $\mathbf{u}$, so it is sufficient to note that if we consider $\mathbf{v} = (v_1, v_2, v_3)$ and $\mathbf{w} = (w_1, w_2, w_3)$ that

$$(\hat{\mathbf{i}} \cdot \mathbf{w})\mathbf{v} - (\hat{\mathbf{i}} \cdot \mathbf{v})\mathbf{w} = w_1\mathbf{v} - v_1\mathbf{w} = (0, w_1v_2 - v_1w_2, w_1v_3 - v_1w_3) = \hat{\mathbf{i}} \times (\mathbf{v} \times \mathbf{w})$$

After two similar calculations for $\mathbf{u} = \hat{\mathbf{j}}$ and $\mathbf{u} = \hat{\mathbf{k}}$, the result follows by linearity. ■

6.1.6 Parametrization of lines in space

Definition 6.1.15: Parametric Form of a Line

We can easily extend the parametric form of a line seen in $\mathbb{R}^2$ to $\mathbb{R}^3$. We work in the orthonormal frame $\mathcal{F} = \{O, \hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\}$. Let $\mathcal{L}$ be the line going through a point $P(x_0, y_0, z_0)$ and oriented by non-zero vector $\mathbf{u} = \alpha\hat{\mathbf{i}} + \beta\hat{\mathbf{j}} + \gamma\hat{\mathbf{k}}$. This line $\mathcal{L}$ in $\mathbb{R}^3$ is the set of points $R(x, y, z)$ such that:

$$\begin{cases} x = x_0 + \lambda\alpha \\ y = y_0 + \lambda\beta \\ z = z_0 + \lambda\gamma \end{cases} \quad \text{with} \quad \lambda \in \mathbb{R}$$

We can write this parametrization vectorially and obtain

$$\mathbf{r}(\lambda) = \mathbf{p} + \lambda\mathbf{u}, \ \lambda \in \mathbb{R} \tag{6.21}$$

Example

Show that (x, y, z) lies on the line $\mathbf{r}(\lambda) = (1, 2, -3) + \lambda(2, 1, 4)$ if and only if $(x - 1)/2 = y - 2 = (z + 3)/4$.

As in the two dimensional case, we can compute the distance between a point M and the line $\mathcal{L}$ going through the point A and parallel to the vector $\mathbf{u}$. If we denote H the orthogonal projection of M on $\mathcal{L}$, we have

$$\mathbf{AM} \times \mathbf{u} = \mathbf{AH} \times \mathbf{u} + \mathbf{HM} \times \mathbf{u} = \mathbf{HM} \times \mathbf{u}$$

As $\mathbf{u}$ and $\mathbf{HM}$ are orthogonal by construction, we have

$$|\mathbf{HM} \times \mathbf{u}| = |\mathbf{HM}||\mathbf{u}| = d(M, \mathcal{L})|\mathbf{u}|$$

so that we can define the distance as:

$$\boxed{d(M, \mathcal{L}) = \frac{|\mathbf{AM} \times \mathbf{u}|}{|\mathbf{u}|}}$$

Proposition 6.1.20: Another vector equation for a line

Consider two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^3$ with $\mathbf{u} \cdot \mathbf{v} = 0$ and $\mathbf{u} \neq \mathbf{0}$. The vectors $\mathbf{r}$ in $\mathbb{R}^3$ which satisfies the following equation $\mathbf{r} \times \mathbf{u} = \mathbf{v}$ form the line parallel to $\mathbf{u}$ and passing through the point $(\mathbf{u} \times \mathbf{v})/|\mathbf{u}|^2$.

Proof. If $\mathbf{v} = 0$, then the equation $\mathbf{r} \times \mathbf{u} = \mathbf{0}$ is satisfied by scalar multiples of $\mathbf{u}$. Now, assume that $\mathbf{v} \neq \mathbf{0}$. As $\mathbf{u} \cdot \mathbf{v} = 0$ then we know that these vectors are linearly independent. Every vector in $\mathbb{R}^3$ can be written as a unique linear combination

$$\mathbf{r} = \lambda\mathbf{u} + \mu\mathbf{v} + \nu\mathbf{u} \times \mathbf{v}$$

with $(\lambda, \mu, \nu) \in \mathbb{R}^3$. Then, we know that $\mathbf{v} = \mathbf{r} \times \mathbf{u}$ if and only if

$$\begin{aligned}\mathbf{v} &= -\mathbf{u} \times (\lambda\mathbf{u} + \mu\mathbf{v} + \nu\mathbf{u} \times \mathbf{v}) \\ &= -\mu\mathbf{u} \times \mathbf{v} - \nu((\mathbf{u} \cdot \mathbf{v})\mathbf{u} - (\mathbf{u} \cdot \mathbf{u})\mathbf{v}) \\ &= -\mu\mathbf{u} \times \mathbf{v} + \nu|\mathbf{u}|^2\mathbf{v}\end{aligned}$$

Because the coordinates are unique, we can compare coefficient in this vectorial equation and we see that λ can take any value, $\mu = 0$ and $\nu = 1/|\mathbf{u}|^2$. Finally, $\mathbf{v} = \mathbf{r} \times \mathbf{u}$ if and only if

$$\mathbf{r} = \frac{\mathbf{u} \times \mathbf{v}}{|\mathbf{u}|^2} + \lambda\mathbf{u}$$

where λ is a real. ■

6.2 Space Curves and Kinematics

In the previous section, we have provided parametrizations of very simple geometric objects in space, mainly straight lines and planes. In this section, we will introduce the concept of vector functions (i.e. functions whose values are vectors) as such functions will allow us to describe more general curves and surfaces in space.

6.2.1 Vector Functions and Space Curves

Thus far, you have been used to dealing with real-valued functions, i.e. rules of the following kind:

$$\begin{aligned}f : \mathcal{D} &\to \mathcal{I} \\ x &\mapsto f(x)\end{aligned}$$

where $\mathcal{D}$ is the **domain of definition of the function** and $\mathcal{I}$ is the **image/range of the function**. For real-valued functions, we have that $\mathcal{D} \subset \mathbb{R}$ and $\mathcal{I} \subset \mathbb{R}$.

A vector function is a function whose domain is a subset of $\mathbb{R}$ and range is a set of vectors. Here, we are most interested in vector functions $\mathbf{r}$ whose values are three-dimensional vectors. If one denotes $\mathcal{D} \subset \mathbb{R}$ the domain of definition of the vector-valued function $\mathbf{r}$, then for every element $t \in \mathcal{D}$ there is a unique vector in $\mathbb{R}^3$ denoted by $\mathbf{r}(t)$. If $f(t)$, $g(t)$ and $h(t)$ are the components of the vector $\mathbf{r}(t)$, then f, g and h are real-valued functions called the **component functions** of $\mathbf{r}$ and we can write in Cartesian coordinates

$$\mathbf{r}(t) = (f(t), g(t), h(t)) = f(t)\hat{\mathbf{i}} + g(t)\hat{\mathbf{j}} + h(t)\hat{\mathbf{k}} \tag{6.22}$$

Remark. *One commonly uses the letter t to denote the independent variable as it oftentimes represents time in applications of vector functions.*

Definition 6.2.1: Limit of a vector function

The limit of a vector function $\mathbf{r}$ is defined by taking the limits of its component functions. If $\mathbf{r}(t) = (f(t), g(t), h(t))$, then

$$\lim_{t\to a} \mathbf{r}(t) = \Big(\lim_{t\to a} f(t), \lim_{t\to a} g(t), \lim_{t\to a} h(t)\Big) \tag{6.23}$$

provided the limits of the component functions exist.

Limits of vector functions obey the same rules as limits of real-valued functions.

Example

Consider the vector function defined by

$$\mathbf{r}(t) = (1 + t^3)\hat{\mathbf{i}} + te^{-t}\hat{\mathbf{j}} + \frac{\sin t}{t}\hat{\mathbf{k}}$$

The limit of $\mathbf{r}$ is the vector whose components are the limits of the component functions, i.e.

$$\lim_{t\to 0} \mathbf{r}(t) = \left[\lim_{t\to 0}(1 + t^3)\right]\hat{\mathbf{i}} + \left[\lim_{t\to 0} te^{-t}\right]\hat{\mathbf{j}} + \left[\lim_{t\to 0} \frac{\sin t}{t}\right]\hat{\mathbf{k}}$$

So we conclude that

$$\lim_{t\to 0} \mathbf{r}(t) = \hat{\mathbf{i}} + \hat{\mathbf{k}}$$

Proposition 6.2.1: Continuity

A vector function $\mathbf{r}$ is **continuous at a** if

$$\lim_{t\to a} \mathbf{r}(t) = \mathbf{r}(a)$$

i.e. that $\mathbf{r}$ is continuous at a if and only if its component functions f, g and h are continuous at a.

Suppose that f, g and h are continuous real-valued functions on an interval $\mathcal{I}$. Then the set $\mathcal{C}$ of all points (x, y, z) in space, where

$$x = f(t) \quad y = g(t) \quad z = h(t) \tag{6.24}$$

and t varies throughout the interval $\mathcal{I}$, is called a **space curve**. Equations (6.24) provide **parametric equations of** $\mathcal{C}$ and t is called the **parameter**. One can think of curve $\mathcal{C}$ as

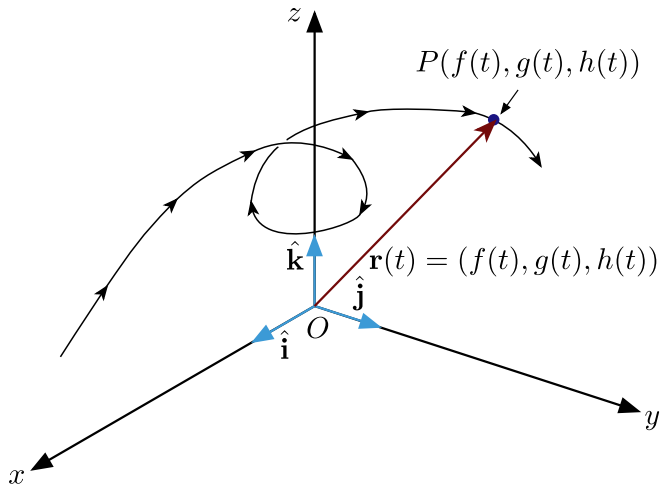


Figure 6.8 Example of space curve.

being traced out by a moving point or particle whose position at time t is $(f(t), g(t), h(t))$. The vector function $\mathbf{r}(t)$ is the position vector of the point $P(f(t), g(t), h(t))$ on $\mathcal{C}$ (see Figure 6.8).

Example

Sketch the curve defined by the following vector equation

$$\mathbf{r}(t) = \cos t\hat{\mathbf{i}} + \sin t\hat{\mathbf{j}} + t\hat{\mathbf{k}}$$

The parametric equations for this curve are given by

$$x = \cos t \quad y = \sin t \quad z = t$$

First, we know that $\cos^2 t + \sin^2 t = 1$ so this curve must lie on the circular cylinder $x^2 + y^2 = 1$. Since $z = t$, we can say that the curve spirals upward around the cylinder in a counterclockwise manner as t increases. This leads to the following **helix** (see Figure 6.9).

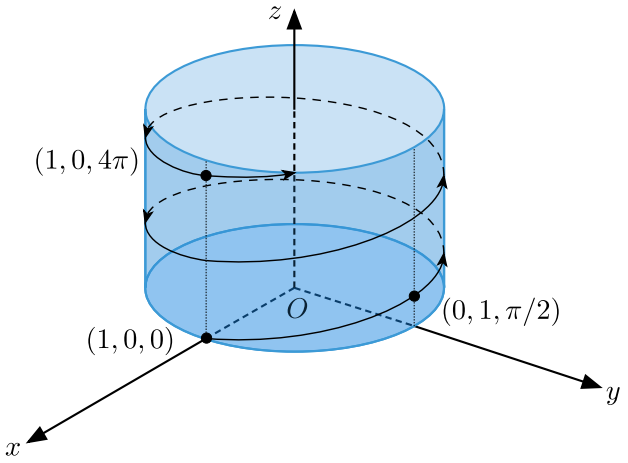


Figure 6.9 Helix

Space curves can be very beautiful and intricate geometric objects; two examples

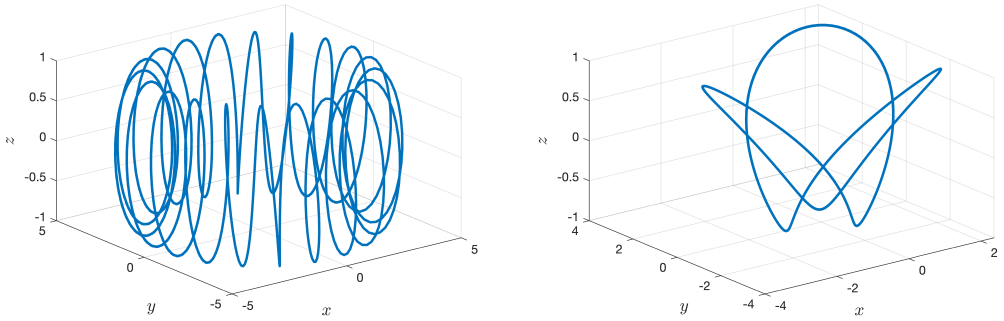


Figure 6.10 *(Left)* A toroidal spiral with parametric equations $x = (4 + \sin 20t)\cos t$, $y = (4 + \sin 20t)\sin t$, $z = \cos 20t$; *(Right)* A trefoil knot with parametric equations $x = (2 + \sin 1.5t)\cos t$, $y = (2 + \sin 1.5t)\sin t$, $z = \sin 1.5t$.

can be found on Figure 6.10. In the next section, we will use vector functions to describe the motion of objects in space. But first, we need to develop some ideas related to the calculus of vector functions.

Derivatives of vector functions

We define the derivative $\mathbf{r}'$ of a vector function $\mathbf{r}$ in much the same way as for real-valued functions.

$$\boxed{\mathbf{r}'(t) = \frac{d\mathbf{r}}{dt} = \lim_{h\to 0} \frac{\mathbf{r}(t+h) - \mathbf{r}(t)}{h}} \tag{6.25}$$

if this limit exists.

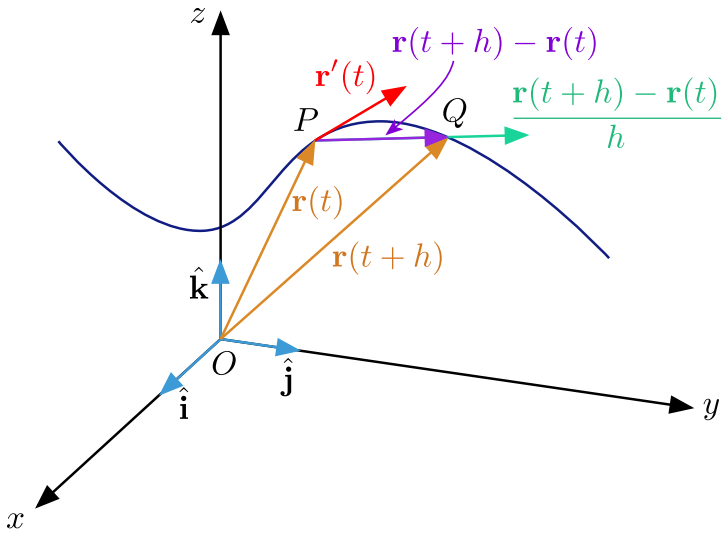


Figure 6.11 Derivatives of vector functions: tangent and secant vectors

As seen on Figure 6.11, we consider two points P and Q with respective position vectors $\mathbf{r}(t)$ and $\mathbf{r}(t+h)$. The vector $\mathbf{r}(t+h) - \mathbf{r}(t)$ is the vector going from P to Q and is called the **secant vector**. As $h \to 0$, if the limit exists, the vector $(\mathbf{r}(t+h) - \mathbf{r}(t))/h$ approaches a vector tangent to the curve. The vector $\mathbf{r}'(t)$ is thus called the **tangent vector** to the curve defined by $\mathbf{r}$ at point P (provided that $\mathbf{r}'(t)$ exists and is not zero). The **tangent line** to the curve $\mathcal{C}$ at P is thus defined as the line parallel to vector $\mathbf{r}'(t)$ and

passing through P. One often denotes the unit tangent vector as:

$$\mathbf{T}(t) = \frac{\mathbf{r}'(t)}{|\mathbf{r}'(t)|} \tag{6.26}$$

Proposition 6.2.2

If $\mathbf{r}(t) = f(t)\hat{\mathbf{i}} + g(t)\hat{\mathbf{j}} + h(t)\hat{\mathbf{k}}$ with f, g and h differentiable functions then

$$\mathbf{r}'(t) = f'(t)\hat{\mathbf{i}} + g'(t)\hat{\mathbf{j}} + h'(t)\hat{\mathbf{k}} \tag{6.27}$$

Example

We consider the following vector equation $\mathbf{r}(t) = (1+t^3)\hat{\mathbf{i}}+te^{-t}\hat{\mathbf{j}}+\sin 2t\hat{\mathbf{k}}$. The derivative of this vector function is given by

$$\mathbf{r}'(t) = 3t^2\hat{\mathbf{i}} + (1-t)e^{-t}\hat{\mathbf{j}} + 2\cos 2t\hat{\mathbf{k}} \tag{6.28}$$

From this, we can easily obtain for instance the unit tangent vector at the point where $t = 0$:

$$\mathbf{T}(0) = \frac{\mathbf{r}'(0)}{|\mathbf{r}'(0)|} = \frac{\hat{\mathbf{j}} + 2\hat{\mathbf{k}}}{\sqrt{1+4}} = \frac{1}{\sqrt{5}}\hat{\mathbf{j}} + \frac{2}{\sqrt{5}}\hat{\mathbf{k}} \tag{6.29}$$

Vector functions admit differentiation rules which are very similar to the differentiation rules for real-valued functions.

Theorem 6.2.1: Differentation rules

Let $\mathbf{u}$ and $\mathbf{v}$ be differentiable vector functions, c a scalar and f a real-valued function. Then we have

1 — $\frac{d}{dt}[\mathbf{u}(t) + \mathbf{v}(t)] = \mathbf{u}'(t) + \mathbf{v}'(t)$

2 — $\frac{d}{dt}[c\mathbf{u}(t)] = c\mathbf{u}'(t)$

3 — $\frac{d}{dt}[f(t)\mathbf{u}(t)] = f'(t)\mathbf{u}(t) + f(t)\mathbf{u}'(t)$

4 — $\frac{d}{dt}[\mathbf{u}(t) \cdot \mathbf{v}(t)] = \mathbf{u}'(t) \cdot \mathbf{v}(t) + \mathbf{u}(t) \cdot \mathbf{v}'(t)$

5 — $\frac{d}{dt}[\mathbf{u}(t) \times \mathbf{v}(t)] = \mathbf{u}'(t) \times \mathbf{v}(t) + \mathbf{u}(t) \times \mathbf{v}'(t)$

6 — $\frac{d}{dt}[\mathbf{u}(f(t))] = f'(t)\mathbf{u}'(f(t))$

Example

Show that if $|\mathbf{r}(t)| = c$ (a real constant), then $\mathbf{r}'(t)$ is orthogonal to $\mathbf{r}(t)$ for all t. Interpret geometrically.

To show this, we make use of the following relation

$$\mathbf{r}(t) \cdot \mathbf{r}(t) = |\mathbf{r}(t)|^2 = c^2$$

From the previous theorem, we can write that

$$0 = \frac{d}{dt}[\mathbf{r}(t) \cdot \mathbf{r}(t)] = \mathbf{r}'(t) \cdot \mathbf{r}(t) + \mathbf{r}(t) \cdot \mathbf{r}'(t) = 2\mathbf{r}'(t) \cdot \mathbf{r}(t)$$

which leads to $\mathbf{r}'(t) \cdot \mathbf{r}(t) = 0, \forall t$.

Geometrically, this result shows that if a curve lies on a sphere with center the origin, then the tangent vector $\mathbf{r}'(t)$ is always orthogonal to the position vector $\mathbf{r}(t)$.

Integration of vector functions

After defining differentation rules, we can define the **definite integral** of a vector function $\mathbf{r}(t)$. We express the integral of $\mathbf{r}$ in terms of the integrals of its real-valued component functions as follows

$$\boxed{\int_a^b \mathbf{r}(t)dt = \left(\int_a^b f(t)dt\right)\hat{\mathbf{i}} + \left(\int_a^b g(t)dt\right)\hat{\mathbf{j}} + \left(\int_a^b h(t)dt\right)\hat{\mathbf{k}}} \tag{6.30}$$

We evaluate an integral of the vector function by integrating each component function. We can thus extend the Fundamental Theorem of Calculus to continuous vector functions and write:

$$\int_a^b \mathbf{r}(t)dt = \mathbf{R}(t)|_a^b = \mathbf{R}(b) - \mathbf{R}(a) \tag{6.31}$$

where $\mathbf{R}$ is an antiderivative of $\mathbf{r}$, i.e. $\mathbf{R}'(t) = \mathbf{r}(t)$. As for real-valued functions, we use the notation $\int \mathbf{r}(t)dt$ for the indefinite integrals (or antiderivatives).

Example

If $\mathbf{r}(t) = 2\cos t\hat{\mathbf{i}} + \sin t\hat{\mathbf{j}} + 2t\hat{\mathbf{k}}$, then the indefinite integral of $\mathbf{r}$ is given by

$$\int \mathbf{r}(t)dt = \left(\int 2\cos t dt\right)\hat{\mathbf{i}} + \left(\int \sin t dt\right)\hat{\mathbf{j}} + \left(\int 2t dt\right)\hat{\mathbf{k}} \tag{6.32}$$

$$= 2\sin t\hat{\mathbf{i}} - \cos t\hat{\mathbf{j}} + t^2\hat{\mathbf{k}} + \mathbf{c} \tag{6.33}$$

where $\mathbf{c}$ is a vector constant of integration. A definite integral is, for instance, given by

$$\int_0^{\pi/2} \mathbf{r}(t)dt = \left[2\sin t\hat{\mathbf{i}} - \cos t\hat{\mathbf{j}} + t^2\hat{\mathbf{k}}\right]_0^{\pi/2} = 2\hat{\mathbf{i}} + \hat{\mathbf{j}} + \frac{\pi^2}{4}\hat{\mathbf{k}} \tag{6.34}$$

Arc length and curvature

The integral of vector functions allows us to introduce another important concept: the **arc length** or **length of a space curve**. Let a curve $\mathcal{C}$ be parametrized by the vector function $\mathbf{r}(t) = (f(t), g(t), h(t))$ or equivalently the parametric equations $x = f(t)$, $y = g(t)$ and $z = h(t)$. We assume that f', g' and h' are continuous. If the curve is traversed only once as t increases from a to b ($a, b \in \mathbb{R}$), then its length is defined as

$$L = \int_a^b |\mathbf{r}'(t)|dt \tag{6.35}$$

which reads

$$L = \int_a^b \sqrt{[f'(t)]^2 + [g'(t)]^2 + [h'(t)]^2}dt \tag{6.36}$$

$$= \int_a^b \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2}dt \tag{6.37}$$

Example

We previously defined the circular helix as the curve associated to the following vector function $\mathbf{r}(t) = \cos t\hat{\mathbf{i}} + \sin t\hat{\mathbf{j}} + t\hat{\mathbf{k}}$. We can compute the length of the arc from the point $(1, 0, 0)$ to the point $(1, 0, 2\pi)$. First, we compute the derivative of the vector function to obtain

$$\mathbf{r}'(t) = -\sin t\hat{\mathbf{i}} + \cos t\hat{\mathbf{j}} + \hat{\mathbf{k}}$$

leading to

$$|\mathbf{r}'(t)| = \sqrt{(-\sin t)^2 + \cos^2 t + 1} = \sqrt{2}$$

The arc from point $(1, 0, 0)$ to point $(1, 0, 2\pi)$ is described by the parameter interval $0 \leq t \leq 2\pi$; so we have

$$L = \int_0^{2\pi} |\mathbf{r}'(t)|dt = \int_0^{2\pi} \sqrt{2}dt = 2\sqrt{2}\pi$$

This also allows us to define the **arc length function** s by

$$s(t) = \int_a^t |\mathbf{r}'(u)|du = \int_a^t \sqrt{\left(\frac{dx}{du}\right)^2 + \left(\frac{dy}{du}\right)^2 + \left(\frac{dz}{du}\right)^2}du \tag{6.38}$$

where $s(t)$ is the length of the part of curve $\mathcal{C}$ between $\mathbf{r}(a)$ and $\mathbf{r}(t)$. The Fundamental Theorem of Calculus gives us that

$$\frac{ds}{dt} = |\mathbf{r}'(t)| \tag{6.39}$$

Remark. *It is often useful to parametrize a curve with respect to the arc length s rather than the independent parameter t; this leads to the development of what are called* **intrinsic coordinates**.

Definition 6.2.2

Let $\mathbf{r}(t)$ be a parametrization for curve $\mathcal{C}$ on an interval $\mathcal{I}$. If $\mathbf{r}'$ is continuous and $\mathbf{r}'(t) \neq 0$, then the parametrization $\mathbf{r}(t)$ is **smooth** and in that case, curve $\mathcal{C}$ is also called smooth. A smooth curve has no sharp corners or cusps.

Remember that if a smooth curve is defined by the vector function $\mathbf{r}(t)$, the unit tangent vector $\mathbf{T}(t)$ is given by

$$\mathbf{T}(t) = \frac{\mathbf{r}'(t)}{|\mathbf{r}'(t)|} \tag{6.40}$$

It indicates the local direction of the curve. As one moves along the curve, $\mathbf{T}(t)$ will change direction as the curve bends: sharper bends of the curve will lead to faster changes of the direction of the unit tangent vector. The **curvature** at a given point is a measure of how quickly the curve changes direction, we define it as the amplitude of the rate of change of the unit tangent vector with respect to arc length:

$$\boxed{\kappa = \left|\frac{d\mathbf{T}}{ds}\right|} \tag{6.41}$$

Using the Chain Rule, we can express the curvature as a function of t instead of the arc length. Indeed, we have

$$\frac{d\mathbf{T}}{dt} = \frac{d\mathbf{T}}{ds}\frac{ds}{dt} \tag{6.42}$$

which leads directly to

$$\boxed{\kappa(t) = \frac{|\mathbf{T}'(t)|}{|\mathbf{r}'(t)|}} \tag{6.43}$$

Theorem 6.2.2

The curvature of the curve defined by the vector function $\mathbf{r}$ is

$$\kappa(t) = \frac{|\mathbf{r}'(t) \times \mathbf{r}''(t)|}{|\mathbf{r}'(t)|^3} \tag{6.44}$$

Proof. You should attempt this proof, solutions will be given in Part III - Problem Sheet 3. ■

Example

Consider the twisted cubic defined by $\mathbf{r}(t) = t\hat{\mathbf{i}} + t^2\hat{\mathbf{j}} + t^3\hat{\mathbf{k}}$. The curvature at any point can be obtained by first calculating the derivatives of the vector function $\mathbf{r}(t)$

$$\mathbf{r}'(t) = \hat{\mathbf{i}} + 2t\hat{\mathbf{j}} + 3t^2\hat{\mathbf{k}} \quad \text{and} \quad \mathbf{r}''(t) = 2\hat{\mathbf{j}} + 6t\hat{\mathbf{k}} \tag{6.45}$$

and

$$\mathbf{r}'(t) \times \mathbf{r}''(t) = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ 1 & 2t & 3t^2 \\ 0 & 2 & 6t \end{vmatrix} = 6t^2\hat{\mathbf{i}} - 6t\hat{\mathbf{j}} + 2\hat{\mathbf{k}} \tag{6.46}$$

So we can conclude that

$$|\mathbf{r}'(t)| = \sqrt{1 + 4t^2 + 9t^4} \quad \text{and} \quad |\mathbf{r}'(t) \times \mathbf{r}''(t)| = 2\sqrt{9t^4 + 9t^2 + 1} \tag{6.47}$$

and finally

$$\kappa(t) = \frac{2\sqrt{9t^4 + 9t^2 + 1}}{\sqrt{1 + 4t^2 + 9t^4}} \tag{6.48}$$

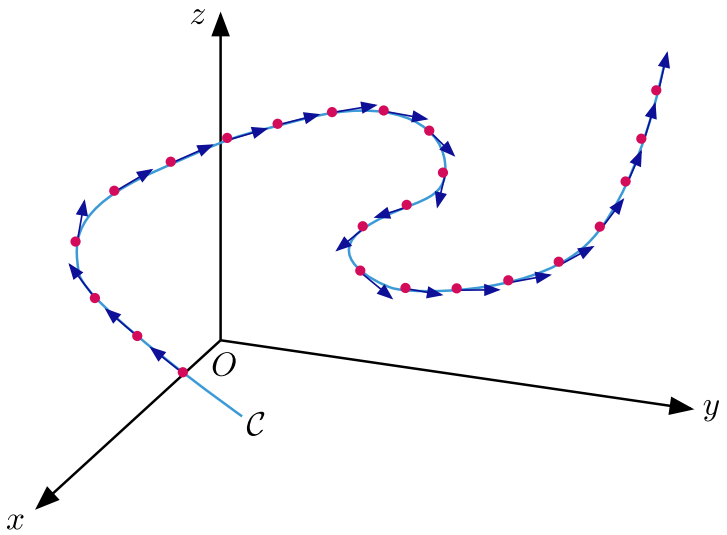


Figure 6.12 Curvature: unit tangent vectors at evenly spaced point on curve $\mathcal{C}$.

In particular, the curvature at the origin $(0,0,0)$ where $t=0$ is $\kappa(0)=2$.

Normal and binormal vectors

At any given point on a smooth curve $\mathbf{r}(t)$, there are many vectors which are orthogonal to the unit tangent vector $\mathbf{T}(t)$. As $|\mathbf{T}(t)|=1$, we know that $\mathbf{T}(t)\cdot\mathbf{T}'(t)=0$ and so that $\mathbf{T}'(t)$ is orthogonal to the tangent vector. If $\mathbf{r}$ is smooth, we then define the (principal) **unit normal vector** $\mathbf{N}(t)$ as

$$\mathbf{N}(t)=\frac{\mathbf{T}'(t)}{|\mathbf{T}'(t)|} \tag{6.49}$$

The **binormal vector** is defined by $\mathbf{B}(t)=\mathbf{T}(t)\times\mathbf{N}(t)$ such that $\{\mathbf{T}(t),\mathbf{N}(t),\mathbf{B}(t)\}$ forms a right-handed orthonormal basis.

Example

Let us consider once again the circular helix defined by $\mathbf{r}(t)=\cos t\hat{\mathbf{i}}+\sin t\hat{\mathbf{j}}+t\hat{\mathbf{k}}$.

First, we compute $\mathbf{r}'(t)=-\sin t\hat{\mathbf{i}}+\cos t\hat{\mathbf{j}}+\hat{\mathbf{k}}$ and get that $|\mathbf{r}'(t)|=\sqrt{2}$. So we obtain the following tangent vector

$$\mathbf{T}(t)=\frac{\mathbf{r}'(t)}{|\mathbf{r}'(t)|}=\frac{1}{\sqrt{2}}(-\sin t\hat{\mathbf{i}}+\cos t\hat{\mathbf{j}}+\hat{\mathbf{k}}) \tag{6.50}$$

which we can differentiate to obtain $\mathbf{T}'(t)=(-\cos t\hat{\mathbf{i}}-\sin t\hat{\mathbf{j}})/\sqrt{2}$ and $|\mathbf{T}'(t)|=1/\sqrt{2}$. Thus, we obtain the following normal vector:

$$\mathbf{N}(t)=\frac{\mathbf{T}'(t)}{|\mathbf{T}'(t)|}=-\cos t\hat{\mathbf{i}}-\sin t\hat{\mathbf{j}} \tag{6.51}$$

Therefore, the normal vector to an helix at any point on the helix points toward the z-axis. The binormal vector is then obtained as follows:

$$\mathbf{B}(t)=\mathbf{T}(t)\times\mathbf{N}(t)=\frac{1}{\sqrt{2}}\begin{vmatrix}\hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}}\\ -\sin t & \cos t & 1\\ -\cos t & -\sin t & 0\end{vmatrix}=\frac{1}{\sqrt{2}}(\sin t\hat{\mathbf{i}}-\cos t\hat{\mathbf{j}}+\hat{\mathbf{k}}) \tag{6.52}$$

Remark. *The plane containing the normal vector $\mathbf{N}$ and binormal vector $\mathbf{B}$ at a point P on a curve $\mathcal{C}$ is called the* **normal plane** *of $\mathcal{C}$ at P. The plane determined by the vectors $\mathbf{T}$ and $\mathbf{N}$ is called the* **osculating plane** *of $\mathcal{C}$ at point P.*

6.2.2 Kinematics in Cartesian coordinates

The concepts we have just developed can be used to describe the motion of objects in space. For a point particle, there are three key kinematic quantities:

- Position: $\mathbf{r}(t)$
- Velocity: $\mathbf{v}(t)$
- Acceleration: $\mathbf{a}(t)$

In general, we will consider $\mathbf{r}$, $\mathbf{v}$, $\mathbf{a}\in\mathbb{R}^3$. These are thus vector functions as described above; the use of variable t takes all its meaning here as it actually represents time. To describe the motion of point particles in space, we can use different coordinate systems to express these quantities. Here, we will start with kinematics in Cartesian coordinates.

Let us write the position of the particle at time t as the following vector function

$$\mathbf{r}(t)=x(t)\hat{\mathbf{i}}+y(t)\hat{\mathbf{j}}+z(t)\hat{\mathbf{k}} \tag{6.53}$$

The magnitude of $r = |\mathbf{r}(t)| = \sqrt{x^2 + y^2 + z^2}$ measures the distance from the origin.

Remark. *The time derivative of $x(t)$ can be denoted dx/dt, $x'(t)$ or $\dot{x}(t)$. Here, we will use the latter notation in the specific case of kinematic quantities.*

As the basis vectors $\{\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\}$ are constant in time, the derivative of the position vector in Cartesian coordinates is given by

$$\mathbf{v}(t) = \frac{d\mathbf{r}}{dt} = \frac{d}{dt}\left(x(t)\hat{\mathbf{i}} + y(t)\hat{\mathbf{j}} + z(t)\hat{\mathbf{k}}\right) = \frac{dx}{dt}\hat{\mathbf{i}} + \frac{dy}{dt}\hat{\mathbf{j}} + \frac{dz}{dt}\hat{\mathbf{k}} \tag{6.54}$$

Using the dot notation, we finally obtain

$$\boxed{\mathbf{v}(t) = \dot{x}\hat{\mathbf{i}} + \dot{y}\hat{\mathbf{j}} + \dot{z}\hat{\mathbf{k}} = v_x\hat{\mathbf{i}} + v_y\hat{\mathbf{j}} + v_z\hat{\mathbf{k}}} \tag{6.55}$$

This defines the instantaneous velocity of the point particle. The magnitude of the velocity vector is given by

$$v(t) = |\mathbf{v}(t)| = \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} \tag{6.56}$$

this quantity is called the **speed** of the particle. Thus, the direction of motion is defined as

$$\hat{\mathbf{v}} = \mathbf{v}/v, \quad \text{with } |\hat{\mathbf{v}}| = 1 \tag{6.57}$$

it is the **unit tangent vector to the trajectory of the particle**. Taking a second time derivative, we define what we call the **acceleration of the particle**

$$\mathbf{a}(t) = \frac{d\mathbf{v}}{dt} = \frac{d^2\mathbf{r}}{dt^2} = \ddot{x}\hat{\mathbf{i}} + \ddot{y}\hat{\mathbf{j}} + \ddot{z}\hat{\mathbf{k}} \tag{6.58}$$

where $\mathbf{a}(t)$ tells us how the velocity changes with time t. As we can write the velocity $\mathbf{v}(t) = v(t)\hat{\mathbf{v}}(t)$, then we obtain

$$\mathbf{a}(t) = \frac{d}{dt}(v\hat{\mathbf{v}}) = \underbrace{\frac{dv}{dt}\hat{\mathbf{v}}}_{\text{change in speed}} + \underbrace{v\frac{d\hat{\mathbf{v}}}{dt}}_{\text{change in direction}} \tag{6.59}$$

Starting from the acceleration $\mathbf{a}(t)$, we can integrate in time to obtain the velocity and a second time to obtain the position.

$$\mathbf{v}(t) = \mathbf{v}(t_0) + \int_{t_0}^{t} \mathbf{a}(t')dt' \tag{6.60}$$

$$\mathbf{r}(t) = \mathbf{r}(t_0) + \int_{t_0}^{t} \mathbf{v}(t')dt' \tag{6.61}$$

Thus, to define velocity and position uniquely, we need to be provided with some initial conditions $\mathbf{r}(t_0)$ and $\mathbf{v}(t_0)$.

Example

> As a first example, we consider the motion of a projectile in the xy-plane (as depicted on Figure 6.13). We consider that the projectile is launched from the origin O with initial velocity $\mathbf{v}_0$ at angle α with the x-axis. Newton's second law provides a link between the acceleration of a object and the forces it is subjected to. In the case of our projectile (in the absence of air resistance), we obtain that $\mathbf{a}(t) = -g\hat{\mathbf{j}}$, with $g = 9.81\text{m.s}^{-2}$ the acceleration due to gravity close to the surface of the earth. At $t = 0$, we release the particle. We know that our initial conditions are given by
>
> $$\mathbf{r}(0) = \mathbf{0} \quad \text{and} \quad \mathbf{v}(0) = \mathbf{v}_0 = v_0\cos\alpha\hat{\mathbf{i}} + v_0\sin\alpha\hat{\mathbf{j}} \tag{6.62}$$

First, we integrate the acceleration to obtain the instantaneous velocity

$$\mathbf{v}(t) = \mathbf{v}(0) + \int_0^t -g\hat{\mathbf{j}}dt' = v_0\cos\alpha\hat{\mathbf{i}} + (v_0\sin\alpha - gt)\hat{\mathbf{j}} \tag{6.63}$$

We can now integrate the velocity to obtain the trajectory

$$\mathbf{r}(t) = \mathbf{r}(0) + \int_0^t \mathbf{v}(t')dt' = v_0 t\cos\alpha\hat{\mathbf{i}} + (v_0 t\sin\alpha - \frac{1}{2}gt^2)\hat{\mathbf{j}} \tag{6.64}$$

This allows us for instance to find the launch angle which maximizes the range R of the projectile (see Figure 6.13). We can first find the time at which the object hits the ground, t_H; we know that we have

$$y(t_H) = \mathbf{r}(t_H)\cdot\hat{\mathbf{j}} = v_0 t_H\sin\alpha - \frac{1}{2}gt_H^2 = 0 \Longrightarrow t_H = 0 \text{ or } t_H = \frac{2v_0\sin\alpha}{g} \tag{6.65}$$

Thus, the range of the projectile is given by

$$R = x(t_H) = v_0\cos\alpha\left[\frac{2v_0\sin\alpha}{g}\right] = \frac{v_0^2}{g}\sin 2\alpha \tag{6.66}$$

We can then easily see that in the interval $0 \leq \alpha \leq \pi/2$, the range is maximal for $\alpha = \pi/4$.

Example

As a second example, we consider a particle whose position in time is described by the following vector function

$$\mathbf{r}(t) = R\sin(\Omega t)\hat{\mathbf{i}} + R\cos(\Omega t)\hat{\mathbf{j}} \tag{6.67}$$

The distance from the origin is given by

$$r = |\mathbf{r}(t)| = \sqrt{R^2\sin^2(\Omega t) + R^2\cos^2(\Omega t)} = R \tag{6.68}$$

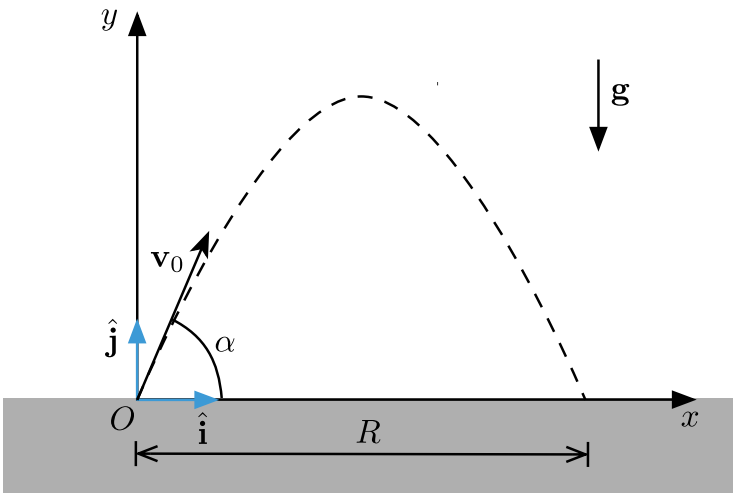


Figure 6.13 Projectile motion

which is constant. Thus, the particle has a circular motion; the path is a circle centered at the origin of radius R. We now determine the velocity of the particle by differentiating $\mathbf{r}(t)$ and obtain

$$\mathbf{v}(t) = R\Omega\cos(\Omega t)\hat{\mathbf{i}} - R\Omega\sin(\Omega t)\hat{\mathbf{j}} \tag{6.69}$$

It is easy to see that the speed is a constant $v = |\mathbf{v}(t)| = R\Omega$. The next question would be to determine in which direction the particle is moving; the direction of motion is given by

$$\hat{\mathbf{v}}(t) = \frac{\mathbf{v}(t)}{v} = \cos(\Omega t)\hat{\mathbf{i}} - \sin(\Omega t)\hat{\mathbf{j}} \tag{6.70}$$

where in particular, $\hat{\mathbf{v}}(t=0) = \hat{\mathbf{i}}$, so the particle moves in a clockwise direction and Ω is then called the angular speed. Finally, we can obtain the acceleration by differentiating the velocity and obtain

$$\mathbf{a}(t) = -R\Omega^2\sin(\Omega t)\hat{\mathbf{i}} - R\Omega^2\cos(\Omega t)\hat{\mathbf{j}} = -\Omega^2\mathbf{r}(t) \tag{6.71}$$

showing that the acceleration is radial and pointing towards the center of the circular trajectory.

6.2.3 Kinematics in polar coordinates

What made things particularly easy in Cartesian coordinates is the fact that the basis vectors $\{\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\}$ are constant in time. Nonetheless, we have seen that depending on the symmetries of your problem, it may be advantageous to work in other coordinates systems. Here, we do not aim at being exhaustive but simply give the example of motion in the plane described in polar coordinates.

Consider a point M with Cartesian coordinates (x, y). If (r, θ) is a system of polar coordinates of the point M, recall that

$$x = r\cos\theta \quad \text{and} \quad y = r\sin\theta \tag{6.72}$$

conversely,

$$r = \sqrt{x^2 + y^2} \quad \text{and} \quad \theta = \arctan(y/x) \tag{6.73}$$

and the unit vectors defining the polar system of coordinates are defined as follows

$$\hat{\mathbf{r}} = \cos\theta\hat{\mathbf{i}} + \sin\theta\hat{\mathbf{j}} \quad \text{and} \quad \hat{\boldsymbol{\theta}} = -\sin\theta\hat{\mathbf{i}} + \cos\theta\hat{\mathbf{j}} \tag{6.74}$$

where $\{\hat{\mathbf{i}}, \hat{\mathbf{j}}\}$ is the canonical basis of $\mathbb{R}^2$; i.e. that $\hat{\mathbf{r}}$ is the unit vector in the direction of increasing r along lines of constant θ and $\hat{\boldsymbol{\theta}}$ points in the direction of increasing θ tangent to the curves of constant r. These unit vectors form an orthonormal basis of the plane but are dependent on θ and so depend on the position of point M. As the point moves in space, these unit vectors will change; thus, this basis is not constant in time.

The kinematic quantities in polar coordinates can be defined as follows:

- **Position:** the vector position in Cartesian coordinates was given by

$$\mathbf{r}(t) = x(t)\hat{\mathbf{i}} + y(t)\hat{\mathbf{j}} \tag{6.75}$$

  using the relations above, we can write that in polar coordinates

$$\boxed{\mathbf{r}(t) = r(t)\hat{\mathbf{r}}(t)} \tag{6.76}$$

  (Note the explicit time dependence of the unit vector $\hat{\mathbf{r}}(t)$.)

- **Velocity:** one can differentiate the position vector to obtain the velocity

$$\mathbf{v}(t) = \frac{d\mathbf{r}}{dt} \tag{6.77}$$
$$= \dot{r}\hat{\mathbf{r}} + r\frac{d\hat{\mathbf{r}}}{dt} \tag{6.78}$$

where we can write that

$$\frac{d\hat{\mathbf{r}}}{dt} = \frac{d}{dt}\left[\cos\theta\hat{\mathbf{i}} + \sin\theta\hat{\mathbf{j}}\right] \tag{6.79}$$
$$= -\dot{\theta}\sin\theta\hat{\mathbf{i}} + \dot{\theta}\cos\theta\hat{\mathbf{j}} \tag{6.80}$$
$$= \dot{\theta}[-\sin\theta\hat{\mathbf{i}} + \cos\theta\hat{\mathbf{j}}] \tag{6.81}$$
$$= \dot{\theta}\hat{\boldsymbol{\theta}} \tag{6.82}$$

We conclude that

$$\boxed{\mathbf{v} = \dot{r}\ \hat{\mathbf{r}} + r\dot{\theta}\ \hat{\boldsymbol{\theta}}} \tag{6.83}$$

- **Acceleration:** to obtain the acceleration of the particle, we differentiate the velocity

$$\mathbf{a} = \frac{d\mathbf{v}}{dt} \tag{6.84}$$
$$= \frac{d}{dt}\left[\dot{r}\ \hat{\mathbf{r}} + r\dot{\theta}\ \hat{\boldsymbol{\theta}}\right] \tag{6.85}$$
$$= \ddot{r}\ \hat{\mathbf{r}} + \dot{r}\frac{d\hat{\mathbf{r}}}{dt} + \dot{r}\dot{\theta}\ \hat{\boldsymbol{\theta}} + r\ddot{\theta}\ \hat{\boldsymbol{\theta}} + r\dot{\theta}\frac{d\hat{\boldsymbol{\theta}}}{dt} \tag{6.86}$$

knowing that

$$\frac{d\hat{\boldsymbol{\theta}}}{dt} = \frac{d}{dt}\left[-\sin\theta\hat{\mathbf{i}} + \cos\theta\hat{\mathbf{j}}\right] \tag{6.87}$$
$$= -\dot{\theta}\cos\theta\hat{\mathbf{i}} - \dot{\theta}\sin\theta\hat{\mathbf{j}} \tag{6.88}$$
$$= -\dot{\theta}[\cos\theta\hat{\mathbf{i}} + \sin\theta\hat{\mathbf{j}}] \tag{6.89}$$
$$= -\dot{\theta}\hat{\mathbf{r}} \tag{6.90}$$

We conclude that

$$\boxed{\mathbf{a} = (\ddot{r} - r\dot{\theta}^2)\ \hat{\mathbf{r}} + (2\dot{r}\dot{\theta} + r\ddot{\theta})\ \hat{\boldsymbol{\theta}}} \tag{6.91}$$

Remark. *Knowing these relations by ♡ is not necessary, it will be much more useful to you to know and understand how to quickly rederive them. Can you derive the expressions of he kinematic quantities in spherical coordinates?*

Example

In the last section, we finished with the example of a circular motion. Due to the symmetries of the particle trajectory, polar coordinates are much more adapted to solving this problem than Cartesian coordinates. Starting from the vector position in Cartesian coordinates:

$$\mathbf{r}(t) = R\sin\Omega t\hat{\mathbf{i}} + R\cos\Omega t\hat{\mathbf{j}} \tag{6.92}$$

we easily obtain that a system of polar coordinates for this particle is given by

$$r(t) = R \quad \text{and} \quad \theta(t) = \frac{\pi}{2} - \Omega t \tag{6.93}$$

i.e.

$$\boxed{\mathbf{r} = R\ \hat{\mathbf{r}}} \tag{6.94}$$

we have just seen that $\mathbf{v} = \dot{r}\ \hat{\mathbf{r}} + r\dot{\theta}\ \hat{\boldsymbol{\theta}}$, leading to

$$\boxed{\mathbf{v} = -R\Omega\ \hat{\boldsymbol{\theta}}} \tag{6.95}$$

and finally, as $\ddot{r} = \ddot{\theta} = 0$, we obtain that

$$\boxed{\mathbf{a} = -R\Omega^2\ \hat{\mathbf{r}}} \tag{6.96}$$

You should be comfortable with deriving kinematic relations in various coordinate systems including the cylindrical and spherical systems of coordinates in $\mathbb{R}^3$ or using intrinsic coordinates as was developed in Section 6.2.1.