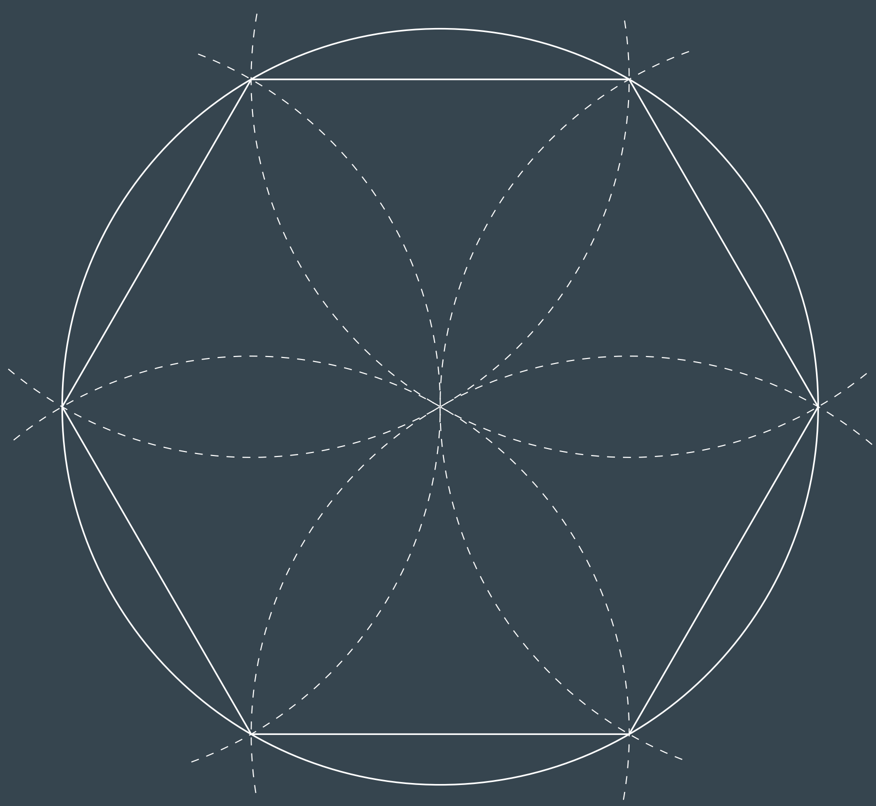
Mathematics for the IGCSE

And a bit more



Lucas Virgili

Mathematics for the IGCSE

Lucas Virgili

Updated 05/06/2022

Introduction

This book may be different from what you are used to. It does not have exercises! What is its goal, you should ask?

As a teacher, I found that many textbooks with a good selection of exercises are available already. However, I believe they miss two things: a more structured explanation of the topics and some extension for those that are interested in understanding more about *why* things work.

Thus, I started writing this book, which tries to emulate how I approach the topics while teaching them to my students. Not only that, I have added, whenever possible, justification and more mathematical theory for the facts used. This extension is usually at the end of the chapter in a section called "Formality after taste", which you can skip if you are not interested, as it will not be accessed.

I hope this book can help you in achieving your goals.

Acknowledgments

I would like to thank all the students that proof read and gave feedback on initial versions of this project. In particular to Laura, Sofia, Victoria, Gabriela and Luiza. I also thank all the motivation my students have given me to continue writing!

More importantly, I thank my wife for, well, everything. Not only the book would not be here if not for you.

Updates on this version

- New techniques to factorise quadratic expressions.

Contents

[1] I highly recommend you to ignore this. It's always better to think than to memorize.

Part I.

Number

1. Number sets

1.1. Why learn number sets (types of numbers)

Every area of knowledge has its *jargon*, its technical terms. It would be very hard for professionals to discuss with one another without them. Mathematicians also have them, of course. Perhaps most important, we have ways to refer to certain types of numbers, as we are interested in numbers in math. Hence, this chapter will give some names to certain sets of numbers, in a way that we can refer to them with ease later on.

1.2. The main ones: natural, integer, rational and irrational numbers

A very important German mathematician, Leopold Kronecker, said

> "God made the natural numbers; all else is the work of man."[1]

What are those "natural numbers", and what did man create? Let us find out.

I have always thought that it is very "human" to have the numbers we use to count: $1, 2, 3$ and so on. That because it is quite human to want to *boast*. I can imagine our ancestors in a discussion and saying "I have more sheep than you". My silly notions aside, counting seems to be very *natural*: it is quite logical, from a survival standpoint, that being able to assess quantities would be useful.

In that sense, the numbers we use to count

$$0, 1, 2, 3, 4, 5, \dots$$

are quite *natural*. Hence, I always found it reasonable that they were called *natural numbers*[2]. We denote sets in maths using curly brackets, $\{\}$, and sometimes we give them a symbol. The symbol for the natural numbers is a weird N, $\mathbb{N}$. Thus, we write

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

to refer to the set of the natural numbers, the numbers we use to count.

[1] `https://en.wikipedia.org/wiki/Pre-intuitionism`. Accessed on 26/12/2019.

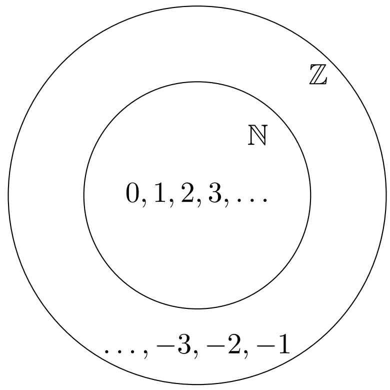
[2] You may be wondering about 0, as it is not really used to count. Some people think 0 is a natural number, some people don't. As long as you are consistent within your work it does not make any difference. I myself do not think 0 is natural, as a symbol (something) to denote nothing is a weird notion. Even so, not all societies which developed numbers had a symbol for nothing, so I think it is reasonable not to assume 0 to be very natural. But if you want to, do it. IGCSE wants to, so we will as well.

According to Kronecker, then, we have finished investigating the work of God. Let us now investigate the work of man.

One way to "extend" the natural numbers is to have the negatives. We call those numbers (usually called whole numbers), by *integers*, and their set is denoted by a weird letter Z, $\mathbb{Z}$:

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

Notice that the integers set contains the natural numbers. This is usually represented using a Venn diagram:



The innermost circle represents the natural numbers. The circle around it represents the integers, but only when we consider the natural numbers with the negatives and 0, which are not natural numbers! The diagram makes it clear that every natural number is an integer, but not every integer is a natural number.

Now, we extend our sets by adding numbers which are not integers. The most general way to represent those numbers is by using fractions[3]. Let us define what a fraction is.

A fraction is a number of the form

$$\frac{a}{b}$$

where a can be any integer and b any natural (see how having those names help?). We call a the *numerator* of the fraction and b the *denominator*. For instance,

$$\frac{1}{2}$$

is read "one over two" or "one half", and it is commonly represented using a decimal, 0.5.

Notice that b can be 1, so we can represent integers as fractions:

$$3 = \frac{3}{1}$$

$$-4 = \frac{-4}{1}$$

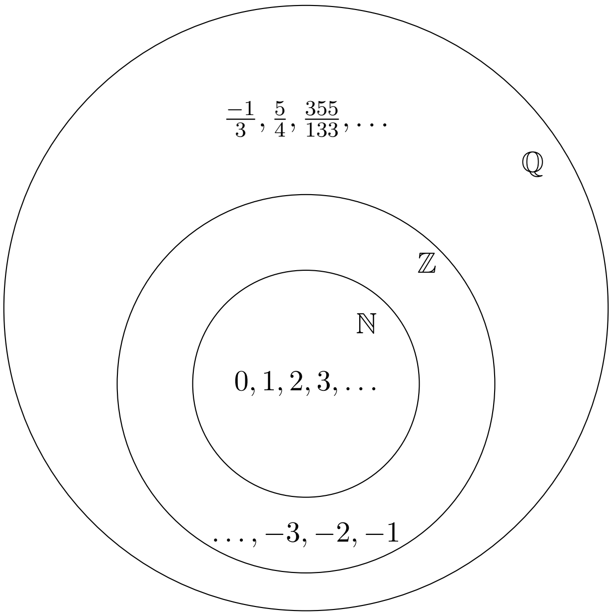
[3]You may be more used to decimal representation of fraction, and that is fine. Fractions are better, though!

Thus, the integers are "inside"[4] the fraction sets. When we combine the fractions with the integers we obtain a set called the *rational* numbers, as they can express divisions, which were called ratios. Hence the name rational. We denote this set by a weird Q:

$$\mathbb{Q} = \left\{\ldots, \frac{-3}{1}, \ldots, \frac{-7}{3}, \ldots, \frac{-2}{1}, \ldots, 0, \ldots, \frac{1}{2}, \ldots, \frac{1}{1}, \ldots\right\}$$

which needs a lot of ... as there are infinite fractions between any two numbers.

We can add the fractions to our Venn diagram:



Some numbers, however, can not be written as fractions. What that means is that there is no way to write them as a division of an integer and a natural number. These numbers are called *irrational*, which comes from the Latin for not rational. We denote this set using the symbol $\mathbb{IQ}$.

The most famous irrational number in school is π, but there are infinitely more[5]. The ones we have to know are irrational are:

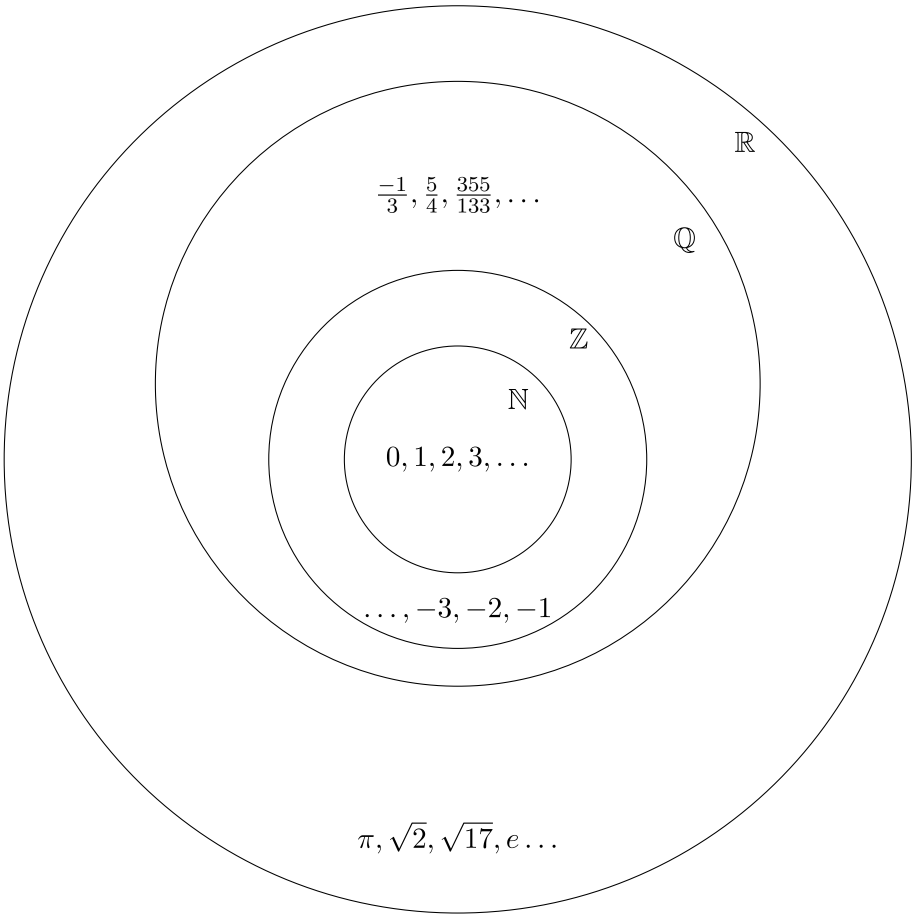
- $\pi = 3.1415\ldots$,
- any square root which is not exact (the correct term would any non-square number, but we have not defined those yet): $\sqrt{2}, \sqrt{3}, \sqrt{5}, \ldots$
- $e = 2.71828\ldots$

[4] the correct term is "contained", and we will see them better when we cover sets.

[5] Actually, the infinity of the irrational number is actually bigger than the infinity from the rational ones. Amazing, no?

Unfortunately, we cannot simply add another circle around the others in our Venn diagram, as the irrational numbers do not “overlap” with the rationals, they are fundamentally different. Thus, the irrational numbers are very different from everyone else: the preceding sets would always contain the one that came before (integers contained the naturals, rationals contained the integers), but the irrational numbers do not contain the rationals. This is obvious from the definition of the irrational numbers: they are the numbers which are *not rational.*

However, when we join together the rational numbers with the irrational numbers we obtain the set of the *real numbers*, which we denote by $\mathbb{R}$. With this new definition we can finally finish our diagram:



Most people stop here, as the real numbers are not only very interesting, they have this sense of *completeness* to them. There are other number sets, and depending on your choices you may learn them later on. The youtube channel Numberphile has a great video on this called “All the numbers”, which I highly recommend if you are interested in learning more.

1.3. Multiples and factors

If we take any number (we will mainly focus on integers, though) and multiply it by an integer, we obtain what we call a *multiple* of our first number. For instance, if we start with 17, we can multiple it by 3:

$$17 \times 3 = 51$$

the result, 51, is called a *multiple* of 17. Given that we have an infinite amount of integers to choose from, we reach the conclusion that any number has an infinite number of multiples. If we choose 2, we can write them as:

$$\ldots 2 \times -2, 2 \times -2, 2 \times -1, 2 \times 0, 2 \times 1, 2 \times 2, \ldots$$

$$\ldots, -4, -2, 0, 2, 4, \ldots$$

and we could continue adding multiples of 2 to either side of our list.

Some interesting properties of multiples of any number are:

- 0 is a multiple of any number, as any number n multiplied by 0 is 0: $n \times 0 = 0$
- any number is a multiple of itself as if we pick a number n and multiply it by 1, we obtain n: $n \times 1 = n$

Hence, multiples are basically a "general times-table", which we could extend forever both towards positive infinity or backwards towards negative infinity. Thus, to find multiples of a number, simply multiply it by integers to your heart content.

Multiples of a number are obtained by multiplying it by another number. It is logical to extend this to division. To do this, we define a *factor* of a number as any integer which divides the first number with remainder 0. For instance, 5 is a factor of 15, as $\frac{15}{3} = 5$ and the remainder is 0. Interestingly, the same relationship also tells us that 3 is a factor of 15, as $\frac{15}{5} = 3$. Normally we just consider the positive integers which divide the given number, but the negatives are totally fine as well. If we think on 15 again, its factors are

$$-15, -5, -3, -1, 1, 3, 5, 15$$

Factors are very different from multiples if we consider the fact that a number has an infinite number of multiples, whereas it always has a finite number of factors. This is easy to understand why: a number cannot be divided with remainder 0 by any integer greater than itself[6]. So, if we think again on 15, any number greater than 15 (or smaller than −15) will always yield a remainder of 15 if we divide 15 by it:

$$15 \div 17 = 0r15$$

Some interesting properties of factors are:

[6] And if you consider the negatives, it also cannot be divided by any integer smaller than minus itself. If you already know the notion of absolute value, the factors of a number have to be smaller than the absolve value of the number itself.

- 0 is not a factor of any number, as we cannot divide by 0[7]
- A number n is always divisible by 1 and n itself (and -1 and $-n$ if we are taking the negative integers into consideration). Hence 1 and n are always factors of n.

One way to find all the factors of a number is by dividing it in a systematic way. For instance, to find all factors of 48, we start by dividing it 1:

$$48 \div 1 = 48$$

With this, we already found two factors of 48: 1 and 48. We now do it with the next positive integer which divides 48, 2:

$$48 \div 2 = 24$$

which gives us two extra factors: 2 and 24. The next one to divide 48 is 3:

$$48 \div 3 = 16$$

which again gives us two factors: 3 and 16. Now with 4:

$$48 \div 4 = 12$$

adding 4 and 12 to the list. Finally we have 6:

$$48 \div 6 = 8$$

which are the last two factors of 48. We know we are done because the next number that divides 48 is 8, and it is already on our list. Hence, the factors of 48 are:

$$1, 48$$

$$2, 24$$

$$3, 16$$

$$4, 12$$

$$6, 8$$

Notice factors always come in pairs. Another example, 36:

[7] As curiosity, we cannot divide by 0 because it is inconsistent. One way to see this is considering the equation $\frac{a}{0} = b$, in which we divide a non-zero number a by 0 and obtain a number b as the answer. We can rearrange it to $a = b \times 0 = 0$, which means $a = 0$. But we started assuming a was *not* 0, and here is the inconsistency.

$$36 \div 1 = 36 \rightarrow 1, 36$$

$$36 \div 2 = 18 \rightarrow 2, 18$$

$$36 \div 3 = 12 \rightarrow 3, 12$$

$$36 \div 4 = 9 \rightarrow 4, 9$$

$$36 \div 6 = 6 \rightarrow 6, 6$$

Sometimes, as with 36, you get a “fake pair” of factors, in this case 6 and 6. But that is fine. And, if you want to find the negative factors as well, just add a negative sign in front of all the positive factors.

1.4. Interesting ones: primes, squares, cubes and triangular numbers

1.4.1. Prime numbers

As we saw in the factors section, all numbers have at least two factors: 1 and itself. The definition of a prime number comes from that:

Definition. A natural number is called *prime* if it has **exactly two distinct factors**.

Notice that this definition excludes the most common mistake for a prime number, 1, as 1 only has **one** distinct factor. Also notice that it avoids the need to say 1 is not prime, as it is with the more common definition that a prime is a number that can only be divided by 1 and itself (this being a consequence of our definition, as all numbers are divisible by 1 and themselves, and prime numbers can only be divided by two numbers).

The first prime numbers are:

$$2, 3, 5, 7, 11, 13, 17, 19, 23, \ldots$$

There are infinite prime numbers (see a proof at the end of the chapter). We will see why prime numbers are so interesting on chapter (add ref to factorisation chapter).

Some interesting things to know about primes and finding them:

- 2 is the only even prime number, as all other even numbers can be divided by 2
- to know if a number is divisible by 3, you can add its digits and see if the result is divisible by 3: if it is, the number itself also is. Example: 255, adding its digits $2 + 5 + 5 = 12$. As 12 is divisible by 3, 255 is also divisible by 3
- any number with 0 or 5 as its last digit is divisible by 5

- to find if a number is prime, just divide it smaller numbers between 2 and the number minus 1 (you can even try the prime ones). If you manage to find a divisor, the number is not prime[8]

1.4.1.1. An algorithm to find prime numbers: the sieve of Eratosthenes

In everything we will be doing in this chapter, we need to know at least some prime numbers. If you forget, however, there is an ancient way of finding all the prime numbers smaller than a limit. It is called the sieve of Eratosthenes, know since the Greek times.

To start it, you first write all the numbers between 1 and your limit, let us say 30:

$$1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15$$

$$16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30$$

Now we follow the steps:

1. We first exclude 1, as it is not a prime number:

$$\not{1}, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15$$

$$16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30$$

2. Now we repeat the following:
   a) Find the first number in our list we has not been excluded, in this case 2:

$$\not{1}, \boxed{2}, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15$$

$$16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30$$

   This number is always a prime. I will colour the primes blue.

   b) What we do now is to eliminate from our list every multiple of the prime we just found. In our case, all the multiples of 2:

$$\not{1}, \boxed{\color{blue}{2}}, 3, \not{4}, 5, \not{6}, 7, \not{8}, 9, \not{10}, 11, \not{12}, 13, \not{14}, 15$$

$$\not{16}, 17, \not{18}, 19, \not{20}, 21, \not{22}, 23, \not{24}, 25, \not{26}, 27, \not{28}, 29, \not{30}$$

   c) Finally, we check if we all numbers of the list are either eliminated or are prime. If yes, we are done. If not, start from 2(a) again.

[8] The faster way is to actually trying dividing the number by the prime numbers until you reach the square root of the number, but no need to know that.

As we are not done yet, we continue from 2(a), selecting 3:

$$\cancel{1}, {\color{blue}2}, \boxed{\color{blue}3}, \cancel{4}, 5, \cancel{6}, 7, \cancel{8}, 9, \cancel{10}, 11, \cancel{12}, 13, \cancel{14}, 15$$

$$\cancel{16}, 17, \cancel{18}, 19, \cancel{20}, 21, \cancel{22}, 23, \cancel{24}, 25, \cancel{26}, 27, \cancel{28}, 29, \cancel{30}$$

Now we eliminate all the multiples of 3 (some of them are already eliminated, but that is not a problem, just ignore them (or eliminate them again!)):

$$\cancel{1}, {\color{blue}2}, \boxed{\color{blue}3}, \cancel{4}, 5, \cancel{6}, 7, \cancel{8}, \cancel{9}, \cancel{10}, 11, \cancel{12}, 13, \cancel{14}, \cancel{15}$$

$$\cancel{16}, 17, \cancel{18}, 19, \cancel{20}, \cancel{21}, \cancel{22}, 23, \cancel{24}, 25, \cancel{26}, \cancel{27}, \cancel{28}, 29, \cancel{30}$$

We still have numbers which have not been eliminated nor are prime, so we continue with 5 and eliminate its multiples:

$$\cancel{1}, {\color{blue}2}, {\color{blue}3}, \cancel{4}, \boxed{\color{blue}5}, \cancel{6}, 7, \cancel{8}, \cancel{9}, \cancel{10}, 11, \cancel{12}, 13, \cancel{14}, \cancel{15}$$

$$\cancel{16}, 17, \cancel{18}, 19, \cancel{20}, \cancel{21}, \cancel{22}, 23, \cancel{24}, \cancel{25}, \cancel{26}, \cancel{27}, \cancel{28}, 29, \cancel{30}$$

Continuing with 7 (we do not eliminate any numbers as they have already been, but if the limit was greater we would):

$$\cancel{1}, {\color{blue}2}, {\color{blue}3}, \cancel{4}, {\color{blue}5}, \cancel{6}, \boxed{\color{blue}7}, \cancel{8}, \cancel{9}, \cancel{10}, 11, \cancel{12}, 13, \cancel{14}, \cancel{15}$$

$$\cancel{16}, 17, \cancel{18}, 19, \cancel{20}, \cancel{21}, \cancel{22}, 23, \cancel{24}, \cancel{25}, \cancel{26}, \cancel{27}, \cancel{28}, 29, \cancel{30}$$

And we would repeat this with $11, 13, 17, 19, 23$ and 29, which are all primes. Their multiples have all been eliminated in our list, and we would get:

$$\cancel{1}, {\color{blue}2}, {\color{blue}3}, \cancel{4}, {\color{blue}5}, \cancel{6}, {\color{blue}7}, \cancel{8}, \cancel{9}, \cancel{10}, {\color{blue}11}, \cancel{12}, {\color{blue}13}, \cancel{14}, \cancel{15}$$

$$\cancel{16}, {\color{blue}17}, \cancel{18}, {\color{blue}19}, \cancel{20}, \cancel{21}, \cancel{22}, {\color{blue}23}, \cancel{24}, \cancel{25}, \cancel{26}, \cancel{27}, \cancel{28}, {\color{blue}29}, \cancel{30}$$

The numbers in blue are all prime numbers smaller than 30:

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29$$

I suggest you to memorise those at least, and should the need arise, either do the sieve with a larger limit or try dividing the numbers to find more primes.

1.4.2. Square numbers

If we think of any natural number, we can multiply it by itself. Take 4 for instance:

$$4 \times 4 = 16$$

We can denote multiplication by itself using the *squared* symbol, a tiny 2 on top of a number:

$$4 \times 4 = 4^2 = 16$$

16, the answer to 4^2, is called a *square number*, as it can be obtained by squaring a natural. The first square numbers are:

$$1^2 = 1 \times 1 = 1$$

$$2^2 = 2 \times 2 = 4$$

$$3^2 = 3 \times 3 = 9$$

$$4^2 = 4 \times 4 = 16$$

$$5^2 = 5 \times 5 = 25$$

and so on. As we have infinite naturals to choose from, there are also infinite square numbers.

1.4.3. Cube numbers

We can also multiply a number by itself 3 times. Say we start with 5:

$$5 \times 5 \times 5 = 125$$

much like squaring a number, this is called cubing it, and we denote it by adding a small 3 on top of the number:

$$5 \times 5 \times 5 = 5^3 = 125$$

Again, 125 is called a cube number, as it is the result of a natural number cubed. The first cube numbers are:

$$1^3 = 1 \times 1 \times 1 = 1$$

$$2^3 = 2 \times 2 \times 2 = 8$$

$$3^3 = 3 \times 3 \times 3 = 27$$

$$4^3 = 4 \times 4 \times 4 = 64$$

$$5^3 = 5 \times 5 \times 5 = 125$$

and so on. Again, we have an infinite numbers to cube, so we have an infinite number of cube numbers.

1.4.4. Triangular numbers

Triangular numbers (also called triangle numbers) are one type of *figurate numbers*: numbers which can be used to form a shape using dots.

Triangular numbers are the numbers which can be used to form these triangles (obviously):

and so on. The next triangle would be obtained by adding a row of 5 circles to the bottom, the next after 6 and following this pattern until we wanted. The *triangle* numbers are the quantity of circles in those triangles:

$$1, 3, 6, 10, 15, 21, 28, \dots$$

They have a formula to find:

$$\frac{n(n-1)}{2}$$

which you will learn how to find in the Sequences chapter.

Also, as curiosity, square numbers are also figurate numbers, and they can be used to make squares:

1.5. Exam hints

This topic is more a foundation to others than a topic in itself, so there is not much to say about studying it for the exams themselves. It is important to memorize all the definitions, of course, and to be able to identify to which set a number belongs to. You also need to know how to find the factors of a number.

Summary

- The *natural* numbers are the **counting** numbers: $\mathbb{N} = \{0, 1, 2, 3, 4, 5, \dots\}$;
- The *integers* are the **naturals and the negative whole numbers**:
  $\mathbb{Z} = \{\dots - 3, -2, -1, 0, 1, 2, 3, \dots\}$;
- The *rationals* are the numbers which **can be written as fractions**:
  $\mathbb{Q} = \left\{\dots, \frac{-3}{1}, \frac{-7}{3}, \frac{-2}{1}, \dots, 0, \frac{1}{2}, \frac{1}{1}, \dots\right\}$;
- The *irrationals* are the numbers which **cannot be written as fractions**. We need to know that

  - π
  - e
  - square roots of non-square numbers

  are all irrational;

- The *real* numbers are **all** the above together: **rationals and irrationals**. We denote them by $\mathbb{R}$;
- A *multiple* of a number is any number which can be obtained by **multiplying the original number by an integer**;
- A *factor* of a number is any integer (normally we only want the positive ones, though) which **divide the original number with remainder** 0;
- A *prime* number is any number which **has exactly two distinct factors**. By consequence, these two factors will always be 1 and the number itself. The first prime numbers are
$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, \ldots$$
- A *square* number is a number obtained by **squaring a natural number** (multiplying the number by itself). The first square numbers are
$$1, 4, 9, 16, 25, 36, \ldots$$
- A *cube* number is a number obtained by **cubing a natural number** (multiplying the number by itself three times). The first cube numbers are
$$1, 8, 27, 64, 125, \ldots$$
- A *triangle* number is any number which can be used to represent triangles in the sequence

  The first triangle numbers are
$$1, 3, 6, 10, 15, 21, \ldots$$

Formality after taste

A general way of writing even and odd numbers and some properties

In school you probably learnt that even numbers are either the numbers which finish in $0, 2, 4, 6$ or 8 (which is a horrible definition, as this is a property of the even numbers) or number which you can divide by 2 with remainder 0 (which is a good way to define them). Let us use this definition to write a general even number.

If any even number can be divided by 2 with remainder 0, it means that it has at least one 2 in its prime decomposition. Let us see some examples:

$$6 = \mathbf{2} \times 3$$

$$10 = \mathbf{2} \times 5$$

$$12 = \mathbf{2} \times 2 \times 3$$

We can then safely say that any even number is equal to 2 times some other number. Let us use algebra to write that: if we assume that a number n is even, we are sure that n is equal to 2 times another number (which we also do not know). Let us call this other number m. Thus, we can write

$$n \text{ is even}$$

$$n \text{ is equal to 2 times another number } m$$

$$n = 2 \times m$$

$$n = 2m$$

That gives us a way to write "a general" even number: $2m$.

What about odd numbers? In the same way, an odd number is a number which we cannot divide by 2 with remainder 0. Now, the only possibilities when we divide a number by 2 is that either the remainder is 0 or 1. Hence, odd numbers, when divided by 2, have remainder 1. This means any odd number can be written as an even number plus 1. Some examples:

$$3 = 2 + 1$$

$$17 = 16 + 1$$

$$35 = 34 + 1$$

We know that any even number can be written as $2m$, so any odd number is an even number plus 1:

$$o \text{ is odd}$$

$$o \text{ is equal to an even number plus 1}$$

$$o = 2m + 1$$

Hence, a "general" odd number is given by $2m+1$.

What can we do with these? Let us see some examples. We can prove, for instance, that an even number added to an even number is even:

$$\underbrace{2m}_{\text{1st even number}} + \overbrace{2n}^{\text{2nd even number}} = 2\underbrace{(m+n)}_{m+n=k} = 2k$$

(here we called $m+n$ as k). As $2k$ is an even number, we know that $2m+2n$ is also even. We can also show that an odd number added to another odd number is even:

$$\underbrace{2m+1}_{\text{1st odd number}} + \overbrace{2n+1}^{\text{2nd odd number}} = 2m+2n+2 = 2\underbrace{(m+n+1)}_{m+n+1=j} = 2j$$

As $2j$ is even, we know that $2m+1+2n+1$ is also even!

We can also show that the multiplying an even number by another even number is even:

$$2m \times 2n = 4mn = 2 \times \underbrace{2mn}_{x=2mn} = 2x$$

Again, as $2x$ is even, $2m \times 2n$ is also even. The fact that an even times another even is even also implies that the square of an even number is also even.

Using the same technique we can prove that and odd number multiplied by another odd number is odd:

$$(2m+1)(2n+1) = 4mn + 2m + 2n + 1 = 2\underbrace{(2mn+m+n)}_{y=2mn+m+n} + 1 = 2y+1$$

As $2y+1$ is odd, we have that the product $(2m+1)(2n+1)$ is also odd. Again, this implies that the square of an odd number is odd.

A proof that $\sqrt{2}$ is irrational

Let us use what we learned in the above to show that $\sqrt{2}$ is irrational.

This is a classic proof, and I do believe everyone should see it. The proof is simple, and it uses a technique called 'contradiction': we are going to assume that $\sqrt{2}$ is rational and we are going to reach a contradiction. This means that our assumption of $\sqrt{2}$ being rational has to be wrong (as everything else we will be doing is certainly correct). A great mathematician, Godfrey H. Hardy said that a proof by contradiction "is a far finer gambit than any chess gambit: a chess player may offer the sacrifice of a pawn or even a piece, but a mathematician offers the game"[9]. He meant that we assume the *opposite* of what we want to prove, in the hopes of reaching our contradiction and showing we were actually wrong. A fine weapon indeed, Mr Hardy.

Let us begin!

[9] In a book called "A Mathematician's Apology", which I highly recommend.

As I mentioned, our first step is assuming that $\sqrt{2}$ is rational (which is the opposite of what we want!). That means we can write $\sqrt{2}$ as a fraction:

$$\sqrt{2} = \frac{p}{q}$$

Our fraction has p on its numerator and q on its denominator. Let us also assume that $\frac{p}{q}$ is already in its simplest terms, that is, there is no number (other than 1) which divides p and q. In mathematicianese, the greatest common factor of p and q is 1. This is not a problem, as if it were not true, we could simplify the fraction until it reached its simplest form.

Now let us do some algebraic manipulation:

$$\begin{aligned}
\frac{p}{q} &= \sqrt{2} && \text{Our assumption} \\
\left(\frac{p}{q}\right)^2 &= \left(\sqrt{2}\right)^2 && \text{Squaring both sides} \\
\frac{p^2}{q^2} &= 2 \\
p^2 &= 2q^2 && \text{Multiplying both sides by } q^2
\end{aligned}$$

Let us think on what we have here: after our initial assumption, we used basic algebra to show that

$$p^2 = 2q^2$$

As we saw above, numbers of the form $2k$ are even. Hence, we can say for sure that p^2 is an even number, as we can obtain it by multiplying the number q^2 by 2. This also implies that p itself has to be an even number, as the square of an odd number is also odd. Thus, we can write p as two times a number x:

$$p = 2x$$

Let us do some more algebra now:

$$\frac{p^2}{q^2} = 2 \qquad \text{One of our steps in our previous derivation}$$

$$\frac{(2x)^2}{q^2} = 2 \qquad \text{Substituting } p = 2x$$

$$\frac{4x^2}{q^2} = 2$$

$$4x^2 = 2q^2$$

$$q^2 = \frac{4x^2}{2}$$

$$q^2 = 2x^2$$

Again, using basic algebra, we have shown that

$$q^2 = 2x^2$$

which implies that q^2 is an even number. One more time, this implies that q itself is even. We can, then, write q as a product of 2 and some number y:

$$q = 2y$$

Thus, we have that *both* p and q are even. However, we have assumed that our fraction representation of $\sqrt{2}$ was already in its simplest form! This is our contradiction: if we could write $\sqrt{2}$ as a fraction, it wouldn't have a simplest form, as we would always be able to divide it by 2. This is impossible, hence $\sqrt{2}$ is an irrational number.

Euclid's proof of the infinitude of primes

There are many proofs that there are infinite prime numbers. The excellent book *Proof from the Book*, by Aigner and Ziegler starts with six of them. This is the first, attributed to the Greek mathematician Euclid.

The proof, like the irrationality of $\sqrt{2}$, also goes by contradiction. As we want to prove there are infinite primes, we will assume, hoping to reach a contradiction, that there is a finite number of prime numbers, which we will be lazy enough to call p_1, p_2 and so on until we reach a final prime, p_n:

$$p_1, p_2, p_3, \ldots, p_n \qquad \text{Our finite prime number list}$$

Now, we will create a new number, x, by multiplying all our prime numbers and adding 1 to the result:

$$x = p_1 \times p_2 \times p_3 \cdots \times p_n + 1$$

We have some possibilities for x now. If x is prime, we just found a new prime not on our list, so our assumption of that list containing all primes is already wrong, and we have our contradiction. So, let us assume x is not prime. If x is not a prime number, it means we can divide it by at least one prime number. However, x cannot be divided by p_1, as the result of the division of x by p_1 will have remainder 1 (hence why the plus 1 we added to x!). The same reasoning can be applied to any prime in our list. Hence, we have a contradiction: x is not prime, but it does not have any prime factors. Thus, our list must be incomplete, and there are infinite prime numbers.

2. Primes, factorisation and friends

2.1. Why learn how to find the prime factorisation of a number

If you are reading this section in order to find some real use application of prime numbers and their use, sorry. I know of none.

Perhaps, though, you would like to find out *why* we have this fascination with prime numbers. You will discover one very important reason in this chapter. It the *Elements*, Euclid proves[1] a very important result called the *fundamental theorem of arithmetic*. A theorem is a mathematical statement which we need to prove to be true. Euclid proved that any integer which is bigger than 1 is either a prime number or it can be written as a product of prime numbers. For instance, if we take the number 42, we can write it as

$$42 = 2 \times 3 \times 7$$

or, if we take 78:

$$78 = 2 \times 3 \times 13$$

or 36:

$$36 = 2 \times 2 \times 3 \times 3$$

and so on. Not only that, the theorem also states that this way to write the number as product of primes is *unique*: there is only one way to write each number as a product of primes, apart from the order we multiply those primes!

Now, you may ask, why is that important? Because we can use prime numbers as "building blocks" for every integer. Hence their importance, why they are "prime", the best. Also, this theorem is so important that it makes sense for 1 not to be a prime: if it were, the uniqueness part of the theorem would break

$$42 = 2 \times 3 \times 7 = \underbrace{2 \times 3 \times 7 \times 1}_{\text{different way}} = \underbrace{2 \times 1 \times 3 \times 1 \times 7 \times 1 \times 1}_{\text{another way!}}$$

so we would have infinite different prime products which would give any number, which is not as nice as having a single one.

2.2. Prime factors

This chapter basically uses the fundamental theorem arithmetic, but we give it a more practical spin: we want to know *how* to find the prime numbers which multiply to any

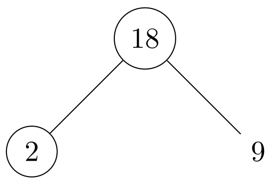
[1] Not the strongest case, according to wikipedia: `https://en.wikipedia.org/wiki/Fundamental_theorem_of_arithmetic`.

given number, and after that, use those primes to find some interesting things. Let us begin.

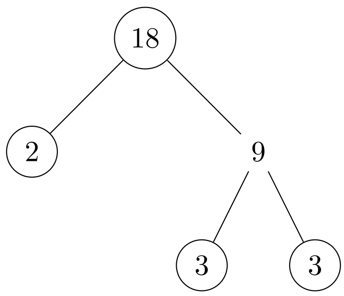
2.2.1. Prime trees

Both prime trees and ladders are based on dividing the number which we are given by prime numbers until we reach 1. Basically, then, we start with our number and divide it by the smallest prime number which divides it. It is a good idea to start with the smallest prime so that we do not forget any. Then you continue until you have only prime numbers as leafs of the tree: any branch finishes in a prime number.

Let us see an example. If we take then number 18, the smallest prime number which divides it is 2. To represent this, we do a little drawing:



We circle 2 as it a prime, so we are done on that branch of the tree. However, 9 is not a prime, so continue dividing it by 3, which is the smallest prime number which divides 9:

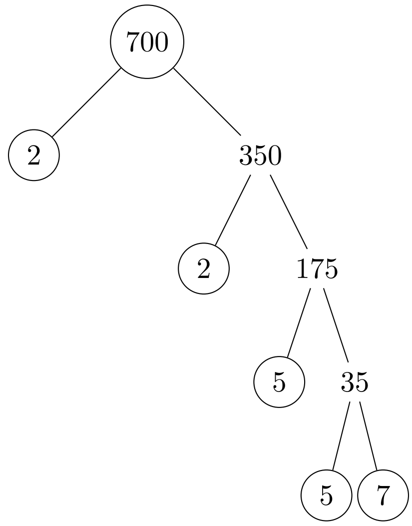


As 3 is prime, all our leafs (the ends of branches) are prime, so we are done. Thus, we can write 18 as

$$18 = 2 \times 3 \times 3$$

Notice that all circled numbers appear in our prime factorisation.

Another example, with 700:



Thus, we can write 700 as

$$700 = 2 \times 2 \times 5 \times 5 \times 7$$

2.2.2. Prime ladders

A second way to set up our search for the prime factors of a number is to draw a prime ladder. It goes pretty much like a prime tree, it is just a different set up. We do exactly the same: we keep dividing the current number by the smaller prime number which divides it, until we reach a 1. Let us see an example and find all the prime factors of 60:

60	2	2 divides 60
30	2	2 divides 30
15	3	3 divides 15
5	5	5 divides 5
1	done!	

Thus, we know we can write 60 as

$$60 = 2 \times 2 \times 3 \times 5$$

Notice that just like with the trees, we don't need necessarily to divide first by 2, then by 3 and so on. However, I highly recommend you do to it, so that you do not forget any prime factor.

Another example, with 6825:

6825	3	$6+8+2+5=21$ so 3 divides 6825
2275	5	$2+2+7+5=16$ so 3 does not divides 2275
455	5	
91	7	
13	13	
1	done!	

Hence, we can write 6825 as its product of prime factors:

$$6825 = 3 \times 5 \times 5 \times 7 \times 13$$

2.2.3. Index form

Take the last example we saw, 6825. We will write it as a product of its prime factors in index form. Starting with its prime factorisation:

$$6825 = 3 \times 5 \times 5 \times 7 \times 13$$

Notice that we have a 5×5 in its prime factorisation. Whenever we multiply a number by itself, we say we are *squaring* it, or *raising it to the power of* 2, and we denote that by adding a tiny 2 to the top of the number:

$$5 \times 5 = 5^2$$

This is just a notation, which you will investigate more of in (ref to Indices chapter). Using this new way to write 5×5, we can write the prime factorisation of 6825 as

$$6825 = 3 \times \underbrace{5 \times 5}_{2 \text{ times}} \times 7 \times 13 = 3 \times 5^2 \times 7 \times 13$$

We can extend this notation to multiplying the same number by itself more than 2 times. For example, if we multiply 4 by itself 3 times, $4 \times 4 \times 4$, we can also write the same product as

$$\underbrace{4 \times 4 \times 4}_{3 \text{ times}} = 4^3$$

which we call *cubing* or *raising to the power of* 3. Let us see some examples.

If we take the number 48, we can find its prime factorisation:

48	2
24	2
12	2
6	2
3	3
1	done!

Thus, we can write:

$$48 = \underbrace{2 \times 2 \times 2 \times 2}_{4 \text{ times}} \times 3 = 2^4 \times 3$$

Sometimes you can write a lot less using the index form. Say you need to write the prime factorisation of 900. First you would find it normally:

900	2
450	2
225	3
75	3
25	5
5	5
1	done!

and write

$$900 = 2 \times 2 \times 3 \times 3 \times 5 \times 5$$

Notice that you have a lot of "pairs" of a number multiplying themselves, so we can write all of them squared:

$$900 = 2 \times 2 \times 3 \times 3 \times 5 \times 5 = 2^2 \times 3^2 \times 5^2$$

2.3. Highest common factor (HCF) between two numbers

Before we learn to find the HCF of two numbers, we first need to understand what it is. Let us break down the name:

- highest: the largest
- common: which appears for both numbers
- factor: number which divides another

Thus, the highest common factor between two numbers is simply the *largest number which divides both numbers*. For instance, the largest number which divides 8 and 10 is 2, whereas the largest number which divides 18 and 24 is 6.

We will see some techniques to find the HCF between two numbers now.

2.3.1. Using lists of factors

As the HCF of two numbers is simply the largest number which divides both numbers, we can simply write all the factors of both numbers and see who is largest number which appears in both lists.

For instance, let us find the HCF of 24 and 36. First, let us find all the factors of 24:

$$24 \div 1 = 24 \rightarrow 1, 24$$

$$24 \div 2 = 12 \rightarrow 2, 12$$

$$24 \div 3 = 8 \rightarrow 3, 8$$

$$24 \div 4 = 6 \rightarrow 4, 6$$

So the factors of 24 are $1, 2, 3, 4, 6, 8, 12$ and 24. Let us find the factors of 36 now:

$$36 \div 1 = 36 \rightarrow 1, 36$$

$$36 \div 2 = 18 \rightarrow 2, 18$$

$$36 \div 3 = 12 \rightarrow 3, 12$$

$$36 \div 4 = 9 \rightarrow 4, 9$$

$$36 \div 6 = 6 \rightarrow 6, 6$$

Hence the factors of 36 are $1, 2, 3, 4, 6, 9, 12, 18$ and 36. Let us now compare the lists:

$$1, 2, 3, 4, 6, 8, 12, 24$$

$$1, 2, 3, 4, 6, 9, 12, 18, 36$$

My favourite way to find the HCF when looking at the list is starting at the largest factor of the smallest number, in our case 24, and from there searching for it in the other list. If we find it, that is our HCF. If we don't, we continue with the next biggest factor. So we start with 24:

$$1, 2, 3, 4, 6, 8, 12, \boxed{24}$$

$$1, 2, 3, 4, 6, 9, 12, 18, 36$$

there is no 24 in the list of factors of 36, so we continue with 12:

$$1, 2, 3, 4, 6, 8, \boxed{12}, 24$$

$$1, 2, 3, 4, 6, 9, \boxed{12}, 18, 36$$

as 12 is also a factor of 36, we have found the HCF of 24 and 36, which is 12. We can denote that as

$$HCF(24, 36) = 12$$

Another example, let us find the HCF between 36 and 108. The factors of 36 are still the same $1, 2, 3, 4, 6, 9, 12, 18$ and 36. Let us find the factors of 108:

$$108 \div 1 = 108 \rightarrow 1, 108$$

$$108 \div 2 = 54 \rightarrow 2, 54$$

$$108 \div 3 = 36 \rightarrow 3, 36$$

$$108 \div 4 = 27 \rightarrow 4, 27$$

$$108 \div 6 = 18 \rightarrow 6, 18$$

$$108 \div 9 = 12 \rightarrow 9, 12$$

thus the factors of 108 are $1, 2, 3, 4, 6, 9, 12, 18, 27, 36, 54$ and 108. From this list we can see that the largest number which divides both 36 and 108 is 36 itself (which is totally fine). Hence:

$$HCF(36, 108) = 36$$

Even though listing factors is fine, it is very time intensive. The next methods are quicker, particularly for larger numbers.

2.3.2. Using a ladder

Another method to find the HCF is somewhat a “generalisation” of the prime ladders. The difference is that we divide *both* numbers by the smallest prime number which is a factor of both. A small example first, the HCF between 24 and 36:

$24, 36$	2	2 divides both 24 and 36
$12, 18$	2	2 divides both 12 and 18
$6, 9$	3	3 divides both 6 and 9
$2, 3$		1 is the only number which divides both 2 and 3

We know we are done when the only number which divides both numbers on the left column is 1 (these numbers are called *coprime*, as curiosity). Now, to find the HCF, we simply multiply the numbers we divided by:

$$HCF(24, 36) = 2 \times 2 \times 3 = 12$$

Another example, the HCF between 22440 and 900:

$22440, 900$	2
$11220, 450$	2
$5610, 225$	3
$1870, 75$	5
$374, 15$	

which we know we are done because the only prime factors of 15 are 3 and 5 and we cannot divide 374 by any. Hence,

$$HCF(22440, 900) = 2 \times 2 \times 3 \times 5 = 60$$

I use this method myself, as it is quite quick.

2.3.3. HCF of more than two numbers

If you need to find the HCF of more than 2 numbers, I recommend either using the same ladder as the previous section, but with more numbers to divide. However, it is totally fine to list all the factors of the numbers and get the largest common one. You can also find the HCF between the first two numbers, and find the HCF of this first result with the next number until you are done.

For example, using the ladder, let us find the HCF of $48, 144$ and 80:

$$\begin{array}{l|l} 48, 144, 80 & 2 \\ 24, 72, 40 & 2 \\ 12, 36, 20 & 2 \\ 6, 18, 10 & 2 \\ 3, 9, 5 & \end{array}$$

we are done as 5 is prime, and the other numbers do not have a factor of 5. Thus

$$HCF(48, 144, 80) = 2 \times 2 \times 2 \times 2 = 16$$

2.3.4. Using the prime factorisation

Another method to find the HCF is to use the prime factorisation of the numbers. I left it to the end because is the more "theoretical" one, but it is definitely something I consider important to know.

Let us say we want to find the HCF between 1820 and 11550. It would take a while to find all their factors. As this method involves using their prime factorisation, let us find them:

$$\begin{array}{l|l} 1820 & 2 \\ 910 & 2 \\ 455 & 5 \\ 91 & 7 \\ 13 & 13 \\ 1 & \end{array}$$

thus $1820 = 2 \times 2 \times 5 \times 7 \times 13 = 2^2 \times 5 \times 7 \times 13$. Now for 11550:

$$
\begin{array}{l|l}
11550 & 2 \\
5775 & 3 \\
1925 & 5 \\
385 & 5 \\
77 & 7 \\
11 & 11 \\
1 &
\end{array}
$$

so we have $11550 = 2 \times 3 \times 5 \times 5 \times 7 \times 11 = 2 \times 3 \times 5^2 \times 7 \times 11$.

Now remember that the HCF of two numbers is the biggest number which divides both numbers; also *remember that any number can be written as a product of its prime factors* (the fundamental theorem of arithmetic!). As the HCF of two numbers is a number, it also can be written as a product of prime factors, and these need to be somehow *embedded* in the prime factorisation of the original two numbers. We just need to find it there.

Thinking that the HCF is what is *common* to the two numbers, let us find what is *common* in the factorisation of both numbers:

$$1820 = \boxed{2} \times 2 \times \quad \boxed{5} \times \quad \boxed{7} \times \quad 13$$

$$11550 = \boxed{2} \times \quad \times 3 \times \boxed{5} \times 5 \times \boxed{7} \times 11$$

Notice that we have in common the prime factors $2, 5$ and 7. If we multiply them all

$$2 \times 5 \times 7 = 70$$

70 is indeed the HCF of 1820 and 11550, which makes sense: we multiplied the *common* prime factors of the original two numbers, which gave a *common* factor. As we multiplied *all* common primes, we got the *highest* common factor between 1820 and 11550.

So, basically what we do to find the HCF of two numbers is find their prime factorisation and multiply everything which is common. Another example: to find the HCF of 22440 and 900. Their prime factorisation gives:

$$22440 = 2 \times 2 \times 2 \times 3 \times 5 \times 11 \times 17$$

$$900 = 2 \times 2 \times 3 \times 3 \times 5$$

Highlighting what is common:

$$22440 = \boxed{2} \times \boxed{2} \times 2 \times \boxed{3} \times \boxed{5} \times 11 \times 17$$

$$900 = \boxed{2} \times \boxed{2} \times \boxed{3} \times 3 \times \boxed{5}$$

thus:

$$HCF(22440, 900) = 2 \times 2 \times 3 \times 5 = 60$$

I suggest using this method is you already have factorised the numbers.

2.4. Lowest common multiple (LCM) between two numbers

Again, let us first understand of what the LCM of two numbers is by breaking down the name:

- lowest / least: the smallest
- common: which appears for both numbers
- multiple: a number obtained by multiplying the original by an integer

Hence, the LCM of two numbers is the smallest number which appears in the times table of both numbers. For instance, the LCM of 10 and 35 is 70, and the LCM of 24 and 16 is 48. Again, let us see some techniques to find the lowest common multiple between numbers.

2.4.1. Using lists of multiples

As with finding the HCF, the most straightforward way of finding the LCM of two numbers is by listing multiples of both until you find a common one. For instance, to find the LCM of 10 and 15, we can start by listing some multiples of 10:

$$10, 20, 30, 40, 50, 60 \dots$$

and then list some multiples of 15:

$$15, 30, 45, 60, 75, \dots$$

and noticing 30 is the smallest common multiple of 10 and 15. Hence, 30 is the LCM of 10 and 15 and we can write:

$$LCM(10, 15) = 30$$

Another example, just to prove a point: let us find the LCM of 5 and 13. First let us write some multiples of 13:

$$13, 26, 39, 52, 65, 78, \dots$$

and now some multiples of 5:

$$5, 10, 15, 20, 25, 30, 35, 40, 50, 55, 60, 65, 70, \dots$$

Notice that 65 is the smallest number which appears in both lists, so

$$LCM(5, 13) = 65$$

We had to write many multiples of 5 this time, and sometimes you have write *a lot* of multiples of both numbers to find their LCM. Therefore, let us learn some faster techniques.

2.4.2. Using the HCF

My favourite way of finding the LCM between two numbers is by using this result: given two positive integers x and y, they satisfy

$$LCM(x,y) = \frac{xy}{HCF(x,y)}$$

that is, the LCM between two numbers is equal to the product of both numbers divided by their HCF.

For instance, if we want to find the LCM of 10 and 15, first we find their HCF:

$$HCF(10,15) = 5$$

which is quite easy. Then we just use the formula above

$$LCM(10,15) = \frac{10 \times 15}{5} = \frac{\cancel{10}^{2} \times 15}{\cancel{5}^{1}} = 2 \times 15 = 30$$

For the 5 and 13 example, it is even easier. 5 and 13 are what we called coprime[2] above, which means that

$$HCF(5,13) = 1$$

Therefore, their LCM is simply their product:

$$LCM(5,13) = \frac{5 \times 13}{1} = 5 \times 13 = 65$$

2.4.3. LCM of more than two numbers

To find the LCM of more than two numbers, I usually find the LCM of the first two and then find the LCM of the third with the LCM of the first two (and so on if more than three numbers).

For instance, to find the LCM between $10, 15$ and 20, first I would find the LCM of 10 and 15:

$$LCM(10,15) = \frac{10 \times 15}{5} \frac{\cancel{10}^{2} \times 15}{\cancel{5}^{1}} = 2 \times 15 = 30$$

and now find the LCM of 30 and 20 (which have HCF of 10):

$$LCM(20,30) = \frac{20 \times 30}{10} = \frac{\cancel{20}^{2} \times 30}{\cancel{10}^{1}} = 2 \times 30 = 60$$

Hence,

$$LCM(10,15,20) = 60$$

[2] Two prime numbers, as 5 and 13, are always coprime.

2.4.4. Using the prime factorisation

There is also a way to find the LCM of two numbers using their prime factorisation.

Let us first understand the method and then explain why it works. If we want to find the LCM of 60 and 126, first we factorise them and obtain:

$$60 = 2 \times 2 \times 3 \times 5 = 2^2 \times 3 \times 5$$

$$126 = 2 \times 3 \times 3 \times 7 = 2 \times 3^2 \times 7$$

To find the LCM between them, we do the following:

- Take any number which appeared in either factorisation: in this example $2, 3, 5$ and 7
- Use the *highest power* that appeared for each prime and multiply them all:
  - For 2, we have 2^2;
  - For 3, we also have 3^2;
  - For 5 we only have 5 (5^1, but no need to write the one);
  - For 7 we also only have 7;
  - Multiplying all of them:

$$2^2 \times 3^2 \times 5 \times 7 = 1260$$

Hence,

$$LCM(60, 126) = 1260$$

Another example. To find the LCM of 13230 and 48600, which are lovely numbers, first you factorise them and obtain:

$$13230 = 2 \times 3 \times 5 \times 7^2 \times 9$$

$$340200 = 2^3 \times 3 \times 5^2 \times 9^2$$

Then we, for every prime number which appeared in either list, copy it with the highest power:

- 2 appeared as 2^3;
- 3 as 3 itself;
- 5 as 5^2;
- 7 as 7^2;
- 9 as 9^2

Finally, we multiply these numbers:

$$LCM(13230, 48600) = 2^3 \times 3 \times 5^2 \times 7^2 \times 9^2 = 2,381,400$$

Cute number, no? But why does this work? It again has to do with the fundamental theorem of arithmetic. Any number can be written as a product of prime factors; hence, if a prime factor appears in the original numbers, it will have to appear in their LCM.

Let us use an example to see this better. To find the LCM of 45 and 30 we would factorise them

$$45 = 3^2 \times 5$$

$$30 = 2 \times 3 \times 5$$

Their LCM will be the smallest number which appears in both their times-table. Hence, their LCM is the smallest number which can be divided by both 30 and 45. But, in order to be divisible by 45, for instance, a number needs to have the same $3^2 \times 5$ in *their* prime factorisation; and to be divisible by 30 they need to have $2 \times 3 \times 5$ in *their* factorisation. This explains why we would take each prime number in each of 30 and 45 factorisation, but why the higher power? Think on the 3^2 in 45 factorisation and the 3 in the 30 factorisation. If we put 3^2 in our common multiple, we *already put* a 3, so we satisfy both the 45 multiple needs (of having a 3^2) and 30 needs (of having a 3). Hence why we take each prime with their highest power.

2.5. A way to find both the HCF and LCM at the same time: Venn diagram

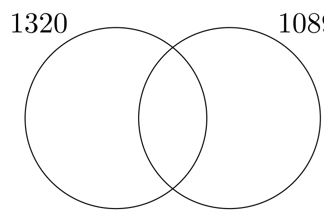
Now, a method that joins on both prime factorisation methods, the Venn diagram one.

Let us say we need to find the HCF and LCM of 1320 and 10890. For this method we first find their prime factorisation:

$$1320 = 2 \times 2 \times 2 \times 3 \times 5 \times 11 = 2^2 \times 3 \times 5 \times 11$$

$$10890 = 2 \times 3 \times 3 \times 5 \times 11 \times 11 = 2 \times 3^2 \times 5 \times 11$$

What we do now is put these numbers in a Venn diagram:



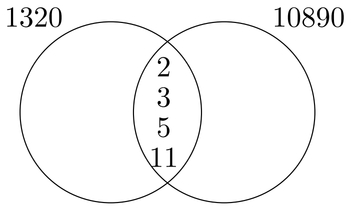
This diagram has three parts: the middle, where both circles intersect, is where we are going to put all the prime factors which appear in *both* 1320 and 10890 factorisation. The left part, which only "belongs" to 1320 will have the "leftovers" from its factorisation, every number we did not put in the intersection goes there. The same goes for the right region, but for 10890.

On the intersection, then, we put everything which is common (it is easier to see, I think, without using the index form):

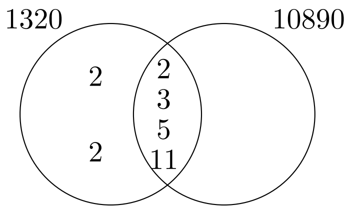
$$1320 = \boxed{2} \times 2 \times 2 \times \boxed{3} \times \boxed{5} \times \boxed{11}$$

$$10890 = \boxed{2} \times \boxed{3} \times 3 \times \boxed{5} \times \boxed{11} \times 11$$

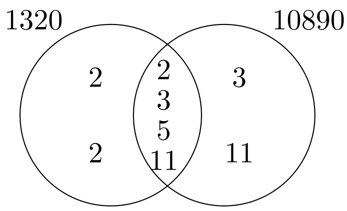
thus the common ones are: $2, 3, 5$ and 11. Adding them to the intersection:



Now, thinking on 1320, we have already "dealt" with some of its factors, and we are left with another two 2's (notice what was not framed in the previous lists). So we add those 2's to the left part of the diagram:



Finally, for 10890, we are missing one 3 and one 11:



Now comes the magic (which is simply using what we learned in Sections 2.3.4 and 2.4.4 of this chapter). To find the HCF of 1320 and 10890, we simply multiply the numbers in the intersection of the diagram, in our case, $2, 3, 5$ and 11:

$$HCF(1320, 10890) = 2 \times 3 \times 5 \times 11 = 330$$

To find the LCM, we multiply every number in the diagram:

$$LCM(1320, 10890) = \underbrace{2 \times 2}_{\text{Left part}} \times \underbrace{2 \times 3 \times 5 \times 11}_{\text{Intersection}} \times \underbrace{3 \times 11}_{\text{Right part}} = 43560$$

Thus, to summarise, after drawing and filling the diagram:

- The HCF is obtained by multiplying the numbers in the intersection
- The LCM is obtained by multiplying all the numbers in the diagram

Another example. To find the LCM and HCF of 2520 and 180. First we factorise them:

$$2520 = 2 \times 2 \times 2 \times 3 \times 3 \times 5 \times 7$$

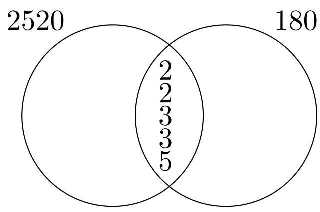
$$180 = 2 \times 2 \times 3 \times 3 \times 5$$

Identify what is common:

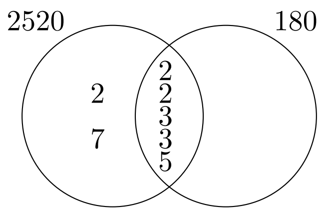
$$2520 = \boxed{2} \times \boxed{2} \times 2 \times \boxed{3} \times \boxed{3} \times \boxed{5} \times 7$$

$$180 = \boxed{2} \times \boxed{2} \times \boxed{3} \times \boxed{3} \times \boxed{5}$$

We know put those numbers in the intersection:



For the left part, we are missing the 7 and one 2 from 2520 factorisation:



We are not missing anything from 180, so we either leave that part with nothing or we can put a 1 (as it does not change anything). I will not put anything because I am lazy, and that means we are done. Now, to find the HCF, we multiply the numbers in the intersection:

$$HCF(2520, 180) = 2 \times 2 \times 3 \times 3 \times 5 = 180$$

and to find the LCM we multiply everything we wrote:

$$LCM(2520, 180) = 2 \times 7 \times 2 \times 2 \times 2 \times 3 \times 3 \times 5 = 2560$$

2.6. Exam hints

Unfortunately, exam wise, this topic is a regurgitation of technique. Know well how to find the LCM and HCF of two numbers, and the definition of prime factors, factors, and index form that you will be fine.

Summary

- A *prime factor* is a **factor of a number which is also prime**;
- The first prime numbers are:

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41 \ldots$$

- The *prime factorisation* of a number is a **product (multiplication) of prime numbers which is equal to the number**. Each of the primes in the prime factorisation is called a *prime factor* of the original number;
- You can use either a *prime tree* or a *prime ladder* to find the prime factorisation of a number;
- The *index form* of a factorisation is simply a "compact" representation of the prime factorisation using indices:

$$24 = \underbrace{2 \times 2 \times 2 \times 3}_{\text{“normal”}} = \underbrace{2^3 \times 3}_{\text{index-form}}$$

- The *highest common factor (HCF)* of two numbers is the **largest number which divides both numbers;**
- You can find the HCF by **listing all factors of the numbers, using a ladder or using the prime factorisation of the numbers;**
- The *lowest common multiple (LCM)* of two numbers is the **smallest number which appears in both numbers' times-tables;**
- You can find the LCM of two numbers by **listing some multiples of both until you find a common one, using the prime factorisation** or using the formula:

$$LCM(x, y) = \frac{xy}{HCF(x, y)}$$

- You can use a **Venn diagram** to find both the LCM and the HCF of two numbers if you have their prime factorisation:
  - The HCF is the **product of all numbers in the intersection**
  - The LCM is **product of all numbers in the diagram**

Formality after taste

A cute fact

Given two integers x and y and a natural number k, we have

$$LCM(kx, ky) = kLCM(x, y)$$

An example: if $x = 3$ and $y = 8$ and $k = 2$, we have:

$$kx = 2 \times 3 = 6$$

$$ky = 2 \times 8 = 16$$

The LCM of 6 and 16 is 48, so we have

$$LCM(kx, ky) = 48$$

The LCM of 3 and 8 is 24, so we have

$$LCM(x, y) = 24$$

which gives

$$LCM(2 \times 3, 2 \times 8) = 2 \times LCM(3, 8)$$

Another cute fact

Given x and y positive integers and coprime, we have that

$$LCM(x, y) = xy$$

Deriving the formula for the LCM

We want to show that the LCM between two integers x and y is given by

$$LCM(x, y) = \frac{xy}{HCF(x, y)}$$

which is equivalent to

$$xy = LCM(x, y) \times HCF(x, y)$$

For ease of typing, I will be writing HCF for $HCF(x, y)$ and LCM for $LCM(x, y)$.

We know that HCF is a factor of x and y. That means there are numbers p and q which

$$x = HCF \times p$$

$$y = HCF \times q$$

and that $HCF(p, q) = 1$.

Multiplying both equations, so that we obtain a side with xy:

$$xy = HCF^2 \times pq$$

$$xy = HCF \times pq \times \underbrace{HCF}_{\text{we like this}}$$

Notice that we have a bit of where we want to get, the only thing ruining is that we do not have LCM on the right side, instead we have $HCF \times pq$. Let us show this is equal to LCM:

$$\begin{aligned} LCM = LCM(x, y) = \\ & LCM(p \times HCF, q \times HCF) \\ & HCF \times LCM(p, q) && \text{1st cute fact above} \\ & HCF \times pq && \text{2nd cute fact above} \end{aligned}$$

Thus we can substitute $HCF \times pq$ by LCM in $xy = HCF \times pq \times HCF$ and obtain

$$xy = LCM \times HCF$$

3. Fractions

3.1. Why learn fractions?

Historically, you can think fractions are very "natural": if you have $1, 2, 3, 4, \ldots$, and you can use those numbers to *measure*, why not measure *something* which is bigger than 2 but not as big as 3? That is the context that fractions appeared. You also may not believe this, but it's much easier to operate with fractions than decimals when you don't have a calculator. In that way, fractions are the lazy way to do calculations when we don't have a calculator (think tests!).

I really want you to ditch your probably prejudice towards fractions: they are helpful and they are not hard to understand. That's why this chapter is divided into two parts. The first tells you the *how*, the algorithms about fractions. The second, much more important, tells you the *why*. It's great for you to understand why the algorithms (the methods) work, as it will help you in mathematics in general. Not only that, should you forget the algorithm, if you understand the idea, you can always figure the method on your own!

3.2. Names

In this book, we will consider a **fraction** to be a number composed of two parts:

$$\frac{1}{2}$$

The top number will be called the *numerator* and the bottom number *denominator*. The numerator can be any integers, and the denominator any natural number (to exclude 0!).

If the numerator of a fraction is bigger than its denominator, the fraction is called *improper*. Such as

$$\frac{23}{4}$$

An improper fraction represents a quantity of "wholes" and a fraction. That's why they are called improper, as fractions, originally, represented only quantities smaller than the "whole" used.

There are also *mixed numbers*, which represent a quantity of whole numbers and a fraction, such as

$$5\frac{3}{4}$$

this means there are 5 "wholes" and 3 quarters together.

3.3. The algorithms

3.3.1. A "fake one": whole numbers are fractions

Sometimes we will need to operate with a mix of integers and fractions. It is important to remember that an integer *is* also a fraction, it just has 1 as its denominator:

$$2 = \frac{2}{1}$$

or

$$-5 = \frac{-5}{1}$$

and this sometimes will help.

3.3.2. Converting improper fractions to mixed numbers and the way back

3.3.2.1. From mixed number to improper fraction

The algorithm to convert a mixed number to an improper fraction is the following:

1. Multiply the denominator of the mixed number by the amount of wholes ("multiply the bottom by the big number in front");

2. Add the result of step 1 to the numerator;

3. The result of step 2 is the numerator of the improper fraction;

4. Copy the denominator of the mixed number to the improper fraction.

Easier to understand with an example:

$$3\frac{2}{5} = 3\frac{2}{5} = \frac{17}{5}$$

3.3.2.2. From improper fraction to mixed number

To convert an improper fraction to a mixed number you:

1. Divide the numerator by the denominator like you did before you knew decimals existed: you see "how many of the denominator fit into the numerator". You are going to obtain the result (how may fit) and the remainder;

2. The result you obtained in step 1 is the amount of wholes you get in the mixed number;

3. The remainder you obtained in step1 is the numerator of the fractional part;

4. The denominator is the same as the improper fraction.

Again, an example is useful. Let's say we want to convert

$$\frac{34}{3}$$

to a mixed number. Following the steps above:

1. Dividing 34 by 3: $34 \div 3 = 11R1$ (we can fit 11 threes into 34, and there is 1 remaining);

2. We know there will be 11 wholes in the mixed number;

3. The numerator of the fraction part is 1;

4. The denominator is the same as the original fraction, therefore 3.

Joining all together:

$$\frac{34}{3} = 11\frac{1}{3}$$

3.3.3. Equivalent fractions

The most fundamental idea when dealing with fractions, at least in the operational sense, is to obtain equivalent fractions to a given fraction. The basic idea is very simple, and to understand it I am sure you have seen one of those geometrical representation of fractions, such as squares divided into pieces. You can also see that using Lego bricks, which you can google to see many examples of[1]

Operationally, to obtain equivalent fractions we simply either multiply both the numerator and denominator by the same number, or divide both by the same number. Examples:

$$\overset{\times 4}{\frac{1}{2} = \frac{4}{8}}\atop{\times 4}$$

So, as you can see, we multiplied both the numerator and denominator of the original fraction by 4. Therefore, $\frac{1}{2}$ and $\frac{4}{8}$ are equivalent fractions, as you can obtain one from the other by multiplying both the numerator and the denominator by the same number. Of course, you can go the other way:

$$\overset{\div 4}{\frac{4}{8} = \frac{1}{2}}\atop{\div 4}$$

Let's us see some more examples:

$$\frac{4}{7} = \frac{8}{14}$$

[1] Such as `https://www.youtube.com/watch?v=ILUJdSsT32c`.

$$\overset{\div 5}{\frac{5}{15} = \frac{1}{3}}\atop{\div 5}$$

3.3.4. Negative fractions

A very useful procedure you can do to negative fractions is changing where the negative is:

$$-\frac{2}{3} = \frac{-2}{3} = \frac{2}{-3}$$

You can even add a positive sign to help sometimes:

$$-\frac{2}{3} = +\frac{-2}{3} = +\frac{2}{-3}$$

3.3.5. Adding and subtracting fractions

3.3.5.1. When the fractions have the same denominator

To add or subtract fractions with the same denominator is very simple: **copy the denominator and add or subtract the numerators**. Examples:

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

As you can see, the denominator stays the same and you just do the operation on the numerators. Same thing for subtraction:

$$\frac{6}{7} - \frac{4}{7} = \frac{6-4}{7} = \frac{2}{7}$$

If you have more than two fractions you do exactly the same:

$$\frac{1}{9} + \frac{2}{9} + \frac{4}{9} = \frac{1+2+4}{9} = \frac{7}{9}$$

And if you have a mix of addition and subtraction you also do the same:

$$\frac{1}{8} + \frac{5}{8} - \frac{3}{8} + \frac{7}{8} - \frac{4}{8} = \frac{1+5-3+7-4}{8} = \overset{\div 2}{\frac{6}{8} = \frac{3}{4}}\atop{\div 2}$$

By the way, always simplify your results. You prefer simple things, don't you?

×2

×2

3.3.5.2. When the fractions do not have the same denominator

When the fractions we're adding or subtracting do not have the same denominator, we use the classical mathematics technique of changing the problem to something we know how to solve. In this case, we change the fractions to equivalent ones, but that have a common denominator. This common denominator can be any common multiple of the denominators of the fractions.

A very simple way to obtain such a multiple is by multiplying all the denominators of the fractions and proceed accordingly:

1. Multiply all the denominators of the fractions;
2. The number obtained in step 1 is going to be the denominator of the equivalent fractions;
3. Do the same operation you did on the denominator of each fraction to the numerator;
4. Add or subtract the fractions, which now have the same denominator.

Let's see an example:

$$\frac{1}{2} + \frac{2}{3}$$

The first thing is to multiply the denominators: $2 \times 3 = 6$. Therefore, you know that both fractions will be converted to equivalent fractions with 6 on the denominator:

$$\frac{1}{2} + \frac{2}{3} = \frac{?}{6} + \frac{?}{6}$$

To obtain the new numerators, you just have to apply the same operation you did on the denominator of each fraction to the numerator. For instance, for $\frac{1}{2}$, we have to multiply the denominator by 3 to obtain 6. Therefore, we also have to multiply the numerator by 3:

$$\underset{\times 3}{\overset{\times 3}{\frac{1}{2} = \frac{3}{6}}}$$

We repeat the process with $\frac{2}{3}$. To make the denominator become a 6 we have to multiply it by 2, and we do the same operation on the numerator:

$$\underset{\times 2}{\overset{\times 2}{\frac{2}{3} = \frac{4}{6}}}$$

You can just write this in the same line:

$$\frac{1}{2} + \frac{2}{3} = \frac{3}{6} + \frac{4}{6}$$

Now we just add as we know how: keep the denominators and add the numerators:

$$\frac{1}{2} + \frac{2}{3} = \frac{3}{6} + \frac{4}{6} = \frac{3+4}{6} = \frac{7}{6}$$

This method also works for more than 3 fractions:

$$\frac{3}{4} + \frac{5}{7} + \frac{7}{10}$$

We start by multiplying the denominators: $4 \times 7 \times 10 = 280$. Now we obtain the equivalent fractions:

$$\frac{3}{4} \overset{\times 70}{=} \frac{210}{280} \quad (\times 70)$$

$$\frac{5}{7} \overset{\times 40}{=} \frac{200}{280} \quad (\times 40)$$

$$\frac{7}{10} \overset{\times 28}{=} \frac{196}{280} \quad (\times 28)$$

One line version:

$$\frac{3}{4} + \frac{5}{7} + \frac{7}{10} = \frac{210}{280} + \frac{200}{280} + \frac{196}{280} \quad (\times 70 \quad \times 40 \quad \times 28)$$

Finally, we can just add and simplify the result:

$$\frac{3}{4} + \frac{5}{7} + \frac{7}{10} = \frac{210}{280} + \frac{200}{280} + \frac{196}{280} = \frac{210+200+196}{280} = \frac{606}{280} \overset{\div 2}{=} \frac{303}{140}$$

This also works when you have subtraction:

$$\frac{4}{5} - \frac{2}{3}$$

First obtain the common denominator: $5 \times 3 = 15$. Now, obtain the equivalent fractions:

$$(\times 3 \quad \times 2)$$

$$\frac{4}{5} \overset{\times 3}{=} \frac{12}{15} \quad (\times 3)$$

$$\frac{2}{3} \overset{\times 5}{=} \frac{10}{15} \quad (\times 5)$$

Usual one liner:

$$\frac{4}{5} - \frac{2}{3} = \frac{12}{15} - \frac{10}{15} \quad (\times 3,\ \times 5)$$

We now proceed as we know how:

$$\frac{4}{5} - \frac{2}{3} = \frac{12}{15} - \frac{10}{15} = \frac{12 - 10}{15} = \frac{2}{15}$$

The problem of finding the common multiple by multiplying every denominator is that the number get too big. There is a better strategy: finding the lowest common multiple of the denominators[2]. Let's solve one as an example:

$$\frac{4}{5} - \frac{1}{10} + \frac{1}{2}$$

First, find the LCM of 5, 10, and 2, which is 10 itself. Now obtain the equivalent fractions:

$$\frac{4}{5} \overset{\times 2}{=} \frac{8}{10} \quad (\times 2)$$

$$\frac{1}{2} \overset{\times 5}{=} \frac{5}{10} \quad (\times 5)$$

Note that we don't have to change the second fraction, which is even less work! Now, for the one line version:

$$\frac{4}{5} - \frac{1}{10} + \frac{1}{2} = \frac{8}{10} + \frac{1}{10} + \frac{5}{10} \quad (\times 2,\ \times 5)$$

[2] Please check future reference here.

We can just finish now:

$$\frac{4}{5}-\frac{1}{10}+\frac{1}{2}=\frac{8}{10}-\frac{1}{10}+\frac{5}{10}=\frac{8-1+5}{10}=\frac{12}{10}\overset{\div 2}{\underset{\div 2}{=}}\frac{6}{5}$$

3.3.6. Comparing fractions

3.3.6.1. When the fractions have the same denominator

When comparing fractions in order to say which is the biggest, if they have the same denominators you just compare the numerators: whoever has the biggest numerator is the biggest fraction. For instance, if we wanted to know which is bigger $\frac{4}{5}$ or $\frac{3}{5}$, given that they have the same denominator, we can just compare their numerators. Therefore, $\frac{4}{5}$ is larger than $\frac{3}{5}$.

3.3.6.2. When the fractions do not have the same denominator

When the fractions do not have the same denominator, you must make them! That is, you convert both fractions to equivalent fractions, but that have the same denominator. You can then compare the numerators. Basically, you proceed in the same way as when adding fractions, but stop before! Let's say we wanted to compare $\frac{5}{7}$ with $\frac{7}{9}$. They have different denominators, so we first must find a common denominator. We can multiply 7 and 9 to obtain 63. Now the usual equivalent moment:

$$\frac{5}{7}\overset{\times 9}{\underset{\times 9}{=}}\frac{45}{63}$$

$$\frac{7}{9}\overset{\times 7}{\underset{\times 7}{=}}\frac{49}{63}$$

Now we can just compare $\frac{45}{63}$ with $\frac{49}{63}$. We know that 49 is bigger than 45, so $\frac{49}{63}$ is bigger than $\frac{45}{63}$. Finally, $\frac{7}{9}$ is bigger than $\frac{5}{7}$.

In the same way as when adding, this works for more than two fractions and it is always easier to find the LCM of the fractions instead of multiplying their denominators.

3.3.7. Multiplying fractions

Multiplying fractions is very simple: multiply all the numerators to obtain the new numerator, and multiply all the denominators to obtain the new denominator. Let's see an example:

$$\frac{2}{3}\times\frac{4}{5}=\frac{8}{15}$$

There is a "trick", though, to make the calculations smaller. You can simplify the numbers before doing the multiplication, even if they are in different fractions! For example:

$$\frac{3}{8}\times\frac{16}{19}$$

Notice that you have 16 on the numerator of the second fraction and 8 on the denominator of the first. You can divide both these numbers by 8:

$$\frac{3}{\cancelto{1}{8}}\times\frac{\cancelto{2}{16}}{19}=\frac{3}{1}\times\frac{2}{19}=\frac{6}{19}$$

Much better than doing 8×19, right? Another example:

$$\frac{3}{4}\times\frac{28}{15}=\frac{\cancelto{1}{3}}{\cancelto{1}{4}}\times\frac{\cancelto{7}{24}}{\cancelto{5}{15}}=\frac{1}{1}\times\frac{7}{5}=\frac{7}{5}$$

One nice example to consider is

$$2\times\frac{4}{3}$$

For this, remember that 2 can be written as a fraction:

$$2\times\frac{4}{3}=\frac{2}{1}\times\frac{4}{3}$$

and not we can use the algorithm, top times top, bottom times bottom:

$$\frac{2}{1}\times\frac{4}{3}=\frac{8}{3}$$

3.3.8. Dividing fractions

To divide fractions you just follow the steps:

1. Copy the first fraction;
2. Change the $\div$ symbol to $\times$;
3. "Flip" the second fraction (that is, the numerator becomes the denominator and vice-versa). This is "flipped" fraction is called the *reciprocal*;
4. Multiply the fractions.

In all, you multiply the first fraction by the reciprocal of the second.

Example:

$$\frac{2}{3} \div \frac{7}{10}$$

change

$$\frac{2}{3} \div \frac{7}{10} = \frac{2}{3} \times \frac{10}{7} = \frac{20}{21}$$

copy flip

One more:

$$\frac{7}{40} \div \frac{21}{10} = \frac{7}{40} \times \frac{10}{21} = \frac{\cancel{7}^{1}}{\cancel{40}^{4}} \times \frac{\cancel{10}^{1}}{\cancel{21}^{3}} = \frac{1}{4} \times \frac{1}{3} = \frac{1}{12}$$

And, to make very clear about the "dangers" or mixing whole numbers with fractions:

$$\frac{3}{7} \div 4$$

remember that $4 = \frac{4}{1}$, and then you we just use the normal algorithms:

$$\frac{3}{7} \div \frac{4}{1} = \frac{3}{7} \times \frac{1}{4} = \frac{3}{28}$$

3.3.9. Calculating a fraction of a number

To find fractions of a number, you simply multiply the fraction by the number. Example: find $\frac{3}{5}$ of 50. Just multiply both:

$$\frac{3}{5} \times 50 = \frac{3}{5} \times \frac{50}{1} = \frac{3}{\cancel{5}^{1}} \times \frac{\cancel{50}^{10}}{1} = \frac{3}{1} \times \frac{10}{1} = 30$$

3.4. Mixing everything

Let us do some more complicated questions. For these, we need to know all the facts above and the order of operations (BIDMAS, PEMDAS, GEMA, whoever you want to call them).

First one, when we have both mixed numbers and fractions:

$$1\frac{2}{5} + \frac{3}{4} \times \frac{1}{2}$$

the first thing I do is converting all mixed numbers to fractions:

$$\frac{7}{5} + \frac{3}{4} \times \frac{1}{2}$$

I do understand that you can operate with mixed numbers, but I just do not like to. Now, we need to the multiplication first:

$$\frac{7}{5}+\frac{3}{4}\times\frac{1}{2}$$

$$\frac{7}{5}+\frac{3}{8}$$

and, finally, we add the two fractions:

$$\frac{7}{5}+\frac{3}{8}=\frac{56}{40}+\frac{15}{40}=\frac{71}{40}$$

which we can, if needed, convert to a mixed number:

$$\frac{71}{40}=1\frac{31}{40}$$

Thus,

$$1\frac{2}{5}+\frac{3}{4}\times\frac{1}{2}=1\frac{31}{40}$$

A more complicated example:

$$\left(2\frac{3}{5}-1\frac{2}{3}\right)\div\left(\frac{2}{3}\times 5\right)$$

$$\left(\frac{13}{5}-\frac{5}{3}\right)\div\left(\frac{2}{3}\times 5\right) \qquad \text{mixed to improper}$$

$$\left(\frac{39}{15}-\frac{25}{15}\right)\div\left(\frac{2}{3}\times 5\right) \qquad \text{1st bracket common bottom}$$

$$\frac{14}{15}\div\left(\frac{2}{3}\times\frac{5}{1}\right) \qquad 5=\frac{5}{1}$$

$$\frac{14}{15}\div\frac{10}{3} \qquad \text{2nd bracket}$$

$$\frac{14}{15}\times\frac{3}{10} \qquad \text{division algorithm}$$

$$\frac{42}{150} \qquad \text{multiplication algorithm}$$

$$\frac{7}{25} \qquad \text{simplify (div by 6)}$$

3.5. Understanding the algorithms

3.5.1. Interpretations of a fraction

Much of the confusion with fractions resides in the fact that we assign different meanings to a fraction depending on the context. There are two interpretations of a fraction which are very important and help us to understand the operations. If you understand these two, then the algorithms actually make sense!

3.5.1.1. Fractions as divisions

The first interpretation of a fraction is to understand them as divisions. In that way, if we say

$$\frac{3}{5}$$

we are just writing "3 divided by 5".

3.5.1.2. Fraction as numbers

This may sound silly, but fractions are just numbers which are made in reference to a *unity*, which your teachers probably refer to as "whole". The easiest way to understand this is to remember those classic geometric representations of fractions.

For instance, let's say the following rectangle is our unit:

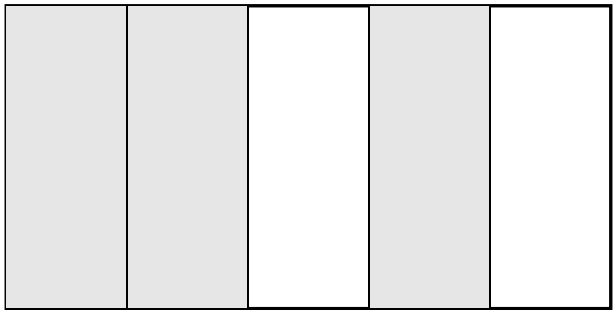
What we mean by that is the rectangle represents 1, or a unit. Now, if we divide the rectangle into 2 equal parts:

the grey part represents half of the unit, that is, we divided that 1 into 2 parts. We represent that by

$$\frac{1}{2}$$

We could also divide that unit into different amounts, for instance 5 equal parts:

The 1, our unit, the rectangle, has now been divided into 5 equal parts, each representing $\frac{1}{5}$ of the original rectangle. The grey part is made of 3 of those $\frac{1}{5}$, which means we have chosen $\frac{3}{5}$ of the unit. Note that we don't have to pick adjacent $\frac{1}{5}$:



this also represents $\frac{3}{5}$.

3.5.2. Converting improper fractions to mixed numbers and the way back explanation

For this one to make sense both interpretations of fractions are useful. Let's start with from improper fraction to a mixed number.

Suppose you have

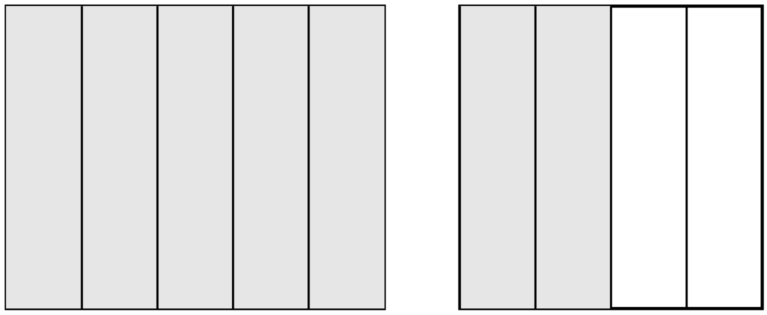
$$\frac{7}{5}$$

Interpreting the fraction as a division, this is the same as $7 \div 5$, which we know, from the past, that is equal to 1 R 2. That means we have 1 "wholes", or units, and 2 extra parts of this whole remaining. Well, we know that 2 parts of a whole of 5 parts is $\frac{2}{5}$, so that explains the

$$\frac{7}{5} = 1\frac{2}{5}$$

as it is just a different way of writing the result of the division. Using this interpretation it is also easy to understand the way back of multiplying the denominator by the amount of wholes and adding to the result, as we are just doing the inverse operation.

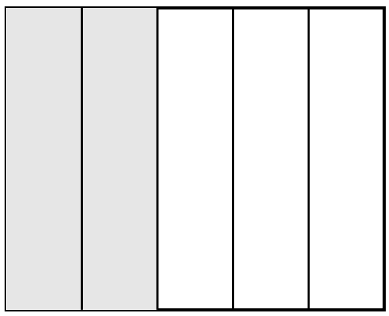
Now, if we look as fractions as the geometric representations, we could see $\frac{7}{5}$ like this:



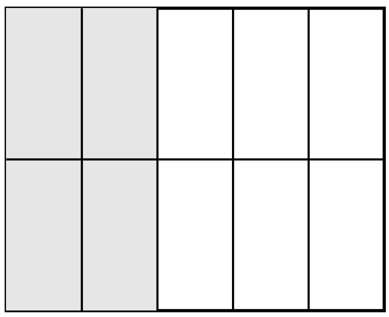
We needed two units, as 1, if divided in 5 gives only 5 pieces (obviously). We then take only 2 parts of the other unit. This is a way to represent $\frac{7}{5}$. As you can see, it is clearly 1 whole and $\frac{2}{5}$. Therefore, the algorithm makes sense: you see how many wholes or units you have, and then get some remaining from a next unit. To go back, you just count how many pieces we shaded. This explains why you multiply the denominator by the big number in front: you are counting how many pieces the wholes have, and then you just add the pieces in the "incomplete" whole.

3.5.3. Equivalent fractions

To understand why equivalent fractions exist we can use the geometric model again. Take $\frac{2}{5}$ as an example:



We could take the same "unit rectangle" and divide it in half:



This second drawing represents $\frac{4}{10}$, given that we have divided the unit into 10 equal parts and picked 4. But as you can see, both $\frac{2}{5}$ and $\frac{4}{10}$ are exactly the same shaded part! This must mean they are equal:

$$\frac{2}{5} = \frac{4}{10}$$

Now, to obtain $\frac{4}{10}$ from $\frac{2}{5}$ we doubled both the total number of parts and the number of parts we shaded. This explains why we must multiply (or divide) both numerator and denominator by the same number to obtain equivalent fractions. The converse reasoning explains the division.

3.5.4. Negative fractions

This one is very simple to understand using the division interpretation. Take $-\frac{3}{4}$ as an example. We can see this as minus the result of 3 divided by 4. This is exactly the same as -3 divided by 4 and the same as 3 divided by -4. That explains the

$$-\frac{3}{4} = \frac{-3}{4} = \frac{3}{-4}$$

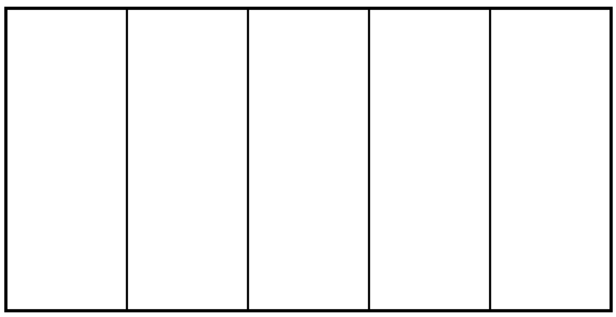
3.5.5. Adding and subtracting fractions

3.5.5.1. When the fractions have the same denominator

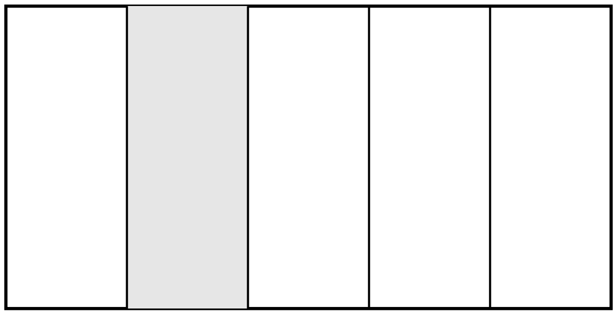
The geometric representation helps again in this case. Let's see how to interpret

$$\frac{1}{5} + \frac{2}{5}$$

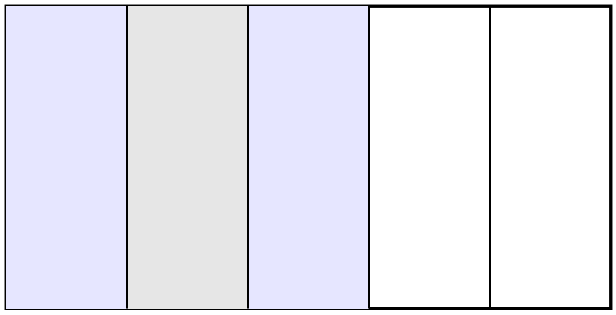
First, let's get one whole and divide it into 5 parts:



now, $\frac{1}{5}$ means "picking" 1 of those 5 parts:



adding $\frac{2}{5}$ means "picking" 2 more of those parts:



now we just count: we have 3 parts picked out of 5, so we have $\frac{3}{5}$. That explains why

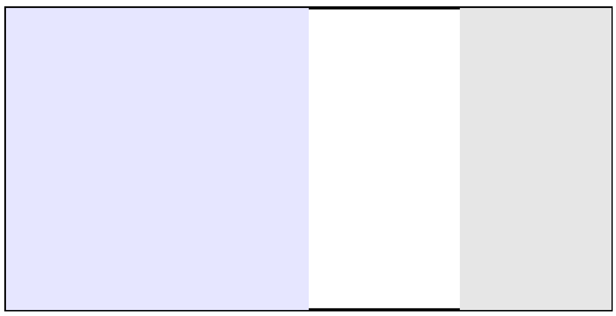
$$\frac{1}{5} + \frac{2}{5} = \frac{3}{5}$$

3.5.5.2. When the fraction do not have the same denominator

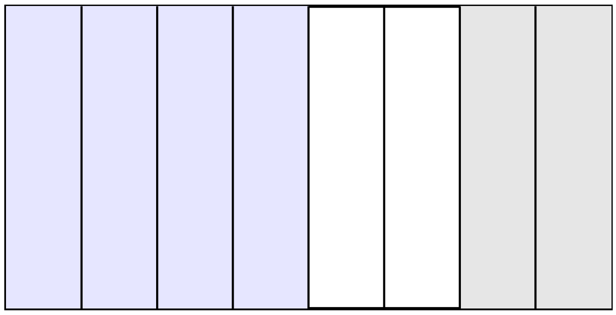
What happens when we have

$$\frac{1}{2} + \frac{1}{4}$$

The idea is the same as in the last case, actually: we want to divide the whole in parts and count them. However, look what happens if you divide one whole into $\frac{1}{2}$ and $\frac{1}{4}$:



as you can see, it's hard to count anything! That's why we change the fractions to have a common denominator, because we can then easily count! We do like easier things, don't we? Let's divide the rectangle in 8 equal parts (as $2 \times 4 = 8$):



in this drawing the same regions are shaded, but now we can literally count how many of the parts we have: 6! Given that we had divided the whole into 8 pieces, we have $\frac{6}{8}$. Therefore:

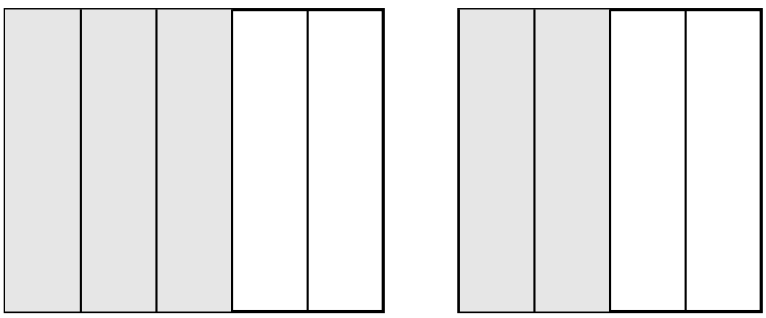
$$\frac{1}{2} + \frac{1}{4} = \frac{6}{8}$$

Clearly you can simplify $\frac{6}{8}$ to $\frac{3}{4}$, but that's not important now: what matters is to understand that we change the fractions to have the same denominator because it is *easier*! I know you probably think this is contradictory, but it is not. It is hard (impossible?) to count how many parts of the whole we have if we are picking different sized parts! That's the idea behind it.

3.5.6. Comparing fractions

3.5.6.1. When the fractions have the same denominator

When comparing fractions with the same denominator, we can just picture them geometrically to "see" which one is bigger. Let's us compare $\frac{3}{5}$ and $\frac{2}{5}$:

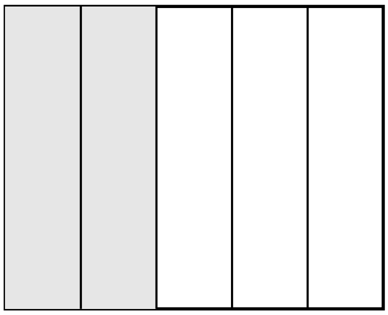


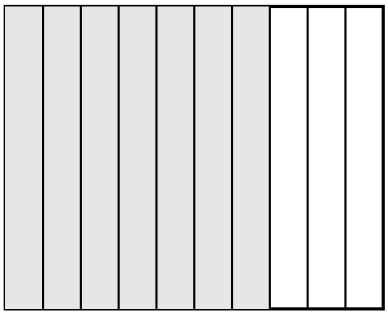
it makes sense that $\frac{3}{5}$ is bigger than $\frac{2}{5}$, doesn't it? You have more space shaded! So we are basically just checking which fraction is bigger by comparing how many parts we picked in each.

3.5.6.2. When the fractions do not have the same denominator

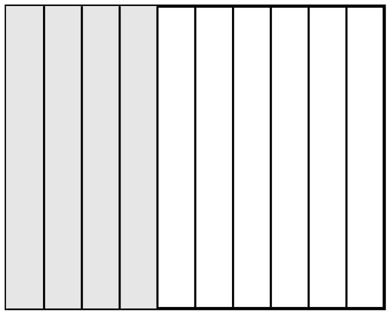
In the same way when we are adding fractions, to compare fractions with different denominators we convert them to equivalent fractions that have the same denominator because it makes our life easier.

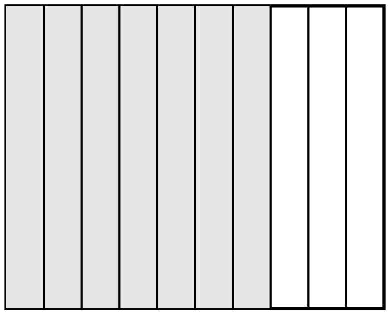
Let's try comparing $\frac{2}{5}$ and $\frac{7}{10}$. If we take a whole of the same size and represent both $\frac{2}{5}$ and $\frac{7}{10}$ we obtain:





it is easy to see that $\frac{7}{10}$ is bigger in "area" than $\frac{2}{5}$, right? But imagine doing drawings like this whenever you wanted to compare two or more fractions? Imagine if they had enormous denominators? It would be, well, boring! If we, however, change them to equivalent fractions with the same denominator, we can just compare numerators, and we do like being lazy, don't we? Therefore, if we just divide each piece of the first whole (which represents $\frac{1}{5}$) into 2 parts:





that's much easier to compare now by just counting! That's the idea behind it.

3.5.7. Multiplying fractions

To understand the "top times top over bottom times bottom" we can just rewrite the fractions as divisions. For instance:

$$\frac{2}{3} \times \frac{5}{7} = (2 \div 3) \times (5 \div 7)$$

Now, remember that multiplication and division have the same priority, so we can rearrange that:

$$(2 \div 3) \times (5 \div 7) = 2 \div 3 \times 5 \div 7 = (2 \times 5) \div 3 \div 7$$

We must now deal with that "divided by 3 divided by 7" part. What does it mean to do one division followed by the other? Let's us think by example. Calculating

$$16 \div 4 \div 2$$

adding some brackets:

$$(16 \div 4) \div 2 = 4 \div 2 = 2$$

The answer is 2. This operation also gives 2 as the answer:

$$16 \div (4 \times 2)$$

Thus, we know that

$$16 \div 4 \div 2 = 16 \div (4 \times 2)$$

Using the same idea on $(2 \times 5) \div 3 \div 7$:

$$(2 \times 5) \div 3 \div 7 = (2 \times 5) \div (3 \times 7)$$

We can know just rewrite the division as a fraction:

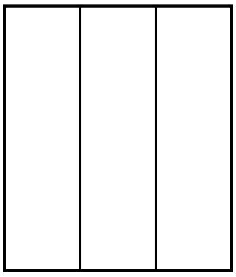
$$(2 \times 5) \div (3 \times 7) = \frac{2 \times 5}{3 \times 7}$$

This explains that

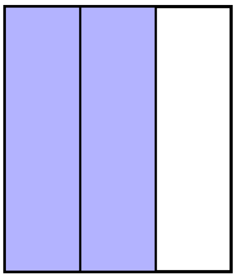
$$\frac{2}{3} \times \frac{5}{7} = \frac{2 \times 5}{3 \times 7}$$

or, more generally, that you simply "multiply tops and divide by the product of the bottoms".

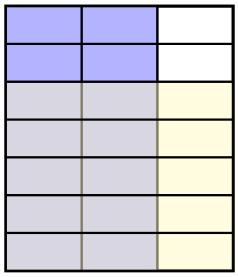
We can also think of this geometrically (if you do not remember how to find areas of rectangles, see Chapter RefHere). Let us take a nice rectangular whole and divide it into thirds:



of those thirds, let us take two, representing $\frac{2}{3}$:



Now, let us split the same whole in sevenths and take 5 parts, representing $\frac{5}{7}$:



Notice, in the drawing, that we have 10 little rectangles which have been shaded both blue *and* yellow. And, in total, we have shaded 21 rectangles. What fraction we have, if we take the number of rectangles that were double shaded as a part, and the total number of shaded rectangles as a whole? You guessed:

$$\frac{10}{21}$$

which is the result of $\frac{2}{3} \times \frac{5}{7}$. Geometrically, you can see that we are dividing a whole into "bottom times bottom" parts, and we are selecting "top times top" parts of them.

3.5.8. Dividing fractions

There are lots of ways to understand why multiplying by the reciprocal works, but I believe my favourite is when you think about it in terms of inverse operations. Let's start with naturals. What does it mean to calculate

$$8 \div 2 =?$$

This means there is a number, let's call it ?, which

$$? \times 2 = 8$$

correct? This is an equation, and you can figure out that ? is 4, which is the result of $8 \div 2$. With fractions is the same. Let's say we wanted to calculate:

$$\frac{5}{7} \div \frac{3}{4} =?$$

This means there is a number, let's again call it ?, which

$$\frac{3}{4} \times ? = \frac{5}{7}$$

Now, to solve this equation, you can remember that multiplying a whole number by a fraction is the same as dividing it by the denominator and multiplying by the numerator. So we are actually doing two operations on the left side: multiplying ? by 3 and dividing it by 4. To solve an equation we do inverse operations on both sides, so we need to multiply both sides by 4 (the inverse of dividing by 4) and divide both sides by 3 (the inverse of multiplying by 3). Therefore we would obtain:

$$\frac{3}{4} \times ? \times 4 \div 3 = \frac{5}{7} \times 4 \div 3$$

The left side will just have the ?, and if we interpret $4 \div 3$ as $\frac{4}{3}$, we have:

$$? = \frac{5}{7} \times \frac{4}{3}$$

There you go, the multiplication by the reciprocal appears!

A less algebraic way to understand this is to the following. Let's start with

$$\frac{4}{5} \div \frac{4}{5}$$

well, that's a number divided by itself. That is always equal to 1:

$$\frac{4}{5} \div \frac{4}{5} = 1$$

What about

$$\frac{4}{5} \div \frac{2}{5} = ?$$

We have 4 parts out of 5, and we want to divide those 4 parts into 2 groups of parts of 5 (sounds complicated, but read it slowly!). This means that each groups has 2 parts:

$$\frac{4}{5} \div \frac{2}{5} = 2 = \frac{2}{1}$$

(Notice the $2 = \frac{2}{1}$ part, it's important!)

And what if we do

$$\frac{4}{5} \div \frac{1}{5} = ?$$

Same reasoning: we have to divide 4 parts (out of 5) into 1 group of parts of 5. That gives 4 parts per groups:

$$\frac{4}{5} \div \frac{1}{5} = 4 = \frac{4}{1}$$

You notice that we are just dividing the numerators and "completely ignoring the denominators" (when they are equal we just divide one by the other and get 1, which doesn't change anything!)

What about

$$\frac{4}{5} \div \frac{3}{5}$$

If we just "ignore" the denominators and divide the numerators:

$$\frac{4}{5} \div \frac{3}{5} = \frac{4}{3}$$

What if we want to do

$$\frac{2}{3} \div \frac{4}{5}$$

If they had the same denominator we could just divide numerators... Let's do it then!

$$\frac{10}{15} \div \frac{12}{15}$$

we can now just divide numerators:

$$\frac{10}{15} \div \frac{12}{15} = \frac{10}{12}$$

You could, however, add some intermediate steps:

$$\begin{aligned}
&\frac{2}{3} \div \frac{4}{5} && \\
&\frac{2 \times 5}{3 \times 5} \div \frac{4 \times 3}{5 \times 3} && \text{Equivalent fractions with the same denominator} \\
&\frac{2 \times 5}{4 \times 3} && \text{We can just divide the numerators!} \\
&\frac{2 \times 5}{3 \times 4} && \text{Multiplication is commutatitve} \\
&\frac{2}{3} \times \frac{5}{4} && \text{Split the division into two fractions}
\end{aligned}$$

There we go! In all, the "multiply by a reciprocal" is just a shortcut, because it would be very annoying to keep changing the denominators!

3.5.9. Calculating a fraction of a number

Remember that to do this you simply "divide the number by the denominator and multiply the result by the numerator". This is easy to understand when we think of fractions as parts of a whole, but with a twist. Let's think about calculating

$$\frac{2}{5} \text{ of } 15$$

As we understand fractions as parts of a whole, this means: if we take a whole that has "size" 15 and divide it into 5 groups, how many "things" will there be in two groups? So, if we take 15 "things" and divide it into 5 groups, each group has 3 "things". Finally, in two groups we have $2 \times 3 = 6$ "things".

3.6. Exam hints

Nothing big to say here, except for you to remember how to operate with the fractions!

Summary

- A **fraction** is a number of the form $\frac{a}{b}$ where a is any integer and b is any integer except 0. The top of the fraction is called the **numerator** and the bottom is called the **denominator;**
- When the numerator of a fraction is bigger than its denominator, we call the fraction **improper**;
- A **mixed number** is a fraction associated with a quantity of "wholes", and we write them like $2\frac{2}{5}$;
- To convert a mixed number to an improper fraction we multiply the denominator by the amount of "wholes", add this result to the numerator. This is the numerator of the improper fraction, and the denominator is the same as the fractional part of the mixed number;
- To convert an improper fraction to a mixed number, we divide the numerator by the denominator, but in the "old way": how many of the denominator fit into the numerator? The result is the number in front of the mixed number, the remainder is the numerator of the fractional part and the denominator is again the same;
- To obtain **equivalent fractions** to a given fraction, we can either *multiply* or *divide* both the numerator and the denominator of the fraction by the same number;
- These are all equivalent
$$-\frac{a}{b} = \frac{-a}{b} = \frac{a}{-b}$$
- To add or subtract fractions that *have the same denominator*, we keep the denominator the same and add or subtract the numerators;
- To add or subtract fractions that *do not have the same denominator*, we first convert each fraction to an equivalent fraction that have the same denominator, and then add or subtract them;
- To compare fractions that *have the same denominator* we just compare numerators: whoever has the biggest numerator is the largest fraction;
- To compare fractions that *do not have the same denominator* we again convert them to equivalent fractions that have the same denominator, and then compare their numerators;
- To multiply fractions we just multiply the numerators to obtain the new numerator, and multiply the denominators to obtain the new denominator;
- To divide fractions, we multiply the first fraction by the reciprocal (the "flipped") of the second;

- To calculate a fraction of a number, we divide the number by the denominator of the fraction and multiply the result by the numerator.

4. Rounding

4.1. Why learn rounding and estimation

Rounding is the art of disregarding information that is too precise. Sometimes you do not care that there were 9898 people in a stadium, 10000 conveys the same information. Paying 29.99 is the same as paying 30 to many of us. Even in science, sometimes the value of the 8th decimal place is not important to your calculations. Rounding is simply a way to "ignore" what we do not want to know in a consistent way.

Estimation is a way to do calculations and get answers which are close to the correct, but using simpler numbers. Very useful when we do not have a calculator!

4.2. Place value

The basic notion we want to remember when rounding is that of place value. Place value is the idea that each digit of a number represents a different quantity in the number. For instance, in the number

$$1234$$

the 4 represents 4 units, the 3 represents $3 \times 10 = 30$, the 2 represents $2 \times 100 = 200$ and the 1 represents $1 \times 1000 = 1000$. In all, 1234 is just lazy notation to

$$1234 = 1 \times 1000 + 2 \times 100 + 3 \times 10 + 4 \times 1$$

As we use base 10 (we have digits 0 to 9 in each position of our number), we multiply by powers of 10. If you study computer science, you will learn binary, in which we only use the digits 0 and 1, and we multiply those by powers of 2. The idea of place value is the same, though: each digit represents a number of a power of 2.

Hence, all numbers are written using this idea. Each of the positions a digit can go has a name. For instance, in 54326.9785:

Ten thousands	Thousands	Hundreds	Tens	Ones		Tenths	Hundredths	Thousandths	Ten thousandths
5	4	3	2	6	·	9	7	8	5

Normally, instead of using the name "hundredths", we would say the third decimal place, and we would say the second decimal place for the tenths value.

4.3. Rounding general idea: the closest number we care about

When we are rounding, the most important thing is the precision: how many digits of the number we care about. Knowing that precision, we can round properly.

Say we have the number 1.7. We want to round it to nearest integer (which is the same to the nearest whole number). There are two obvious possibilities: 1, the first integer which is smaller than 1.7, and 2, the first integer which is bigger than 1.7. Let us see who is closer to 1.7:

$$1.7 - 1 = 0.7$$

$$2 - 1.7 = 0.3$$

Hence, the distance of 1.7 to 2 is 0.3, smaller than the distance from 1, which is 0.7. Thus, when rounding 1.7 to the nearest integer, we round it to 2.

Of course, this is boring, so let us see how to do this fast. However, keep the idea in mind: rounding is simply getting the closest number to a given one in a certain precision.

4.4. Rounding basic idea

To round any number, the first thing we do is identify which digit of our number is in the precision we want. For instance, if we were rounding to the second decimal place the number 1234.678, we would first identify 7 as the number in that position. After identifying the digit in our precision, we simply look at the number to its right and do the following:

- if the number to the right of our precision is $0, 1, 2, 3$ or 4, we simply copy our number until our precision, and after it, change every digit to 0;
- if the number to the right of our precision if $5, 6, 7, 8$ or 9, we copy our number until the digit to the left of our precision, add 1 to the number in our precision, and change everything to its right to 0.

The next sections will have some examples.

4.5. Rounding to a given place

Using the number

$$1234.3567$$

let us round it to:

- nearest whole number:

  First we identify the last digit we care about: as we only care for whole numbers, we will focus on the 4 before the decimal point:

  $$123\underset{\uparrow}{4}.3567$$

Now, we look at the digit to the right of our 4: as it is a 3, we simply copy our number until the 4 and do not do anything to it. To the right of it, we change every digit to 0:

$$\underbrace{1234}_{\text{copy}}.\underbrace{0000}_{\text{0 it}}$$

Of course, we could simply write 1234, as the 0s to the right of the decimal point do not change the number;

- to the second decimal place (2 d.p.):

  Again, first we identify the last digit we care about: the 5 at the second decimal place:

  $$1234.3\underset{\uparrow}{5}67$$

  Then we look at the digit to its right: as we have a 6, we copy all digits until our 5, then add 1 to the 5. Everything else becomes a 0:

  $$\underbrace{1234.3}_{\text{copy}}\overbrace{6}^{5+1}\underbrace{00}_{\text{0 it}} = 1234.36$$

  in which I already did not write the zeroes after the last digit as they are after the decimal point.

- nearest 100:

  As usual, we first identify our digit. The number in the hundreds place is the first 3 from the left:

  $$12\underset{\uparrow}{3}4.3567$$

  now we look to the number to its right. As it is a 4, we simply copy everything until our digit, and don't change it. Everything to the right of the 3 becomes a 0:

  $$\underbrace{123}_{\text{copy}}\underbrace{0.0000}_{\text{0 it}} = 1230$$

  It is very important to notice that the 4 becomes a 0, and we *have to write it*, as it is before the decimal point. A classic mistake would be writing 123, but 123 is *very* far from 1234.3567!

- nearest 1000:

  Same idea: the 1 is in the thousands place:

  $$\underset{\uparrow}{1}234.3567$$

  Looking to the number to its right, 2, we just copy everything until our 1 and don't change the 1 itself. Everything else to the right becomes a 0:

  $$\underbrace{1}_{\text{copy}}\underbrace{000.0000}_{\text{0 it}} = 1000$$

Be careful with the zeroes until the decimal point!

Now the number

$$0.9098$$

- to 1 decimal place (1 d.p.):

  Identify the number on the first decimal place:

  $$0.\underset{\uparrow}{9}098$$

  looking to its right, we have a 0, so we copy everything until the 9, do not change it, and every number to the right of 9 becomes a 0:

  $$0.9000 = 0.9$$

- to 3 decimal places (3 d.p.):

  Same drill:

  $$0.90\underset{\uparrow}{9}8$$

  To the right of our 9 we have a 8, so we need to add a 1 to the 9. However, adding 1 to 9 makes it a 10, so we carry it (like when doing addition) to the decimal place to the left of the 9:

  $$0.9\underbrace{1}_{0+1}\overbrace{0}^{9+1}0 = 0.91$$

  This is the only case you have to be careful when rounding.

4.6. Rounding to a number of significant figures

Rounding to a certain number to significant figures is exactly the same, we just need to know how to identify significant figures. The rule is

- From the left of the number, find the first digit which is not 0: this is the first significant figure;
- After the first significant figure, every digit is significant (including possible zeroes).

After identifying the correct number of significant figures, rounding is the same thing.

For instance, with the number

$$0.03809$$

- rounding it to 2 significant figures (2 s.f.):

  The first significant figure is the first non-zero digit from the left, in this case the 3. The second significant figure is the number to the right of the first, so the 8:

  $$0.03\underset{\uparrow}{8}09$$

now we look to the digit to the right of the one we care about, which has a 0. Hence we just copy:

$$\underbrace{0.038}_{\text{copy}}\underbrace{00}_{\text{0 it}} = 0.038$$

- rounding to 3 significant figures (3 s.f.):

  From the left, the first non-zero digit is the 3, after it everything is significant. The third significant figure is, then, the 0 between the 8 and the 9:

$$\underset{\uparrow}{0.03809}$$

  as to its right we have a 9, we copy everything and add one to our 0:

$$\underbrace{0.038}_{\text{copy}}\overbrace{1}^{0+1}0 = 0.0381$$

With the number

$$238$$

- rounding it to 1 significant figure (1 s.f.):

  The first significant figure is the 2:

$$\underset{\uparrow}{2}38$$

  and to its right we have a 3. Hence, we do not change anything until our 2, and everything to its right becomes a 0:

$$\underbrace{2}_{\text{copy}}\underbrace{00}_{\text{0 it}} = 200$$

  Again, be careful not to forget the zeroes before the decimal point.

- rounding it to 2 significant figures (2 s.f.):

  The second significant figure is 3:

$$2\underset{\uparrow}{3}8$$

  and to its right we have an 8. Hence, we copy everything to its left and add 1 to it. To its right, everything becomes a 0:

$$\underbrace{2}_{\text{copy}}\overbrace{4}^{3+1}\underbrace{0}_{\text{0 it}} = 240$$

4.7. Estimation

When we estimate the result of a calculation, we are trying to find a result that is close enough to the real answer.

For instance, if you had to pay \$4.90 for each of your 9 sweets you bought, you could quickly estimate you would spend around $5 \times 10 = 50$. The real value is less, of course, but it gives you information nevertheless.

In the IGCSE exam, there is a general rule to estimate the result of an operation: round every number to one significant figure, *then* do the operation.

For instance, to estimate

$$23.3 \times 4.8$$

you would first round 23.3 to one significant figure, 20. Then you would round 4.8 to one significant figure, 5. Then you do the operation:

$$23.3 \times 4.8 \approx 20 \times 5 = 100$$

Thus, you would estimate the result to be 100. Here I used the symbol $\approx$, which means "approximately". Hence, we can read that 23.3×4.8 is approximately 100.

A "harder" example:

$$\frac{12.3 + 24.82 \times 8.84}{3.2^2}$$

First round every number to 1 significant figure:

$$12.3 \xrightarrow{\text{1 s.f..}} 10$$

$$24.82 \xrightarrow{\text{1 s.f.}} 20$$

$$8.84 \xrightarrow{\text{1 s.f.}} 10$$

$$3.2 \xrightarrow{\text{1 s.f.}} 3$$

$$2 \xrightarrow{\text{1 s.f.}} 2$$

and then do the actual calculation:

$$\frac{12.3 + 24.82 \times 8.84}{3.2^2} \approx \frac{10 + 20 \times 10}{3^2} = \frac{10 + 200}{9} = \frac{210}{9} = 23.333\ldots = 23.3 \text{ (3s.f.)}$$

4.8. Exam hints

Not much to remember here. Just be careful with not forgetting zeroes before the decimal point and that after the first significant figure everything is significant!

The most important thing about rounding is that: if your answer to a question is not exact, you *always have to write at least* 3 *significant figures* as its answer. Unless you are dealing with *angles*, then you should use one decimal place (yes, I know, confusing).

Also remember that to estimate, we always round every number to one significant figure *before* doing the calculation (they usually remind you of it in the paper, though).

Summary

- *Place value* is an important notion in math: each digit in a number represents a different quantity of powers of a certain number, the base. For instance:

$$1234 = 1 \times 10^3 + 2 \times 10^2 + 3 \times 10^1 + 4 \times 10^0$$

- To round a number to a given precision:
    1. Identify the digit at the precision demanded;
    2. If the number to the right is $0, 1, 2, 3$ or 4, copy everything until the digit, and everything to its right becomes a 0;
    3. If the number to the right is $5, 6, 7, 8$ or 9, copy everything until the number to the left of the digit, add 1 to the digit, and everything to its right becomes a 0.

- To estimate the result of an operation, first round every number to *one significant figure*. Then calculate normally.

5. Ratio and proportion

5.1. Why learn ratios

I have to say that ratios are somewhat hard to justify. The one thing I can think of is C-3PO line in *Star Wars*:

> "Sir, the possibility of successfully navigating an asteroid field is approximately three thousand seven hundred and twenty to one!"

C-3PO said, then, the ratio of not surviving the asteroid field to surviving it was:

$$3720 : 1$$

In the end, ratios may not be that exciting, but they do appear in *Star Wars*, so we have a point for them!

5.2. What is a ratio and an important warning

A ratio is simply a way to give information on how a number of objects is divided. For instance, a class may have 3 boys and 5 girls, so we can say that the ratio of boys to girls is

$$\text{boys} : \text{girls}$$

$$3 : 5$$

We read this as "the ratio of boys to girls is 3 to 5". Basically the colon is read as a "to".

What this means is that for every 3 boys in the class, we have 5 girls.

It is important, however, to understand that we could have a class with 6 boys and 10 girls, and the ratio of boys to girls would be the same. This will get clearer in the next section, as we can simplify the ratios to be the same. Actually, the ratio $3 : 5$ of boys to girls means that we can have any multiple of both sides of boys to girls, and we can represent that as a "fancy" ratio using a little bit of algebra:

$$3x : 5x$$

where x can be any positive integer. This is a mathematical way of saying that the number of boys to girls in a class will always follow the same ratio, but can be any of

$$3 : 5$$

$$6 : 10$$

$$9 : 15$$

and so on. This is important as a classic mistake is something like this: a class has boys to girls in a ratio of 5 : 8, but this does not mean we have 13 people in the class! We could, of course, but we could also have 10 boys and 16 girls, or 15 boys and 24 girls.

Another relevant aspect of ratio is that we can divide our objects in as many categories as we want. For instance, let us say we have red, blue, green and yellow marbles. The ratio of red to blue to green to yellow could be:

$$1 : 3 : 2 : 2$$

which means that for every 1 red marble, we have 3 blue ones, 2 green and 2 yellow. But remember this does not mean we have $1 + 3 + 2 + 2 = 8$ marbles, as we could have 2 red, 6 blue, 4 green and 4 yellow, as they would still be in the same ratio.

And here is the warning. A classic mistake involving ratios is around these lines: say the ratio of A to B is 1 : 3. What is the **fraction** of A of the total? The mistake is saying $\frac{1}{3}$, but this is wrong! Remember that 1 : 3 means that for every 1 A we have 3 B, that is, we have "little groups" of 4 objects! Hence, the fraction of A of the total is $\frac{1}{4}$. This is important because, as we will see, ratios and fractions have similarities, but they are not the same. Mathematicians are lazy, we would not have different entities to represent the same object.

5.3. Equivalent ratios - simplifying

Much like fractions in this regard (but remember the difference above), we can obtain equivalent ratios: they represent the same groupings of things, but with few or more elements in each group.

5.3.1. Simplest form

Given a ratio, it is in its simplest form if the highest common factor of all the numbers in the ratio is 1 (exactly like a fraction). For instance, given

$$4 : 10$$

the HCF of 4 and 10 is 2, hence this ratio is not in its simplest form. We can divide both numbers by 2:

$$4 : 10$$

$$2 : 5$$

and we obtain an equivalent ratio to 4 : 10, but in its simplest form. Some people can write this as

$$4 : 10 = 2 : 5$$

but it is not common in the IGCSE.

Of course, you can do this in multiple steps, not necessarily dividing by the HCF. Given

$$30 : 42$$

we can divide first by 2:

$$\div 2 \quad \begin{matrix} 30 : 42 \\ 15 : 21 \end{matrix} \quad \div 2$$

obtaining the first equivalent ratio to 30 : 42. Now, we can take 15 : 21 and further divide it by 3:

$$\div 3 \quad \begin{matrix} 15 : 21 \\ 5 : 7 \end{matrix} \quad \div 3$$

to obtain the simplest form of 30 : 42.

The same technique goes for ratios with more than 2 parts:

$$\div 5 \quad \begin{matrix} 10 : 15 : 20 \\ \div 5 \\ 2 : 3 : 4 \end{matrix} \quad \div 5$$

Just notice you have to divide every single number in the ratio.

If you have a ratio with decimal numbers, such as

$$0.6 : 0.18$$

the first thing is to obtain an equivalent ratio with only integers. To do this, you can keep multiplying all the numbers by 10 until you obtain only whole numbers:

$$\begin{matrix} 0.6 : 0.18 \\ 6 : 1.8 \\ \div 2 \quad 60 : 18 \quad \div 2 \end{matrix}$$

(of course you could multiply by 100 and do it in a single step as well). Now, we just simplify the ratio normally:

$$\div 6 \;\; \begin{array}{c} 60 : 18 \\ \downarrow \\ 10 : 3 \end{array} \;\; \div 6$$

Why not have ratios with fractions as well? Let us say we have

$$\frac{3}{5} : \frac{7}{15}$$

let us get rid of the denominators by multiplying both numbers by 15, the LCM of 5 and 15:

$$\times 15 \;\; \begin{array}{c} \frac{3}{5} : \frac{7}{15} \\ \downarrow \\ 9 : 7 \end{array} \;\; \times 15$$

as the HCF of 9 and 7 is already 1, we are done, but we could have to simplify the ratio normally after it.

5.3.2. Ratios with units

When a ratio has numbers with units, first convert every number to the same unit, then proceed normally. For instance, given

$$50\text{cm} : 1\text{m}$$

To put it in its simplest form, first convert both to centimetres:

$$50\text{cm} : 100\text{cm}$$

and then divide both numbers by 50:

$$\div 50 \;\; \begin{array}{c} 50\text{cm} : 100\text{cm} \\ \downarrow \\ 1\text{cm} : 2\text{cm} \end{array} \;\; \div 50$$

Another example:

$$\times 10 \;\; 0.4\text{kg} : 300\text{g} \;\; \times 10$$

We begin by writing 0.4kg in grams:

$$\times 10 \;\; \downarrow \;\; \times 10$$

$$400\text{g} : 300\text{g}$$

and now we divide both numbers by 100:

$$\div 100 \quad \begin{matrix} 400\text{g} : 300\text{g} \\ 4\text{g} : \ \ 3\text{g} \end{matrix} \quad \div 100$$

5.3.3. Equivalent ratios cross-multiplication

One interesting thing to notice when it comes to equivalent ratio is that, if given two equivalent ratios, such as

$$3 : 5$$

$$9 : 15$$

which are equivalent as you can multiply the top one by 3 to obtain the second, we can "cross-multiply" the numbers in the ratios:

$$\begin{matrix} 3 : 5 \\ 9 : 15 \end{matrix}$$

and we obtain $3 \times 15 = 45$ when we multiply the top left number by the bottom right number, and we obtain $5 \times 9 = 45$. This always happens, and we can use it to solve problems. Using the = notation for equivalent ratios:

$$2 : 7 = 6 : 21$$

we can multiply 2 by 21, and we get 42, which is the same result as 7×6.

5.3.4. $1 : n$ and friends

A common type of question is to put a ratio in the form $1 : n$ or $n : 1$. For instance, given

$$3 : 12$$

we want to put this ratio in the form $1 : n$. That means we would like to start with $3 : 12$ and get to an equivalent ratio with 1 on the left:

$$3 : 12$$

$$1 : \ ?$$

to discover what goes on the right, we can either notice that we divided the left number by 3, so we should do the same to the right number:

$$\div 3 \quad 3 : \quad 12 \quad \div 3$$
$$1 : ? = 4$$

which gives the answer, 4. We can also use the cross-multiplication property and set up an equation:

$$3 : 12$$

$$1 : x$$

as we know that

$$3 \times x = 12 \times 1$$

$$3x = 12$$

$$x = 4$$

The answer is

$$1 : 4$$

The form $n : 1$ is exactly the same. Say we have

$$8 : 5$$

and we would like to have it as

$$? : 1$$

We can use the division method:

$$\div 5 \quad 8 : 5 \quad \div 5$$
$$? = \tfrac{8}{5} = 1.6 : 1$$

or we can use the cross-multiplication:

$$8 : 5$$

$$x : 1$$

cross-multiplying we obtain

$$8 \times 1 = 5 \times x$$

$$8 = 5x$$

$$\frac{8}{5} = 1.6 = x$$

The final answer is, then,

$$1.6 : 1$$

Notice that we could have a question asking us to put a ratio in the form $3 : n$ (or any other number to n), and the same techniques could be used. Say we would like

$$9 : 21$$

to be in the form $3 : n$:

$$9 : 21$$

$$3 : \ ?$$

we could use the division method:

$$\div 3 \quad \begin{matrix} 9 : & 21 \\ 3 : ? = & 7 \end{matrix} \quad \div 3$$

or the cross-multiplication:

$$9 : 21$$

$$3 : x$$

when we cross-multiply:

$$9 \times x = 21 \times 3$$

$$9x = 63$$

$$x = 7$$

Thus the final answer is

$$3 : 7$$

5.4. Ratio problems

5.4.1. Dividing a quantity into a ratio

A classic question using ratios is to divide a quantity into a ratio. Say that you have to divide \$100 between Alice and Bob. If you were to divide the money equally, you would divide the \$100 by 2, and give both Alice and B \$50. Let's divide this into steps:

1. We are dividing the money it into 2 parts;
2. We take the total, \$100, and divide it by the number of parts, and we obtain \$50;
3. We give \$50 to each Alice and Bob

We can say that we divided the money into a ratio of $1 : 1$, as both Alice and Bob got the same amount. What if we wanted to divide the same \$100, but into the ratio $3 : 2$? We can follow the same steps as above:

1. We are dividing the money into 5 parts, as Alice will get 3 of them and Bob 2 $(2 + 3 = 5)$;
2. We take the total, \$100, and divide it by the number of parts, in this case 5: $100 \div 5 = 20$;
3. We know that each part is \$20, but Alice gets 3 parts, so she obtains $3 \times 20 = 60$, and Bob obtains $2 \times 20 = 40$, as he gets 2 parts.

Thus, Alice will get \$60 and Bob \$40.

This method works for any quantity and any ratio:

Dividing a quantity into a ratio algorithm

1. Add the numbers in the ratio, to calculate the total number of parts;
2. Divide the quantity to divide by the total number of parts;
3. Multiply all the numbers in the ratio by the number calculated in step 2.

Dividing a quantity into a ratio - example 1

Divide 250 marbles into the ratio $17 : 8$.

Solution: Let us follow the above algorithm:

1. We take our ratio and add them together: $17 + 8 = 25$. Thus, we have 25 parts in total;
2. We divide 250 by 25, to get 10: this is our quantity per part;
3. Finally, the ratio is $17 : 8$, so the we multiply all these numbers by 10: $170 : 80$

And we are done. You can check to see if the answer makes sense by adding the numbers in the final ratio: it must be the same as the original quantity: $170 + 80 = 250$, so we are correct.

Dividing a quantity into a ratio - example 2

Divide 1200 into the ratio 1 : 2 : 3 : 4.

Solution: Again, we follow the algorithm:

1. Add all the numbers in the ratio: $1 + 2 + 3 + 4 = 10$. We have 10 parts in total;
2. Divide 1200 by 10: $1200 \div 10 = 120$: this is our quantity per part;
3. Finally, multiply all the numbers in the ratio by 120: 120 : 240 : 360 : 480

To check, let us add the numbers in the final ratio: $120 + 240 + 360 + 480 = 1200$, and we are done.

Dividing a quantity into a ratio - example 3

Luiza and Gabriela have to split 48 candies into the ratio 1 : 5. How many candies does Gabriela receive?

Solution: Follow your algorithm:

1. Add all the numbers in the ratio: $1 + 5 = 6$. We have 6 parts;
2. Divide 48 by 6 to find 8, our quantity per part;
3. Multiply the numbers in the ratio by 8: 8 : 40.

As Gabriela's share is represented by the second number in the ratio, she receives 40 candies.

5.4.2. Given $A : B$ and $B : C$ find $A : C$

In a classroom, we have boys to girls in the ratio of 3 : 5, and we have girls to teachers in the ratio 15 : 2. What is the ratio of boys to teachers? We can represent like this:

$$\begin{array}{ccc} \text{Boys} : & \text{Girls} : & \text{Teachers} \\ 3 : & 5 & \\ & 15 : & 2 \end{array}$$

The method is to make the numbers representing girls to be the same in both ratios, using equivalent ones. Thus, we can take the first ratio and multiply both numbers by 3:

$$\begin{aligned} \text{Boys :}&\text{Girls :}\text{Teachers} \\ 9 :\ & 15 \\ & 15 :\qquad 2 \end{aligned}$$

now that girls have 15 in both ratios, we can just join the ratios together:

$$\begin{aligned} \text{Boys :}&\text{Girls :}\text{Teachers} \\ 9 :\ & 15 \\ & 15 :\qquad 2 \\ 9 :\ & 15 :\qquad 2 \end{aligned}$$

Therefore, we have the ratio of boys to teachers is 9 : 2.

To summarise, when the give you a ratio of $A : B$, another ratio of $B : C$, and want you to find the ratio $A : C$, just write the ratios in a way that B is the same number in both, then just combine.

5.5. Proportionality

5.5.1. Direct proportion

Let us use the classic "real world" problem that makes your maths teacher pretend maths is useful: recipes. Pretend that in a recipe we use potatoes and carrots in the ratio 2 : 1. How many carrots do you need to use if you are using 8 potatoes? It is clear that the more potatoes we use, the more carrots we use as well. Whenever this happens we have a direct proportion problem: if one quantity increases, the other one increases as well.

We can write that as

$$\begin{aligned} \text{Potatoes :}&\text{Carrots} \\ 2 :\ & \quad 1 \\ 8 :\ & \quad x \end{aligned}$$

One way to solve this is to notice that to go from 2 to 8,we multiplied the number of potatoes by 4 (which you can figure out by dividing 8 by 2). So, we just need to multiply the number of carrots by 4 as well:

$$\times 4 \begin{pmatrix} 2 : & 1 \\ 8 : x = & 4 \end{pmatrix} \times 4$$

Therefore, we need 4 carrots. We could also have used the cross-multiplication technique:

$$\begin{matrix} 2 : 1 \\ \times \\ 8 : x \end{matrix}$$

which gives us

$$2 \times x = 1 \times 8$$

$$2x = 8$$

$$\frac{2x}{2} = \frac{8}{2}$$

$$x = 4$$

Another type of problem than can appear is something as: 5 tickets to a concert cost \$18, how much do 7 tickets cost? There are two classic ways to solve this problem. One is called the *unitary method* and the other is just to use cross-multiplication.

Proportion using cross-multiplication

5 tickets to a concert cost \$18, how much do 7 tickets cost?

Solution: We can set up this problem as ratio:

$$\begin{aligned} \text{Tickets} &: \text{Price} \\ 5 &: \quad 18 \\ 7 &: \quad x \end{aligned}$$

we know cross multiply the ratios and obtain:

$$5 \times x = 18 \times 7$$
$$5x = 126$$
$$\frac{5x}{5} = \frac{126}{5}$$
$$x = 25.20$$

Thus, we know that 5 tickets cost \$25.20.

Proportion using the unitary method

5 tickets to a concert cost \$18, how much do 7 tickets cost?

Solution: The unitary method is very creatively called: it revolves around finding how much 1 ticket cost. We can do this by simplifying the ratio given to us:

$$\begin{aligned} \text{Tickets} :& \text{Price} \\ 5 :& \quad 18 \\ 1 :& \quad ? \end{aligned}$$

We know that we have to divide both sides of the ratio by 5 to obtain a 1 on the left:

$$\begin{aligned} &\text{Tickets} : \qquad \text{Price} \\ \div 5 \; &\left(\begin{aligned} &5 : \qquad\qquad 18 \\ &1 : ? = \tfrac{18}{5} = 3.60 \end{aligned} \right) \div 5 \end{aligned}$$

With this we have found how much 1 ticket costs: 3.60. To find how much 7 cost, we just need to multiply this value by 7: $7 \times 3.60 = 25.20$.

There are these two basic methods, then: cross-multiplication or the unitary method. The first one gives you the answer straight away, whereas the unitary method first finds the corresponding value for an unit and then you can get any value you want by multiplying appropriately. I recommend using the unitary method if you need to find more than a value (for instance, in the above examples, if you needed to find the cost of 7, 13 and 17 tickets). For a single value, I would use the cross-multiplication method, as it is faster.

5.5.1.1. Money conversion

Converting money from one currency to the other is an application of direct proportion. Let us solve the same problem using the two techniques above.

Money conversion using cross-multiplication

If 15 USD can buy 13.16 Euros, how many Euros can 12 USD buy?

Solution: Let us set up our ratios

$$\text{USD} : \text{Euros}$$
$$15 : 13.16$$
$$12 : \quad x$$

we now can cross multiply the ratios and solve the equation:

$$15 \times x = 13.16 \times 12$$
$$15x = 157.92$$
$$\frac{15x}{15} = \frac{157.92}{15}$$
$$x = 10.528$$

It is hard to buy that amount, though, so let us round it to 10.53.

Money conversion using the unitary method

If 15 USD can buy 13.16 Euros, how many Euros can 12 USD buy?

Solution: Let us set up our ratios and remember that the unitary method needs to find how many Euros 1 USD can buy:

$$\text{USD} : \text{Euros}$$
$$15 : 13.16$$
$$1 : \quad ?$$

We need to divide both sides of the ratio by 15, then:

$$
\begin{array}{rcl}
 & \text{USD :} & \text{Euros} \\
 & 15 : & 13.16 \\
\div 15 & \downarrow & \downarrow \div 15 \\
 & 1 : ? = \frac{13.16}{15} & = 0.8773
\end{array}
$$

Thus we know that 1 USD can buy 0.8773 Euro. Let us multiply this by 12:

$$12 \times 0.8773 = 10.528$$

Again, let us round it to 10.53, which is our answer.

5.5.2. Inverse proportion

We learned that when a quantity increases with the other following a fixed ratio we have a direct proportion. But there can also be quantities related by a ratio which have an *inverse* relationship: when one of them grows, the other one decreases. This is called *inverse proportion*, and the classic example is painting a wall: if 5 people take 3 days to paint a wall, how long does it take for 10 people to do it? We have *increased* the number of people, so the time taken to paint the wall should *decrease*. Our assumption in these problems is that everyone works the same, of course.

We can still solve these problems using ratios, though. Let us set up the basic:

$$
\begin{array}{rl}
\text{People :} & \text{Time to paint} \\
5 : & 3
\end{array}
$$

now we will increase the number of people to 10, and we do that by multiplying the number of people by 2:

$$
\begin{array}{crl}
 & \text{People :} & \text{Time to paint} \\
 & 5 : & 3 \\
\times 2 & \downarrow & \\
 & 10 : & ?
\end{array}
$$

Now comes the twist: we *multiplied* the number of people by 2, but the quantities are *inversely* related: the time to paint the wall should decrease. And we do that by using the *inverse* operation of multiplying by 2: we *divide* the time to paint by 2:

$$\text{People} : \text{Time to paint}$$

$$\times 2 \begin{array}{rr} 5: & 3 \\ 10: & \frac{3}{2} = 1.5 \end{array} \div 2$$

Thus, it takes 1.5 days to paint the wall.

Speed and time are also inversely related: if you increase your speed, it takes *less* time for you to move the same distance. For all these questions, just use inverse operations: if one quantity increased by multiplying by a number, the other is divided by the same number.

5.6. Map scales and ratios

5.6.1. Map ratios

If we say a map has a ratio of

$$1\text{cm} : 5\text{m}$$

we mean that a line of 1cm in the map represents 5m in the real world. With this kind of ratio, we can either calculate the real distance between two places given the distances between them on the map, or vice-versa: we can calculate their distance on the map given their distance on the real world. These problems can be solved the same direct proportion techniques we have learned above. We can also solve problems relating areas on the map and on the real world, which demand a little manipulation. However, they are also solved by direct proportion.

5.6.1.1. Distance problems

Continuing to use the example above of a map ratio

$$1\text{cm} : 5\text{m}$$

we can first solve problems of the kind: given distance on the map, find real distance. Let us say we know that two points on the map are 3cm apart and we want to find their real distance. We set up the problem as ratios:

$$1\text{cm} : 5\text{m}$$

$$3\text{cm} : \quad x$$

which we can solve using either the unit method or cross-multiplication. Let us use cross-multiplication:

$$1\text{cm} : 5\text{m}$$

$$3\text{cm} : \ x$$

$$1 \times x = 5 \times 3$$

$$x = 15$$

Finally, as the number we wanted to find was on the "real world column", it must be in metres. So the real distances between the two points is 15m.

The other way around is pretty much the same: let us pretend we have this map ratio now

$$2\text{cm} : 10\text{km}$$

and we know that the real distance between two cities is 65km. We want to find the distance on the map using the above ratio. Again, we set up the problem:

$$2\text{cm} : 10\text{km}$$

$$x : 65\text{km}$$

which we can solve using cross-multiplication:

$$2\text{cm} : 10\text{km}$$

$$x : 65\text{km}$$

$$2 \times 65 = 10 \times x$$

$$130 = 10x$$

$$\frac{130}{10} = \frac{10x}{10}$$

$$13 = x$$

This time we are finding a distance on the map, and this is measure in centimetres. Thus, the distance on the map between the two cities is 13cm.

Map ratio example

A map has the following ratio

$$3\text{cm} : 150\text{km}$$

1. Point A is 18cm from point B on the map. What is their real distance?
2. The real distance between points A and C is 37.5km. What is their distance on the map?

Solution:

1. Setting up our ratio problem:

$$\begin{aligned} 3\text{cm} &: 150\text{km} \\ 18\text{cm} &: \quad x \end{aligned}$$

cross-multiplying gives us an equation to solve:

$$\begin{aligned} 3\text{cm} &: 150\text{km} \\ 18\text{cm} &: \quad x \end{aligned}$$

$$\begin{aligned} 3 \times x &= 150 \times 18 \\ 3x &= 2700 \\ \frac{3x}{3} &= \frac{2700}{3} \\ x &= 900 \end{aligned}$$

As the real distance is in kilometres, the real distance between points A and B is 900km.

2. Again, we set up the problem as a ratio:

$$\begin{aligned} 3\text{cm} &: 150\text{km} \\ x &: 37.5\text{km} \end{aligned}$$

cross-multiplying (yet again) gives us another equation to solve:

$$\begin{aligned} 3\text{cm} &: 150\text{km} \\ x &: 37.5\text{km} \end{aligned}$$

$$\begin{aligned} 3 \times 37.5 &= 150 \times x \\ 112.5 &= 150x \\ \frac{112.5}{150} &= \frac{150x}{150} \\ 0.75 &= x \end{aligned}$$

As we are finding the distance on the map, the answer is in centimetres. Thus, points A and C are 0.75cm apart on the map.

5.6.1.2. Area problems

These problems are a little harder because they need some manipulation, but still use (again) direct proportion.

They go somewhat along the lines of: in a map with ratio

$$1\text{cm} : 3\text{km}$$

the area of a park is 5cm^2. What is the real area of the park in km^2?

What we need to do is convert our linear map ratio to area measures. We do that by squaring both sides of the ratio:

$$\begin{aligned} 1\text{cm} &: \quad 3\text{km} \\ (1\text{cm})^2 &: (3\text{km})^2 \\ 1\text{cm}^2 &: \; 9\text{km}^2 \end{aligned}$$

Now we can compare areas, using exactly the same direct proportion techniques we already know. As we know that 1cm^2 on the map corresponds to 9km^2 on the map, we just set up our ratio problem, which we solve by cross multiplication:

$$\begin{aligned} 1\text{cm}^2 &: 9\text{km}^2 \\ 5\text{cm}^2 &: \quad x \end{aligned}$$

$$1 \times x = 9 \times 5$$

$$x = 45$$

As we found area on the real world, which we are measuring in km^2, the area of the park is $45km^2$.

The exact same technique can be used to solve volume or mass problems, but you have to cube both sides of the ratio (see worked example in Section 7).

5.6.2. Map scales

Map scales are exactly the same as map ratios, but they come **without units**. This is because they are **always in centimetres**.Thus, if you see a question that says a map has a scale

$$1 : 100$$

it is the same as

$$1\text{cm} : 100\text{cm}$$

and you have to memorise it. Apart from that, we just have to do some unit manipulation.

Map scale example question

A map has scale

$$1 : 100000$$

1. Find the real distance between points A and B given that they are 3 centimetres apart in the map. Give your answer in metres.

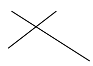
   Solution: The first thing I like doing is remembering that map scales are in centimetres to centimetres:

   $$1\text{cm} : 100000\text{cm}$$

   Now let us convert the real distance to metres, as they want us to. To convert centimetres to metres we divide by 100, so:

   $$1\text{cm} : 1000\text{m}$$

   Finally we just do cross-multiplication:



$$1\text{cm} : 1000\text{m}$$

$$3\text{cm} : x$$

$$1 \times x = 1000 \times 3$$

$$x = 3000$$

Thus the distance between A and B is 3000m.

2. Find the area in the map, in square centimetres, of a field that has 2.5km^2.

   Solution: Again let us convert to the appropriate units. We have km^2, so first we will convert to km. To do this I like first to convert to metres, dividing by 100:

   $$1\text{cm} : 100000\text{km}$$

   $$1\text{cm} : 1000\text{m}$$

   and now to convert to km we divide by 1000:

   $$1\text{cm} : 1\text{km}$$

   As we are comparing areas, we square both sides of the ratio:

   $$(1\text{cm})^2 : (1\text{km})^2$$

   $$1\text{cm}^2 : 1\text{km}^2$$

   Finally, we cross-multiply:

   $$1\text{cm}^2 : 1\text{km}^2$$

   $$x : 2.5\text{km}^2$$

$$1 \times 2.5 = 1 \times x$$

$$2.5 = x$$

Thus, the area on the map is 2.5cm^2.

5.7. Exam hints

- Ratio questions are very common as a part of the first question of paper 3/4 (0580);
- The most common questions are dividing a quantity into a ratio and map scales, particularly with area;
- It is important to master direct proportion, as you can use it to solve many different questions;
- Remember that map scales are always centimetres to centimetres.

Summary

- A ratio tells us how a number is divided into parts;
- We can obtain equivalent ratios by either multiplying or dividing all the numbers of the ration by the same number;
  - A ratio is in its simplest form when all its numbers are integers and they have HCF of 1;
  - Ratios with units need to have the same unit on both sides to be in their simplest form;
- Given two equivalent ratios, we can cross-multiply their values to obtain an equality;
- To divide a quantity into a ratio use the following algorithm:
  1. Add the numbers in the ratio, to calculate the total number of parts;
  2. Divide the quantity to divide by the total number of parts;
  3. Multiply all the numbers in the ratio by the number calculated in step 2.
- To solve proportion problems you can either use the unitary method or cross-multiplication;
- Money conversion is just direct proportion;
- Inverse proportion is when a quantity increases and the other decreases: if one side gets multiplied by a number, the other side gets divided by the same number;

- Map ratios are proportion problems. If they are about areas remember to square both numbers of the ratio; if they have volumes cube them;
- Map scales are always in centimetres to centimetres: convert the units to what you need first and then apply the same techniques as map ratios.

6. Percentages

6.1. Why learn percentages

If I were to choose one single topic that has "real world" application to teach, I would choose percentages. They are everywhere. News are filled with percentage change in relevant values to society, companies publish their percentage increase (or decrease) in profits. I was told they are also used in businesses (who'd guess!). In all, percentages are a plague, so it's very important that you understand them.

6.2. What are percentages

I don't want to be the one that ruins the fun, but percentages are just numbers referring to a certain total. The origin of the world is probably from Latin, *per centum.* We basically will take a quantity and say it corresponds to 100%. Smaller quantities of this original one are referred by a percentage of the original, how many parts of that one. Bigger quantities of the original are also percentages, but greater than 100%. Thus, you already know percentages: they are fractions!

6.3. Fractions, decimals and percentages equivalency

6.3.1. Percentages to fractions

A percentage is simply that, a quantity *per cent.* *Cent* is just a fancy word for a hundred (a *cent*ury, for instance, has a hundred years). Therefore, a percentage is simply something divided by 100. For example, 15% is

$$15\% = \frac{15}{100}$$

as we have 15 *per cent*, or 15 *per hundred.* So we already know how to convert a percentage to a fraction: we divide the percentage by 100. We should always simplify the fraction after, because everybody likes simpler things. Therefore, in our 15% example:

$$15\% = \frac{15}{100} = \frac{3}{20}$$

as we can divide both numerator and denominator by 5. Sometimes we cannot simplify, say 17%:

$$17\% = \frac{17}{100}$$

and sometimes it will be a fraction greater than 1, for instance 225%:

$$225\% = \frac{225}{100} = \frac{9}{4}$$

as we can divide both top and bottom by 25.

To summarise, to convert a percentage to a fraction we simply divide the percentage by 100.

6.3.2. Fractions to percentages

In my opinion, the easiest way to convert fractions to percentages is to remember that a percentage is just something divided by 100. Therefore, if we obtain an equivalent 'fraction'[1] to the original with denominator 100, we already have a percentage.

For instance, say we have the fraction

$$\frac{4}{25}$$

If we multiply both the numerator and the denominator by $100 \div 25 = 4$, we obtain:

$$\frac{4}{25} = \frac{16}{100}$$

Thus, we have that $\frac{4}{25}$ is equivalent to $\frac{16}{100}$, which is 16%.

Another example is to convert $\frac{9}{16}$ to a percentage. If we multiply both numerator and denominator by $100 \div 16 = 6.25$, we obtain:

$$\frac{9}{16} = \frac{56.25}{100}$$

so we have that $\frac{9}{16}$ is the same as 56.25%.

In summary, to convert a fraction to a percentage, obtain an equivalent fraction which has denominator 100.

6.3.3. Percentages to decimals

Well, this one is easy: a percentage is a number divided by 100, so to convert a percentage to a decimal you divide the percentage by 100. For example,

$$17\% = 0.17$$

Remember that you can have percentages greater than 100, so

$$235\% = 2.35$$

[1] I have added the quotation marks because a fraction should have integers on the numerator and the denominator. We will see in a bit that is not necessarily true when we are converting to a percentage.

6.3.4. Decimals to percentages

Also very easy: multiply the decimal by 100 and add the percentage sign! So we have

$$0.035 = 3.5\%$$

or

$$1.23 = 123\%$$

6.3.5. Fractions to decimals

Converting fractions to decimals is quite easy: simply divide the numerator by the denominator. For instance,

$$\frac{30}{50} = 0.6$$

6.3.6. Decimals to fractions

6.3.6.1. Finite decimals

By a finite decimal we refer to a number which has a finite number of places, such as

$$3.2$$

or

$$1.12345$$

The easiest way, in my opinion, of converting these decimals to fractions is to first convert them to percentages and them simplifying the fraction:

$$3.2 = 320\% = \frac{320}{100} = \frac{16}{5}$$

Another example, 0.08:

$$0.08 = 8\% = \frac{8}{100} = \frac{2}{25}$$

6.3.6.2. Recurring decimals

Definition and notation of recurring decimals We will define a recurring decimal to be a decimal number with a pattern that repeats forever, such as:

$$0.33333333333\ldots$$

where 3 is repeating forever, indicated by the ... at the end. But we can have "harder" recurring parts, as

$$0.1717171717\ldots$$

in which 17 repeats forever. The recurring section does not even have to start right from the decimal point:

$$0.1234444444\ldots$$

here 4 eventually starts repeating forever, but is preceded by 123 after the decimal point. Finally, we can have a number greater than 1, such as

$$1234.869321321321\ldots$$

in which 321 starts repeating forever after the whole 1234.869 part.

It is quite annoying to keep writing the recurring part of the number, so we usually denote that it is repeating using a little dot:

$$0.\dot{3} = 0.333333\ldots$$

or

$$0.\dot{1}\dot{7} = 0.171717\ldots$$

If we have a recurring part of more than 2 digits, we just put the little dot above the first and the last:

$$0.\dot{1}23\dot{4} = 0.123412341234\ldots$$

The last example we had:

$$1234.869\dot{3}2\dot{1} = 1234.869321321321\ldots$$

We will call the number of digits in the recurring part of the number its recurring length.

A very important fact is that all recurring decimals are rational numbers, which means they can be written as fractions[2] (add ref to number sets chapter). Before we continue, notice that recurring decimals are completely different from irrational numbers, such as π: in π and all other irrational numbers, there is *never* a pattern that eventually repeats forever. However, it is important to know that not all decimals with an infinite representation, such as $0.\dot{3} = 0.3333\ldots$ are irrational.

Converting recurring decimals to fractions

Method 1: using an equation This is my favourite method, and actually one of my favourite algorithms in school. The whole idea is based on the simple fact that

$$0.\dot{3} - 0.\dot{3} = 0.3333\ldots - 0.3333\ldots = 0$$

which simply means that we can cancel out the recurring part of a number by subtracting it at the appropriate place. For instance,

$$0.4\dot{5} - 0.2\dot{5} = 0.45555\ldots - 0.25555\ldots = 0.2$$

Because of this, we try to obtain two numbers with the same recurring part, so that we can subtract them and get rid of it!

Let us see the best example, $0.\dot{9}$.

[2] See the proof in the formality after taste

The first thing we do is give $0.\dot{9}$ a name, such as x:

$$x = 0.9999\ldots$$

which is also our first equation with only the recurring part after the decimal point. We need another, and we can obtain it by multiplying both sides of the equation above by 10:

$$x = 0.9999\ldots$$

$$10x = 9.9999\ldots \text{Multiplying both sides by 10}$$

Now we have two equations with only the recurring part after the decimal point. We can now subtract one from the other:

$$10 - x = 9.9999\ldots - 0.9999\ldots$$

and we can simply solve this equation for x:

$$10x - x = 9.999\ldots - 0.9999\ldots$$

$$9x = 9$$

$$\frac{9x}{9} = \frac{9}{9}$$

$$x = 1$$

So we have shown that

$$x = 0.\dot{9} = 1$$

This method works for any recurring decimal. Say we had $0.\dot{1}\dot{7}$. First we give it a name:

$$x = 0.171717\ldots$$

which already has only the recurring part after the decimal point. We need another. To obtain it, we need to multiply both sides by 100, as the recurring part has two digits:

$$x = 0.171717\ldots$$

$$100x = 17.171717\ldots$$

and now we subtract one equation from the other:

$$100x - x = 17.171717\ldots - 0.171717\ldots$$

$$99x = 17$$

$$\frac{99x}{99} = \frac{17}{99}$$

$$x = \frac{17}{99}$$

Hence,

$$0.\dot{1}\dot{7} = \frac{17}{99}$$

Another example, $3.4\dot{1}2\dot{3}$. First, give it a name

$$x = 3.4123123123\ldots$$

This time, we cannot use this as an equation as we have a 4 ruining the decimal part (remember we want only the recurring part!). To get rid of that 4, let us multiply both sides by 10:

$$x = 3.4123123123\ldots$$

$$10x = 34.123123123\ldots$$

Now we have one equation. To obtain another, we have to multiply this last equation by 1000, as the recurring part has 3 digits:

$$10x = 34.123123132\ldots$$

$$10000x = 34123.123123123\ldots$$

Finally, we subtract the $10x$ equation from the $10000x$ one:

$$10000x - 10x = 34123.123123\ldots - 34.123123\ldots$$

$$9990x = 34089$$

$$\frac{9990x}{9990} = \frac{34089}{9990}$$

$$x = \frac{34089}{9990}$$

Thus,

$$x = 3.4\dot{1}2\dot{3} = \frac{34089}{9990}$$

Method 2: using a formula For simple decimals (and if you really want for all of them), we can use a simple formula.

If you have a recurring decimal, which can be written as

$$0.\underbrace{}_{\text{recurring part}}$$

and the recurring part has length n, which means we have n digits in it, its fractional representation will be

$$\frac{\overbrace{}^{\text{recurring part}}}{\underbrace{999\cdots9}_{n\text{ 9s}}}$$

This may seem complicated because of the generality, but see some examples.
For instance, in

$$0.\dot{3} = 0.3333\ldots$$

we have the recurring part with a single digit, 3. So our fraction will be only 3 in the numerator and a single 9 at the denominator, as we have a single digit repeating:

$$0.\dot{3} = 0.333\ldots = \frac{\overbrace{3}^{\text{recurring}}}{\underbrace{9}_{\text{1digit in recurring part}}} = \frac{1}{3}$$

For $0.\dot{4}\dot{5}$, we have 2 digits on the recurring part, so our fraction becomes

$$0.\dot{4}\dot{5} = 0.454545\ldots = \frac{\overbrace{45}^{\text{recurring}}}{\underbrace{99}_{\text{2 digits in recurring part}}} = \frac{5}{11}$$

You can use this formula for any number, but sometimes you have to be clever. For instance:

$$1.178\dot{2} = 1.17822222\ldots$$

we first write this as

$$1.178\dot{2} = 1.178 + 0.000\dot{2}$$

and now we convert each term separately:

$$1.178 = 117.8\% = \frac{117.8}{100} = \frac{589}{500}$$

and

$$0.000\dot{2} = 0.\dot{2} \div 1000 = \frac{2}{9} \div 1000 = \frac{2}{9 \times 1000} = \frac{2}{9000} = \frac{1}{4500}$$

Finally, you add these fractions:

$$1.178\dot{2} = \frac{589}{500} + \frac{1}{4500} = \frac{2651}{2250}$$

6.4. Percentage of a number

If you remember that 100% corresponds to a 'total amount', it is easy to understand the percentage of a number 'trick'. Let's say we want to calculate 35% of 60. If we write this a proportion:

$$60 : 100\% \qquad\qquad 60 \text{ is } 100\%$$

$$x : 35\% \qquad\qquad x \text{ is } 35\%$$

We can solve this by cross multiplying:

$$100\% \times x = 35\% \times 60$$

Now, remember that 100% is equal to 1, as $100\% = \frac{100}{100} = 1$, and that $35\% = 0.35$:

$$x = 0.35 \times 60 = 21$$

Thus, we have that 35% of 60 is equal to 21. To reach that, in the end, we simply multiplied 60 by 35%, and that is the 'general way':

To calculate a percentage of a number, multiply the number by the percentage in either decimal or fraction form.

For instance, to calculate 12.5% of 350:

$$12.5\% \times 350 = \frac{12.5}{100} \times 350 = 43.75$$

Or, to calculate 234% of 750:

$$234\% \times 750 = \frac{234}{100} \times 750 = 1755$$

6.5. Number as a percentage of another number

Again, to understand the 'trick' let us write a ratio problem. Let's say we want to calculate 35 as a percentage of 80. That means that 80 is 100%, and we want to find out how much 35 is:

$$80 : 100\% \qquad\qquad 80 \text{ is } 100\%$$

$$35 : x\%$$

we cross multiply to obtain

$$80x = 100 \times 35$$

Dividing both sides by 80 and rearranging the order of multiplication on the right-hand side:

$$x = \frac{35}{80} \times 100$$

Thus:

To calculate a number x as a percentage of y, calculate $\frac{x}{y}$ and multiply the result by 100

As an example, let us calculate 42 as a percentage of 28:

$$\frac{42}{28} \times 100 = 150$$

Therefore, 42 is 150% of 28.

6.6. Percentage change

Let us say we a country had a population of 100 in 2000 and it had a population of 125 in 2010. We want to calculate how much the population changed from 2000 to 2010, in percentage. This 'in percentage' is the trick: we need to calculate the change as a percentage of the original value.

If we call the initial value i and the new value n, we can calculate the change in value:

$$\text{change} = \text{new} - \text{initial} = n - i$$

We can now calculate this change as a percentage of the initial value:

$$\text{percentage change} = \frac{\text{change}}{\text{initial}} \times 100 = \frac{n-i}{i} \times 100$$

Thus, in our country example:

$$\text{percentage change} = \frac{125 - 100}{100} \times 100 = 25\%$$

We can also have a quantity that decreases, and then we have to be a bit more careful. Say that we had \$350 and now we have \$275 and we need to calculate the percentage change. As usual we calculate the change:

$$\text{change} = \text{new} - \text{initial} = 275 - 350 = -75$$

Finally, we calculate the change as a percentage of the initial value:

$$\frac{\text{change}}{\text{initial}} \times 100 = \frac{-75}{350} \times 100 = -0.214 \times 100 = -21.4\%$$

Therefore, we have a decrease of 21.4% of our money. Notice that if they ask you to calculate the percentage decrease you can ignore the sign, as decrease already denotes it. If they just ask you for the percentage change, the sign is important. Be careful!

6.7. Percentage increase and decrease

Now we will focus on solving problems such as 'increase 75 by 12%' and 'decrease 250 by 90%'. The first kind of problem is finding a percentage increase in a quantity, and the second a percentage decrease.

For both percentage increase and decrease we will be using the same technique. This technique is based on finding the *multiplier* of the change and then, surprisingly, multiplying the initial quantity by the multiplier.

Let's first understand what these problems are asking.

If we have to increase 75 by 12%, that means we start with 75, which corresponds to 100%, and then we need to *add* (as we are increasing) 12% of 75 to our total. That is

$$75 + 12\% \text{ of } 75 = 75 + 12\% \times 75$$

we have 75 as a common factor:

$$75 + 12\% \times 75 = 75(1 + 12\%)$$

and we know how to convert percentages to decimals (we divide them by 100):

$$75(1 + 0.12) = 75 \times 1.12$$

This 1.12 we obtained is the *multiplier* of an increase of 12%! If we just do the multiplication now:

$$75 \times 1.12 = 84$$

and we have increased 75 by 12%. As you can see, there is no need to calculate 10% or any other percent of 75 and then combine them, we just do a single multiplication!

The multiplier method also works for decreasing a quantity by a percentage. Let's say we want to decrease 135 by 21.5%. That means we start with 135 and take away 21.5% of 135 from it:

$$135 - 21.5\% \text{ of } 135 = 135 - 21.5\% \times 135$$

again we have a common factor of 135:

$$135 - 21.5\% \times 135 = 135(1 - 21.5\%)$$

and we can convert 21.5% to decimal by dividing it by 100:

$$135(1 - 0.215) = 135 \times 0.785$$

Finally, we multiply

$$135 \times 0.785 = 105.975$$

and we are done: we have reduced 135 by 21.5%.

Notice that in both examples we either did 1 plus the percentage (when we were increasing the quantity) or 1 minus the percentage (when we were decreasing the quantity). This step finds the multiplier of the change. After this, we just have to multiply the original quantity by the multiplier and we are done.

In all, the steps to find a percentage change are:

1. Identify if the change is an increase (+) or a decrease (−) and the percentage we are changing, $p\%$;
2. Calculate the multiplier of the change:
   a) If we have an increase: $100\% + p\%$
   b) If we have a decrease: $100\% - p\%$
3. Change the multiplier to a decimal by dividing it by 100
4. Multiply the original quantity by the multiplier.

Let's do some examples.

- Increase 325 by 37%.

  We have an increase of 37%. To calculate the multiplier of a percentage increase we add 100% to our percentage:

  $$100\% + 37\% = 137\%$$

  now we convert this percentage to a decimal:

  $$137\% = 1.37$$

  Finally, we multiply our original quantity, 325, by 1.37:

  $$1.37 \times 325 = 445.25$$

- Decrease 525 by 53.5%

  We have a decrease of 53.5%. To find the multiplier of a percentage decrease we subtract our percentage from 100%:

  $$100\% - 53.5 = 46.5\%$$

  now we divide this percentage by 100 to convert it to a decimal:

  $$46.5\% = 0.465$$

  finally, we multiply our quantity by the multiplier:

  $$0.465 \times 525 = 244.125$$

- Increase 321 by 123%

  To find the multiplier of a percentage increase we add our percentage to 100%:

  $$100\% + 123\% = 223\%$$

  now we convert it to a decimal:

  $$223\% = 2.23$$

  To end it, we multiply our original quantity by the multiplier:

  $$2.23 \times 321 = 715.83$$

- Decrease 100 by 150%

  The multiplier of a percentage decrease is found by subtracting our percentage from 100%:

  $$100\% - 150\% = -50\%$$

we now convert this to a decimal:

$$-50\% = -0.5$$

To finish, we multiply our original quantity by the multiplier:

$$-0.5 \times 100 = -50$$

Notice that we have a negative quantity, which will only make sense if we are talking abstractly or or quantities that have meaning when negative, such as money.

In all, notice that the multiplier is great: with a single multiplication we calculate all the types of percentage increase or decrease, no matter the type.

6.8. Reverse percentage

Reverse percentage problems are a classic, and usually look like: 'you buy shoes at a sale of 15% for \$65. What is the original price of the shoes?'.

Basically, they are saying that you have an unknown quantity that went through a reduction of 15% and is now \$65. We need to find that quantity. This is where the multiplier really shines! Let's make a 'drawing' to represent our problem. Calling our unknown quantity x, we have:

$$x \xrightarrow[\text{15\%}]{\text{reduction of}} 65$$

We know how to apply a reduction of 15% to a quantity. We start with 100% and take away 15%, that is

$$100\% - 15\% = 85\%$$

now we convert this percentage to a decimal to obtain our multiplier:

$$85\% = 0.85$$

To reduce a quantity by this we simply multiply it. Our quantity is x, so we have

$$x \times 0.85$$

Now, we know the result of this: it is the price on sale of our shoes, 65:

$$x \times 0.85 = 65$$

we just have to solve this equation now:

$$\begin{aligned} 0.85x &= 65 \\ \frac{0.85x}{0.85} &= \frac{65}{0.85} && \text{Dividing both sides by 0.85} \\ x &= 76.47 \end{aligned}$$

Thus, the original price of the shoes was \$76.47.

Solved exercise: finding the original value before a percentage increase

The population of Neverland in 2100 is 258.125 million after an increase of 3.25% in the last 100 years. Calculate the population of Neverland in 2000.

Solution: First, we can ignore the 'million' part and just add it to the answer in the end. Now, we can call the population in 2000 x, as usual. We know that from 2000 to 2100 x increased by 3.25%. The multiplier for such an increase is:

$$100\% + 3.25\% = 103.25\% = 1.0325$$

Thus, we know that

$$x \times 1.0325 = 258.125$$

Solving for x:

$$\frac{1.0325x}{1.0325} = \frac{258.125}{1.0325}$$

$$x = 250$$

The population of Neverland in 2000 was 250 million.

Solved exercise: finding the original value before a percentage decrease

I bought the *Pokemon X* on sale of 25% for \$49. Find the price before the discount.

Solution: Again, let's call the original price x. Given that I bought the game at a discount, its original price was reduced by 25%. The multiplier for a reduction of 25% is

$$100\% - 25\% = 75\% = 0.75$$

If we multiply x by 0.75 we obtain the price on sale:

$$x \times 0.75 = 49$$

now we just solve for x:

$$\frac{0.75x}{0.75} = \frac{49}{0.75}$$

$$x = 60$$

The original price was \$60.

6.9. Interest

Say we have a given quantity of money and we put into a savings account. Because the bank can use that money, we are basically lending the bank some money. Thus, the bank pays us *interest*, which is basically a percentage of the money we have into the savings account each month.

In a certain country, the population grows by 1.5% each year. A car devalues by 5% per year.

All those situations are basically the same: we have an initial value (be that money, population or whatever) and this value increases or decreases by a fixed percentage every period of time.

To discuss interest, we will be using certain notation. We have:

P: stands for the *principal*, the initial quantity invested (or the initial population, or anything that will grow by a constant percentage);

r: stands for the *interest rate*, the percentage that increases (or decreases) our principal;

n: the *number of periods* the interest is going to be paid (basically the time of the investment, loan, population growth, you will see)

I: the value of the *interest* after n periods (only the 'extra', not including the principal)

T: the *total value of the investment* after n periods (the value of the interest added to the principal, that is, $P + I$)

6.9.1. Simple interest

You have some left over money, \$200, and you put into my amazing account. This account works like this: it looks at your principal (initial amount invested) and pays 5% of your principal every month. Thus, every month it will pay you 5% of 200:

$$5\% \text{ of } 200 = \frac{5}{100} \times 200 = 10$$

After the first month, therefore, you will have

$$\begin{aligned} \text{Value in the account } &+ 5\% \text{ of the principal} \\ 200 + 10 &= 210 \end{aligned}$$

After the second month you get another \$10:

$$\begin{aligned} \text{Value in the account } &+ 5\% \text{ of the principal} \\ 210 + 10 &= 220 \end{aligned}$$

and so on: you will get \$10 every month. This is a situation that has *simple interest*: the interest paid (the extra money you win) every period only depends on the *initial amount invested* (the principal), not on the current amount in the account.

In our example, if you leave the money in the account during one year (12 months), you will receive

$$12 \times 5\% \text{ of } 200 = 12 \times 10 = 120$$

as interest of your investment. I said I would refer to this by I, so we have

$$I = 12 \times 5\% \text{ of } 200 = 120$$

we can rearrange this to

$$I = 200 \times 5\% \times 12$$

in this equation, 200 is the principal (P), 5% is the interest rate (r) and 12 is the number of periods you left your money in the account (n). Hence, we have a formula to calculate the interest given by an investment using simple interest:

$$\boxed{I = P \times \frac{r}{100} \times n}$$

It is very important to pay attention to the fact that this formula only calculates the *interest*, the extra amount you receive. If you want to know the full value of your investment you have to add the principal (as it is still there!):

$$T = P + I$$

$$\boxed{T = P + P \times \frac{r}{100} \times n}$$

Also keep in mind that you can factorise this equation (and I believe it will help you to memorise it, as compound interest has a similar one):

$$\boxed{T = P\left(1 + \frac{r}{100} \times n\right)}$$

Solved exercise: finding the total value of an investment with simple interest

Luiza invests \$540 at a rate of 3.5% per year simple interest for 15 years. Calculate the total amount of Luiza's investment.

Solution: As it is simple interest, each year Luiza will receive 3.5% of her principal, 540:

$$3.5\% \text{ of } 540 = \frac{3.5}{100} \times 540 = 18.90$$

She invests for 15 years, so she will receive that value 15 times:

$$15 \times 18.90 = 283.50$$

Hence, her profit is 283.50. The total value of its investment is her profit added to her principal:

$$283.50 + 540 = 823.50$$

We could also have used the formula:

$$T = P\left(1 + \frac{r}{100} \times n\right)$$

Luiza has $T = 540$, $r = 3.5$ and $n = 12$. Thus:

$$T = 540\left(1 + \frac{3.5}{100} \times 15\right) = 823.50$$

which gives the same answer.

Solved exercise: finding the profit of an investment with simple interest

Gabriela invests \$150 at a rate of 4% simple interest per month for 3 years. Calculate Gabriela's profit.

Solution: Again, given that Gabriela is investing using simple interest, she will receive 4% of her principal every month. That is:

$$4\% \text{ of } 150 = \frac{4}{100} \times 150 = 6$$

She is investing for 3 years, or 36 months, so she will receive:

$$6 \times 36 = 216$$

This is her profit. Using the formula:

$$I = P \times \frac{r}{100} \times n$$

we have $P = 150$, $r = 4$ and $n = 36$:

$$I = 150 \times \frac{4}{100} \times 36 = 216$$

Solved exercise: finding the rate of an investment with simple interest
Laura invested \$250 for 50 months at simple interest. She profited \$375. Find the rate of the investment.

Solution: Given that Laura receives the same amount every month, she received

$$\frac{375}{50} = 7.50$$

per month. Now, we can obtain the rate by multiplying the principal, 250, by r, as we know that must equal 7.50:

$$250 \times r = 7.50$$

$$r = 0.03$$

Multiplying it by 100 we have that $r = 3\%$.
I think it is actually much easier, in this case, to use the formula:

$$I = P \times \frac{r}{100} \times n$$

$$375 = 250 \times \frac{r}{100} \times 50$$

$$375 = 125r$$

$$r = \frac{375}{125} = 3$$

This already gives the answer in percentage as well.

Solved exercise: finding the number of periods of an investment with simple interest

Valentina invests \$1000 at a rate of 3.75% simple interest per month. Her total amount after the investment is \$1300. How many months has she invested her money for?

Solution: First, we can find out Valentina's profit: $1300 - 1000 = 300$. Now, we know that she received 3.75% of her principal, 1000, every month. That is:

$$1000 \times \frac{3.75}{100} = 37.50$$

We can, finally, find the number of months she invested her money by dividing her profit, 300, by how much she received each month:

$$\frac{300}{37.50} = 8$$

Thus, Valentina invested for 8 months.
We could also have used the formula:

$$I = P \times \frac{r}{100} \times n$$

$$300 = 1000 \times \frac{3.75}{100} \times n$$

$$300 = 37.5n$$

$$n = \frac{300}{37.5} = 8$$

6.9.2. Compound interest

Again, let's say you have \$200 to invest. This time, however, you save it in an account that pays you 5% of the *current value in the account* every month. This is the difference between simple interest and compound interest: simple interest only looks at your principal, whereas compound looks at the current value in the account.

After the first month you get 5% of the value in the account. This time that is just the principal:

$$\text{Value in the account} + 5\% \text{ of the value in the account}$$

$$200 + 5\% \text{ of } 200$$

$$200 + 10 = 210$$

On the next month, however, you get 5% on the *new value*:

$$\text{Value in the account} + 5\% \text{ of the value in the account}$$
$$210 + 5\% \text{ of } 210$$
$$210 + 10.5 = 220.5$$

On the third month, again, you get another 5% on the value in the account:

$$\text{Value in the account} + 5\% \text{ of the value in the account}$$
$$220.5 + 5\% \text{ of } 220.5$$
$$220.5 + 11.025 = 231.525$$

And that would continue for the number of periods you left your money in the account. I will put the derivation of the formula at the end of the chapter, but here it is:

$$\boxed{T = P\left(1 + \frac{r}{100}\right)^n}$$

Notice that the formula for compound interest calculates the total amount after the n periods, not the interest. If you want to calculate only the interest, we need to take away the principal:

$$I = T - P$$

Solved exercise: finding the total value of an investment with compound interest

Victoria invested $1250 at a rate of 1.5% per month in a bank that pays compound interest. Calculate the total value of her investment after 12 months.

Solution: Here we have the principal, $P = 1250$, $r = 1.5$ and $n = 12$. Substituting those values in the formula:

$$T = P\left(1 + \frac{r}{100}\right)^n$$
$$T = 1250\left(1 + \frac{1.5}{100}\right)^{12} = 1494.52$$

Thus, Victoria's investment is worth $1494.52 after 12 months.

Solved exercise: finding the profit of an investment with compound interest

Cecilia invests \$3400 at a rate of 0.9% compound interest per year for 4 years. Find her profit.

Solution: We have $P = 3400$, $r = 0.9$ and $n = 4$. Let's find the total value of the investment first:

$$T = P\left(1 + \frac{r}{100}\right)^n$$

$$T = 3400\left(1 + \frac{0.9}{100}\right)^4 = 3524.06$$

The question wants us to find Cecilia's profit, so we have to take away what she invested:

$$3524.06 - 3400 = 124.06$$

therefore, she profited \$124.06.

Solved exercise: finding the rate of an investment with compound interest

Bia invested \$1000 for 17 years at compound rate. Her total value at the end of the investment was \$1947.90. Find the rate of the investment.

Solution: Here we have $P = 1000$, $n = 17$ and $T = 1947.90$. Substituting it:

$$T = P\left(1 + \frac{r}{100}\right)^n$$

$$T = 1000\left(1 + \frac{r}{100}\right)^{17} = 1947.90$$

We have a lovely equation to find r:

$$1000\left(1 + \frac{r}{100}\right)^{17} = 1947.90$$

$$\left(1 + \frac{r}{100}\right)^{17} = \frac{1947.90}{1000} \qquad \text{Dividing both sides by 1000}$$

$$\sqrt[17]{\left(1 + \frac{r}{100}\right)^{17}} = \sqrt[17]{\frac{1947.90}{1000}} \qquad \text{Taking root 17 on both sides}$$

$$1 + \frac{r}{100} = 1.03999$$

$$\frac{r}{100} = 1.03999 - 1 \qquad \text{Subtracting 1 on both sides}$$

$$\frac{r}{100} = 0.03999$$

$$r = 0.03999 \times 100 \qquad \text{Multiplying both sides by 100}$$

$$r = 3.999 \approx 4$$

Therefore Bia invested her money at a rate of 4% per year.

Solved exercise: finding the number of periods of an investment with compound interest

Julia invested her money at a rate of 3% per year compound interest. She invested \$2500 and, at the end of the investment, received \$3074.68. For how many years did Julia invest her money?

Solution: We have $P = 2500$, $r = 3$ and $T = 3074.68$ and need to find n. Substituting this information:

$$T = P\left(1 + \frac{r}{100}\right)^n$$
$$T = 2500\left(1 + \frac{3}{100}\right)^n = 3074.68$$

One way of solving this is guessing possible values for n on your calculator:

$$n = 5 \rightarrow \qquad 2500\left(1 + \frac{3}{100}\right)^5 = 2898.19$$
$$n = 6 \rightarrow \qquad 2500\left(1 + \frac{3}{100}\right)^6 = 2985.13$$
$$n = 7 \rightarrow \qquad 2500\left(1 + \frac{3}{100}\right)^7 = 3074.68$$

Thus, Julia invested for 7 years.
The second way is to use logarithms:

$$2500\left(1 + \frac{3}{100}\right)^n = 3074.68$$
$$\left(1 + \frac{3}{100}\right)^n = \frac{3047.68}{2500} \qquad \text{Dividing both sides by 2500}$$
$$1.03^n = \frac{3047.68}{2500} \qquad \text{Simplifying the brackets}$$
$$\log_{1.03} 1.03^n = \log_{1.03}\left(\frac{3047.68}{2500}\right) \qquad \text{Taking } \log_{1.3} \text{ on both sides}$$
$$n \log_{1.03} 1.03 = \log_{1.03}\left(\frac{3047.68}{2500}\right) \qquad \log_a b^n = n \log_a b$$
$$n = \log_{1.03}\left(\frac{3047.68}{2500}\right) \qquad \log_b b = 1$$
$$n = 6.70$$

So we know that exactly 6.70 years Julia receives that money. But the answer has to be a whole number of years, thus she invested for 7 years.

6.10. Exam hints

- Always check your conversion of recurring decimals in your calculator;
- Be careful not to acquire *interestites*, a pathology which makes you think every percentage question is about interest;
- Simple interest questions have simple interest written on them! If not, do not use simple interest;
- Population growth is usually compound interest, but can also be percentage change;
- Percentage change and reverse percentage questions are usually worth 3 marks.

Summary

- Percentages are just fractions with 100 on the denominator;
- To convert decimals to percentages, multiply by 100;
- To convert percentages to decimals, divide by 100;
- To convert fractions to percentages, obtain an equivalent fraction with 100 on the denominator;
- To convert fractions to decimals, simply divide the numerator by the denominator;
- To convert percentages to fractions, just write a fraction with the percentage number on the numerator and 100 on the denominator;
- All recurring decimals are rational numbers, and there is an algorithm (recipe) to convert them to fractions;
- To calculate a percentage of a number, just multiply the number by the percentage in decimal or fraction form;
- To calculate a number as a percentage of another, just divide the first by the second and multiply by 100;
- To calculate percentage change, use the formula

$$\frac{\text{new-initial}}{\text{initial}} \times 100$$

- To calculate percentage increase, first calculate the multiplier by adding 100% to the percentage you are increasing. Then, multiply the original quantity by the multiplier;

- To calculate percentage decrease, first calculate the multiplier by subtracting 100% to the percentage you are decreasing. Then, multiply the original quantity by the multipliers;

- Reverse percentage problems are also solved by the multiplier: just call the original quantity x and solve the equation you obtain when you multiply x by the change multiplier equals the value they give you;

- Just use simple interest when they explicitly say so. The formula for the interest (the extra) is:
$$I = P \times \frac{r}{100} \times n$$
and the formula for the total amount is
$$T = P\left(1 + \frac{r}{100} \times n\right)$$

- For compound interest we have only a formula for the total amount:
$$T = P\left(1 + \frac{r}{100}\right)^n$$

Formality after taste

Proof that recurring decimals are rational numbers - and a formula for free

We will only prove for decimals of the form 0.recurring, as the "harder" cases can be built on this.

A general recurring decimal of this type can be written as

$$0.\dot{a}_1 a_2 a_3 \cdots a_{n-1}\dot{a}_n = 0.a_1 a_2 a_3 \cdots a_{n-1} a_n a_1 a_2 a_3 \cdots a_{n-1} a_n \ldots$$

Now we use the same technique we used to find the fraction representation of a number: the little equation method.

$$x = 0.a_1 a_2 a_3 \cdots a_n a_1 a_2 a_3 \cdots a_n \cdots \quad \text{Begin calling the number } x \tag{6.1}$$

$$10^n x = a_1 a_2 a_3 \cdots a_n . a_1 a_2 a_3 \cdots a_n a_1 a_2 a_3 \cdots a_n \cdots \quad \text{Multiply } x \text{ by } 10^n \tag{6.2}$$

$$10^n x - x = a_1 a_2 a_3 \cdots a_n . a_1 a_2 a_3 \cdots a_n \cdots - 0.a_1 a_2 a_3 \cdots a_n \cdots \quad \text{Subtract (1) from (2)}$$

$$x(10^n - 1) = a_1 a_2 a_3 \cdots a_n \quad \text{Put } x \text{ as common factor}$$

$$x = \frac{a_1 a_2 a_3 \cdots a_n}{10^n - 1} \quad \text{Divide both sides by } 10^n - 1$$

Thus, we have shown that any recurring decimal of the form $0.\dot{a_1}a_2a_3\cdots a_{n-1}\dot{a_n}$ can be written as a fraction of the form

$$0.\dot{a_1}a_2a_3\cdots a_{n-1}\dot{a_n} = \frac{a_1a_2a_3\cdots a_n}{10^n - 1}$$

Which is also the formula we learned above: the numerator has the recurring part and the denominator has n 9s, as $10^n - 1$ is always n digits 9 (for instance, for $n = 3$, $10^3 - 1 = 1000 - 1 = 999$).

Compound interest formula derivation

Remember we have r as the interest rate, n as the number of periods and T as the total value after n periods. Let's increment the notation and have P_n, the value invested after n periods. Thus, we start with our principal, P_0. After one period, we have P_1, after two periods P_2, and so on until all n periods.

Now, let's assume we are increasing our principal each period (everything is the same with a minus sign if we are decreasing). After one period, we will increase P_0 by $r\%$. We know that the multiplier for a $r\%$ increase is $1 + \frac{r}{100}$, so we have that P_1 is:

$$P_1 = P_0 \times \left(1 + \frac{r}{100}\right)$$

To calculate P_2, we need to increase P_1 by $r\%$ again, that is, multiply P_1 by $1 + \frac{r}{100}$ again:

$$P_2 = P_1 \times \left(1 + \frac{r}{100}\right)$$

We know, however, that $P_1 = P_0 \times \left(1 + \frac{r}{100}\right)$. If we substitute this in the equation for P_2 we obtain:

$$P_2 = \underbrace{P_0 \times \left(1 + \frac{r}{100}\right)}_{P_1} \times \left(1 + \frac{r}{100}\right)$$

we can now multiply the $1 + \frac{r}{100}$ to obtain

$$P_2 = P_0 \times \left(1 + \frac{r}{100}\right)^2$$

If we repeat the increase for P_3:

$$P_3 = P_2 \times \left(1 + \frac{r}{100}\right)^2$$

and substitute the value for P_2:

$$P_3 = \underbrace{P_0 \times \left(1 + \frac{r}{100}\right)^2}_{P_2} \times \left(1 + \frac{r}{100}\right)$$

we obtain:

$$P_3 = P_0 \times \left(1 + \frac{r}{100}\right)^3$$

notice the pattern: for each P_n, we have the initial principal, P_0, multiplied by the multiplier to the nth power:

$$P_n = P_0 \times \left(1 + \frac{r}{100}\right)^n$$

if we call $P_0 = P$ and $P_n = T$ we have our formula:

$$T = P\left(1 + \frac{r}{100}\right)^n$$

7. Roots and surds

7.1. Why learn roots and surds

If I were a magician, and were to learn how to do magic, I would find it quite useful to learn how to *undo* my magic. Imagine if I were to do something wrong!

Mathematics is the magic anyone can do, and we learned how to do the indices tricks. Now we will work a bit with the *undoing* of indices, which are called roots. Let us do some desentanching.

7.2. Roots

Roots are the inverse of powers[1], so I highly suggest revising indices if you feel the need.

7.2.1. Square roots

We know how to square numbers, for instance

$$3^2 = 9$$

To denote the *inverse operation*, we use the *root* symbol, $\sqrt{}$

$$\sqrt{9} = 3$$

and you can see the full pathway:

$$3 \to 3^2 = 9 \to \sqrt{9} = 3 \to 3$$

where we start with 3 and end up with 3.

As we are undoing the *squaring* operation, we call this root the *square root*. The square root of any number is equal to the number which, when squared, gives you the number inside the square root. For instance

$$\sqrt{25}$$

is asking "what is the number that when we multiply by itself gives 25?". The answer is 5, as $5 \times 5 = 25$, so we have

$$\sqrt{25} = 5$$

[1] They are actually indices, see Indices chapter, section 3, property 8. For instance, $\sqrt{4} = 4^{\frac{1}{2}}$.

and

$$\sqrt{289}$$

is asking "what is the number that when squared gives 289?". As $17 \times 17 = 289$, we have

$$\sqrt{289} = 17$$

Unfortunately, though, there is a smaller complication. Think about $\sqrt{289}$ again: we are searching for a number that, when squared, gives 289. 17 is definitely *one of them*, but what about -17:

$$(-17)^2 = -17 \times -17 = 289$$

Because of this, when we ask "what is the square root of a number?", we are normally asking for the *positive* answer, which is called the *principal* square root.

As we have *two* possible answers (the negative and the positive) to any square root, we sometimes use the symbol $\pm$, which is read as "plus or minus", in front of the number:

$$\sqrt{16} = \pm 4$$

which means that the square roots of 16 are either 4 or -4.

7.2.2. General roots

In the same way we can calculate "what is the number that we square to give" another number using the square root, we can ask "what is the number we have to *cube* to obtain a given number". For instance, we would like to know what number cubed gives 8. We denote that using the cube root symbol:

$$\sqrt[3]{8}$$

which is the same "base" root symbol with the added 3 to the left. Thus, when we write

$$\sqrt[3]{64}$$

we are asking "what number cubed gives 64?". The answer is 4, as $4^3 = 4 \times 4 \times 4 = 64$, so

$$\sqrt[3]{64} = 4$$

In general, we can calculate the n-th root of a number x, which means asking "what is the number that, when raised to the n-th power, gives x?". We denote this by

$$\sqrt[n]{x}$$

For example,

$$\sqrt[5]{32}$$

is saying "find the number that raised to the power of 5 gives 32". As $2^5 = 2 \times 2 \times 2 \times 2 \times 2 = 32$, we have

$$\sqrt[5]{32} = 2$$

Another example

$$\sqrt[6]{729}$$

is telling us to find the number which when raised to the power of 6 gives 729. As $3^6 = 729$ and $(-3)^6 = 729$,

$$\sqrt[6]{729} = \pm 3$$

To summarise, the n-th root of a number x is a number y which, when raised to the power of n, gives the original number. In symbols:

$$\sqrt[n]{x} = y \iff y^n = x$$

and in such a way we can clearly see the relation between indices and roots. Be careful with even roots (square root, fourth root, sixth root, and so on), as they have both the positive and the negative values as solutions.

7.2.3. Numerical roots

If numerical roots appear in your exam and you can use a calculator, just use it. If you cannot use the calculator, then you have to do some trial and error. Say we wanted to find

$$\sqrt[3]{125}$$

which means we need to find a number cubed which is equal to 125. We can create a table:

$$1^3 = 1 \times 1 \times 1 = 1$$

$$2^3 = 2 \times 2 \times 2 = 8$$

$$3^3 = 3 \times 3 \times 3 = 27$$

$$4^3 = 4 \times 4 \times 4 = 64$$

$$5^3 = 5 \times 5 \times 5 = 125$$

and find out that

$$\sqrt[3]{125} = 5$$

The same technique works for any root. To find

$$\sqrt[6]{4096}$$

we calculate powers of 6:

$$1^6 = 1$$

$$2^6 = 64$$

$$3^6 = 729$$

$$4^6 = 4096$$

Therefore,

$$\sqrt[6]{4096} = 4$$

7.2.4. Algebraic roots

To simplify expressions such as

$$\sqrt[3]{x^6}$$

let us first think about what they mean: we are searching for a number that, when cubed, is equal to x^6. Let us call this number y. We can write

$$y^3 = x^6$$

which is a simple equation. One way to make y the subject is to raise both sides to the power of $\frac{1}{3}$:

$$\left(y^3\right)^{\frac{1}{3}} = \left(x^6\right)^{\frac{1}{3}}$$

and we can now use indices properties:

$$\left(y^3\right)^{\frac{1}{3}} = \left(x^6\right)^{\frac{1}{3}}$$

$$y^{3\times\frac{1}{3}} = x^{6\times\frac{1}{3}}$$

$$y = x^2$$

Hence, we have

$$\sqrt[3]{x^6} = x^2$$

which we could also have found by dividing the power of x by 3. We could do the same as above with the n-th root of x (try it!), to obtain the result

$$\sqrt[n]{x^p} = x^{\frac{p}{n}}$$

This means that the n-th root of x^p is simply x to the power of p divided by n.

7.2.5. Properties of roots

7.2.5.1. $\sqrt[n]{1} = 1$

This means that any root of 1 is always 1. This follows from the fact that any power of 1 is 1:

$$1^n = 1$$

7.2.5.2. The product of roots is the root of the product

This simply means that

$$\sqrt{a} \times \sqrt{b} = \sqrt{a \times b}$$

For instance[2]:

$$\sqrt{2} \times \sqrt{8} = \sqrt{2 \times 8} = \sqrt{16} = 4$$

We can also use this to simplify some algebraic expressions, such as

$$\sqrt{16x^4}$$

by writing it as

$$\sqrt{16x^4} = \sqrt{16 \times x^4} = \sqrt{16} \times \sqrt{x^4} = 4 \times x^2 = 4x^2$$

Or

$$\sqrt[4]{81y^{20}}$$

$$\sqrt[4]{81} \times \sqrt[4]{y^{20}}$$

$$3 \times y^{\frac{20}{4}}$$

$$3y^5$$

We can do exactly the same with more terms:

$$\sqrt[3]{8x^{12}y^{21}}$$

$$\sqrt[3]{8} \times \sqrt[3]{x^{12}} \times \sqrt[3]{y^{21}}$$

$$2 \times x^{\frac{12}{3}} \times y^{\frac{21}{3}}$$

$$2x^4y^7$$

7.2.5.3. A warning: $\sqrt[n]{x+y} \neq \sqrt[n]{x} + \sqrt[n]{y}$

I am putting this as it is a classic mistake: a root of a sum is *not* equal to the sum of the roots.

It is easy to see that with a simple example:

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

If we could "split" the sum, we would get

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

and, as $\sqrt{2} + \sqrt{2}$ is definitely not 4, we can see we would be wrong.

[2] If you already studied indices: $\sqrt{2} \times \sqrt{8} = 2^{\frac{1}{2}} \times \left(2^3\right)^{\frac{1}{2}} = 2^{\frac{1}{2}} \times 2^{\frac{3}{2}} = 2^{\frac{1+3}{2}} = 2^2 = 4$.

7.3. Surds

7.3.1. Definition

A *surd* is simply a root which has not been calculated, such as

$$\sqrt{2} \text{ or } 5\sqrt[3]{17}$$

Surds are also called radicals in some texts.

7.3.2. Simplifying square roots

What do we mean by simplifying a square root is "removing" the largest square number we can from it. This may sound weird, but an example would be to write

$$\sqrt{24} = 2\sqrt{6}$$

which you can see are equal by putting them on the calculator. By the way, some calculator already do this simplification for you, so if you are allowed to use it, do it.

Let us first do these kind of simplifications using a longer method, but understand what is going on. If we wanted to simplify $\sqrt{24}$, first we factorise 24:

$$24 = 2^3 \times 3$$

from this, we can write

$$\sqrt{24} = \sqrt{2^3 \times 3}$$

and the trick is to now write everything we can as a power of 2, as it is the inverse operation of a square root. In our case, the only thing we can do is to write

$$2^3 = 2^2 \times 2$$

which gives

$$\sqrt{24} = \sqrt{2^2 \times 2 \times 3}$$

Finally, using property 7.2.5.2 above, we can write

$$\sqrt{24} = \sqrt{2^2} \times \sqrt{2 \times 3}$$

and we can simplify this last to

$$\sqrt{24} = 2 \times \sqrt{6} = 2\sqrt{6}$$

Another example of this longer method:

$$\sqrt{45} = \sqrt{3^2 \times 5} = \sqrt{3^2} \times \sqrt{5} = 3\sqrt{5}$$

The "faster way" is to find the largest square number which is a factor of the number inside the root (also called radicand, for curiosity), and go from there. For instance, to simplify

$$\sqrt{108}$$

you would first figure out the largest square number which is a factor of 108:

$$108 \div 4 = 27 \to 4 \text{ could work}$$

$$108 \div 9 = 12 \to 9 \text{ could work}$$

$$108 \div 16 = 6.75 \to 16 \text{ does not}$$

$$108 \div 25 = 4.32 \to 25 \text{ does not}$$

$$108 \div 36 = 3 \to 36 \text{ could work}$$

$$108 \div 49 = 2.2 \to 49 \text{ does not}$$

$$\text{all above 64 will not work}$$

Hence, we have that 36 is the largest square number which is a factor of 108:

$$\sqrt{108} = \sqrt{36 \times 3} = \sqrt{36} \times \sqrt{3} = 6\sqrt{3}$$

Notice that after you put the largest square factor, everything else is simple manipulation.

This method may look long, but with practice you will not need to try all of the square numbers in order like this, and you will get really good at identifying them.

Some other examples

$$\sqrt{175} = \sqrt{25 \times 7} = \sqrt{25} \times \sqrt{7} = 5\sqrt{7}$$

$$\sqrt{294} = \sqrt{49 \times 6} = \sqrt{49} \times \sqrt{6} = 7\sqrt{6}$$

$$\sqrt{8} = \sqrt{4 \times 2} = \sqrt{4} \times \sqrt{2} = 2\sqrt{2}$$

7.3.3. Simple manipulations

We can treat surds similarly to what we do with variables in algebraic expressions. For instance, we can calculate

$$\sqrt{2} + 3\sqrt{2} = 4\sqrt{2}$$

much in the same way that $x + 3x = 4x$. Also, using the property above, we can multiply surds:

$$\sqrt{5} \times \sqrt{3} = \sqrt{5 \times 3} = \sqrt{15}$$

or

$$2\sqrt{3} \times 5\sqrt{14} = 2 \times 5 \times \sqrt{3} \times \sqrt{14} = 10\sqrt{3 \times 14} = 10\sqrt{42}$$

7.3.4. Expressions with surds

Using the above, we can simplify expressions with surds. Let us see some examples:

$$\sqrt{3}+5\sqrt{3}-7\sqrt{3}+2\sqrt{3}=\sqrt{3}$$

in this one we simply have an "analogue" of $x+5x-7x+2x=x$.

We can do some multiplications and expansion of brackets:

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

in which we expand the brackets normally and just use the property above to join roots, when needed.

A classic is to do something as

$$\left(2+\sqrt{5}\right)^2$$

$$\left(2+\sqrt{5}\right)\left(2+\sqrt{5}\right)$$

$$4+2\sqrt{5}+2\sqrt{5}+5$$

$$9+4\sqrt{5}$$

Sometimes we also need to first simplify the surds and later collect like terms:

$$\sqrt{8}+3\sqrt{2}+5\sqrt{20}$$

$$\sqrt{4\times 2}+3\sqrt{2}+5\sqrt{4\times 5}$$

$$2\sqrt{2}+3\sqrt{2}+5\times 2\sqrt{5}$$

$$2\sqrt{2}+3\sqrt{2}+10\sqrt{5}$$

$$5\sqrt{2}+10\sqrt{5}$$

7.3.5. Rationalisation of the denominator

The name says everything in this particular part: we have a denominator with a surd, such as

$$\frac{17}{\sqrt{5}+1}\text{ or }\frac{1}{\sqrt{2}}$$

and we know that surds are irrational numbers. Hence, the denominator is not rational. Our goal is to remove the surd from the denominator, in order to *rationalise* it.

Both cases below are solved using a similar technique, which basically revolves around the fact that we can multiply the numerator and the denominator of a fraction by the same number.

7.3.5.1. Case 1: surd "alone" in the denominator

Let us use the example above:

$$\frac{1}{\sqrt{2}}$$

We want to get rid of the $\sqrt{2}$. Remember that

$$\sqrt{2} \times \sqrt{2} = 2$$

so if we multiply both the numerator and the denominator of the fraction by $\sqrt{2}$ we will obtain a cuter number on the denominator:

$$\frac{1}{\sqrt{2}} = \frac{1 \times \sqrt{2}}{\sqrt{2 \times \sqrt{2}}} = \frac{\sqrt{2}}{2}$$

and we are done. The denominator is not an irrational number anymore, so it was *rationalised*. Another example:

$$\frac{5\sqrt{3}}{\sqrt{5}}$$

Here we multiply both the numerator and the denominator by $\sqrt{5}$:

$$\frac{5\sqrt{3} \times \sqrt{5}}{\sqrt{5} \times \sqrt{5}} = \frac{5\sqrt{3 \times 5}}{5} = \frac{5\sqrt{15}}{5} = \frac{\cancelto{1}{5}\sqrt{15}}{\cancelto{1}{5}} = \sqrt{15}$$

It does not matter if you have any rational number multiplying the surd on the denominator, or the numerator is more sophisticated, the technique is the same:

$$\frac{2 - \sqrt{5}}{3\sqrt{7}} = \frac{\left(2 - \sqrt{5}\right) \times \sqrt{7}}{3\sqrt{7} \times \sqrt{7}} = \frac{2\sqrt{7} - \sqrt{5} \times \sqrt{7}}{7} = \frac{2\sqrt{7} - \sqrt{5 \times 7}}{7} = \frac{2\sqrt{7} - \sqrt{35}}{7}$$

Therefore, just remember:

> To rationalise fractions of the form
>
> $$\frac{a}{\sqrt{b}}$$
>
> simply multiply both the numerator and the denominator by $\sqrt{b}$ (multiply top and bottom by the bottom).

7.3.5.2. Case 2: surd in an addition or subtraction

Now we are interested in rationalising fractions such as

$$\frac{17}{\sqrt{3}-1}$$

The technique before of multiplying both the numerator and the denominator by the denominator does not work:

$$\frac{17\times\left(\sqrt{3}-1\right)}{\left(\sqrt{3}-1\right)\left(\sqrt{3}-1\right)} = \frac{17\sqrt{3}-17}{3\underbrace{-\sqrt{3}-\sqrt{3}}_{\text{Bad part}}+1} = \frac{17\sqrt{3}-17}{4-2\sqrt{3}}$$

We continue with a surd in the denominator. The trick here is remembering that

$$(a+b)(a-b) = a^2 - ab + ab - b^2 = a^2 - b^2$$

as this cancels out the "bad part" from above! So, instead of multiplying by the denominator, we multiply by *almost* the denominator, we only change the sign of the second term:

$$\frac{17}{\sqrt{3}-1} = \frac{17\times\left(\sqrt{3}+1\right)}{\left(\sqrt{3}-1\right)\left(\sqrt{3}+1\right)} = \frac{17\sqrt{3}+17}{3+\sqrt{3}-\sqrt{3}-1} = \frac{17\sqrt{3}+17}{2}$$

Another example:

$$\frac{2-\sqrt{7}}{3-\sqrt{5}}$$

we multiply by $3+\sqrt{5}$:

$$\frac{2-\sqrt{7}}{3-\sqrt{5}} = \frac{\left(2-\sqrt{7}\right)\left(3+\sqrt{5}\right)}{\left(3-\sqrt{5}\right)\left(3+\sqrt{5}\right)} = \frac{6+2\sqrt{5}-3\sqrt{7}-\sqrt{7}\times\sqrt{5}}{9+3\sqrt{5}-3\sqrt{5}-5} = \frac{6+2\sqrt{5}-3\sqrt{7}-\sqrt{35}}{4}$$

This technique even works when you have a sum of surds in the denominator:

$$\frac{9\sqrt{2}}{2\sqrt{2}+\sqrt{5}} = \frac{9\sqrt{2}\left(2\sqrt{2}-\sqrt{5}\right)}{\left(2\sqrt{2}+\sqrt{5}\right)\left(2\sqrt{2}-\sqrt{5}\right)} = \frac{18\times 2 - 9\sqrt{2\times 5}}{4\times 2 - 5} = \frac{36-9\sqrt{10}}{3} = \frac{\cancel{36}^{12} - \cancel{9}^{3}\sqrt{10}}{\cancel{3}^{1}} = 12-3\sqrt{10}$$

Thus:

> To rationalise fractions of the form
>
> $$\frac{a}{\sqrt{b}\pm c} \text{ or } \frac{a}{c\pm\sqrt{b}}$$
>
> multiply both numerator and denominator by "1st term of the denominator change sign 2nd term of the denominator".

7.4. Exam hints

One suggestion that I give: if your calculator can work with surds, let it do its job. It helps you to avoid early rounding mistakes. Apart from this, very algorithmic topic.

Summary

- **Roots** are the i*nverse operation of powers.* We denote roots using the symbol $\surd$, called a *radical* sometimes;
- The **square root** of a number is any number which *squared* gives the number inside the root, which is sometimes called the *radicand.* For instance

  $$\sqrt{9} = \pm 3$$

  means that both 3 and -3, when squared, give 9, and they are both the square root of 9;
- An important property of algebraic roots is that

  $$\sqrt[n]{x^p} = x^{\frac{p}{n}}$$

  such as

  $$\sqrt[8]{x^{16}} = x^{\frac{16}{8}} = x^2$$
- Any root of 1 is 1:

  $$\sqrt[n]{1} = 1$$
- The root of a product is the product of roots:

  $$\sqrt[n]{xy} = \sqrt[n]{x} \times \sqrt[n]{y}$$
- The root of a sum **is not** the sum of roots:

  $$\sqrt[n]{x+y} \neq \sqrt[n]{x} + \sqrt[n]{y}$$
- A **surd** is a root which we do not calculate, such as $\sqrt{3}$ or $\sqrt[5]{18}$;
- We can simplify square roots by finding the largest square number which is a factor of the radicand:

  $$\sqrt{56} = \sqrt{4 \times 14} = \sqrt{4} \times \sqrt{14} = 2\sqrt{14}$$
- To **rationalise** the denominator of a fraction means *to remove the roots from it*;
  - For fractions of the form

    $$\frac{a}{\sqrt{b}}$$

    simply multiply both numerator and denominator by $\sqrt{b}$;

- For fractions of the form

$$\frac{a}{\sqrt{b} \pm c} \text{ or } \frac{a}{c \pm \sqrt{b}} \text{ or } \frac{a}{\sqrt{b} \pm \sqrt{c}}$$

  multiply both the numerator and the denominator by the first term of the denominator followed by the *opposite* (change sign) of the second term.

8. Standard form

8.1. Why learn standard form (scientific notation)

The world if full of important quantities. The speed of light, the mass of an electron, a bunch of constants like Plank's, the gravitational constant and the list goes on. All these quantities share an important feature: they are either too big or too small, and it is annoying, to say the least, to write them. Take G, the gravitational constant (without units):

$$G = 0.0000000000667$$

which is a very small number (as the gravitational force between two bodies is quite small relatively to other forces). How about one light-year, in metres:

$$\text{light-year} = 9460000000000000\,\text{m}$$

We have to agree that writing these numbers, or similar ones, would eventually get boring to write so much. Standard form, also called scientific notation, is a way to write these numbers in a more convenient way. And, as a bonus, we can operate with numbers in standard form very easily. Basically, standard form is a form of laziness, as it allows us to be write things faster. It is great, then.

8.2. Definition of standard form

A number is said to be in standard form when it is written following this pattern:

$$\underbrace{[-]}_{\text{optional}}\ a \times 10^n$$

in which a is any real number between 1 (including) and 10 (excluding) and n is any integer. In symbolic notation, $a \in \mathbb{R}, 1 \leq a < 10$ and $n \in \mathbb{Z}$. We can write any number using this definition, but before we learn how to do that, let us see some examples of numbers in standard form:

$$5 \times 10^2 \to \quad a = 5, n = 2$$

$$1.23 \times 10^{-3} \to a = 1.23, n = -3$$

$$9.999 \times 10^{17} \to a = 9.999, n = 17$$

$$6.67 \times 10^{-11} \to a = 6.67, n = -11$$

$$1 \times 10^3 \to \quad a = 1, n = 3$$

Notice that on the definition there is an "optional" negative sign, as it does not matter: the number may be negative as well [1], and the methods to convert it to standard form will not change. Here are some numbers in standard form which have a negative in front (I am not considering the negative part of a, but we could):

$$-3.5 \times 10^7 \to \quad a = 3.5, n = 7$$

$$-2.3 \times 10^{-3} \to a = 2.3, n = -3$$

Now, let us some examples of numbers *not* in standard form, and understand why:

$$0.23 \times 10^5 \to a = 0.23 \text{ is not between 1 and 10}$$

$$7.77 \times 10^{-1.5} \to \quad n \text{ is not an integer}$$

$$10 \times 10^{17} \to \quad a = 10$$

We could also write 2 as

$$2 \times 10^0$$

and it would be in standard form, as $10^0 = 1$. However I don't know why you would write that instead of 2.

8.3. Converting to and from standard form

8.3.1. Numbers larger than 10 and smaller than −10

The unifying characteristic of a "large" (either bigger than 10 or smaller than −10) number written in standard form is that the power of 10 is *positive*. Think on the definition of a number in standard form

$$a \times 10^n$$

[1] If you already know absolute value, the better definition would be $1 \leq |a| < 10$.

If n is positive, we are simply multiplying a by a number greater or equal to 10, which means it is getting larger itself. The trick here is precisely to get the ”large part” of the number “contained” in the power of 10.

Some examples of large numbers written in standard form: just remember that the power of 10 *must be positive*:

$$8.1 \times 10^{17}$$

$$1.1 \times 10^{99}$$

$$-2 \times 10^{2}$$

$$7.1 \times 10^{79}$$

All these numbers are meaningless, except the last: it is an estimate of the number of atoms in the observable universe[2].

Let us see now how to convert to and from standard form for “large” numbers.

8.3.1.1. From standard form to “regular” number

To convert a number in standard form back to its “regular” form, we simply do the multiplication. Before we continue, remember that multiplying by a power of 10 is same as “shifting” the decimal point to the right; how much we shift depends on the power. For example, if we multiply a number by 10^5 we would shift the decimal point 5 places to the right.

For instance, if we have

$$1.234 \times 10^{2}$$

we just have to multiply 1.234 by 10^2. This is equivalent to shifting the decimal point two place to the right:

$$1.234 \times 10^{2} = 123.4$$

$\times 10^{2}$

Sometimes we have to add some zeroes just to have a place to put the decimal place. For instance, in

$$7.1 \times 10^{6}$$

we can add zeroes as

$$7.10000000 \times 10^{6}$$

and then shift the decimal point 6 places to the right:

$$7.10000000 \times 10^{6} = 7100000.0$$

$\times 10^{6}$

[2]https://physics.stackexchange.com/questions/47941/dumbed-down-explanation-how-scientists-know-the-number-
Accessed on 14/01/2020.

8.3.1.2. "Regular" number to standard form

To convert a number written in "regular" form to standard form we divide it and multiply by a power of 10 at the same time. This may sound strange, but the idea is that we will move the decimal point of the number to the left, which is the result of dividing a number by a power of 10; however, as we want the number still to be the same, we cancel out our division (the shift) by multiplying by a power of 10.

The explanation is harder than the actual technique. Just remember that, to be in standard form, a number must be between 1 and 10, not including 10. Hence, we will move the decimal point until we have a number between 1 and 10! Look how easy it is. Let's start with

$$12345$$

First let us add the decimal point at the end

$$12345.0$$

Now, 12345.0 is definitely not in standard form: it is bigger than 10 and we don't have our power of 10. Let us solve the first part by shifting the decimal point to the left:

12345.0

4 places

1.2345

This changes the number, but we will fix that. Moving the decimal point to the left 4 times is the same as dividing the number by 10^4 (or multiplying it by 10^{-4}). The inverse operation of a division by 10^4 is a multiplication by 10^4, so after the shift we multiply the number by 10^4:

4 places 12345.0

1.2345 $\underbrace{\times 10^4}$

to cancel out the shift to the left

This is the justification for the method: we simply shift the decimal point to the left until we reach the position we want. As this is equivalent to a division, we multiply the number by 10 to the power of the number of places we shifted the point, and that's it.

Some more examples:

$$890$$

First add the decimal point:

$$890.0$$

then shift to the left and multiply by the appropriate power of 10 to cancel out:

890.0

8.90

As we shifted 2 places to the left, we multiply by 10^2:

$$890.0 = 8.90 \times 10^2$$

Finally, the example of one light-year in metres:

$$9460000000000000$$

Adding the decimal point and counting how many places to the left we shift:

$$9460000000000000.0$$

15 places

$$9.460000000000000$$

as we shifted left 15 places, we multiply by 10^{15}:

$$9460000000000000 = 9.46 \times 10^{15}$$

Thankfully we do not have to write trailing 0s, so I didn't. But you could have written

$$9460000000000000 = 9.460000000000000 \times 10^{15}$$

had you wanted to.

With negative numbers nothing changes, except that you have a negative in front. For instance:

$$-7135$$

to convert this to standard form, we ignore the negative and follow the same steps:

$$-7135.0$$

3 places

$$-7.135$$

and we then multiply by 10^3:

$$-7135 = -7.135 \times 10^3$$

8.3.2. Numbers between -1 and 1

When a number is between -1 and 1, its standard form has a *negative* power of 10. Analogue to the positive power of 10, multiplying by a negative power of 10 is the same as *dividing* by a power of 10, hence the number gets smaller. Here are some numbers of this type written in standard form

$$-3.15 \times 10^{-5}$$

$$9.9 \times 10^{-20}$$

$$6.63 \times 10^{-34}$$

Again these numbers are meaningless, with the exception of the last: it is *Planck constant*, denoted by h.

8.3.2.1. From standard form to "regular" number

Converting from standard form back is the same to these numbers: we do the multiplication. Just remember that, for instance

$$10^{-3} = \frac{1}{1000}$$

and, because of this, multiplying by 10^{-3} is the same as dividing by 1000. Thus, when we multiply by 10^{-n}, we simply use the trick of shifting the decimal point to the left n places. The only difference from the previous case is that always had to add some zeroes before the number (the precise number is not important, I add one more than the absolute value of the power):

$$2.1415 \times 10^{-2} = 0002.1415 \times 10^{-2}$$

$$0002.1415 \times 10^{-2} = 0.021415$$

$$\times 10^{-2}$$

Here we shifted the decimal point twice to the left, as the power of 10 is -2.

With negative numbers, again, we simply ignore the negative and keep it there:

$$-8.99 \times 10^{-3} = -00008.99 \times 10^{-3}$$

$$-00008.99 \times 10^{-3} = -0.00899$$

$$\times 10^{-3}$$

8.3.2.2. "Regular" number to standard form

As for converting to standard form, the technique is the same. However, as we shift the decimal point to the right, this is the same as multiplying by a power of 10. To cancel this multiplication, we have to divide by the same power of 10, but instead of writing this as a division, we write as a product by a negative power of 10. For instance:

3 places

$$0.0037$$

$$0003.7 \quad \underbrace{\times 10^{-3}}_{\text{to cancel out the shift to the right}}$$

Hence the techniques for both types of standard form are similar: we either multiply or divide by a power of 10, and cancel it by multiplying by the appropriate power of 10. Some more examples:

$$-0.4$$
$$\text{1 place}$$
$$-04.0 \times 10^{-1}$$

thus we have $-0.4 = -4 \times 10^{-1}$;

$$0.000012$$
$$\text{6 places}$$
$$000001.2 \times 10^{-6}$$

8.3.3. How I set up the conversion to standard form

To end this, here is how I set up the conversion to standard form. For instance:

$$8765432$$

First I add the decimal point:

$$8765432.0$$

followed by counting how many times I have to shift the decimal point until reaching the correct place:

$$8.\underbrace{7\,6\,5\,4\,3\,2}_{\text{6 places}}.0$$

after counting I multiply by the correct power of 10. As we shifted the decimal point to the left, we "divided" the number, hence we have to multiply by a positive power of 10

$$8.765432 \times 10^6$$

Of course, you can always memorise:

$$\text{Shifted left} \rightarrow \text{Positive power}$$

$$\text{Shifted right} \rightarrow \text{Negative power}$$

For instance, a small negative number:

$$-0.00423$$

Again, counting the number of places to shift:

$$-0.\underbrace{0\,0\,4}_{\text{3 places}}.23$$

finally, as we shifted the decimal point right, we "multiplied": to cancel we multiply by a negative power of -3:

$$-4.23 \times 10^{-3}$$

8.3.4. Converting to, and from, "almost" standard

When we start operating with numbers in standard form, we will sometimes want for the numbers to be "almost" in standard form, in the sense that it will break the part of the definition that a must be between 1 and 10. Other times, the answer we will obtain will not be in standard form, but "almost" there, in the same sense. So it is important to know how to deal with those cases.

It is important to remember that, due to indices properties,

$$10^n \times 10^m = 10^{n+m}$$

so we can do things like

$$10^{15} = 10^1 \times 10^{14} = 10 \times 10^{14}$$

or

$$10^{-6} = 10 \times 10^{-7}$$

Thus, we can "remove" one 10 from the compact index notation, and we will use it to our advantage.

As I said, sometimes our answer is "almost" in standard form, such as

$$13.5 \times 10^5$$

here, the number 13.5 is larger than 10, so the number is not in standard form. To fix that, we ignore the 10^5 and convert only the 13.5 to standard form:

$$13.5 \times 10^5 = \underbrace{1.35 \times 10}_{13.5} \times 10^5$$

and, to finish, we combine the powers of 10 using the property that we can add the powers:

$$13.5 \times 10^5 = 1.35 \times 10^{1+5} = 1.35 \times 10^6$$

Thus, if we end up with a number not in standard form, we simply transform the part which is not correct to standard form and use indices properties to join the power of 10.

Other times, we will want to do precisely the inverse, so that we have a specific power of 10 in the number. For instance, say we have

$$3.14 \times 10^4$$

and we want this power of 10 to be 3 (you will see why in the next section). To do this, we "extract" a 10 from 10^4:

$$3.14 \times 10^4 = 3.14 \times \underbrace{10}_{\text{extract}} \times 10^3$$

which is the same, as we could simplify 10×10^3 to $10^{1+3} = 10^4$. Finally, we could multiply 3.14 by 10:

$$3.14 \times 10^4 = 3.14 \times 10 \times 10^3 = 31.4 \times 10^3$$

and we could use this number to our purposes. We can go to any power we want, say from

$$2 \times 10^{-8}$$

we wanted to have the same number with 10^{-10}:

$$2 \times 10^{-8} = 2 \times 10^{2} \times 10^{-10}$$

as $10^2 \times 10^{-10} = 10^{2-10} = 10^{-8}$, and simplify it:

$$2 \times 10^{-8} = 2 \times 100 \times 10^{-10} = 200 \times 10^{-10}$$

Hence, if you need to change the power to do a calculation, just be careful to use the property correctly and do the correct calculation to obtain an equivalent power.

8.4. Operations with numbers in standard form

8.4.1. Multiplication and division

Multiplying and dividing numbers in standard form is very convenient, and it is definitely one of the reasons we use it!

For instance:

$$2.4 \times 10^{13} \times 4 \times 10^{8}$$

The first important thing here is that in your exams this will normally be written as

$$\left(2.4 \times 10^{13}\right) \times \left(4 \times 10^{8}\right)$$

which does not make any difference, but makes it easier to see we are multiplying two numbers in standard form. If they appear like this, simply ignore the brackets!

The basic idea is to use commutative property of multiplication: the order does not matter, thus

$$2.4 \times 10^{13} \times 4 \times 10^{8} = 2.4 \times 4 \times 10^{13} \times 10^{8}$$

and what we do now is to multiply 2.4 by 4 and use the property on the indices part of 10:

$$\begin{aligned} 2.4 \times 10^{13} \times 4 \times 10^{8} =& \overbrace{2.4 \times 4}^{\text{do this}} \times \underbrace{10^{13} \times 10^{8}}_{\text{property}} \\ =& 9.6 \times 10^{13+8} \\ =& 9.6 \times 10^{21} \end{aligned}$$

Notice, then, that multiplying two numbers in standard form is very straightforward: multiply the numbers in front and add the powers of 10. Or, if you feel more formal:

$$(a \times 10^{n}) \times (b \times 10^{m}) = ab \times 10^{n+m}$$

There is one possible catch: sometimes the resulting number, ab, may be larger than 10, so we have to adjust the result to be in "standard" standard form. For instance,

$$\left(3 \times 10^2\right) \times \left(5 \times 10^{-8}\right)$$

$$3 \times 5 \times 10^{2+-8}$$

$$15 \times 10^{-6}$$

in which 15 is larger than 10. So we do what we learned how to do in the previous section: ignore the 10^{-6} and write 15 in standard form:

$$15 \times 10^{-6} = \underbrace{1.5 \times 10}_{15} \times 10^{-6}$$

finally, simplify the power of 10:

$$15 \times 10^{-6} = 1.5 \times 10 \times 10^{-6} = 1.5 \times 10^{-5}$$

With division the idea is very similar. If we wanted to calculate

$$\frac{8 \times 10^6}{4 \times 10^2} = \left(8 \times 10^6\right) \div \left(4 \times 10^2\right)$$

we could rearrange it to

$$\frac{8 \times 10^6}{4 \times 10^2} = \frac{8}{4} \times \frac{10^6}{10^2}$$

and simply divide 8 by 4 and use the division property of indices (subtract the powers) on the second term:

$$\frac{8 \times 10^6}{4 \times 10^2} = \frac{8}{4} \times \frac{10^6}{10^2} = 2 \times 10^{6-2} = 2 \times 10^4$$

Thus, to divide numbers in standard form, we simply divide the numbers in front and subtract the powers of 10. Or, with a little bit of flair

$$\frac{a \times 10^n}{b \times 10^m} = \frac{a}{b} \times 10^{n-m}$$

When dividing, we can get a result smaller than 1 for $\frac{a}{b}$, and again we need to fix the standard form:

$$\frac{3 \times 10^{17}}{6 \times 10^{10}} = \frac{3}{6} \times 10^{17-10} = 0.5 \times 10^7$$

in which we write 0.5 as 5×10^{-1} and simplify the powers of 10:

$$\frac{3 \times 10^{17}}{6 \times 10^{10}} = \frac{3}{6} \times 10^{17-10} = 0.5 \times 10^7 = \underbrace{5 \times 10^{-1}}_{0.5} \times 10^7 = 5 \times 10^6$$

Of course, if you have negative numbers, sign rules are still valid:

$$\left(1.3 \times 10^{-9}\right) \times \left(-3 \times 10^{10}\right) = 1.3 \times -3 \times 10^{-9} \times 10^{10} = -3.9 \times 10^{1}$$

as positive times negative is negative[3].

And we can solve expressions:

$$\frac{\left(2 \times 10^{3}\right) \times \left(4.5 \times 10^{8}\right)}{\left(3 \times 10^{-2}\right)} = \frac{9 \times 10^{3+8}}{3 \times 10^{-2}} = \frac{9 \times 10^{11}}{3 \times 10^{-2}} = 3 \times 10^{11--2} = 3 \times 10^{13}$$

as usual, be careful with signs when using the properties, particularly the division one!

8.4.2. Addition and subtraction

8.4.2.1. Case 1: same powers of 10

Consider $2x + 3x = 5x$. In this equation, x can take any value we wish. In particular, we could make $x = 10^3$:

$$2 \times 10^{3} + 3 \times 10^{3} = 5 \times 10^{3}$$

and this needs to be correct. Notice that we could have simply added 2 and 3 and copied the $\times 10^3$. That is precisely how we add (or subtract) numbers in standard form when they have the same power of 10: simply add or subtract the numbers in front of the power of 10.

For instance,

$$9 \times 10^{4} - 3 \times 10^{4} = \underbrace{6}_{9-3} \times 10^{4}$$

$$2.5 \times 10^{17} + 4 \times 10^{17} = \underbrace{6.5}_{2.5+4} \times 10^{17}$$

Sometimes, though, we will need to adjust the number to be in proper standard form:

$$4 \times 10^{-3} + 8 \times 10^{-3} = 12 \times 10^{-3}$$

as 12 is larger than 10, we fix the result:

$$12 \times 10^{-3} = \underbrace{1.2 \times 10}_{12} \times 10^{-3} = 1.2 \times 10^{1+-3} = 1.2 \times 10^{-2}$$

This can also happen with numbers smaller than 1:

$$2 \times 10^{2} - 1.8 \times 10^{2} = 0.2 \times 10^{2} = \underbrace{2 \times 10^{-1}}_{0.2} \times 10^{2} = 2 \times 10^{-1+2} = 2 \times 10$$

Apart from this extra step, if the powers of 10 are equal, simply add or subtract the numbers in front.

[3] Have you ever wondered why? If you are curious, check the formality after taste!

8.4.2.2. Case 2: different powers of 10

If the powers of 10 are different, we cannot simply add or subtract the numbers in front. However, we can *make* the powers equal using the idea in section 8.3.4.

For instance, when calculating

$$8 \times 10^4 + 9 \times 10^5$$

we cannot add 8 with 9, as the powers of 10 are different (the idea of $8x+9x = 17x$ does not work, as we have different values for x!). However, we can write 9×10^5 differently by extracting a 10 from 10^5:

$$9 \times 10^5 = 9 \times 10 \times 10^4 = 90 \times 10^4$$

After doing this, we can simply do the addition:

$$8 \times 10^4 + 9 \times 10^5 = 8 \times 10^4 + 90 \times 10^4 = 98 \times 10^4$$

and, finally, correct the result to proper standard form, as 98 is larger than 10:

$$98 \times 10^4 = \underbrace{9.8 \times 10}_{98} \times 10^4 = 9.8 \times 10^{1+4} = 9.8 \times 10^5$$

You could also rewrite 8×10^4 with a power of 5:

$$8 \times 10^4 = 0.8 \times 10 \times 10^4 = 0.8 \times 10^5$$

and then add

$$8 \times 10^4 + 9 \times 10^5 = 0.8 \times 10^5 + 9 \times 10^5 = 9.8 \times 10^5$$

which in this case is already in proper standard form.

A final example:

$$3 \times 10^{-5} - 3.5 \times 10^{-6}$$

Let us rewrite first 3×10^{-5} with a power of -6:

$$3 \times 10^{-5} = 3 \times \underbrace{10 \times 10^{-6}}_{10^{-5}} = 30 \times 10^{-6}$$

and now we can subtract:

$$3 \times 10^{-5} - 3.5 \times 10^{-6} = 30 \times 10^{-6} - 3.5 \times 10^{-6} = 26.5 \times 10^{-6}$$

which needs to be written in proper standard form:

$$26.5 \times 10^{-6} = \underbrace{2.65 \times 10}_{26.5} \times 10^{-6} = 2.65 \times 10^{1+-6} = 2.65 \times 10^{-5}$$

Now let us rewrite 3.5×10^{-6} with a power of -5:

$$3.5 \times 10^{-6} = 3.5 \times 10^{-1} \times 10^{-5} = 0.35 \times 10^{-5}$$

and now subtract:

$$3 \times 10^{-5} - 3.5 \times 10^{-6} = 3 \times 10^{-5} - 0.35 \times 10^{-5} = 2.65 \times 10^{-5}$$

which is already correct.

Sometimes you need to "extract" even more powers of 10:

$$5 \times 10^7 + 9 \times 10^5$$

Let us write 9×10^5 with a power of 7:

$$9 \times 10^5 = 9 \times 10^{-2} \times 10^7 = 0.09 \times 10^7$$

and now add:

$$5 \times 10^7 + 9 \times 10^5 = 5 \times 10^7 + 0.09 \times 10^7 = 5.09 \times 10^7$$

From these examples, I suggest always rewriting the number with the smallest power of 10 with the largest power you have, as the answer will probably be already in standard form.

Sometimes we have to solve equations with numbers in standard from

8.5. Exam hints

Be careful when counting to convert to standard form! It is very easy to lose marks by miscounting. Avoid the temptation to write numbers back in regular form, it will make your life harder.

If you have to solve equations or expressions with numbers in standard form, remember that everything that is true for numbers is true for them written in standard form as well, particularly BIDMAS.

Summary

- A number is said to be written in **standard form** if it is the form

  $$[-]a \times 10^n$$

  where $1 \leq a < 10$ and n is an integer;

- To convert a number to standard form, first ignore its sign. Then,
  - if a is larger than 10, shift the decimal place of the number to *the left* until you reach the first digit. Multiply the result of this step by 10 to the power of the number of places you shifted the decimal place to the left:

$$12345.0$$

4 places

$$1.2345 \times 10^4$$

- if a is smaller than 1, shift the decimal place of the number to *the right* until you reach the last digit. Multiply the result of this step by 10 to the power of minus the number of places you shifted the decimal place to the left:

$$-0.00423$$

$$-0\underbrace{.0.0.4}_{\text{3 places}}.23 = -4.23 \times 10^{-3}$$

- When *multiplying or dividing* numbers in standard form, simply multiply or divide the numbers in front and use indices properties on the powers of 10:
  - For multiplication,

$$(a \times 10^n) \times (b \times 10^m) = ab \times 10^{n+m}$$

  - For division,

$$\frac{a \times 10^n}{b \times 10^m} = \frac{a}{b} \times 10^{n-m}$$

- When *adding or subtracting* numbers in standard form:
  - If they have the same powers of 10, simply add or subtract the numbers in front;
  - If they do not have the same powers of 10, rewrite all the terms with the same power of 10 and add or subtract the numbers in front;
- Remember to always write the result of your calculations in proper standard form.

Formality after taste

Justification for the sign rules

Way 1

A very simple way to show that

$$a \times -b = -a \times b = -ab$$

is the following:

$$
\begin{aligned}
0 &= 0 && \text{Which is true!}\\
a \times 0 &= 0 a \times 0 = 0\text{, so still true. Here } a > 0\\
a \times (b + -b) &= 0 && b + -b = 0\text{, so still true}\\
a \times b + a \times -b &= 0 && \text{Expanding brackets}\\
a \times b - a \times b + a \times -b &= 0 - a \times b && \text{Subtracting } a \times b \text{ on both sides}\\
a \times -b &= -a \times b\\
a \times -b &= -(ab)
\end{aligned}
$$

and a similar argument shows that

$$-a \times -b = ab$$

Way 2

This is a more formal way, using the axioms of the real numbers. I suggest reading Moise, *Elementary Geometry from an Advanced Standpoint* to learn more about this.

What we want to show is that

$$(-a)b = -(ab)$$

for every a and b. This statement is the same as showing that

$$(-a)b + ab = ab + (-a)b = 0$$

as this is the definition of the *negative* of a number ($a + -a = -a + a = 0$). Hence, we will be showing that $-(ab)$ is the negative of $(-a)b$, as we want.

We just need to show that $(-a)b + ab = 0$ (commutative property):

$$
\begin{aligned}
(-a)b + ab =&\\
=&((-a) + a)b\text{Distributive property}\\
=&0b \qquad -a + a = 0\\
=&0 \qquad 0b = 0
\end{aligned}
$$

Of course, we first needed to prove that $0b = 0$, but let us assume that it true.

Thus, we have proved that a negative times a positive is negative (and that positive times a negative is negative).

To prove that negative times negative is positive we use some properties of numbers and the fact that $-(-x) = x$ for all x:

$$
\begin{aligned}
(-a)(-b) =& \\
&= -\left(a(-b)\right) \\
&= -\left((-b)a\right) \\
&= -\left(-(ba)\right) \\
&= ba = ab \quad -(-x) = x \text{ for all } x
\end{aligned}
$$

9. Time, distance and speed

9.1. Why learn about time, distance and speed

Distance, speed and time are concepts which deal with important fundamental concepts of our reality. Distance is about our position in space, speed how our position changes, and time is about... I don't really know. Defining time is tricky.

The important thing is that we can measure all time, distance and speed, and we will be learning how to work with those measurements.

9.2. The metric prefixes

It is important to know these by heart:

- kilo means $10^3 = 1000$: 1 kilogram means 1×1000 grams. We write a k for kilo: kg is kilo gram;
- centi means $10^{-2} = \frac{1}{100}$ (dividing by 100): 1 centimetre means $1 \times \frac{1}{100} = 0.01$ centimetres. We write c for centi: cm means centi metre;
- milli means $10^{-3} = \frac{1}{1000}$ (dividing by 1000): 1 millilitre means $1 \times \frac{1}{1000} = 0.001$ litres. We write m for milli: mL means milli litre

There are many others[1], but those are the most important to know.

9.3. Calculations with time

Before we continue, a disclaimer: I will not be defining time. The only important thing to us about it is that we can measure time. Particularly, we are interested in how much time goes on between two events.

9.3.1. Basic units of time

As with all measurements, we can measure time using different unities. Some of them (the most important to us in our IGCSE context) are:

- A second;
- A minute, a span of 60 seconds;

[1] See `https://en.wikipedia.org/wiki/Metric_prefix` for them.

- An hour, a span of 60 minutes[2];
- A day has 24 hours.

With these units we can use the "day to day" conversions we know: to convert minutes to seconds, multiply by 60; to convert seconds to minutes, divide by 60. The others are similar.

A little more interesting is when they want us to convert minutes to hours. For instance, 850 minutes to hours. To do this, first divide 850 by 60, and just take the whole part: $850 \div 60 = 14.16\ldots$. Thus, we have 14 hours in 850 minutes, with some remaining minutes. To obtain those, we can do $14 \times 60 = 840$, which is the number of minutes in 14 hours, and see how much we are missing to 850: $850 - 840 = 10$. Thus, 850 minutes is the same as 14 hours and 10 minutes.

You can also do that using "integer" division:

$$850 \div 60 = 14\text{r}10$$

which already gives you the answer nicely.

9.3.2. Time differences

A common type of question is when they give you two times, and want you to calculate the difference, how much time there is between, them. For instance, how many hours and minutes are there between $10:34$ and $18:10$? Even though there are techniques to to do this faster, I like doing it using the following steps:

$$10:34 \xrightarrow{+26\text{min}} 11:00 \xrightarrow{+7\text{h}} 18:00 \xrightarrow{+10\text{min}} 18:10$$

and then I add the times above the arrows: $26 + 10 = 36$min and 7h with those gives 7h36min.

Another example: the difference between $08:42$ and $03:51$ of the next day:

$$08:42 \xrightarrow{+18\text{min}} 09:00 \xrightarrow{+15\text{h}} \underbrace{00:00}_{\text{same as 24h}} \xrightarrow{+3\text{h}} 03:00 \xrightarrow{+51\text{min}} 03:51$$

When we add the hours we get $15 + 3 = 18$h. The minutes: $18 + 51 = 69$, which is bigger than 60, so we can get an hour from it: 69min is the same as 1h and 9min. In total, then, we have $18 + 1 = 19$ hours and 9 minutes.

9.3.3. Converting between decimal and "usual time"

Sometimes we need to convert times in standard use, such as 2 hours and 18 minutes, to a decimal representation. By decimal representation I mean something as 2.5 hours. This is a classic cause of mistakes: many would say that 2h18min is equal to 2.18, and that is very wrong!

[2] See `https://www.scientificamerican.com/article/experts-time-division-days-hours-minutes/` for the history behind these values.

The problematic part is always the minutes: the hours are just themselves. To convert 18 minutes to decimal time, simply *divide it by* 60. We do that because the minutes are a division of an hour:

$$18\text{min} = \frac{18}{60} \text{ of an hour} = 0.3\text{h}$$

Thus, 2h18 minutes is the same as 2.3 hours.

To convert from decimal hours to usual, subtract the whole number of hours and multiply the 0.something obtained by 60. In our case, 2.3 hours is the same as 2hours plus 0.3 of an hour:

$$0.3 \text{ of an hour} = 0.3 \times 60 = 18 \text{ minutes}$$

and you obtain the 2h18minutes from the start.

9.4. Distance, speed and acceleration

9.4.1. Distance

Distance is a measure of how "much space" there is between two points. Interestingly enough, in math we can generalize distance using a metric function[3], but we are only concerned with usual distance here.

We usually measure distance using the metric system: metres and its derivations, such as kilometres, centimetres and friends.

9.4.2. Speed

To define it poorly[4], speed is a measure of how fast we are moving through space: how much has our position changed in a given span of time. This is very important: speed is how much space changes with time.

Normally, in the IGCSE, we are more concerned with average speed. We define average speed as the distance traveled divided by the time the journey took:

$$\text{average speed} = \frac{\text{distance traveled}}{\text{time it took}}$$

Unfortunately to us, laziness is common here, and it is very usual to write

$$\text{speed} = \frac{\text{distance}}{\text{time}}$$

or, using short-hand notation

$$s = \frac{d}{t}$$

[3]See `https://en.wikipedia.org/wiki/Metric_(mathematics)`.

[4]Speed is a part of velocity: when we move through space we have a direction and how fast we do it, called its magnitude. The magnitude of velocity, which is the rate of change of space with time, is called speed.

I will be using this equation throughout this chapter, but remember that when I say speed I mean average speed. A pet-peeve of mine.

We can use this equation to find the speed of an object. For instance, if I take 50 seconds to move 300 metres, my speed is given by

$$s = \frac{d}{t} = \frac{300\text{m}}{50\text{s}} = 6\frac{\text{m}}{\text{s}}$$

Thus, my speed is 6m/s. Units are very important in this context, so be careful with them.

The same equation can be used to find the distance we move given our speed and time, simply be rearranging the equation[5]. And, as we are dealing with physical quantities, the equations have to make sense.

For example, say we are moving at 5km/h and we need to go somewhere which is 15km away, and we want to calculate how long it will take. Thus, we have a distance and a speed, and we want to find the time of the movement. Speed is how much distance we move in a given span of time, so the time a journey take should be found by dividing the distance by the speed. The equation gives that as well:

$$\begin{aligned} s &= \frac{d}{t} && \text{Definition of speed} \\ s \times t &= \frac{d}{t} \times t && \text{Multiplying both sides by } t \\ st &= d \\ \frac{st}{s} &= \frac{d}{s} && \text{Dividing both sides by } s \\ t &= \frac{d}{s} \end{aligned}$$

Hence, time can be found by dividing the distance by the speed. In our case:

$$t = \frac{d}{s} = \frac{15\text{km}}{5\text{km/h}} = 3\text{h}$$

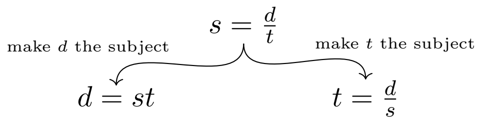
If we make distance the subject:

$$\begin{aligned} s &= \frac{d}{t} && \text{Definition of speed} \\ s \times t &= \frac{d}{t} \times t && \text{Multiplying both sides by } t \\ st &= d \\ d &= st \end{aligned}$$

[5]Or you can use one of those disgusting triangles. I highly recommend actually learning how to manipulate an equation, though. Useful skill!

we obtain another equation, which lets us find the distance traveled: we multiply speed by time. Which again, makes sense: speed is how our position is changing with time, so if we multiply it by time, we obtain the distance traveled.

In summary, we have "3 equations", but we only have to really know the definition of speed, and we can obtain the others:



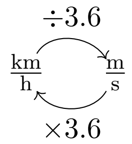
9.4.3. Units of speed conversion

9.4.3.1. Kilometres per hour to metres per second

In this section we will reach a very simple way to convert kilometres per hour to, and from, metres per second.

$$1\frac{\text{km}}{\text{h}} = \frac{1000\text{m}}{\underbrace{60}_{\text{minutes}} \times \underbrace{60}_{\text{seconds}}} = \frac{1000\text{m}}{3600\text{s}} = \frac{1}{3.6}\frac{\text{m}}{\text{s}}$$

Hence, 1km/h is the same as $\frac{1}{3.6}$m/s. Thus, to convert from km per hour to metres per second, we simply divide the quantity by 3.6. The opposite goes to convert from metres per second to kilometres per hour: multiply by 3.6:



Let us see some examples:

- Convert 72km/h to m/s:

  We could redo the whole reasoning above (I recommend that you try), we can simply divide 72 by 3.6. As $\frac{72}{3.6} = 20$, we have:

  $$72\text{km/h} = 20\text{m/s}$$

- Convert 5m/s to km/h:

  Now we just have to multiply by 3.6. As $5 \times 3.6 = 18$, we have:

  $$5\text{m/s} = 18\text{km/h}$$

9.4.3.2. "Weird" conversions

If you have to convert anything else, you have to know the definition of the units given and manipulate them.

For instance, say we would like to convert 12km/h to metres per minute:

$$12\frac{\text{km}}{\text{h}} = 12\frac{1000\text{m}}{60\text{min}} = \frac{12 \times 1000\text{m}}{60\text{min}} = 200\frac{\text{m}}{\text{min}}$$

Hence, 12km/h is the same as 200m/minute.

9.4.4. Acceleration

Acceleration is a measure of how much speed is changing with time. Not very common in the exams, but average acceleration can be calculated by

$$\text{average acceleration} = \frac{\text{final speed} - \text{initial speed}}{\text{time interval for the change}}$$

9.5. Distance-time graphs

It is time for a story.

Everyday, at 06:30, I unfortunately wake up (I hate waking up early!). For 30 minutes I prepare myself to go to school. I then take my bike and go to school, where I arrive at 07:40. The distance to the school from my house is around 2.5km. On this particular day, at 12:30, I go to the market near the school to buy chocolate, as no one deserves apple for desserts. Also I have a coffee. This market is 2km from my house. I arrive at the market at 12:45, stay there for 30 minutes and return to school at 13:30. I then work until 16:00, and I go back home walking as my legs cannot take anymore. I need 45 minutes to arrive at my home, at 17:00.

This most exciting tale of my day can be put in a *distance-time* graph: a graph which associates my distance from a particular point at each moment in time. In my case, we can measure the distance from my house. Thus, when I am at my house I am "at distance 0" and when I am at school, for instance, my distance from my house is 2.5km.

Thus, we have the following set of coordinate points:

$$(\text{time}, \text{position})$$

$$(06:30, 0) \rightarrow \text{at home}$$

$$(07:00, 0) \rightarrow \text{at home still}$$

$$(07:20, 2.5) \rightarrow \text{at school}$$

$$(12:30, 2.5) \rightarrow \text{leave school}$$

$$(12:45, 2) \rightarrow \text{at market}$$

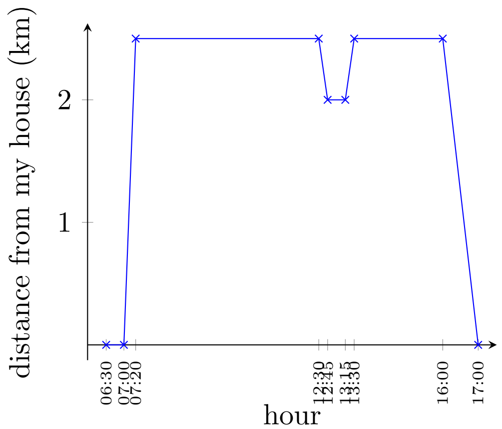
$$(13:15, 2) \rightarrow \text{leave market}$$

$$(13:30, 2.5) \rightarrow \text{at school again}$$

$$(16:00, 2.5) \rightarrow \text{leave school}$$

$$(17:00, 0) \rightarrow \text{back at home}$$

which we can put in a graph:



We can now ask questions about my adventure. What was the total distance travelled during my day? To find that, we simply add the distance I moved for each stage of the journey. By stage we refer to each different "part", whenever the line changes its behaviour:

Stage	Distance (km)
06:30-07:00	0
07:00-07:20	2.5
07:20-12:30	0
12:30-12:45	0.5
12:45-13:15	0
13:15-13:30	0.5
13:30-16:00	0
16:00-17:00	2.5

and we can now add all those distances:

$$2.5 + 0.5 + 0.5 + 2.5 = 6\text{km}$$

We could calculate my average speed on the day. From 06:30 to 17:00 there are 10h30, but we need to write it as 10.5 hours:

$$\text{average speed} = \frac{\text{total distance}}{\text{total time}} = \frac{6}{10.5} = \frac{4}{7}\text{km/h}$$

Finally, we can find what was my speed at any particular stage. And here enters the interesting part: as speed is the rate of change of position with time, we can find speed on a distance-time graph by finding gradients!

For instance, on the stage from 12:45-13:15, I was at the market. Hence, as I was not moving, my speed was 0. Looking at the graph, you can see the gradient of the line is 0, as it is horizontal! Let us find my speed on my back, during the 07:00-07:20 stage:

$$\text{speed} = m = \frac{2.5 - 0}{20 \text{ min}} = \frac{2.5}{\frac{1}{3}\text{h}} = 7.5\text{km/h}$$

It is very important to remember what we can find using a distance-time graph:

- We can find the distance travelled (obvioULy);
- We can find the average speed by dividing the total distance traveled by the time it took to move that distance;
- We can find the speed of a particular stage by finding the gradient of the line on that stage.

9.6. Speed-time graphs

In the same way we have a graph of our distance from a point over time, we can have a graph of our speed with time.

Say we will run a bit. From rest, we speed up until we reach a speed of 8m/s in 10 seconds. We then maintain our speed for anoter 10 seconds and I start to stop, which takes another 5 seconds. Hence, we have the following points:

$$\text{(time (s), speed (m/s))}$$

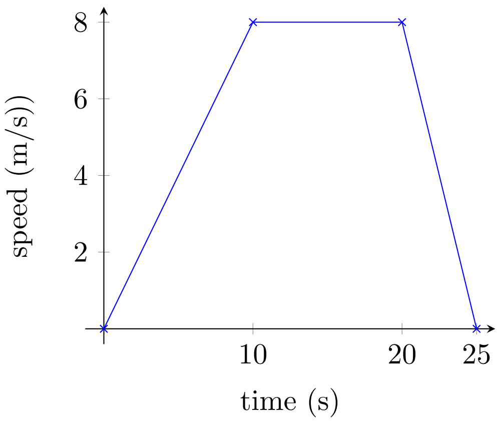
$$(0, 0) \to \text{at rest}$$

$$(10, 8) \to \text{after accelerating}$$

$$(20, 8) \to \text{my endurance}$$

$$(25, 0) \to \text{I give up}$$

which we can graph:



Again, we can ask some questions, such as "what is our highest speed" and so. But those are simple interpretation.

The interesting questions are:

- what was the distance we travelled?
- what was our acceleration?

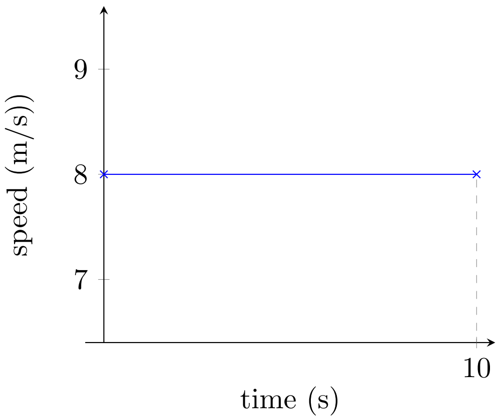
To answer the first, we can make a very loose argument. Remember average speed is given by

$$\text{average speed} = \frac{\text{distance}}{\text{time}}$$

and we can rearrante this to

$$\text{distance} = \text{average speed} \times \text{time}$$

If we had a very simple speed graph, such as



the average speed is 8m/s, as the speed does not change, and the time passed is 10 seconds. To calculate the distance traveled, we could multiply speed by time, which is geometrically equivalent to the *area of the rectangle on the graph.* This same idea is true for any speed graph: *the area of a speed-time graph is equal to the distrance traveled.*

In our case, then, we have a trapezium with bases of length 25 and 10, and height 8:

$$\text{distance traveled} = \text{area} = \frac{(25+10)\times 8}{2} = 140\text{m}$$

As for the acceleration, in the same way that speed is the change of rate between position and time, acceleration is rate of change of speed with time. Hence, to find acceleration in a speed-time graph, we find the gradient of the line! In the second stage, for instance, our speed did not change, and the gradient of a horizontal line is 0. The first stage has acceleration given by

$$\text{acceleration} = \text{gradient} = \frac{8-0}{10-0} = \frac{4}{5} = 0.8\text{m/s}^2$$

and the last stage

$$\text{acceleration} = \text{gradient} = \frac{0-8}{25-20} = \frac{-8}{5} = -1.6\text{m/s}^2$$

Sometimes negative acceleration is called deceleration, and you could say the deceleration on the last stage was 1.6m/s^2.

In summary, in a speed-time graph:

- The area under the curve is equal to the total distance traveled;
- The gradient of the line is equal to the acceleration.

9.7. Exam hints

This is a very common topic, and luckily it also has some overlap with Physics. Be very careful with units in graphs, remember to make sure they are compatible (seconds in time, metres per second in speed, for instance).

Summary

- To calculate *time differences* (how much time between two moments), use the "arrow method". For instance, betweem 10:34 and 18:10 therea are

$$10:34 \xrightarrow{+26\text{min}} 11:00 \xrightarrow{+7\text{h}} 18:00 \xrightarrow{+10\text{min}} 18:10$$

  $26 + 10 = 36$ minutes and 7 hours, hence a total of 7h36;

- Be careful when working with time. Usual time is not in decimal base, so to convert 2h45 minutes you have to

$$2 + \frac{45}{60} = 2\frac{3}{4} = 2.75\text{hours}$$

- Distance is a measure of space between two points;
- Speed is how fast our distance is changing with time: speed is the rate of change of space with time;
- Average speed is given by

$$\text{average speed} = \frac{\text{total distance traveled}}{\text{time it took}}$$

- Acceleration is the rate of change of speed with time, that is, how speed is changing. You can find average acceleration with the formula

$$\text{average acceleration} = \frac{\text{final speed} - \text{initial speed}}{\text{time interval for the change}}$$

- To convert from km/h to m/s divide by 3.6;
- To convert from m/s to km/h multiply by 3.6;
- In a distance-time graph:
  - The gradient of the curve is equal to the speed;
- In a speed-time graph:
  - The area under the curve is equal to the total distance traveled;
  - The gradient of the line is equal to acceleration.

10. Set notation and Venn diagrams

10.1. Why learn set notation

Regarding the "real world", I am sorry to inform that I have no knowledge of this topic being "useful".

However, if you consider mathematics by itself, sets have a fundamental role in our communication and concepts. In the preface to his amazing book *Naive Set Theory*, Paul Halmos claims that

> "Every mathematician agrees that every mathematician must know some set theory; the disagreement begins in trying to decide how much is some"

Hence, we definitely should learn some set theory, as well versed humans in mathematics. How much, though? For now, let us focus on the contents of the IGCSE syllabus.

10.2. Sets: mathematical bags

We are not going to define what sets are here, and we will use the general notion we all have of sets: a collection of things, which we will call elements. Bags, in a way: you can put things in the bags, including other bags. Then, you can ask very interesting questions such as "Is there an elephant in your bag?" or "how many things do you have in your bag?", which are the type of questions we need to learn how to answer.

What I want you to take from this brief discussion on what sets are is that sets are collections of elements, and the most important question we would like to answer will be "is there some particular element in your bag?". You will see soon why this is important.

10.3. How to write sets

10.3.1. Lists

Before we begin, some conventions. We will denote sets by capital letters, A, B and so on.

A very straightforward way to say that set A has some elements is to *list* those elements. Say that A has the elements blue and red. We list that by using curly braces with the elements within:

$$A = \{\text{blue, red}\}$$

If set B contains all even numbers smaller than 10, we can write

$$B = \{0, 2, 4, 6, 8\}$$

And it is totally fine to have a set with elements that are sets. Say set C has as its elements the sets A and B above:

$$C = \{A, B\}$$

Of course, it is somewhat boring to list sets with many elements. Say set D is the set with all positive whole numbers from 1 to 1000. Instead of listing all those numbers, we can use ellipsis:

$$D = \{1, 2, 3, \ldots, 998, 999, 1000\}$$

where the ... indicate that the pattern continues.

10.3.2. Some special sets: empty set and universal set

We need to start with the very important concept of the "universal set". This set contains all possible elements we are interested in. For instance, think on the people in your maths class: you form a set whose elements are people, and the universal set would be a set containing every single one of you. Or, if we are thinking about the whole numbers from 1 to 10, the universal set would be those numbers.

We denote the universal set in the IGCSE using a weird E letter: $\mathscr{E}$. Therefore, if our universal set contains the whole numbers from 1 to 10, we write

$$\mathscr{E} = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$$

or

$$\mathscr{E} = \{1, 2, 3, \ldots, 9, 10\}$$

if we are feeling a little lazy.

The empty set, on the other hand, would be a set which has no elements. It is denoted by

$$\varnothing$$

or sometimes by the curly braces without anything inside:

$$\varnothing = \{\}$$

10.3.3. Descriptions using rules

Sometimes, however, it is much easier to define a set by stating a rule that its elements must follow. Say you want a set $\mathbb{P}$ with all prime numbers below 1000. Writing that would be very tedious. You could, definitely, do something as

$$\mathbb{P} = \{\text{primes below 1000}\}$$

However, there is a more "formal" way. To do that, we use what is called "set-builder notation" or "set comprehension[1]". This is a fancy name to say that we will describe

[1] If you are studying computer science and learning Python, for instance, you can define lists in much the same way. Ask your teacher!

the properties that the elements of the set must have. It works like this:

$$\underbrace{X}_{\text{name of the set}} = \left\{ \underbrace{x}_{\text{variable}} \middle| \text{ properties that } x \text{ much satisfy to be part of } X \right\}$$

or, instead of using the vertical bar as a separator, a colon

$$\underbrace{X}_{\text{name of the set}} = \left\{ \underbrace{x}_{\text{variable}} : \text{properties that } x \text{ much satisfy to be part of } X \right\}$$

Don't despair, some examples will make it very clear.

Say you want the set of even numbers bellow 100, and let's call the set A. We could write:

$$A = \{x : 1 \leq x \leq 100, x \text{ is even}\}$$

this can be read as: "A is the set that contains all numbers x such that x is between 1 and 100 and is even". To figure out if a number is part of A, you would substitute values for x and check if they satisfy the conditions. Say, for instance, if $x = 8$: then it is a number between 1 and 100 and is even, so 8 is an element of A. If $x = 102$, on the other hand, is greater than 100, so it is not part of A.

Another example:

$$B = \{d : d \text{ is a day of the week starting with T}\}$$

here we could substitute all seven days of the week for d, and the ones that start with T would be part of B:

$$B = \{\text{Tuesday, Thursday}\}$$

Final example, one that defines a straight line:

$$L = \{(x, y) : x + y = 5\}$$

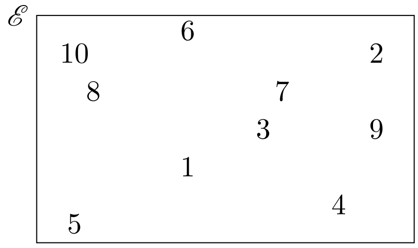
here we have a point in the coordinate plane as our variable, and to be an element of set L the x and y coordinates need to add up to 5.

To summarize, just read this notation as "all numbers x which satisfy whatever is to the right of the | or :".

10.3.4. Representing sets using Venn diagrams

A Venn diagram is a visual representation of a set, and it really helps to visualize some of the operations on the next session.

Say that our universe set is $\mathscr{E} = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$, all integers between 1 and 10. We can represent this universe by the following diagram:

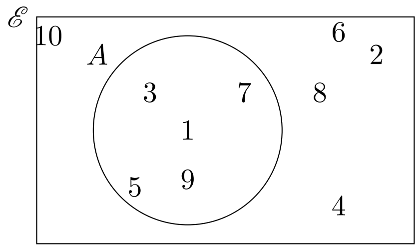


The rectangle represents the universal set, $\mathscr{E}$, and its elements are spread around within the rectangle. It sounds stupid, but the rectangle, being the "whole universe" in this context, must have all the elements of $\mathscr{E}$.

Let us add a set to our universe, $\mathscr{E} = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$. Say set A is the set

$$A = \{1, 3, 5, 7\}$$

that is, all odd numbers of our universe. We can represent this in a Venn diagram by:

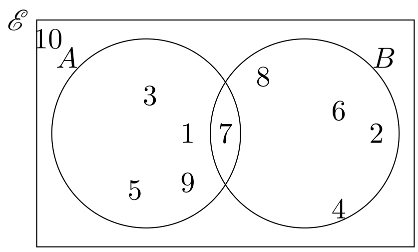


Now, the elements of A are written inside the circle representing A, whereas the numbers of the universe which are not part of A are outside the circle (but still inside the rectangle).

We can represent as many sets within our universal set rectangle as we want (and they do not need to be represented by circles, it is just more common). For example, let us add another set, B:

$$B = \{2, 4, 6, 7, 8\}$$

which has some of the even numbers and 7, because why not. We represent our sets using a diagram such as this:



Now we have two circles, the left one representing A and the right one representing B. Notice that 10, which is not an element of neither A nor B is still outside both circles; 7, on the other hand, is part of *both* A and B, so it appears in the section of the diagram where the circles overlap.

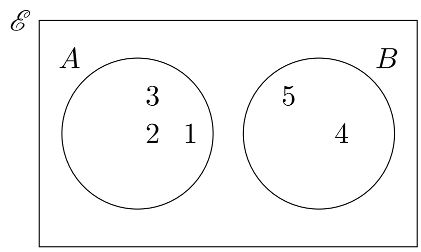
We normally draw the sets like this, with overlap, but it is totally fine for them to have different "arrangements". For instance, given the sets

$$\mathscr{E} = \{1, 2, 3, 4, 5\}$$

$$A = \{1, 2, 3\}$$

$$B = \{4, 5\}$$

the Venn diagram would look like this:



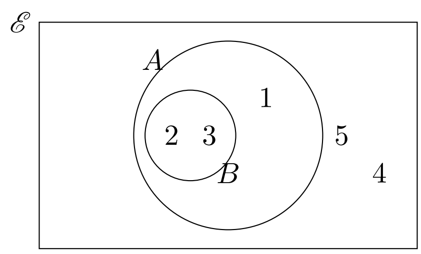
A and B have no common elements, so they do not overlap. Also, there is no number outside A and B as they contain all elements of the universe.

Another interesting arrangement is when a set is within the other. Say we have

$$\mathscr{E} = \{1, 2, 3, 4, 5\}$$

$$A = \{1, 2, 3\}$$

$$B = \{2, 3\}$$



Here, as B is a "part" of A (you will learn on the next section the correct terminology), B is completely within A.

Let us do a final example with 3 sets to practice. Say we have the following sets:

$$\mathscr{E} = \{\text{whole numbers from 1 to 10}\}$$

$$A = \{x : 2 \leq x \leq 5\}$$

$$B = \{1, 3, 5, 7, 9\}$$

$$C = \{2, 4, 6\}$$

Whenever you see sets like this, my suggestion is to write explicitly all the sets, by writing them as lists of elements, when they are not already that. So, here, we write the universal set and set A as lists:

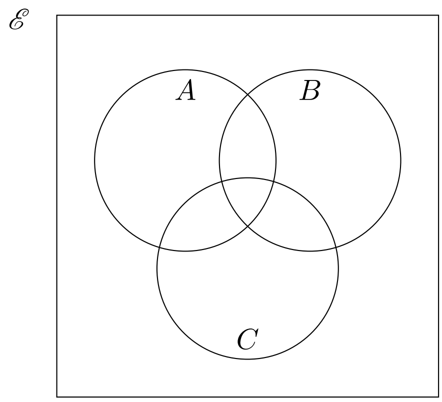
$$\mathscr{E} = \{\text{whole numbers from 1 to 10}\} = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$$

$$A = \{x : 2 \leq x \leq 5\} = \{2, 3, 4, 5\}$$

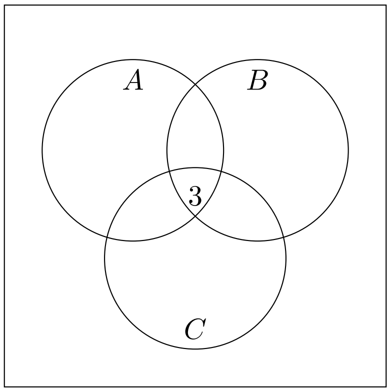
$$B = \{1, 3, 5, 7, 9\}$$

$$C = \{2, 3, 4, 6\}$$

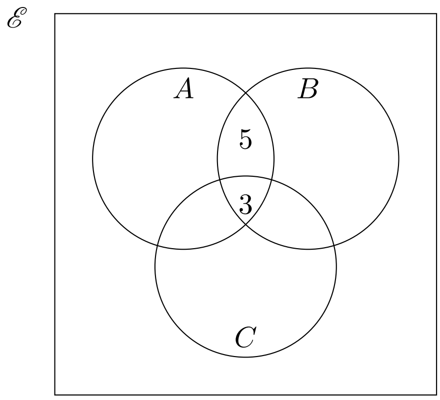
Now we need to fill a 3 set Venn diagram, which looks like this:



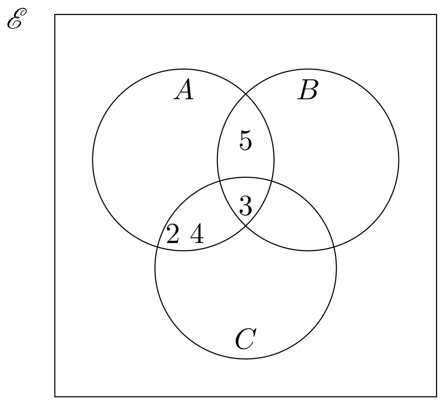
Best way is to *always start by the regions with more sets.* Here, it is the middle section, which is where the elements in all sets should go. In this case, the only element in all A, B and C is 3, so let's add it there:



Now we move to the overlaps of sets two by two (which means the overlaps of each set with only one other set; 3 is in the overlap of the sets with the two other sets). Sets A and B have in common the elements 3 and 5, and as 3 was already put, we put only 5 in their overlap:

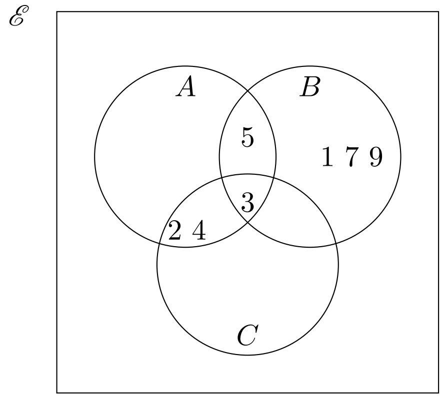


For the intersection of A and C, we have that they have in common 2, 3 and 4. As 3 was already put, we add 2 and 4 to their overlap:

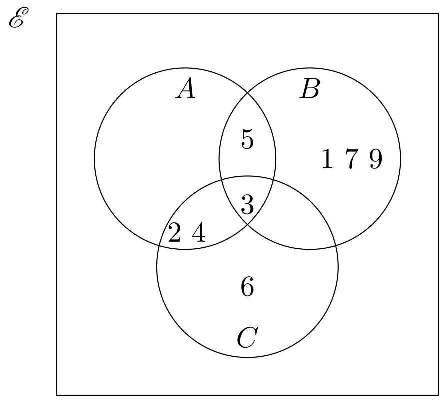


Last intersection, of B and C: they have only 3 in common, which is already in the diagram. So, we leave that part blank.

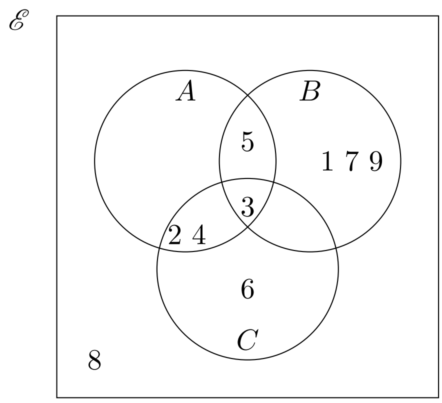
We now have left only the numbers which are exclusively in a single set. In A, nobody is left: $2, 3, 4$ and 5 are already in the diagram. For B, we are missing 1,7 and 9, which we add to the section of the disk representing B that does not overlap with any other disk:



C still needs its 6:



We are only missing 8, which is part of the universal set but not part of A, B or C:



An important thing to keep in your mind is that 3, for instance, is still part of the "overlap" of A and B. Remember that if a number is inside a disc, it is an element of the set represented by that disk.

10.4. Set operations

10.4.1. The most important one: pertinence or belonging (an element is in a set)

(Well, at least in my opinion!)

Remember when I said that a set is like a "mathematical bag"? The most important thing you can ask of a set is "is something a part of you?" We use a special symbol to state that something is part of a set, $\in$.

For instance, say we have the set

$$A = \{2, 3, 5, 7\}$$

The number 2 is part of A, so we write

$$2 \in A$$

which is read as "2 is in A" or "2 is a member of A" or "2 belongs to A". The number 4, on the other hand, is not part of A, so we write

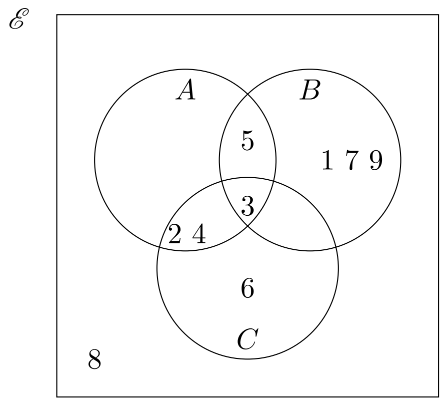
$$4 \notin A$$

and we read that as "4 is not in A", or "4 is not a member of A" or "4 does not belong to A" .

One thing that is important to reflect on is that given belonging is the most important question we can ask of elements of a set, it does not make sense to have repeated elements in a set: having more than one 2 in A above would not change the fact that $2 \in A$. This is one justification to why sets cannot have equal elements. Also, this is why the order in which we write the elements of a set is irrelevant: $\{1, 2, 3\}$ is the same set as $\{3, 1, 2\}$.

In summary, the symbol $\in$ is read as "is in" or "belongs", and we use it to show belonging to a set. Its opposite is the symbol $\notin$, which shows not belonging to a set.

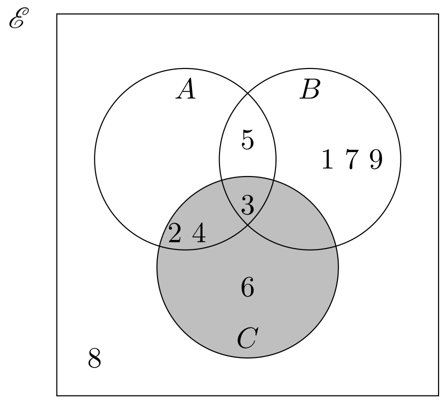
In a Venn diagram, we can represent belonging by writing the element within the disc that represents a set. For instance, in the last example of the previous section:



we can write

$$6 \in C$$

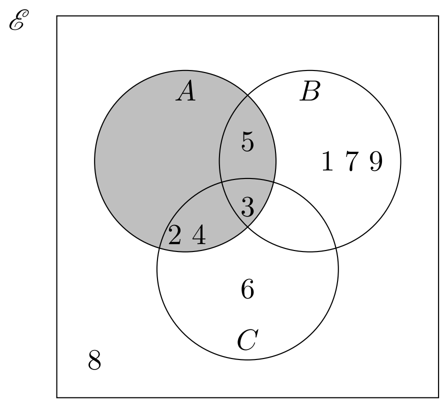
as 6 is within the disc that represents set C, as you can see by shading all C:



We can also write

$$3 \in A$$

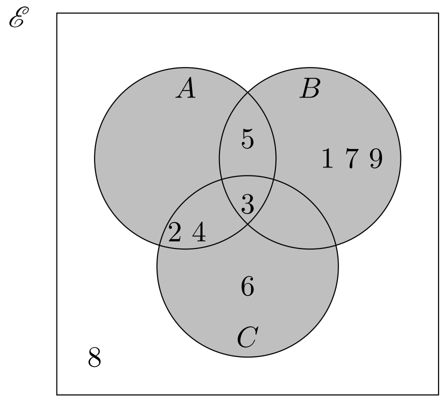
as 3 is part of the disc that represents A, and it does not matter that 3 also belongs to sets B and C, which can be seen by shading A:



Another true statement is

$$8 \notin B$$

as 8 is not part of any set but the universal one, which we can see by shading A, B and C:



10.4.2. Number of elements in a set

Another very "intuitive" question we can ask of a set is "how many elements does it have?" If we have a set A, we use this notation to write "number of elements":

$$n(A) : \text{number of elements of set } A$$

So, if we have the set

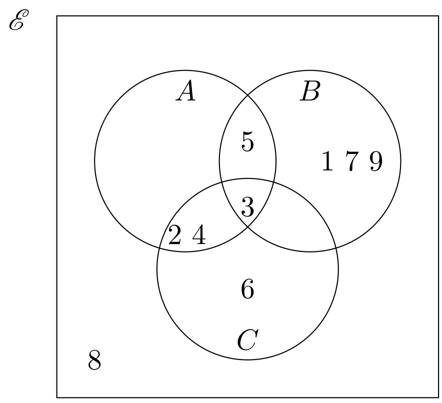
$$X = \{1, 3, 5, 7, 9, 17\}$$

we can write

$$n(X) = 6$$

as there are 6 elements in set X.

In a Venn diagram, to find the number of sets in a set, we simply count the elements within each disk. Again, in the example:



We can find $n(A)$ by counting all the numbers within the A disk, which are $5, 3, 2$ and 4. Hence,

$$n(A) = 4$$

Be careful not to think that just because an element is on an overlap it is not part of the individual sets.

10.4.3. The complement of a set

The complement of a set is its "opposite": if an element is not part of A, it is part of the complement of A. Another way to think of how to refer to the complement of A is by saying *not* A. We denote the complement of a set A in the IGCSE by

$$A'$$

So, let us say we have

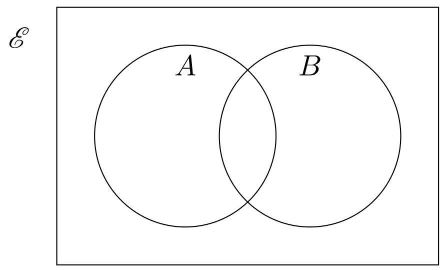
$$\mathscr{E} = \{1, 2, 3, 4, 5\}$$

$$A = \{1, 2, 3\}$$

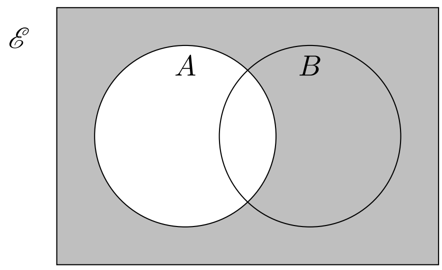
To find A', the complement of A, we simply look at the universe and get all elements there *which are not part of* A. Here, those are 4 and 5. Thus:

$$A' = \{4, 5\}$$

Using a Venn diagram, the complement of a set is the region *outside* the shape that represents the set. For instance, for the classic two set Venn diagram:



the complement of A is the region shaded:



10.4.4. Subsets

Now, say that we have the sets

$$A = \{1, 2, 3, 4, 5\} \quad \text{and} \quad B = \{2, 3\}$$

Notice that every element of B is also part of A. Whenever this happens, we say that B is a *subset* of A, and we denote that by

$$B \subset A$$

(To help you notice the pattern, this symbol works exactly as the $<$ one: $2 < 3$ means 2 is smaller than 3, that is, the "open part" of the symbol is near to the biggest number. For subsets is the same, the "open part" of $\subset$ is to "larger set", the one the other is part of).

We can also change the order and say that "A is a *superset* of B" or that "A contains B":

$$A \supset B$$

but this is much less standard.

The similarities to the inequalities symbols do not stop there. We can also say that a set is a subset of another *or equal to* it by using the $\subseteq$ symbol. For instance,

$$A \subseteq A$$

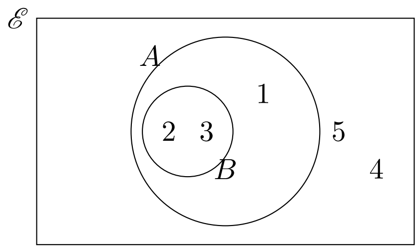
as A is equal to itself (we have not defined set equality, but two sets are equal it they have the same elements).

Whenever a set is a subset of another and is not equal to the "larger" set, such as

$$B \subset A$$

we also say that B is a *proper subset* of A.

In a Venn diagram, subsets are represented when a set disc is within another disc, such as this example from the last session:



here $B \subset A$.

We can also say that a set is not a subset of another:

$$A \not\subset B$$

which means that A is *not a subset of* B. Same thing with a proper subset:

$$A \not\supseteq B$$

means that set B is not a proper superset of A.

10.4.5. Union of sets

Let us say we have two sets of students. Set A represents those that take arts, and set B those who do biology:

$$A = \{\text{Victoria, Laura, Sophia, Victor}\} \text{ and } B = \{\text{Sophia, Victor, Carolina}\}$$

It would be quite useful to make a set of all students above, somewhat like “adding” them together. We do have to define what we mean by that, as we cannot just write all the names in a big set as there would be repetitions. Hence, the proper way to *unite* both sets is to just add any element which appear in A *or* in B, which we denote by

$$A \cup B = \{\text{Victoria, Laura, Sophia, Victor, Carolina}\}$$

and read as “A union B” (sometimes people say “A or B” as well).

Thus, the *union* of sets A and B is another set $A \cup B$ which includes all elements which are either in A or in B. Notice that it is totally fine to be in both (such as Sophia and Victor in our example): to be an element of the union you have to be a member of at least one of the sets.

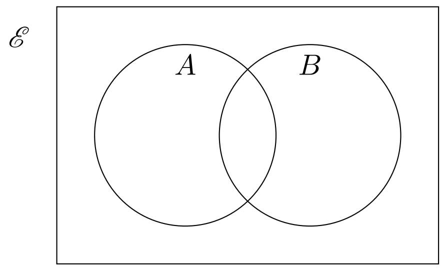
Another example. We have the sets

$$X = \{1, 3, 5, 7, 17\} \text{ and } Y = \{2, 4, 7, 17\}$$

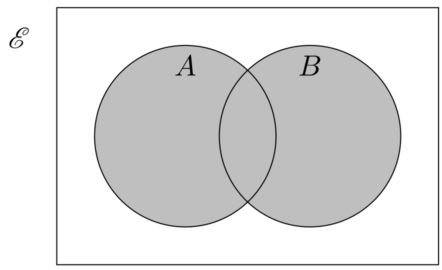
To find their union we just put any element that appears either in X or in Y in the result:

$$X \cup Y = \left\{ \underbrace{1, 3, 5}_{\text{in } X}, \underbrace{2, 4}_{\text{in} Y}, \underbrace{7, 17}_{\text{both} X \text{ and } Y} \right\}$$

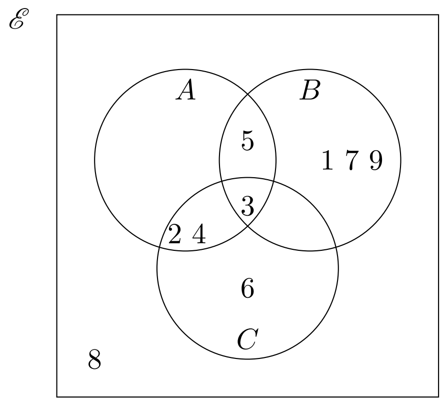
In a Venn diagram, the union on two sets is the region within both sets discs. In a “standard” two sets Venn diagram, such as:



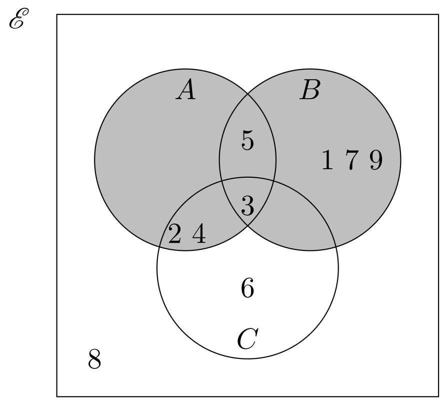
the union of A and B is the region inside both circles:



Thus, any element which is inside the shaded area is part of $A \cup B$.
For instance, in this example:



the set $A \cup B$ is the shaded region:



hence,

$$A \cup B = \{2, 4, 5, 3, 1, 7, 9\}$$

10.4.6. Intersection of sets

Say we have a set C which has as elements some people that have a cat:

$$C = \{\text{Lyra, Aluvian, Elaine}\}$$

and set D has as elements some people that have a dog:

$$D = \{\text{Aluvian, Rand, Mat}\}$$

and all these 5 people make up our universal set:

$$\mathscr{E} = \{\text{Lyra, Aluvian, Elaine, Rand, Mat}\}$$

A question we could ask is if there is anyone in our universe which has *both* a cat and a dog. Thinking on our sets, that would be the same as finding an element which is both in set C and in set D. In this case, we have Aluvian as our answer. Whenever we want to find the elements which are part of two sets *at the same time*, we are finding a new set called the *intersection* of the original sets. We denote the intersection of C and D by

$$C \cap D = \{\text{Aluvian}\}$$

and we read this as "the intersection of C and D is the set which contains Aluvian". Sometimes we also say "C and D is the set that contains Aluvian".

Let us do a numeric example. Say we have the sets

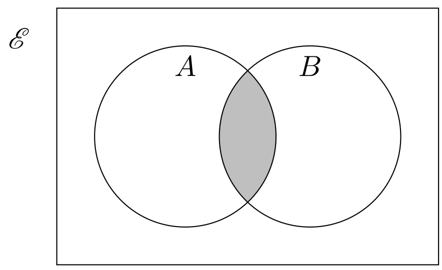
$$A = \{1, 2, 3, 4\}$$

$$B = \{2, 3\}$$

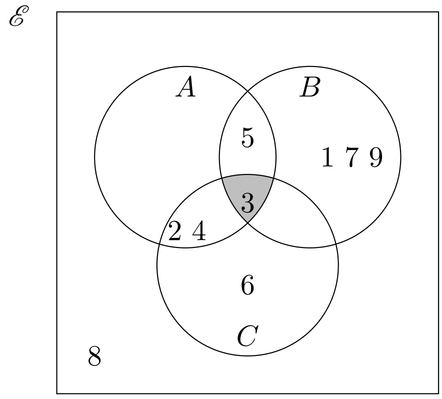
and we want to find the intersection of A and B, that is, the elements which are *both in* A and B:

$$A \cap B = \left\{ \underbrace{2, 3}_{\text{both in } A \text{ and } B} \right\}$$

In a Venn diagram, the intersection of two sets is the overlap between the circles that represent the sets. In a two set Venn diagram, for instance:



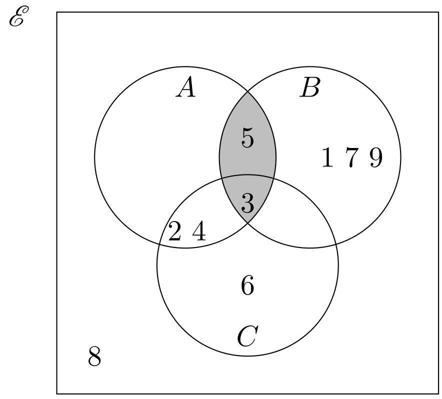
When we have 3 sets, we have two different types of intersections. The first one is where all the 3 sets overlap:



which we can write:

$$A \cap B \cap C = \{3\}$$

However, we also have the intersections of sets two-by-two. For example, we can find $A \cap B$:



which gives us

$$A \cap B = \{3, 5\}$$

Be careful, though: sometimes the intersection of two (or more) sets is the empty set. For instance, the sets

$$X = \{17\}$$

$$Y = \{0\}$$

have no elements in common, hence

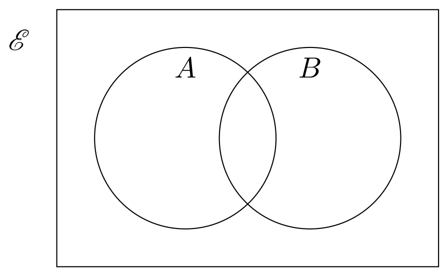
$$X \cap Y = \varnothing$$

10.5. Venn diagram shading

10.5.1. Shading the result of an operation in Venn diagrams

A very common type of question in the IGCSE is to, given a Venn diagram, to shade the region of it which is the result of a given operation.

For instance, given the Venn diagram:



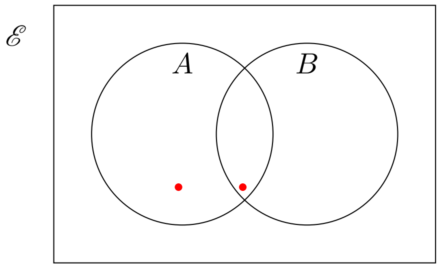
we need to shade the region which is the result of

$$\left(A \cap B'\right)$$

I think the easiest way to do this is by using what I call the "little ball method" (it probably has a proper name). This method works as:

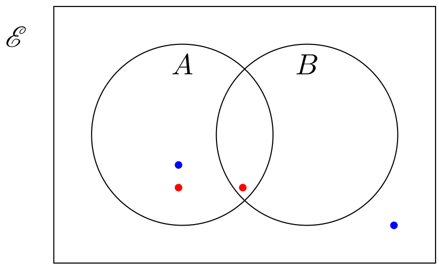
1. For each set in the operation (here A and B'), draw a little ball (of different colours helps!) in the diagram in every region that set is;
2. Look at the operation we need to do (in this case the intersection) and "join" what the little balls tell us.

Let me exemplify with the problem above. First, for set A, add a little red ball in all regions of the diagram where A is:



There are two regions where A lives in the diagram: the part which is "just A" and where A intersects B, hence the two little balls.

Now, for B', we draw a little blue ball everywhere which is *not* B, that is, outside of B:

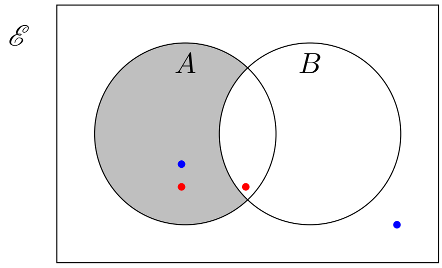


Not B is found outside both A and B and within the regions which only has A.

Finally, we have to "join" what the little balls tells us. We want to find

$$A \cap B'$$

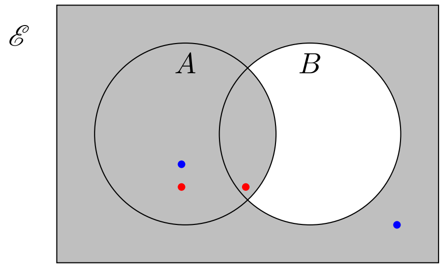
the intersection of A and B'. We have little balls on the regions these sets are, we need the regions which they *intersect*, where *both* a red ball and a blue ball are, that is, the regions where we have *two* little balls, one red one representing A, and a blue one, representing B':



What if, instead of the intersection, we wanted to find

$$A \cup B'$$

Then, we need the regions which have *either* A *or* B', that is, any region that has at least one little ball:



So, for the union, *wherever there is a little ball* (and the colours do not matter!), then you are inside the region you want.

Now, let us recap the rules of the "little ball method":

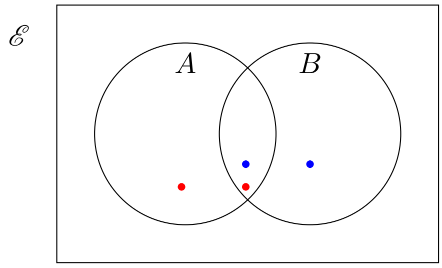
1. For each set in the expression, add a little ball to the regions that set is (possibly of different colours for each set);
2. Depending on the operations:
   a) If the operation between the sets is *union*, any region that has a little ball is what you want;
   b) If the operation between the sets is *intersection*, any region that has a little ball of each colour (or two balls of the same colour) is what you want.
   c) If you have a complement of an operation, you want the places which *do not satisfy* your restriction.

Let us do more examples.

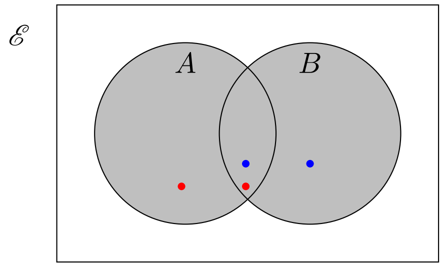
Say we need to shade the regions of the Venn diagram which satisfy:

$$(A \cup B)'$$

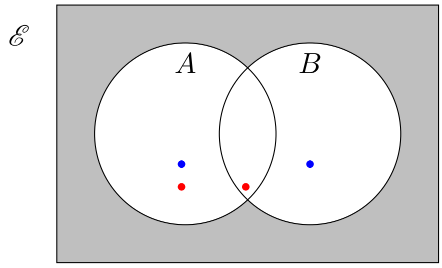
In this case, we first have to do the operation inside the brackets, and then the region will be the opposite because of the complement. Let us start by adding a red little ball to the regions of A, and a little blue ball to the regions of B:



for the union of A and B, we would select any region that has a little ball:



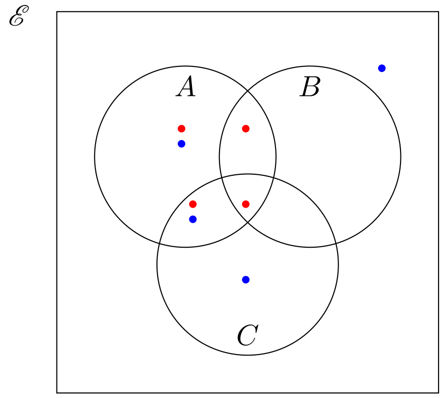
But we want the complement of this, that is, everything which is not shaded. Thinking in terms of the little balls, the "outside" region is the only one that has no little balls, so we shade that part:



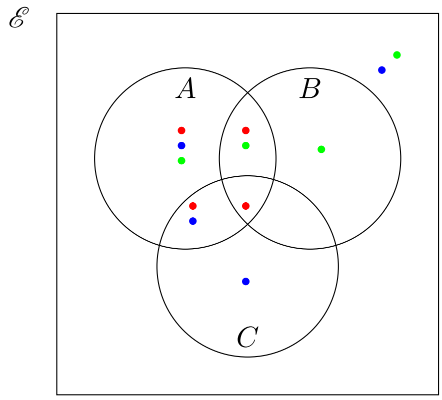
This technique also works when we have 3 sets. Say we want to shade the region that corresponds to

$$(A \cup B') \cap C'$$

We start by the brackets, as usual. Let us add a red little ball to A and a blue little ball to B' (remember that B complement is everything outside of B):



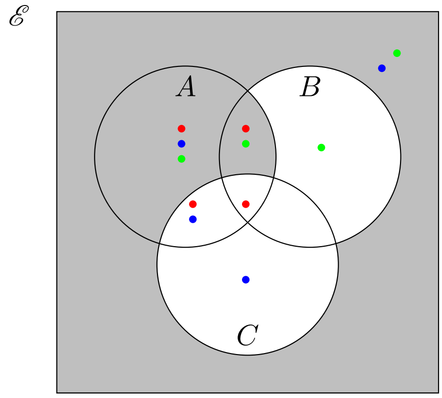
As we want to find $A \cup B'$, anywhere we have a little ball is part of the solution to the brackets. Now, let us add green little balls to the regions corresponding to C':



Finally, we need to combine our lovely little balls. Remember we were finding the region which represents

$$(A \cup B') \cap C'$$

The brackets are any region which has *either* a red or a blue ball (union is either). We need, then, to intersect those regions with the green balls, which means finding all regions which have *at least one blue or red ball* ***and*** *a green ball.* Thus, the region we want is:

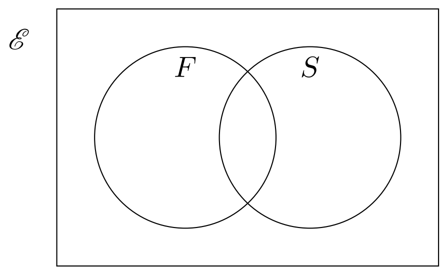


10.6. Solving problems with sets and Venn diagrams

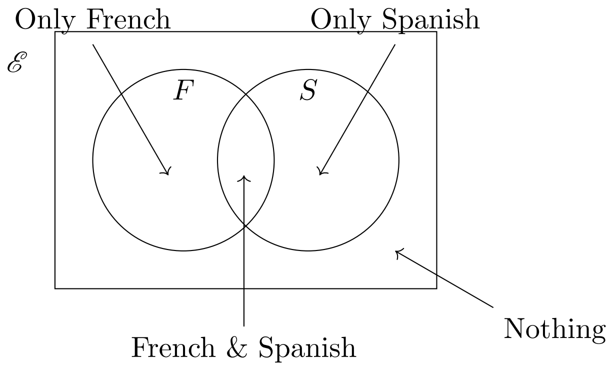
Sets are "useful" to solve problems such as this, though:

> In a school, 10 students take French, 8 students take Spanish and 4 take both French and Spanish. There are also 3 students that take neither. Find the total number of students in the school.

We can draw a Venn diagram representing the situation, where F represents the set of students that take French, and S represents the students that take Spanish. However, be careful: we will not be putting the students themselves in the regions, we will put *how many* are in each region. This will always be clear from context (either the numbers in the diagram are the elements themselves or the number of elements in that region). Here is our basic two sets Venn diagram:

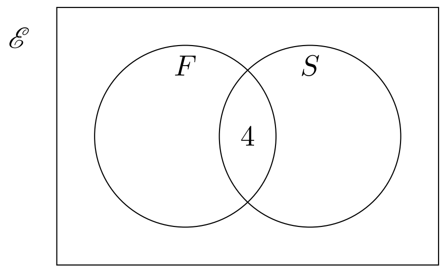


Before we start filling the diagram, here is what region in the set represents:



It is always a good idea to do this annotating of the diagram until you get used to what each region represents.

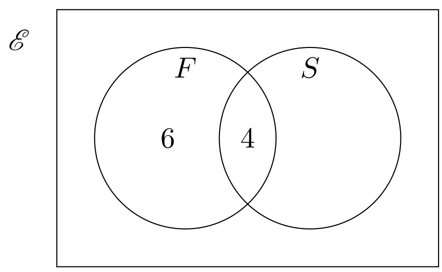
Now, to start filling the diagram, remember: *always start at the intersection.* So, here, we need to add the 4 students that take both languages to the intersection of the two circles:



Let us find now the number of students that take *only* French. Notice that when the problem states that 10 people take French, they are saying that 10 people *in total* take French, including those which take both. As we already counted the 4 people which take both, there are

$$\underbrace{10}_{\text{total French}} - \overbrace{4}^{\text{French \& Spanish}} = \underbrace{6}_{\text{only French}}$$

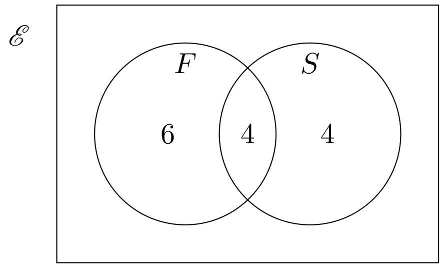
Thus, we can add 6 to the other region of set F:



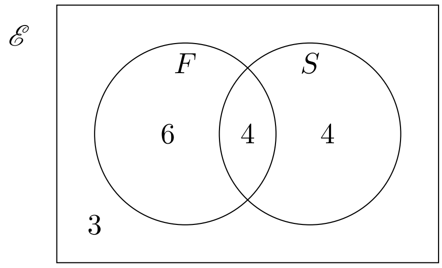
Using the same reasoning for Spanish, we have

$$\underbrace{8}_{\text{total Spanish}} - \overbrace{4}^{\text{French \& Spanish}} = \underbrace{4}_{\text{only Spanish}}$$

which we can add to the diagram:



Finally, we add the number of people that do not study French nor Spanish to the space outside the sets F and S:



We can now add all these values to determine the total number of students:

$$6 + 4 + 4 + 3 = 17$$

Sometimes, the information can be given using set notation:

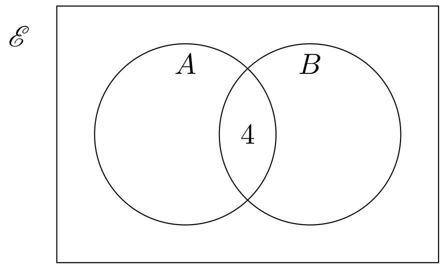
$$n\left(\mathscr{E}\right) = 30$$

$$n\left(A \cap B\right) = 4$$

$$n\left(A \cup B\right)' = 3$$

$$n\left(A \cap B'\right) = 12$$

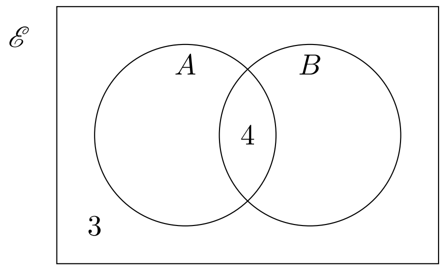
and they ask us to find $n\left(B\right)$. We need to fill the diagram, and remember to *start from the intersection*. In this example, they already told us the number of elements in the intersection: $n\left(A \cap B\right) = 4$, which we can add immediately:



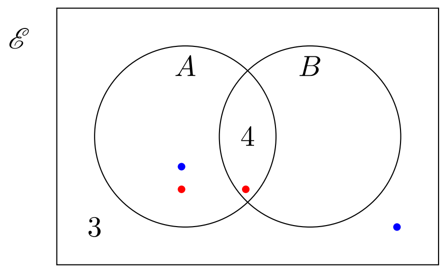
The other information is a little bit harder to add. Let us start with

$$n\,(A \cup B)' = 3$$

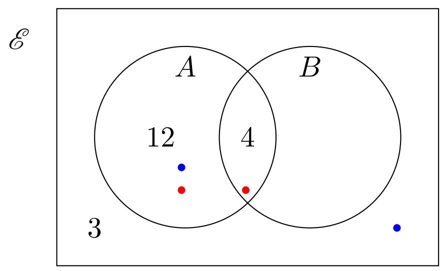
This means that the number of elements in $(A \cup B)'$ is 3, that is, *outside* (the complement) of A union B has 3 elements. The union of A and B is everything inside the two circles, hence its complement is everything *outside* the two circles:



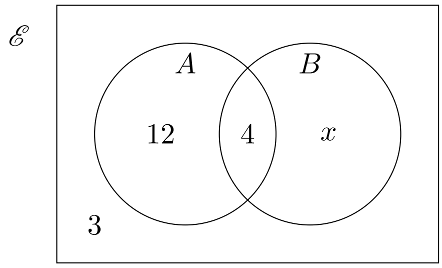
Finally, the $n\,(A \cap B')$. Let us first identify where $A \cap B'$ is represented in the diagram, using the "little balls" method. I'll add red balls to the regions A is, and blue balls to the regions B' is (that is, anything outside B):



the region where we have both a red and a blue ball are in the "only A" part, so we add 12 there, as we know $n\,(A \cap B') = 12$:



We are done adding the information, apart from $n(\mathscr{E}) = 30$, which we will use later. As we want to find $n(B)$, our goal now is to find the number of elements in the part of B which "has only B". I'll add an x there:



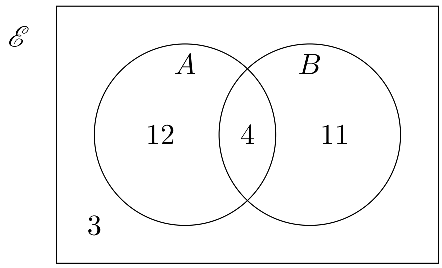
Finally, let us use the information that $n(\mathscr{E}) = 30$. If we add all numbers in the diagram, that must add up to 30. Thus:

$$\underbrace{12}_{\text{only } A} + \underbrace{4}_{A\&B} + \underbrace{x}_{\text{only } B} + \underbrace{3}_{\text{outside}} = \underbrace{30}_{\text{total}}$$

and we just have to solve this equation (see Chapter 14):

$$\begin{aligned} 12 + 4 + x + 3 &= 30 \\ 19 + x &= 30 && \text{Collecting like terms} \\ 19 + x - 19 &= 30 - 19 && \text{Subtracting 19 on both sides} \\ x &= 11 \end{aligned}$$

We can now fill the last region of the diagram:



and find $n(B)$ by adding the two regions that make B up:

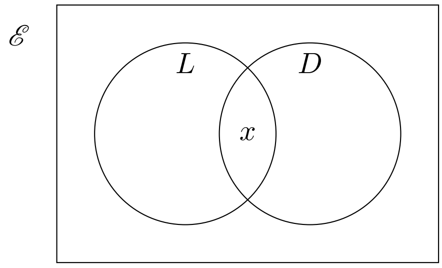
$$n(B) = 4 + 11 = 15$$

One instance these problems get a little bit trickier is when we do not know the value of the intersection. For instance:

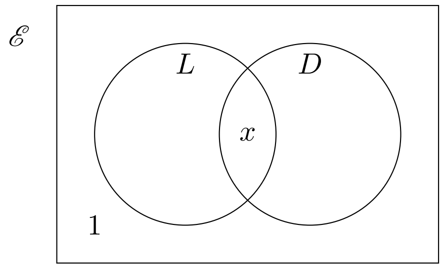
> In an office, there are 17 workers. Some people use a laptop and some workers use a desktop PC. There are also workers which use both. We know that 8 people use laptops, 12 use desktops, and 1 lonely person uses neither. Find how many people use both laptops and desktops.

Notice that if you add all the information given, you obtain *more than* 17 people. That is because you are over-counting the people that use both a Desktop and a Laptop. But worry not, let us use a diagram to help.

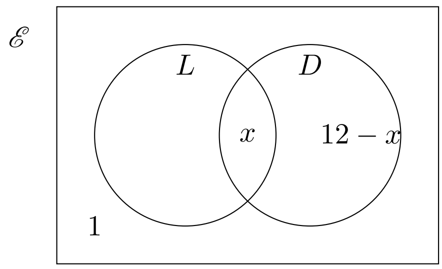
The bigger issue here is that we do not know the intersection, and we always start by filling the intersection! As good mathematicians, whenever we don't know something, we call it x. So let us add an x to the intersection:



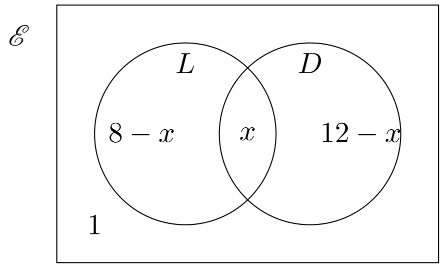
We can also add the lonely guy that uses neither a desktop nor a laptop to the outside region:



To fill the other part that makes up the D set now, we need to be careful: there are 12 people who use desktops, but we already counted x people in it (which use both a desktop and a laptop). Hence, only $12 - x$ people use *only* a desktop:



Using the same reasoning, we only have $8 - x$ people that use *only* laptops:



Not that every region is filled, we can add all of them together. As there are 17 workers in total, that must be the result of our addition:

$$\underbrace{8-x}_{\text{only Laptop}} + \underbrace{x}_{\text{both}} + \underbrace{12-x}_{\text{only Desktop}} + \underbrace{1}_{\text{neither}} = 17$$

and now we solve it:

$$\begin{aligned} 8-x+x+12-x+1 &= 17 \\ 21-x &= 17 && \text{Collecting like terms} \\ 21-x-21 &= 17-21 && \text{Subtracting 21 on both sides} \\ -x &= -4 \\ x &= 4 && \text{Dividing both sides by } -1 \end{aligned}$$

Hence, there are 4 people that use both a desktop and a laptop.

10.7. Exam hints

Unfortunately, you do have to memorize all the notation in this chapter (even in 0607 it is not given in the formulae page).

Remember that the subset symbol is very similar to the inequality symbols you are already familiar with.

The "little balls" technique is really practical for the shading diagrams part, and I highly recommend practising questions of shading diagrams, as they are very common.

Solving problems using sets is very common as well, particularly in questions that also use probability (see Chapter 49). Remember, when filling diagrams, to *always start at the intersection* (if 3 sets, first from the intersection of all 3 sets and after that one the two-by-two intersections).

Summary

- *Sets* are mathematical collections of *elements*;
- You can write sets as:

- Lists: $A = \{1, 2, 3, 4\}$
- Using rules: $A = \{\text{positive whole numbers smaller than } 5\}$
- Builder / comprehension notation: $A = \{x : 1 \leq x \leq 4\}$

- The *universal* set is a set which contains all elements we are interested in. We denote it by $\mathscr{E}$;
- The *empty* set is a set which has no elements. We denote it by $\varnothing$;
- We can represent sets using Venn diagrams. When we need to fill a Venn diagram with elements, we *always start at the intersections*;
- The set operations you must know are:
  - *belonging*: when an element is in a set

$$3 \in A$$

    and when it is not

$$3 \notin A$$

  - *number of elements in a set*: denoted by $n(A)$, which is ready as "the number of elements in set A"
  - *complement of a set*: we denote the complement of A by A', which is the set of all elements which are not in A but are part of the universal set;
  - *subsets*: a set is a subset of another set if all of its elements are also elements of the other set. We denote that by

$$B \subset A$$

    which is read as "B is a subset of A". When a set is a subset of another and they are not equal, we say that B is a *proper* subset of A. When B is not a subset of A we write

$$B \not\subset A$$

    and when a subset is either a subset of another, or is equal to the other, we write

$$B \subseteq A$$

  - *union of sets* is the same as "joining" two sets: every element which either in A *or* B is part of their union. We denote the union of A and B by

$$A \cup B$$

  - *intersection of sets*: an element is in the intersection of two sets if it is part of *both sets at the same time*. Hence, to be part of the intersection of A and B, you have to be a member of both A *and* B. We denote the intersection of A and B by

$$A \cap B$$

- Use the "little ball" method to find regions of the Venn diagram that satisfy the operation given;
- We can use Venn diagrams to solve problems involving number of elements in many situations. Remember to always start filling by the intersection!

Formality after taste

Everything we did in this chapter is quite formal, as sets are a very formal topic. There are two things I would like to show you though, which are quite interesting.

Set equality

Notice that we never defined when two sets are equal! I did say that two sets were equal when they have the same elements, but that is quite problematic to think when we have an infinite number of elements in a set!

We (and by we I mean real mathematicians) say that sets A and B are equal when

$$A \subseteq B \text{ and } B \subseteq A$$

This means that to be equal, A must be a subset of B and B must be a subset of A.

The empty set is a subset of all sets

A fact that I always felt was "weird" is that the empty set is a subset of any set.

To show that, we can use a contradiction: say that the empty set was *not* a subset of any set A. That means that $\varnothing$ has an element which is not in A. But $\varnothing$ has no elements! Hence, we reach a contradiction.

11. Bounds

11.1. Why learn bounds

Whenever we measure something, we have uncertainty: our measuring instruments have a certain precision, and we cannot be sure of the measurement beyond that precision level. For instance, our school rulers use measure all the way to millimetres. So we can be sure that, until the millimetre level, our measurement is correct. Anything smaller than that, though, we cannot know.

Paul Lockhart, a famous mathematician who wrote a piece called *A Mathematician's Lament*[1], says, in another book of his called *Measurement*:

> A blade of grass has no actual length. Any measurement made in this universe is necessarily a rough approximation. It's not bad; it's just the nature of the place. The smallest speck is not a point, and the thinnest wire is not a line.[2]

Bounds are a way for us to work with these uncertainties, and know how wrong our results can be. In this sense, bounds are fundamentally important, as all our measurements are wrong after a certain degree.

11.2. Finding lower and upper bounds

Before going into the methods, let us imagine a situation in which a measurement is wrong. Say we have a scale, and we will weight ourselves. This scale has a precision of 1 decimal place. This means that it can measure 82.1, but not 82.08. Hence, the scale is rounding our weight accordingly to its precision. Thus, we do not know our real weight, only the rounded value our scale gave.

The idea of finding bounds is this: given the precision of our scale (or anything we are measuring with), what is the *smallest possible value* for our real measurement? What about the *largest possible value*?

We call the lowest possible value of our real measure its *lower bound* given our precision. The largest possible value is called the *upper bound.*

Let us see how to find those values now.

[1] A great read. You can annoy your teacher with the arguments and see what it is the answer. You can read here `https://www.maa.org/external_archive/devlin/LockhartsLament.pdf`. Accessed on: 21/01/2020.

[2] His argument is that mathematical reality is much easier to work with, as we imagine it. It can be whatever we want. And it makes sense: if you think mathematics is hard, look through a window.

11.2.1. The "bad ruler" method

This is the way I like using. It works by imagining our measuring tool is always a ruler. In the case of our scale that measures correctly to 1 decimal place, it would "look" that our ruler would have these numbers on it

$$\ldots, 81.9, 82.0, 82.1, 82.2, 82.3, 82.4, \ldots$$

As we know, our "scale" (as we imagine it as a ruler) is rounding the result to its precision of 1 decimal place. Hence, when it says the weight is 82.1, it simply means 82.1 is the *closest value* it actually can measure. We need to find what is the smallest possible value (the lower bound) and the largest possible value (the upper bound) that rounds to 82.1 in this ruler.

This is very easy to do: the smallest value, the lower bound, is found by finding the midpoint (the mean) between the measurement given, 82.1, and the number that comes before it on the rule, 82.0:

$$\text{lower bound} = \frac{82.0 + 82.1}{2} = 82.05$$

The upper bound is found by finding the mean between the measured value, 82.1 and the number that comes after it in the ruler, 82.2:

$$\text{upper bound} = \frac{82.1 + 82.2}{2} = 82.15$$

We are done: the lower bound of 82.1, when measured correctly to 1 decimal place, is 82.05. The upper bound is 82.15. We normally write this using an inequality. Say we call our weight w. We can represent its possible real values by

$$\underbrace{82.05}_{\text{lower bound}} \leq w < \underbrace{82.15}_{\text{upper bound}}$$

An important thing to notice when representing the possible real values of w using this is that we *include* the lower bound, 82.05, in the possible real values, as it would round to 82.1 when rounded to 1 decimal place. We *do not*, however, include the upper bound (denoted by the symbol $<$), as 82.15 would round to 82.2 when rounded to 1 decimal place.

Another example: say we know that the amount of people in a stadium is 10500, correct to nearest 500 people. In this case, our "ruler" would have the following measurements:

$$\ldots, 9500, 1000, 10500, 11000, 11500, \ldots$$

To find the lower bound of our measurement of 10500, we calculate the midpoint between it and the value that came before it on our ruler:

$$\text{LB} = \frac{10000 + 10500}{2} = 10250$$

and the upper bound is found by finding the midpoint of 10500 and the number above it, 11000:

$$\text{UB} = \frac{10500 + 11000}{2} = 10750$$

Thus, if the number of people is denoted by m, we can write

$$10250 \leq m < 10750$$

11.2.2. The "reverse rounding" method

This method involves thinking about the rounding taken. Remember the scale that measured 82.1 correct to the nearest decimal place? This method finds the lower bound by asking "what is the *smallest* number which rounds to 82.1 when rounded to 1 decimal place?" The answer is 82.05. The upper bound is found by asking "what is the *largest* number which rounds to 82.1 when rounded to 1 decimal place?" The answer would be 82.149999..., but as we do not actually include the upper bound, we again say it is 82.15[3]

As you can see, this method is much faster, but it involves the "reverse" reasoning of finding the numbers that would round to the value given. As always, use whichever you like better.

11.2.3. Some more examples of bounds

Here are some examples of common types of questions:

- Find the lower and upper bounds of $a = 14$, correct to the nearest integer.

  If you use the ruler method, our ruler would be measuring $\ldots, 12, 13, 14, 15, \ldots$. The lower bound is found by finding the mean value between the measurement, 14 and the number that comes before it in the ruler, 13:

  $$\text{LB} = \frac{13 + 14}{2} = 13.5$$

  and the upper bound is found by the finding the mean between the measurement, 14, and the number above it, 15:

  $$\text{UB} = \frac{14 + 15}{2} = 14.5$$

  Thus,

  $$13.5 \leq a < 14.5$$

[3] If you already know recurring decimals, $82.14\dot{9}$ is indeed equal to 82.15, hence another reason to not actually including the upper bound.

- Finding the lower and upper bounds of $b = 10$ correct to 1 significant figure.

  Using the "reverse rounding" method, the smallest number which rounds to 10 when rounding to 1 significant figure is 9.5. The largest number which rounds to 10 when rounding to 1 significant figure would be 14.99999..., but we use 15 as the upper bound, as we do not include it. Hence:

$$9.5 \leq b < 15$$

11.3. Calculations with bounds

The most important thing to remember when having to do calculations with bounds is to *always find the bounds first*. What I mean by that is: find the lower and upper bounds of all quantities you will use for the calculation *before* doing the calculation. The classic mistake is first doing the calculation and then doing the bounds!

Thus, remember: bounds first, calculations later.

An important warning: **do not round results of calculations with bounds.** We are interested in the precise value for these calculations, so do not round them.

11.3.1. Addition and multiplication

Let us use an example to understand. Say we have $a = 2.5$ and $b = 5.3$, both correct to one decimal place. We want to find the lower and upper bounds of both $a + b$ and ab. This means we want to find the smallest possible value for their sum and product (the lower bound) and the maximum possible value for their sum and product (the upper bound).

The first thing is to always remember: bounds first, calculations later. So let us write both the lower and upper bounds for both a and b:

$$a = 2.5 \text{ (1 d.p.)} \rightarrow \begin{cases} \text{UB}: & 2.55 \\ \text{LB}: & 2.45 \end{cases}$$

$$b = 5.3 \text{ (1 d.p.)} \rightarrow \begin{cases} \text{UB}: & 5.35 \\ \text{LB}: & 5.25 \end{cases}$$

What we do now is *combine* the bounds adequately.

To find the upper bound of an addition, we want to add the largest possible values. Thus, to find the upper bound of an addition, we add the upper bounds of the individual quantities. The same is valid for multiplication: to find the upper bound of a product, we multiply the upper bounds of the individual quantities.

As for the lower bound of additions and multiplications, we use the lower bounds of individual values. Therefore,

$$a + b \rightarrow \begin{cases} \text{UB: UB of } a\text{+UB of } b \rightarrow & \overbrace{2.55}^{\text{UB of } a} + \overbrace{5.35}^{\text{UB of } b} = 7.9 \\ \text{LB: LB of } a\text{+LB of } b \rightarrow & \underbrace{2.45}_{\text{LB of } a} + \underbrace{5.25}_{\text{LB of } b} = 7.7 \end{cases}$$

$$ab \to \begin{cases} \text{UB: UB of } a \times \text{UB of } b \to & \overbrace{2.55}^{\text{UB of } a} \times \overbrace{5.35}^{\text{UB of } b} = 13.6425 \\ \text{LB: LB of } a \times \text{LB of } b \to & \underbrace{2.45}_{\text{LB of } a} \times \underbrace{5.25}_{\text{LB of } b} = 12.8625 \end{cases}$$

Notice that I did not round the results of the multiplications: do not round the result of calculations with bounds.

11.3.2. Subtraction and division

Again, let us use an example: say we have $a = 10000$ to the nearest thousand and $b = 550$ to the nearest 10. We want to find the lower and upper bounds for $a - b$ and $\frac{a}{b}$. Again, remember: the first thing to do is finding the individual bounds of a and b:

$$a = 10000 \text{ (thousand)} \to \begin{cases} \text{UB}: & 10500 \\ \text{LB}: & 9500 \end{cases}$$

$$b = 550 \text{ (ten)} \to \begin{cases} \text{UB}: & 555 \\ \text{LB}: & 545 \end{cases}$$

Now we need to combine them correctly. When we subtract a quantity from other, for the answer to be as large as possible (the upper bound), we want the quantity we are subtracting from to be as large as possible, and the quantity we are subtracting to be as small as possible. Thus, to find the upper bound of a difference, we subtract the lower bound from the upper bound. To find the lower bound, the smallest possible value, we want the quantity we are subtracting from to be as small as possible, and the one we are subtracting to be as large as possible. Hence, to find the lower bound of a difference, we subtract the upper bound from the lower bound. In our case:

$$a - b \to \begin{cases} \text{UB: UB of } a - \text{LB of } b \to & \overbrace{10500}^{\text{UB of } a} - \overbrace{545}^{\text{LB of } b} = 9955 \\ \text{LB: LB of } a - \text{UB of } b \to & \underbrace{9500}_{\text{LB of } a} - \underbrace{555}_{\text{UB of } b} = 8945 \end{cases}$$

For divisions, the idea is the same. To find the upper bound of a division, its largest value, we want to divide the largest possible quantity by the smallest possible quantity. And the opposite for the lower bound: we want to divide the smallest possible value by the largest possible value:

$$\frac{a}{b} \rightarrow \begin{cases} \text{UB: } \frac{\text{UB of } a}{\text{LB of } b} \rightarrow \quad \frac{\overbrace{10500}^{\text{UB of } a}}{\underbrace{545}_{\text{LB of } b}} = 19.26605505 \\ \\ \text{LB: } \frac{\text{LB of } a}{\text{UB of } b} \rightarrow \quad \frac{\overbrace{9500}^{\text{LB of } a}}{\underbrace{555}_{\text{UB of } b}} = 17.11711712 \end{cases}$$

Again, notice I did not round the results.

11.3.3. Expressions and formulae with bounds

When dealing with expressions and formulae with bounds, the golden rule still is: find the bounds *first*. Apart from that, we still have to do BIDMAS. And, if dealing with formulae, sometimes you have to change the subject. But the bounds part still is the same.

For instance, say we want to find the lower and upper bounds of

$$a - (b + c)$$

when $a = 3$ to the nearest integer, $b = 2.3$ to 2 significant figures and $c = 1.5$ to the nearest 0.5. No matter the expression we first find the bounds of every quantity :

$$a = 3 \text{ (integer)} \rightarrow \begin{cases} \text{UB :} & 3.5 \\ \text{LB:} & 2.5 \end{cases}$$

$$b = 2.3 \text{ (2s.f.)} \rightarrow \begin{cases} \text{UB :} & 2.35 \\ \text{LB:} & 2.25 \end{cases}$$

$$c = 1.5 \text{ (nearest 0.5)} \rightarrow \begin{cases} \text{UB :} & 1.75 \\ \text{LB:} & 1.25 \end{cases}$$

Now, we need to decide which bound we will use for each variable. Let us start with identifying the "biggest" operation in the expression. That would be the "minus": we are subtracting the result of $b + c$ from a. To find the lower bound of a subtraction, we want the quantity we are subtracting from to be as small as possible, and the quantity we are subtracting to be as big as possible. Hence, in this case we will use the lower bound of a and the upper bound of $(b + c)$. To find the upper bound of this subtraction, we will use the upper bound of a and the lower bound of $(b + c)$

$$a-(b+c) \to \begin{cases} \text{UB: UB of } a - \text{LB of } (b+c) \to \overbrace{3.5}^{\text{UB of } a} - \overbrace{\left(\underbrace{2.25}_{\text{LB of } b} + \underbrace{1.25}_{\text{LB of } c}\right)}^{\text{LB of } b+c} = 3.5 - 3.5 = 0 \\ \text{LB: LB of } a - \text{UB of } (b+c) \to \underbrace{2.5}_{\text{LB of } a} - \underbrace{\left(\overbrace{2.35}^{\text{UB of } b} + \overbrace{1.75}^{\text{UB of } c}\right)}_{\text{UB of } b+c} = 2.5 - 4.1 = -1.6 \end{cases}$$

Using the same values, we could find the lower and upper bounds of

$$\frac{b-c}{a}$$

and as we are dividing, to find the lower bound we want to divide the lower bound of $b-c$ by the upper bound of a; to find the upper bound we want to divide the upper bound of $b-c$ by the lower bound of a:

$$\frac{b-c}{a} \to \begin{cases} \text{UB: } \frac{\text{UB of } b-c}{\text{LB of } a} \to \frac{\overbrace{\left(\underbrace{2.35}_{\text{UB of } b} - \underbrace{1.25}_{\text{LB of } c}\right)}^{\text{UB of } b-c}}{\underbrace{2.5}_{\text{LB of } a}} = 0.44 \\ \text{LB: } \frac{\text{LB of } b-c}{\text{UB of } a} \to \frac{\overbrace{\left(\underbrace{2.25}_{\text{LB of } b} - \underbrace{1.75}_{\text{UB of } c}\right)}^{\text{LB of } b-c}}{\underbrace{3.5}_{\text{UB of } a}} = \frac{0.5}{3.5} = \frac{1}{7} \end{cases}$$

Lastly, let us find the lower and upper bounds of the volume for the gravitational force between two objects. The formula for this is

$$F = G\frac{m_1 m_2}{d^2}$$

in which $G = 6.67 \times 10^{-11}$, $m_1 = 25$ to the nearest integer, $m_2 = 15$ to the nearest 5 and $d = 0.1$ to 1 decimal place (assume units are correct).

First, let us find the bounds. As G is a constant we do not find its bound, we simply use the value given. For the other quantities:

$$m_1 = 25 \text{ (integer)} \to \begin{cases} \text{UB}: & 25.5 \\ \text{LB}: & 24.5 \end{cases}$$

$$m_2 = 15 \text{ (nearest 5.)} \rightarrow \begin{cases} \text{UB}: & 17.5 \\ \text{LB}: & 12.5 \end{cases}$$

$$d = 0.1 \text{ (1d.p.)} \rightarrow \begin{cases} \text{UB}: & 0.15 \\ \text{LB}: & 0.05 \end{cases}$$

As we are calculating a division, to find the upper bound we will have to divide the upper bound of $m_1 m_2$ by the lower bound of d^2; to find the lower bound, we will divide the lower bound of $m_1 m_2$ by the upper bound of d^2. As everything else is a product ($m_1 m_2$, $d^2 = d \times d$) we just use the bound we want to find:

$$F = G\frac{m_1 m_2}{d^2} \rightarrow \begin{cases} \text{UB: } G\frac{\text{UB of } m_1 m_2}{\text{LB of } d^2} \rightarrow & G\frac{\overbrace{\left(\underbrace{25.5}_{\text{UB of } m_1} \times \underbrace{17.5}_{\text{UB of } m_2}\right)}^{\text{UB of } m_1 m_2}}{\underbrace{0.05}_{\text{LB of } d} \times \underbrace{0.05}_{\text{LB of } d}} = 1.190595 \times 10^{-5} \\ \text{LB: } G\frac{\text{LB of } m_1 m_2}{\text{UB of } d^2} \rightarrow & G\frac{\overbrace{\left(\underbrace{24.5}_{\text{LB of } m_1} \times \underbrace{12.5}_{\text{LB of } m_2}\right)}^{\text{LB of } m_1 m_2}}{\underbrace{0.15}_{\text{UB of } d} \times \underbrace{0.15}_{\text{UB of } d}} = 9.078611111 \times 10^{-7} \end{cases}$$

11.4. Exam hints

Always remember: bounds first, calculations later. Also, you need to memorise (or think at the time) that:

- UB of $a + b$ or ab: add or multiply the upper bounds of a and b;
- LB of $a + b$ or ab: add or multiply the lower bounds of a and b;
- UB of $a - b$: calculate UB of a minus LB of b;
- LB of $a - b$: calculate LB of a minus UB of b;
- UB of $\frac{a}{b}$: divide UB of a by LB of b;
- LB of $\frac{a}{b}$: divide LB of a by UB of b.

And remember to not round the results of bounds calculations!

Summary

- Bounds are useful because the are *uncertainties* in our measurements;
- The lower and upper bounds of a measurements give us a *range* of values which our measurement could really be;
- To find the lower and upper bounds, you can either use the "bad ruler" or the "reverse rounding" methods:
  - The "bad ruler" method works by imagining a ruler which measures to the given precision, and finding the midpoints of our given value with the numbers that come before and after it in the ruler;
  - The "reverse rounding" method works by finding the smallest and largest values which round to the given measurement with the precision given;
- When doing calculations with bounds, remember to *always find the bounds first* and to *not round the final answers*;
- Remember:
  - UB of $a + b$ or ab: add or multiply the upper bounds of a and b;
  - LB of $a + b$ or ab: add or multiply the lower bounds of a and b;
  - UB of $a - b$: calculate UB of a minus LB of b;
  - LB of $a - b$: calculate LB of a minus UB of b;
  - UB of $\frac{a}{b}$: divide UB of a by LB of b;
  - LB of $\frac{a}{b}$: divide LB of a by UB of b.

Part II.

Algebra

12. Algebraic expressions

12.1. Why learn the basics of algebraic expressions

We will start now our study of algebra. You may have bad memories of it, but as Pythagoras used to say "all is number": everything we will do is about numbers. Of course, we will use variables to speak of any possible number, but we are still talking about numbers. Hence, you can think of algebra as a 'generalization' of our number skills (and that may give you indication you need to revise those concepts!).

So, do not despair: algebra is as natural as numbers; and as numbers are very natural to us, given that we are a pattern seekers, algebra is *necessary*. With it, we are able to investigate patterns and properties, and what could be more satisfying?

Of course, a more prosaic argument to learn the contents of this chapter would be that they are the foundation to many other concepts in your mathematical studies, and you will struggle should you not understand them. But try not to think of algebra (or math) like this. Try to remember these words by Dirk J. Struit, in his amazing *A Concise History of Mathematics*:

> "Mathematics helped to find order in chaos, to arrange ideas in logical chains, to find fundamental principles."

and algebra is a fundamental tool to brave this chaos.

12.2. The most important concept: an unknown (a.k.a. variables)

First of all, variables and unknowns refer to the same idea. I believe that it is more common to use unknown when solving equations or problems, and variables when we are dealing with expressions that have no particular solution. However, I will use them interchangeably.

A variable is a placeholder for a number. This allows us to write general ideas. For instance, if we want to refer to adding 3 to a number (any number), we could write

$$\text{a number} + 3$$

However, it is quite lengthy to say that. Thus, we simply use a symbol, a placeholder, in place of "a number". The most used is our dear x:

$$\text{a number} + 3 \text{ is the same as } x + 3$$

Therefore, we know use any symbol in place of a number. We can decide the number later or, if possible, find the possibilities that satisfy certain restrictions. In any form, we could use any symbol to stand for numbers, say

 +  = 8

and we could later choose 5 and 3 to replace  and , or whatever other pair of numbers. What is important is that they are just placeholders for numbers. Hence the appeal of the "All is number" quote: algebra is just numbers, we just have not decided which numbers yet! Call it lazy numbers!

So just remember: a variable is a symbol used in place of a number we still have not decided (or found). Everything you know about numbers is still true, so we will continue to use them when doing algebra.

12.3. Algebraic expressions

12.3.1. Examples of algebraic expressions

Any expression with a variable is an example of an algebraic expression. For instance

$$x + 17$$

or

$$x + y + z$$

12.3.2. Writing algebraic expressions to represent given situations

One of the most important aspects of algebra is actually going from a situation to a correct mathematical representation of it.

Let us say we want to represent "3 more than a given number". The "given number" part is a variable, as we have not been yet given the number. Hence, let us call it x. Now, to do 3 more than x we write:

$$x + 3$$

or

$$3 + x$$

as the usual properties of addition still work.

Basically, then, follow this algorithm:

1. Identify any quantities you have not been given yet, that you need to use a *place-holder* for. Those you will attribute variables to;
2. Connect the variables with the other information you have.

For instance, say that a given statistic of a car is found by adding its power to its number of doors (yes, I don't understand much about cars). Here, we have to quantities we are unsure of, the power and the number of doors. So let us use variables to represent them:

$$\text{power} \rightarrow p$$

$$\text{doors} \rightarrow d$$

Now we can add them together:

$$p + d$$

One important thing to remember is to use the same variable anytime you refer to the same quantity in the problem. Say you want to find a number which is given by subtracting the number of legs from the number of wings, followed by dividing the result by the number of wings. Here we have two quantities to use variables:

$$\text{legs} \rightarrow l$$

$$\text{wings} \rightarrow w$$

and we want to subtract l from w and divide the result by w itself

$$\frac{w - l}{w}$$

or

$$(w - l) \div w$$

Notice that we use w for all instances of "number of wings" the problem tells us.

12.4. Substitution: attributing a value to an unknown

Now we are familiar with unknowns and expressions, we are ready to do something with them: assigning values to the variables! This will make our algebraic expressions become simple numerical expressions. After that, we just do calculations (and just to start the chorus: we still have to follow BIDMAS / PEMDAS / GEMA!).

Let us say we have the following expression:

$$x + 5$$

in which x is our unknown / variable. As x is a *placeholder* for a number, we can actually "put a number into" x. Say, for instance, we decide that x should be equal to 3:

$$x = 3 \qquad \text{The value we want for } x$$

$$x + 5 \qquad \text{We "put" 3 into } x$$

$$3 + 5 \qquad x \text{ "becomes" } 3$$

$$8 \qquad \text{Calculate}$$

You can see why we call this substitution: we have *substituted* our placeholder, x, by a number, in our case 3. From that we only have numbers, which we manipulate normally.

We could have chosen a different value for x that we wanted. Say we wanted to make $x = -2$

$$\begin{aligned}
x &= -2 && \text{The value we want for } x \\
x &+ 5 && \text{We “put” } -2 \text{ into } x \\
-2 &+ 5 && x \text{ “becomes” } -2 \\
&3 && \text{Calculate}
\end{aligned}$$

Another expression:

$$5y - 4$$

and let us make $y = 4$. Remember that $5y$ is the same as $5 \times y$, and we have to be careful with BIDMAS:

$$\begin{aligned}
y &= -2 && \text{The value we want for } y \\
5y &+ 5 && \text{We “put” } 4 \text{ into } y \\
5 \times 4 &+ 5 && y \text{ “becomes” } 4 \\
20 &+ 5 && \text{Calculate, following BIDMAS} \\
&25 &&
\end{aligned}$$

Be careful when substituting negative numbers. I recommend always putting brackets around them, just to be sure. Say we want to make $y = -1$:

$$\begin{aligned}
y &= -2 && \text{The value we want for } y \\
5y &+ 5 && \text{We “put” } -1 \text{ into } y \\
5 \times (-1) &+ 5 && y \text{ “becomes” } -1 \\
-5 &+ 5 && \text{Calculate, following BIDMAS} \\
&0 &&
\end{aligned}$$

Should we have more than a variable in the expression, we just substitute the value in each accordingly. For instance, if we have the expression

$$2x - 3y + 5$$

and we want to make $x = 3$ and $y = -2$:

$$\begin{aligned} x = 3 \text{ and } -2 & \qquad \text{The values for } x \text{ and } y \\ 2x - 3y + 5 & \qquad \text{Substituting} \\ 2 \times 3 - 3 \times (-2) + 5 & \\ 6 + 6 + 5 & \qquad \text{BIDMAS time} \\ 17 & \end{aligned}$$

Sometimes we have expressions with indices. For instance,

$$x^2 - 5x + 6$$

and we can have $x = 1$. In cases like this, all x "receive" the same value:

$$\begin{aligned} x = 1 & \qquad \text{The value we want for } x \\ x^2 - 5x + 6 & \qquad \text{Substituting } x = 1 \text{ in all } x \\ (1)^2 - 5 \times 1 + 6 & \\ 1 - 5 + 6 & \qquad \text{BIDMAS} \\ 2 & \end{aligned}$$

I recommend being particularly pedantic about brackets when substituting negative numbers into variables with powers. So, if we wanted to have $x = -2$ into the last expressions, add brackets for safety:

$$\begin{aligned} x = -2 & \qquad \text{The value we want for } x \\ x^2 - 5x + 6 & \qquad \text{Substituting } x = -2 \text{ in all } x \\ (-2)^2 - 5 \times -2 + 6 & \\ 4 + 10 + 6 & \qquad \text{BIDMAS} \\ 20 & \end{aligned}$$

In this case, if you have written $-2^2 - 5 \times -2 + 6$ in your calculator, you would have obtained 12 as the answer, because our calculators have a strange way to deal with the negative in front of powers. So, I cannot stress this enough: always add brackets when substituting negative values.[1]

A very useful application of substitution are using formulae. For instance, this is a formula for the volume of a cuboid with lengths a, b and c:

$$V = abc$$

If we have a cuboid with side lengths $a = 1, b = 3$ and $c = 5$, we can substitute those values in the formula for V and find its volume:

$$\begin{aligned} a = 1, b = 3, c = 5 & \qquad \text{Our values} \\ V =& abc \qquad \text{Substituting} \\ V =& 1 \times 3 \times 5 \\ V =& 15 \end{aligned}$$

12.5. Simplifying expressions (collecting like terms)

Our objective now is to take an algebraic expression and rewrite it into an equivalent (but different) expression. What this means is that if we were to substitute a value for each of the variables in both expressions, we would obtain the same value.

To see what we mean by equivalent expressions, let us first use a numerical example. Say we start with the expression

$$2 + 2$$

which we know is equal to 4. We can rewrite it into a different expression, which also is equal to 4, say $5 - 1$. Thus, we know that

$$2 + 2 = 5 - 1$$

To do that with algebraic expressions, you just have to remember that you have been doing this since you learned how to count. Or, as I like to think about it, apple algebra.

[1] The calculator actually has two different interpretations of the minus sign: one is the operation subtraction, the other the negative of a number. The negative interpretation has a very low precedence (imagine that it comes after all the BIDMAS operations), hence when we say -2^2 the calculator first does $2^2 = 4$, and then find the negative of 4, hence -4. Be very careful!

12.5.1. Apple algebra

Say we have an apple and we are given another apple. We could write this as

$$\text{apple} + \text{apple}$$

And we know that we now have 2 apples, so

$$\text{apple} + \text{apple} = 2\text{apple}$$

Another example would be

$$5\text{apple} - 3\text{apple} = 2\text{apple}$$

as I am sure you knew since you were very young. What about apples and bananas? You know that

$$\text{apple} + \text{banana} = \text{apple} + \text{banana}$$

as there is no way to join them both together.

In these examples lies the basic idea of simplifying algebraic expressions: you can join variables together when they are equal, and not when they are different!

12.5.2. Apple algebra to variable algebra

Now, let us say we are too lazy to keep writing apple all the time. Let us simply write a instead:

$$\text{apple} + \text{apple} = 2\text{apple}$$

$$a + a = 2a$$

much better no? Fewer characters. The same could be said if we used the classic variable, x:

$$x + x = 2x$$

as we are still counting: if we have 1 of *something* and are given another *something*, we now have 2 of *something*. This is what is written in

$$\underbrace{x}_{\text{something}} + \overbrace{x}^{\text{something}} = \underbrace{2x}_{\text{2 something}}$$

12.5.3. Like terms and addition

In an expression which has things being added or subtracted together, we call each of the things being added or subtracted a *term*. For instance, in the expression

$$3x + 7x - 4x$$

we have 3 terms: $3x$ is the first, $7x$ the second and $4x$ the third.

Just like in apple algebra, we could add $3x$ with $7x$ to obtain $10x$ (3 something plus 7 something is 10 something, right?):

$$\underbrace{3x + 7x} - 4x$$

$$10x - 4x$$

and we could remove $4x$ (take away 4 something!) from our result, to obtain $6x$:

$$10x - 4x = 6x$$

Notice that to go from $3x + 7x$ to $10x$, we simply added 3 with 7, the coefficients of x, and got 10. We then just put the x at the end (just like counting apples!). The same can be done to go from $10x - 4x$ to $6x$: we just do $10 - 4 = 6$, and the x copies from.

Now, say that we had the expression

$$4x + 3y$$

which could be seen as 4 apples added to 3 bananas. What would be the answer? Well, there is not really a way to combine apples and bananas, they are different fruits. Just like that, we cannot join $4x$ and $3y$, they are different! So there is nothing to do to this expression!

From all this discussion, we can obtain a rule to when we can add or subtract terms in an algebraic expression:

Rule: Two terms can be added (or subtracted) together only if they *have exactly the same variables*[2].

We call terms that have the same exact variables *alike.* We can always add or subtract alike terms. We cannot add or subtract unlike terms.

12.5.3.1. Identifying like terms

Identifying terms which area alike is fundamental to simplifying algebraic expressions. Let us see some examples of terms which are alike:

- $3x$ and $4x$: both terms have a variable x;
- $-xy$ and $3xy$: both terms have a product of the same variables, xy;
- x^2 and $9x^2$: both terms have the same variable, x^2;
- 17 and 8: both terms are numbers[3];
- $-4x^2y$ and $17x^2y$: both terms have the same product of variables, x^2y;

[2] Also known as the same fruit!
[3] They actually have the same variables, but to the power of 0.

- xy and yx: both terms have the same product of variables, xy. This is because multiplication is commutative (the order is not important), hence $yx = xy$.

In all these cases we could add or subtract these terms.

Let us see some terms which are not alike:

- x and y: different variables, x and y;
- x and x^2: different variables, the power changes everything;
- x^2y and xy^2: the powers are different in x and y.

Basically, alike terms have *exactly the same variables (or product of variables) with the same power*. If you remember this you will remember which terms are alike.

12.5.3.2. Adding or subtracting like terms

To actually do the calculation, we just look at the coefficients of the terms and operate with them (with the risk of being annoying: remember we are just counting!).

So let us see some examples.

- Example 1:

$$3p - 2p$$

  Here we have two terms, $3p$ and $2p$. Both have the same variable, p. Hence, they are alike we can simplify this expression by manipulating the coefficients. As $3 - 2 = 1$, we have $3p - 2p = 1p$, but remember we are too lazy to write $1p$, we just write p:

$$3p - 2p = p$$

- Example 2:

$$5x + 3x - 9x$$

  In this expression we have 3 terms: $5x$, $3x$ and $9x$. They all have the same variable, x, hence are all alike and we can simplify the expression. As $5 + 3 - 9 = -1$, we have that $5x + 3x - 9x = -1x$. However, again we are too lazy to write $-1x$, we just write $-x$:

$$5x + 3x - 9x = -x$$

- Example 3:

$$12a - 4b$$

  This expression has 2 terms, $12a$ and $4b$. They have different variables, so are not alike. Thus, we cannot subtract one from the other, and there is nothing to do in this expression.

Simplifying these expressions gets more exciting if they get bigger, such as

$$3x + 2y + 6y + 4x$$

but they are easy to simplify. The idea is to identify alike terms and work with them independently. I suggest using some visual way to mark them. In this expression, for instance, we have 4 terms, the first and last with variable x and the two in the middle with variable y. Thus, $3x$ and $4x$ are alike, and we can simplify those. The same goes for $2y$ and $6y$. To do that, I mark the like terms with different symbols, *including the sign which precedes* them:

$$\boxed{3x}\,(+2y)\,(+6y)\,\boxed{+4x}$$

and we can use these symbols as visual cues: we can simplify the terms which have the same symbol around. We put the signs within the symbols to be sure we are doing the correct operation (as the sign could be seen as part of the number, specially with negatives). Hence we have:

$$\boxed{3x}\,(+2y)\,(+6y)\,\boxed{+4x}$$

$$\boxed{3x + 4x}\,(+2y + 6y)$$

$$7x + 8y$$

as the result of $3x + 4x$ is $7x$ and the result of $+2y + 6y = 8y$.

This method works with subtractions as well (it is actually its strong suit):

$$-2a + 3b - 7a - 14b$$

First we identify the terms with are alike with our shapes, and then we can simply manipulate the coefficients:

$$\boxed{-2a}\,(+3b)\,\boxed{-7a}\,(-14b)$$

$$\boxed{-2a - 7a}\,(+3b - 14b)$$

$$-9a - 11b$$

The best thing about this way to identify the alike terms is that the operations we do appear "for free" within the shape!

Another example:

$$-8x + 3y + 5x - 4z - 7x + 3z$$

$$\boxed{-8x}\,(+3y)\,\boxed{+5x}\,\underline{-4z}\,\boxed{-7x}\,\underline{+3z} \qquad \text{Identify alike terms}$$

$$\boxed{-8x + 5x - 7x}\,(+3y)\,\underline{-4z + 3z} \qquad \text{Write them together}$$

$$-10x + 3y - z \qquad \text{Operate with coefficients}$$

Of course, as soon as you get practice you will stop doing these shapes, but it does help in the beginning.

A final example:

$$4xy + 8x^2y - 3yx + 4yx^2 - 7y^2x$$

The first thing to do here is to remember that each term, such as $-3yx$ is a multiplication. Hence, we can rewrite it to $-3xy$. Let us always write the terms with x first:

$$4xy + 8x^2y - 3xy + 4x^2y - 7xy^2$$

now it is much easier to identify the alike terms. Remember that to be alike we need to have the same variables to the same powers:

$$\boxed{4xy}\,(+8x^2y)\,\boxed{-3xy}\,(+4x^2y)\,\underline{-7xy^2}$$

$$\boxed{4xy - 3xy}\,(+8x^2y + 4x^2y)\,\underline{-7xy^2}$$

$$xy + 12x^2y - 7xy^2$$

12.5.3.3. Multiplying terms

Unlike addition and subtraction, we can *always* multiply algebraic terms (we will hold division until the algebraic fractions chapter, see ADD REF).

Let us start with an example:

$$2x \times 3y$$

The first thing to remember is that $2x$ is the same as $2 \times x$ and $3y$ the same as $3 \times y$:

$$2x \times 3y = 2 \times x \times 3 \times y$$

Second thing is that multiplication is commutative: the order does not matter. Hence, we can rewrite the above:

$$2x \times 3y = 2 \times x \times 3 \times y = 2 \times 3 \times x \times y$$

and now notice that we have 2×3 in our expression. Well, numbers are still numbers, and we know that $2 \times 3 = 6$:

$$2x \times 3y = 2 \times 3 \times x \times y = 6 \times x \times y$$

The only thing left to solve is the $x \times y$ part. But there is nothing for us to do there: we do not know the values of x and y, so the only thing we *can do* is leave them as they are:

$$2x \times 3y = 6 \times x \times y = 6xy$$

Thus, if we have a product of terms, we multiply all the numbers together and just "glue" different variables. Some more examples:

$$-3y \times 4x = -12xy$$

as we can just multiply -3 by 4 and "glue" x and y.

$$\frac{3}{4}a \times \frac{2}{9}bc = \frac{1}{6}abc$$

because we do $\frac{3}{4} \times \frac{2}{9} = \frac{1}{6}$ and just "glue" a with bc.

When we do have the same variable, though, we need to know a bit about indices (Chapter 15). The basic idea is to remember that a number multiplied by itself is equal to the number squared. This is still true for variables:

$$x \times x = x^2$$

and we can generalize this for any power (multiply the same number by itself 3 times means cubed, 4 times means to the power of 4, and so on).

Using this, we can simplify products such as

$$3xy \times 4x = 3 \times 4 \times x \times x \times y$$

in which we just changed the order of the multiplication. We can now do $3 \times 4 = 12$ and as we have $x \times x$, which is the same as x^2:

$$3xy \times 4x = 3 \times 4 \times \underbrace{x \times x}_{2\ x's} \times y = 12x^2y$$

Another example:

$$-2x^2y \times 4x^2y^2 = -2 \times 4 \times x^2 \times x^2 \times y \times y^2$$

which we can continue by either using indices properties (Chapter 15) or we can "split" x^2 into $x \times x$ and the same for y^2 into $y \times y$:

$$-2 \times 4 \times x^2 \times x^2 \times y \times y^2 = -8 \times \underbrace{x \times x}_{x^2} \times \underbrace{x \times x}_{x^2} \times y \times \underbrace{y \times y}_{y^2}$$

and now we count: we have 4 x multiplying themselves, so x^4. We have 3 y multiplying themselves, so y^3:

$$-8 \times \underbrace{x \times x}_{x^2} \times \underbrace{x \times x}_{x^2} \times y \times \underbrace{y \times y}_{y^2} = -8x^4y^3$$

12.6. Exam hints

I believe the most important to remember with this topic is when substituting negative values: remember to always add brackets to that you do not risk getting the

$$-2^2 = -4$$

mistake!

Also, be cautious when using variables to represent quantities: always use the same variable to the same quantity.

Summary

- A *variable* is a placeholder for a value we still do not know;
- An *algebraic expression* is any expression that has at least one variable;
- When using variables to represent a given situation, remember to attribute a variable to each quantity you need a placeholder for, and use the same variable all times that quantity appears;
- *Substituting* a value into a variable means assigning a value to the variable (that is, it stops being a placeholder for a value, you now know the value);
- To *collect like terms* means to add or subtract the terms which have exactly the *same variables to the same power.*

13. Brackets

13.1. Why do we use brackets again?

Brackets are absolutely fundamental in math: they tell us what operations we need to do first! As you know, when solving numerical expressions we follow BIDMAS or PEMDAS, and we always start with the part of the expression *inside* the brackets.

The question is how to proceed when you are solving *algebraic* expressions and they have brackets, such as

$$3(x+y)$$

It's a very common thing to do: given two numbers, x and y, we want to add them together and only **after** multiply the result by 3.

We have already learned that we cannot do anything to the $x+y$ within the brackets, so how could we get rid of them? How can we obtain an equivalent expression without the brackets? Let's find out!

13.2. Developing an idea

Let's think about the example I just gave, $3(x+y)$. First, we need to remember that the "absence" of anything between the bracket and the 3 means we are multiplying one by the other:

$$3(x+y) = 3 \times (x+y)$$

Now, let's think on what we know of multiplication. We know that 3×4, for instance, is the same as doing $4+4+4$, correct? Let's apply the same idea to our brackets:

$$3 \times (x+y) = \underbrace{x+y}_{\text{1st time}} + \underbrace{x+y}_{\text{2nd time}} + \underbrace{x+y}_{\text{3rd time}}$$

Well, we simply used our "multiplication as repeated addition" knowledge here! We can just simplify the resulting expression, as we know how to do it:

$\underline{x}\boxed{+y}\underline{+x}\boxed{+y}\underline{+x}\boxed{+y}$	Identify alike terms
$\underline{x+x+x}\boxed{+y+y+y}$	Collect them!
$3x+3y$	

That makes sense, doesn't it? Let's do some testing to check if we are correct. To do this, we can just choose values for x and y and substitute them in $3(x+y)$ and in $3x+3y$. Hopefully they will be the same!

$3(x+y)$	$3x+3y$
$x=0, y=0 \to 3(0+0) = 3\times 0 = 0$	$3\times 0 + 3\times 0 = 0$
$x=1, y=2 \to 3(1+2) = 3\times 3 = 9$	$3\times 1 + 3\times 2 = 3+6 = 9$
$x=-2, y=4 \to 3(-2+4) = 3\times 2 = 6$	$3\times -2 + 3\times 4 = -6+12 = 6$

They are all the same! That's great, it indicates we are in the right track here!

Now, let's try with a "more complex" case:

$$2(4a-5b)$$

Let's use the same idea: we know that means 2 times the bracket, which means add the brackets to itself:

$$2(4a-5b) = \underbrace{4a-5b}_{\text{1st time}} + \underbrace{4a-5b}_{\text{2nd time}}$$

Simplifying it:

$\underline{4a}\,\boxed{-5b}\,\underline{+4a}\,\boxed{-5b}$	Identify like terms
$\underline{4a+4a}\,\boxed{-5b-5b}$	Collect them!
$8a-10b$	

Let's check if the expressions give them same values for some cases:

$2(4a-5b)$	$8a-10b$
$a=0, b=1 \to 2\times(4\times 0 - 5\times 1) = 2\times -5 = -10$	$8\times 0 - 10\times 1 = 0-10 = -10$
$a=1, b=-1 \to 2\times(4\times 1 - 5\times -1) = 2\times(4+5) = 18$	$8\times 1 - 10\times -1 = 8+10 = 18$

It seems we are correct again!

Now, let's try to see if we can find a pattern. I highly recommend for you to stop reading now and try to figure out on your own!

What I wanted you to notice was this:

$$2\,(4a-5b) = 2\times 4a + 2\times -5b = 8a-10b$$

Isn't this great? We can just *distribute* the "number outside" to each of the parts of the expression inside the brackets! Let's try with our first example:

$$3(x+y) = 3x + 3y$$

Again we obtain the same result! This is our general method!

13.3. Single brackets: the distributive property of multiplication, a.k.a, velociraptor claws

What we have just "discovered" is what is called the distributive property of multiplication. We cannot prove it: it is one of the *basic properties* of multiplication in combination of addition[1].

This property states that, if you have three numbers, x, y and z, then

$$x(y+z) = xy + xz$$

As you can see, it's exactly what we have figured out by comparison on the previous section. I highly suggest you to use the arrows, or velociraptor claws, to remind yourself to distribute the multiplication to **every number inside the bracket**:

$$x(y+z) = xy + xz$$

Even though I am constantly putting only two terms inside the brackets, you can have as many as you want, and you distribute the multiplication to all of them:

$$a(b+c+d+e) = ab + ac + ad + ae$$

Let's see some examples in action:

Example 1:

$$4x(y-3)$$

Always remember to distribute, or draw the velociraptor claws:

$$4x(y-3) = 4x \times y + 4x \times -3 = 4xy - 12x$$

[1] If you want to know more about these properties check the formality after taste of this chapter.

Example 2:

$$-2(3x - 4)$$

This example is **very** important: it is a very common mistake to forget to distribute the negative sign!

$$-2\ (3x - 4) = -2 \times 3x - 2 \times -4 = -6x + 8$$

Do you notice that I always put the sign of the number outside between each term? In that way, you'll never forget to distribute signs and never get any question wrong!

Example 3:

$$-(a - b)$$

This one is a classic. A very good way to remember what to do is to always have in mind that what we are doing is multiplying the bracket by -1:

$$-1(a - b)$$

You won't get it wrong then!

$$-1\ (a - b) = -1 \times a - 1 \times -b = -a + b$$

Example 4:

$$x(x + y)$$

When you have only variables, you need to remember the very important property of indices:

$$x^n \times x^m = x^{n+m}$$

Let's distribute:

$$x\ (x + y) = x \times x + x \times y = x^1 \times x^1 + xy = x^{1+1} + xy = x^2 + xy$$

Example 5

$$-2x(x^5 - 3x^3 + 4x^2 - 3)$$

Look at the size of that beast! Fear not, though, you are well prepared with your velociraptor claws:

$$-2x \quad (x^5 - 3x^3 + 4x^2 - 3) = -2x \times x^5 - 2x \times -3x^3 - 2x \times 4x^2 - 2x \times -3$$

You see, nothing changes! You just distribute the multiplication to everyone inside the brackets! We just need to simplify it now:

$$-2x^1 \times x^5 - 2x^1 \times -3x^3 - 2x^1 \times 4x^2 - 2x^1 \times -3$$

$$-2x^{1+5} + 6x^{1+3} - 8x^{2+1} + 6x$$

$$-2x^6 + 6x^4 - 8x^3 + 6x$$

With practice you will just do these calculations in your head!

Solved exercise: Simplifying an expression with brackets

Simplify the following expression fully:

$$17 - 2x(x^2 - 5) + 14x - (3x^3 - 5)$$

Solution As with any expression, you need to follow BIDMAS. There is nothing for us to inside the brackets, so we need to do the multiplications: let's expand the brackets!

$$17 - 2x(x^2 - 5) + 14x - (3x^3 - 5)$$

$$17 - 2x \times x^2 - 2x \times -5 + 14x - 3x^2 - -5 \quad \text{Be very careful with the negatives!}$$

$$17 - 2x^{1+2} + 10x + 14x - 3x^2 + 5$$

$$17 - 2x^3 + 10x - 14x - 3x^2 + 5$$

$$\underline{17} - 2x^3 \underbrace{+10x}\underbrace{-14x} - 3x^2 \underline{+5} \quad \text{Collecting like terms}$$

$$\underline{17 + 5} - 2x^3 \underbrace{+10x - 14x} - 3x^2$$

$$22 - 2x^3 - 4x - 3x^2$$

And we're done! Normally we write expressions like these from the biggest power to the smallest:

$$\boxed{-2x^3 - 3x^2 - 4x + 22}$$

Double brackets

Now we will focus on expressions which have brackets multiplying brackets, such as

$$(x - 2)(x + 4)$$

Let's use the classic strategy of trying to make this become a problem we have seen before. I highly suggest you trying to figure out on your own.

The basic idea is as follows: we know that, even though they are algebraic expressions, each bracket represents a number. We know how to expand brackets such as $a(b + c)$, so the trick is making a *substitution*:

$\underbrace{(x-2)}_{\text{Let's make } y=(x-2)} \quad (x+4)$	Call the whole bracket something
$y(x+4)$	Marvel at your magic
$y(x+4)$	Expand the brackets
$xy+4y$	
$x(x-2)+4(x-2)$	Change y back to $(x-2)$
$x(x-2)+4(x--2)$	Expand more brackets!
$x^2 \underbrace{-2x+4x} -8$	Collect like terms
x^2+2x-8	Done!

Although this works, it's quite lengthy. In the end, what we are doing is, for each number inside the first bracket pair, distributing the multiplication to every number inside the second pair:

$$(x-2)(x+4)$$

$$x \times x + x \times 4 - 2 \times x - 2 \times 4$$

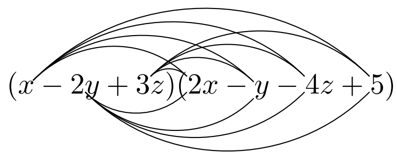
$$x^2 + 4x - 2x - 8$$

$$x^2 + 2x - 8$$

We get the same answer, and it's much faster! Basically it's just distributing everything inside the first brackets to everything inside the second. Therefore, just remember to multiply everything!

It even works when you have more than 2 terms in each bracket:

$$(x - 2y + 3z)(2x - y - 4z + 5)$$



$$2x^2 - xy - 4xz + 5x - 4xy + 2y^2 + 8yz - 10y + 6xz - 3yz - 12z^2 + 15z$$

$$2x^2 - 5xy + 2xz + 5x + 2y^2 + 5yz - 10y - 12z^2 + 15z$$

So many lines! It may get confusing, but we are basically multiplying everything by everything!

Solved exercise: Simplifying an expression with double brackets

Simplify the following expression fully:

$$3x(x-2y)-(x+3)(2x-5)$$

Solution As usual, we follow BIDMAS. We cannot do anything inside the brackets, so we can expand them already. Let's start with the single one on the left:

$$3x\,(x-2y)\underbrace{-(x+3)(2x-5)}_{\text{Just copy this part!}}$$

$$3x^2-6y-(x+3)(2x-5)$$

Now, we need to expand the double brackets. Be **very careful** with the negative in front of them, as we are going to have to distribute it later!

$3x^2-6y-(x+3)(2x-5)$	Velociraptor claws
$\underbrace{3x^2-6y}_{\text{Copy this part}}\;-(2x^2-5x+6x-15)$	Keep the brackets!
$3x^2-6y-(2x^2+x-15)$	Collecting like terms inside
$3x^2-6y-1\,(2x^2+x-15)$	More claws
$3x^2-6y-2x^2-x+15$	Distribute the -1
$x^2-x-6y+15$	Collect like terms

13.4. Notable products

In the previous exercises, did you notice some patterns? For instance,

$$(a+b)^2=(a+b)(a+b)$$

We can now use our velociraptor claws:

$$(a + b)(a + b)$$

$$a^2 + ab + ba + b^2$$

$$a^2 + 2ab + b^2$$

That's one of the "notable products". They are very creatively called, as you can imagine: they are a product, and you notice them! So, we know that:

$$\boxed{(a + b)^2 = a^2 + 2ab + b^2}$$

Let's see more of them:

$$(a - b)^2 = (a - b)(a - b)$$

Again, velociratorize the brackets:

$$(a - b)(a - b)$$

$$a^2 - ab - ba + b^2$$

$$a^2 - 2ab + b^2$$

Another one:

$$\boxed{(a - b)^2 = a^2 - 2ab + b^2}$$

A very important one:

$$(a + b)(a - b)$$

$$a^2 - ab + ab - b^2$$

$$a^2 - b^2$$

The middle terms cancel out! Therefore,

$$\boxed{(a + b)(a - b) = a^2 - b^2}$$

One of my favourites:

$$(a+b)^3 = (a+b)^2(a+b) \quad \text{Indices properties}$$

$$(a+b)^3 = (a^2+2ab+b^2)(a+b) \quad \text{Our newly found power}$$

$$(a+b)^3 = a^3+a^2b+2a^2b+2ab^2+ab^2+b^3 \quad \text{Some fast distribution}$$

$$(a+b)^3 = a^3+3a^2b+3ab^2+b^3 \quad \text{Collecting like terms}$$

What a beautiful product:

$$\boxed{(a+b)^3 = a^3+3a^2b+3ab^2+b^3}$$

You clearly don't need to memorize any of this[2], but they are nice results!

13.5. Exam hints

The only hint I can offer is to be calm: being fast has never meant anything in math, it's not a race. Take your time, be **very careful with the negatives** and everything will be fine!

Summary

- To expand **single brackets** you just need to remember the distributive property of multiplication, aka the velociraptor claws;
- To expand **double brackets** you need to multiply every term in the first bracket by every term in the second bracket. Remember to collect like terms after!;
- You may want to remember the "notable products":
  - $(a+b)^2 = a^2+2ab+b^2$
  - $(a-b)^2 = a^2-2ab+b^2$
  - $(a+b)(a-b) = a^2-b^2$
  - $(a+b)^3 = a^3+3a^2b+3ab^2+b^3$
  - $(a-b)^3 = a^3-3a^2b+3ab^2-b^3$
- Be very careful when distributing negative "signs", such as $-(x-2y) = -x+2y$.

[2] It does help you if you plan on taking the SAT or the ACT, as they require you to do very quick calculations.

Formality after taste

A set of axioms for the real numbers[3]

I won't be as formal as I could be here, but I want to show you the *basic facts* in which the math you are learning stand on (at least for numbers).

Let's say we have some numbers and call them x, y and z. We also need to know how to add them and how to multiply them. Then, all of these are *axioms*, that is, a statement we assume to be true. Here are the axioms:

Axiom 1: commutative laws

This means that:

$$x + y = y + x$$

and that

$$xy = yx$$

The classical "when you are doing addition or multiplication, the order in which you write the numbers does not matter".

Axiom 2: associative laws

This means that:

$$x + (y + z) = (x + y) + z$$

and

$$x(yz) = (xy)z$$

Also known as "if you have many additions or multiplications together, you can do them in any order that you like".

Axiom 3: distributive law

This you know: the fancy name for our velociraptor claws:

$$x(y + z) = xy + xz$$

Axiom 4: existence of identity elements

This axiom asserts that exists a number, which we write as 0, which

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

and there is also a number, which we write as 1, which

$$1 \times x = x \times 1 = x$$

0 and 1 are called "identity" elements (0 of addition, and 1 of multiplication), because they do not "change" the other number if you operate with them: they don't change the other number's identity!

[3] This section is based, as usual, in our lord and savior, Mr Apostol.

Axiom 5: existence of negatives

This axiom tells us that, for every number x, exists a number $-x$ which satisfies

$$x + -x = 0$$

Now the negatives exist!

Axiom 6: existence of reciprocals

This axiom tells us that, for every number x that is not 0, there exists a number $\frac{1}{x}$ which satisfies

$$x \times \frac{1}{x} = 1$$

Now we have division!

With these 6 "basic facts" , we can deduce *everything* you know of algebra. Amazing, isn't it? Let me give you a taste. Let us prove that

$$b - a = b + (-a)$$

The proof goes as follows. First create two new numbers, $x = b - a$ and $y = b + (-a)$. We would like to show that $x = y$. From

$$x = b - a \qquad \text{Add } a \text{ on both sides}$$

$$x + a = b$$

Now that we know that $b = x + a$:

$$y + a = \underbrace{(b + (-a))}_{\text{Definition of } y} + a = b + \underbrace{((-a) + a)}_{\text{Axiom 2}} = b + 0 = b$$

The last equation comes from axiom 4. Notice that we didn't prove that $-a + a = 0$, but it's not hard to prove it as well. So, we have proved that $y + a = x + a$, and from another thing we could prove, called the "cancellation law", we would obtain that $x = y$! Therefore,

$$b - a = b + (-a)$$

This was just a taste, but I hope you liked it!

14. Linear equations

What is an equation?

Before we have our initial discussion, I just want to make **very clear** the difference between an equation and an expression.

The word *equation* comes from the latin word *æquare*, which means "to make equal". This "equal" is the magic word: an equation **needs to have an equal sign**. Therefore, this is an equation

$$x + 2 = 4$$

as it has an equals sign, whereas this is an expression

$$x + 2$$

as it has no equals sign.

As we'll see, we can solve equations. We **cannot solve expressions**, we can only manipulate them to obtain equivalent expressions.

So, in summary: if it has an equals sign, it's an equation. If it does not have one, it's not!

14.1. Why learn how to solve equations?

As humans, we are beings of *telling*. We tell stories all the time, we share our experiences, we marvel at the experiences of others and are inspired by them.

Equations are mathematical stories[1]. By solving them, we are being human, we are sharing with others a story. Not any kind of story, a mystery case: we need to solve the story. An equation (or any good mathematical problem) is the perfect story: it gives you all you need to solve the mystery (good ones would even give you details that *you don't need!*), and from them, by using our logic reasoning, we'll be able to "get to the bottom of it". In our equations, we'll be able to find the *solution*, the (for now) one number that makes everything *true*.

In all, as Robin Williams tells us in *The Dead Poets Society*

> "We don't read and write poetry because it's cute. We read and write poetry because we are members of the human race. And the human race is filled with passion. And medicine, law, business, engineering, these are noble pursuits and necessary to sustain life. But poetry, beauty, romance, love, these are what we stay alive for."

[1] A *proof* would be another type of mathematical story, which is very important!

He was speaking about poetry, yes, but it applies perfectly to equations (and math in general): we don't learn it because of our exams; we don't solve equations because it's cute: we do it because we are members of the human race, and we are filled with the desire for sharing our *truth.* And an equation is just a piece of *truth* we can all share.

14.2. Inverse operations

For us to master the art of telling stories of equations, we need to be masters in the tools we use to solve them. You already know them!, they are inverse operations.

What is an inverse operation? It's just something that *undoes* what a previous operation has done. Let me show you an example.

Imagine you start with the number 17. Now, we can add 5 to it:

$$17 + 5 = 22$$

We have done an operation with the starting 17, we added 5 to it. Now, we can undo that addition by *taking away* 5 from the 22:

$$22 - 5 = 17$$

as you can see, we go back to the intial 17. In this example, the inverse operation of adding 5 is taking away 5, as one undoes the other. You can see this process like this:

$$17 \overset{+5}{\Longrightarrow} 22 \overset{-5}{\Longrightarrow} 17$$

To solve equations, we'll be basically applying inverse operations, so it's of vital importance for you to master them. Let's remember the inverse operations we can do:

Operation	Inverse operation	Example
Addition ($+$)	Subtraction ($-$)	The inverse of doing $+3$ is doing -3
Subtraction ($-$)	Addition ($+$)	The inverse of doing -4 is doing $+4$
Multiplication ($\times$)	Division ($\div$)	The inverse of doing $\times 2$ is doing $\div 2$
Division ($\div$)	Multiplication ($\times$)	The inverse of doing $\div 5$ is doing $\times 5$

Be very careful with particular "pitfalls":

- A number with no sign has a + hidden:

$$2 = +2$$

- Division is very rarely shown using the division symbol. It is usually denoted by a fraction:

$$x \div 4 = \frac{x}{4}$$

- A fraction can be seen as a "mix" of two operations: a multiplication and a division

$$\frac{4}{5}x$$

means "multiply x by 4 and divide the result by 5" (or in the reverse order "divide x by 5 and multiply the result by 4"). Therefore, the inverse operation of multiplying a number by $\frac{4}{5}$ is the same as multiplying it by 5 and dividing it by 4. We can write that as a fraction product as well:

$$\frac{5}{4}x$$

Now, the questio is: why are inverse operations important? You'll see soon!

Solved exercise: finding the initial number in a "story"

Find the initial number in the following story:

> Someone took a number and added 1 to it. After, they took the result of this operations and multiplied by 3. They obtained as a result the number 18. What is the initial number?

We can solve this by writing a diagram telling us what happened:

$$\text{number} \overset{+1}{\Longrightarrow} \text{result } 1 \overset{\times 3}{\Longrightarrow} 18$$

We can "go back" from 18 by doing the reverse operations:

$$\text{number} \overset{-1}{\Longleftarrow} \text{result} \overset{\div 3}{\Longleftarrow} 18$$

So, starting with the 18,we first divide it by 3, and we obtain 6:

$$\text{number} \overset{-1}{\Longleftarrow} 6 \overset{\div 3}{\Longleftarrow} 18$$

Now, we take that 6 and subtract 1from it. The result is 5:

$$5 \overset{-1}{\Longleftarrow} 6 \overset{\div 3}{\Longleftarrow} 18$$

So, the initial number is 5. Let's check our answer: we take the 5 and add 1 to it, the answer is 6. We then multiply 6 by 3 and get 18. We are correct!

14.3. Solving equations

14.3.1. What does it mean to solve an equation?

Before we actually learn how to solve an equation, we need to understand what it means to solve one! Let's take an equation as example:

$$2x - 1 = x + 3$$

This is an equation, as it has an equals sign. Solving it means *finding a value for* x which, if we substitute on the left-hand side, gives the same numerical result if we substitute on the right-hand side.

In this particular case, the unknown we are trying to find is x, but an equation can have any unknown that you want. Sometimes, it can have more than one "variable", such as:

$$3a + b = 5c$$

We can also solve these equations, but we have to **very clear** what is the unkown we are solving for! When there is only one, as in the first example, we can only solve for x, so we don't need to especify. In the second example, we can solve for any of a, b or c, and we have to be very clear for which.

In all, solving an equation for a especific unknown is finding a value for that unknown which, if we substitute on both sides of the equation, gives the same value.

14.3.2. Important terminology

Let's learn some names which will be important when solving equations.

Left-hand and right - hand sides

In all equations, we have a left-hand side and a right-hand side:

$$\underbrace{2a + 3}_{\text{left-hand side}} \quad = \quad \underbrace{a + 3}_{\text{right-hand side}}$$

I will be writing LHS for left-hand side and RHS for right-hand side.

The subject of an equation

This is very important. The subject of an equation is the variable / unknown you want to solve the equation for. For instance, in

$$x + 3 = 5$$

the subject is x, as there is no other possibily. Normally we don't even say you are "solving for x" or "making x the subject". However, for this equation

$$2x - y = 18$$

you **have to** say for which unknown you are solving the equation, x or y. You can either say you are "solving for y" or "making y the subject". Both expressions mean the same.

In summary, if you see a problem with instructions such as "Solve for x" or "Make x the subject", they are asking for exactly the same thing!

14.3.3. General idea of solving an equation

Let's say we are given an equation and we need to solve it for the variable a (also known as making a the subject of the equation). This means we are giving something like

$$\text{algebraic expression} = \text{algebraic expression}$$

Those algebraic expressions need to have at least one variable a for us to solve for. Solving the equation means, after manipulation of the equation, to obtain something like

$$a = \text{algebraic expression}$$

$$\text{or}$$

$$\text{algebraic expression} = a$$

Usually our expression, at the end, will be just a number. However, it can be another algebraic expression **without the variable we are solving for in it**.

Let me give you an example. Let's say we have an equation

$$2x - 5 = 11$$

and we want to solve it for the variable x. In the end, we'll obtain this:

$$x = 8$$

As you can see, one of the sides of of the equation has the variable x **alone**, and the other side **does not have x.** That brings us to the general idea of solving an equation is to manipulate it in a way that *makes the variable we are solving for alone on one of the sides.* Making the variable alone is the same as solving the equation or making that variable the subject of the equation.

That's it! Now, guess how we are going to do that?

14.3.4. Using inverse operations to solve equations

Now that we know what we are aiming for, let's understand the technique with an example. Let's say we need to solve

$$x - 5 = 10$$

First, let's read the story here: there is a number that the result of taking away 5 from is 10. I know you can figure out the answer is 15, but let's try what we learned in the previous section, that is, to find $x =$ expression.

Now, in an equation, we know that the left-hand side is equal to right-hand side. This is a fact. Therefore, if we do **exactly the same operation on both sides of the equation**, both sides will remain being true. Let's see that into work:

$$17 = 17$$

you agree that 17is equal to 17, correct? Now, if we add 2 on both sides of the equation:

$$17 + \underbrace{2}_{\text{add 2 in LHS}} = 17 + \underbrace{2}_{\text{add 2 in RHS}}$$

We obtain

$$19 = 19$$

See, still true! We can do any operation we like, as long as we do it on both sides. Let's try another one, let's square both sides:

$$19^2 = 19^2$$

$$381 = 381$$

again, same number on both sides!

That's the idea when solving an equation: we will progressively *eliminate* everything that is not the variable we want on one of the sides of the equation. Eventually, everything but the variable we are solving will be gone, and we will be done! Now, guess, what operations we are going to be doing? You guessed: inverse operations!

Let's go back to our initial equation:

$$x - 5 = 10$$

If we could kill that annoying -5 on the LHS we would have exactly what we want, x alone. We do know, however, that the inverse operation of -5 is $+5$. Let's, then, do $+5$ on both sides of the equation:

$$x - 5 \underbrace{+5}_{+5 \text{ on the LHS}} = 10 \underbrace{+5}_{+5 \text{ on the RHS}} \qquad \text{Do } +5 \text{ on } \textbf{both sides}$$

$$x \underbrace{-5 + 5}_{\text{Cancel each other}} = 15 \qquad \text{The } -5 + 5 \text{ is } 0$$

$$x + 0 = 15$$

$$x = 15 \qquad \text{Anything } +0 \text{ is the thing itself}$$

There you go! We solved the equation. Now, a great idea is to check our answer by taking the value of x we found and subsituting it back in the original equation:

$$\underbrace{x}_{\text{We know that } x = 15} - 5 = 10 \qquad \text{Subsitute all } x \text{ for } 15$$

$$15 - 5 = 10 \qquad \text{Simplify both sides fully (don't forget BIDMAS!)}$$

$$15 - 5$$

$$10 = 10 \qquad \text{Both sides are equal, so we are correct!}$$

That's one great thing about equation exercises: you can always check your answer. I highly recommend you doing this.

Let's see another example:

$$5d = 85$$

We are searching for a number that we called d which, when me multiply by 5 gives 85. Let's use inverse operations to solve this:

$$5d = 85 \qquad \text{The inverse operation of } \times 5 \text{ is dividing by } 5$$

$$\frac{5d}{\color{red}5} = \frac{85}{\color{red}5} \qquad \text{Divide both sides by } {\color{red}5}$$

$$\frac{\cancel{5}d}{\cancel{5}} = \frac{85}{5}$$

$$d = \frac{85}{5} \qquad 5d \text{ divided by } 5 \text{ is just } d$$

$$d = 17 \qquad 85 \text{ divided by } 5 \text{ is } 17$$

Again, we're done! Let's subsitute in the original equation and check:

$$5 \times 17 = 85$$

$$85 = 85$$

We are correct!

One more example:

$$\frac{p}{5} = 16$$

Again, let's apply the inverse operation that makes p alone. Here, we are dividing p by 5, and the inverse operation of dividing by 5 is multiplying by 5. Let's do that on both sides:

$$5\times\frac{p}{5} = 16\times 5 \qquad \text{Do } \times 5 \text{ on both sides}$$

$$\not{5}\times\frac{p}{\not{5}} = 80$$

$$p = 80$$

So $p = 80$. If we substitute 80 back into p, we will get $\frac{80}{5}$, which is indeed 16.

A bit of a rant: nothing "goes to the other side"

Now that you have seen the basic idea, I have a **very important warning to** give. Take our first example:

$$x - 5 = 10$$

When we were solving it, we did $+5$ on both sides, as it's the inverse operation of -5:

$$x - 5 + 5 = 10 + 5$$

We also know that $-5 + 5 = 0$, so we can write

$$x = 10 + 5$$

Many people look at these two steps in sequence

$$x - 5 = 10 \qquad \text{Initial equation}$$

$$x = 10 + 5 \qquad \text{After } +5 \text{ and doing } x + 0$$

and they say that "the -5 **went over to the other side positive**". This is incorrect, never say that. Look carefully at the steps:

$$x - 5 = 10 \qquad \text{Initial equation}$$

$$x - 5 + 5 = 10 + 5 \qquad +5 \text{ on both sides}$$

$$x + 0 = 15 \qquad -5 + 5 = 0$$

$$x = 15$$

As you can see by the arrow, the -5 never went to right-hand side, it was *killed* by the $+5$ from the inverse operation. As we do have to do the inverse operation on both sides, it *seems* that the -5 "went to the other side with signal changed". The arrow does not lie: it didn't. The -5 dies in the process. It's easy to remember that nothing "goes to the other side", just keep in your mind that

Equations are solved by killing stuff. They are not a stroll in the park, nothing goes over to the other sid

Some more examples:

$$\frac{1}{2}a = 7$$

Remember that $\frac{1}{2}a = \frac{a}{2}$:

$$\frac{1}{2}a = 7$$

$$\frac{a}{2} = 7 \qquad \text{Remember that } \frac{1}{2}a = \frac{a}{2}$$

$$2\times\frac{a}{2} = 7\times 2 \qquad \text{Do } \times 2 \text{ on both sides}$$

$$\not{2}\times\frac{a}{\not{2}} = 14$$

$$a = 14$$

Checking our answer: $LHS = \frac{1}{2} \times 14 = 7 = RHS$.
Another:

$$2 = 4 + a$$

Let's make a alone on the RHS here:

$$2 = +4 + a \qquad \text{Remember that } 4 = +4$$

$$2-4 = +4-4 + a \qquad \text{Do } -4 \text{ on both sides}$$

$$-2 = a$$

Let's check our answer: $RHS = 4 + a = 4 + -2 = 4 - 2 = 2 = LHS$. Again, we are correct!

14.3.5. "Longer" equations

Now that you understand the basic idea of solving equations, let's try solving "longer" equations. Example:

$$2x - 5 = 11$$

This equation is "longer" because we'll need to do more than one inverse operation to solve it. But **the idea is the same**: we are going to use inverse operations to make x alone on one of the sides.

To solve this equation, we can take two different paths. Let's do both.

$$
\begin{aligned}
2x - 5 &= 11 \\
2x - 5 + 5 &= 11 + 5 && \text{Doing } +5 \text{ on both sides} \\
2x &= 16 && \text{Collecting like terms} \\
\frac{2x}{2} &= \frac{16}{2} && \text{Doing } /2 \text{ on both sides} \\
x &= 8
\end{aligned}
$$

We are done! If you subsitute $x = 8$ at the original equation, you'll obtain $LHS = 2 \times 8 - 5 = 16 - 5 = 11 = RHS$, which proves that we're correct.

The other way of solving it is by, instead of doing +5 on both sides first, we choose to divide both sides by 2:

$$
\begin{aligned}
2x - 5 &= 11 \\
\frac{2x - 5}{2} &= \frac{11}{2} && \text{Doing } /2 \text{ on both sides} \\
\frac{\cancel{2}x}{\cancel{2}} - \frac{5}{2} &= 5.5 && \text{Separate the fraction to help} \\
x - 2.5 &= 5.5 && \text{Do the division} \\
x - 2.5 + 2.5 &= 5.5 + 2.5 && \text{Do } +2.5 \text{ on both sides} \\
x &= 8
\end{aligned}
$$

As you can see, we reach the exact same answer. I do think it is a bit harder to this second way, so I highly recommend getting rid of everything that isn't the variable you are solving for, and only on the last step to multiply or divide both sides by something.

A classic mistake is to do this:

$$
\begin{aligned}
2x - 5 &= 11 \\
x - 5 &= \frac{11}{2} && \text{VERY WRONG!}
\end{aligned}
$$

This is very common when, instead of thinking that you solve equations by applying operations to both sides, you think like "sending to the other side". Those people would forget to divide everything on the LHS by 2, and would get a wrong answer! So, choose the path that you prefer, but always remember: **do the inverse operation on both sides**.

Let's see some more examples:

$$\frac{x+2}{3} = 5$$

$$3\times\frac{x+2}{3} = 5\times 3 \qquad \text{Do } \times 3 \text{ on both sides}$$

$$\cancel{3}\times\frac{x+2}{\cancel{3}} = 15$$

$$x+2 = 15$$

$$x+2-2 = 15-2 \qquad \text{Doing } -2 \text{ on both sides}$$

$$x = 13$$

Checking our answer: $LHS = \frac{13+2}{3} = \frac{15}{3} = 5 = RHS$. We are correct!
More:

$$\frac{b}{2} + 5 = 10$$

$$\frac{b}{2} + 5-5 = 10-5 \qquad \text{Doing -5 on both sides}$$

$$\frac{b}{2} = 5$$

$$2\times\frac{b}{2} = 5\times 2 \qquad \text{Do } \times 2 \text{ on both sides}$$

$$\cancel{2}\times\frac{b}{\cancel{2}} = 10$$

$$b = 10$$

checking our answer: $LHS = \frac{10}{2} + 5 = 5 + 5 = 10 = RHS$. We are correct!
How about more:

$$-4x + 5 = 21$$

$$-4x + 5-5 = 21-5 \qquad \text{Doing } -5 \text{ on both sides}$$

$$-4x = 16$$

$$\frac{-4x}{-4} = \frac{16}{-4} \qquad \text{Dividing both sides by } -4$$

$$x = -4$$

checking our asnwer: $LSH = -4 \times -4 + 5 = 16 + 5 = 21 = RHS$. Correct!

More more more

$$
\begin{aligned}
2x - 4 &= 3 \\
2x - 4 + 4 &= 3 + 4 && \text{Doing } +4 \text{ on both sides} \\
2x &= 7 \\
\frac{2x}{2} &= \frac{7}{2} && \text{Dividing both sides by } 2 \\
x &= \frac{7}{2}
\end{aligned}
$$

what a pretty answer, a fraction! Let's check it: $LHS = 2 \times \frac{7}{2} - 4 = \cancel{2} \times \frac{7}{\cancel{2}} - 4 = 7 - 4 = 3 = RHS$. Amazing!

Let's try a "word" problem: I am a years old. My wife is 2 years younger than I am. Together, we have lived 44 years. What is my age?

As my wife is 2 years younger than I am, we know that her age has to be $a - 2$, as I am a years old. Adding our ages together gives 44:

$$\underbrace{a}_{\text{me}} + \underbrace{a - 2}_{\text{my wife}} = 44$$

we just need to solve the equation!

$$
\begin{aligned}
\underline{a + a} - 2 &= 44 && \text{Identifyling like terms} \\
2a - 2 &= 44 && \text{Collecting like terms} \\
2a - 2 + 2 &= 44 + 2 && \text{Doing } +2 \text{ on both sides} \\
2a &= 46 \\
\frac{2a}{2} &= \frac{46}{2} && \text{Dividing both sides by } 2 \\
a &= 23
\end{aligned}
$$

Given that a represents my age, we know that I am 23 years old. It does check out, as if I am 23 my wife has to be $23 - 2 = 21$, and adding our ages together is indeed 44: $23 + 21 = 44$.

14.3.6. Equations with the variable on both sides

Now, we'll solve equations such as

$$2x - 5 = 3x - 2$$

The idea is to choose one of the sides of the equation to have the variables on, and we will kill everything else on that side! Gotta love some killing while doing equations.

Let's put the x on the LSH. You can choose any side that you want, I just picked the left because I wanted. How could we remove that $3x$ from the RHS? Hum, I wonder... just kidding, just do inverse operations!

$$2x - 5 - 3x = 3x - 3x - 2 \qquad \text{The inverse of } +3x \text{ is } -3x$$

$$\underline{2x} - 5\underline{-3x} = -2 \qquad \text{Finding like terms}$$

$$\underline{2x - 3x} - 5 = -2 \qquad \text{Collecting like terms}$$

$$-x - 5 = -2$$

We have arrived at a "regular" equation! We just need to keep going:

$$-x - 5 + 5 = -2 + 5 \qquad \text{The inverse of } -5 \text{ is } +5$$

$$-x = 3$$

$$\frac{-1x}{-1} = \frac{3}{-1} \qquad \text{The inverse of } \times -1 \text{ is } \div -1$$

$$x = -3$$

let's check our answer. Now, we substitute x both on the LHS and RHS and (hopefully!) they will be equal.

$$LHS = 2 \times -3 - 5 = -6 - 5 = -11$$

now the RHS:

$$RHS = 3 \times -3 - 2 = -9 - 2 = -11$$

They are equal, so we are correct!

Let us try more examples:

$$5d - 4 = -3d + 4$$

Same idea, let's put all d variables on the LHS:

$$5d - 4 + 3d = -3d + 3d + 4 \qquad \text{The inverse of } -3d \text{ is } +3d$$

$$\underline{5d} - 4 \underline{+3d} = 4 \qquad \text{Identifying like terms}$$

$$\underline{5d + 3d} - 4 = 4 \qquad \text{Collecting like terms}$$

$$8d - 4 = 4$$

$$8d - 4 + 4 = 4 + 4 \qquad \text{The inverse of } -4 \text{ is } +4$$

$$8d = 8$$

$$\frac{8d}{8} = \frac{8}{8} \qquad \text{The inverse of } \times 8 \text{ is } \div 8$$

$$d = 1$$

Checking our answer: $LHS = 5 \times 1 - 4 = 5 - 4 = 1$ and $RHS = -3 \times 1 + 4 = -3 + 4 = 1$. We are correct!

Let's try

$$4x - 5 = 2x$$

$$4x - 2x - 5 = 2x - 2x \qquad \text{The inverse of } +2x \text{ is } -2x$$

$$\underline{4x - 2x} - 5 = 0 \qquad \text{Collecting like terms}$$

$$2x - 5 = 0$$

$$2x - 5 + 5 = 0 + 5 \qquad \text{The inverse of } -5 \text{ is } +5$$

$$2x = 5$$

$$\frac{2x}{2} = \frac{5}{2} \qquad \text{The inverse of } \times 2 \text{ is } \div 2$$

$$x = \frac{5}{2}$$

Let's check it: $LHS = 4 \times \frac{5}{2} - 5 = 10 - 5 = 5$ and $RHS = 2 \times \frac{5}{2} = 5$. We are correct!

How about

$$\frac{2y}{3} - 5 = 3y + 5$$

The first thing I'd do is multiply both sides by 3 to get rid of that denominator:

$$3 \times (\frac{2y}{3} - 5) = 3 \times (3y + 5)$$

$$2y - 15 = 9y + 15$$

Now we just continue as normal:

$$2y{\color{red}-9y}-15=9y{\color{red}-9y}+15 \quad \text{The inverse of } +9y \text{ is } -9y$$

$$\underline{2y-9y}-15=15 \quad \text{Identifying like terms}$$

$$-7y-15=15 \quad \text{Collecting like terms}$$

$$-7y-15{\color{red}+15}=15{\color{red}+15} \quad \text{The inverse of } -15 \text{ is } +15$$

$$-7y=30$$

$$\frac{-7y}{\color{red}-7}=\frac{30}{\color{red}-7} \quad \text{The inverse of } \times-7\text{is } \div-7$$

$$y=-\frac{30}{7}$$

what a lovely result! Let's check it: $LHS = \frac{2}{3}\times-\frac{30}{7}-5=\frac{-20}{7}-5=-\frac{-55}{7}$; $RHS = 3\times-\frac{30}{7}+5=-\frac{90}{7}+5=\frac{-55}{7}$. Lovely fractions, all correct!

14.3.7. Equations with brackets

If you need to solve an equation with brackets, good news, it's again the same! Let me show you some examples:

$$2(x-5)=8$$

One way is expand the brackets:

$$2\times x+2\times-5=8 \quad \text{Expanding the brackets}$$

$$2x-10=8$$

$$2x-10{\color{red}+10}=8{\color{red}+10} \quad \text{The inverse of } -10 \text{ is } {\color{red}+10}$$

$$2x=18$$

$$\frac{2x}{\color{red}2}=\frac{18}{\color{red}2} \quad \text{The inverse of } \times2 \text{ is } {\color{red}\div2}$$

$$x=9$$

After expanding the brackets it's just a normal equation (not that it wasn't normal before!). Let me just show you a shortcut for the same equation:

$$
\begin{aligned}
2(x-5) &= 8 \\
\frac{2(x-5)}{2} &= \frac{8}{2} && \text{Dividing both sides by } 2 \\
\frac{\cancel{2}(x-5)}{\cancel{2}} &= 4 \\
x-5 &= 4 \\
x-5+5 &= 4+5 && \text{The inverse of } -5 \text{ is } +5 \\
x &= 9
\end{aligned}
$$

as expected, the same result. Sometimes it can be faster to do it like this.

Just a more sophisticated example:

$$
\begin{aligned}
3(x-5)-5 &= 2-2(x-4) \\
3 \times x + 3 \times -5 - 5 &= 2 - 2 \times x - 2 \times -4 && \text{Expanding the brackets} \\
3x - 15 - 5 &= 2 - 2x + 8 \\
3x - 20 &= 10 - 2x && \text{Collecting like terms} \\
3x + 2x - 20 &= 10 - 2x + 2x && \text{The inverse of } -2x \text{ is } +2x \\
5x - 20 &= 10 \\
5x - 20 + 20 &= 10 + 20 && \text{The inverse of } -20 \text{ is } +20 \\
5x &= 30 \\
\frac{5x}{5} &= \frac{30}{5} && \text{The inverse of } \times 5 \text{ is } \div 5 \\
x &= 6
\end{aligned}
$$

Let's check our answer: $LHS = 3(6-5)-5 = 3(1)-5 = 3-5 = -2$ and $RHS = 2-2(6-4) = 2-2(2) = 2-4 = -2$. We're correct!

14.4. Changing the subject of equations

Let's focus for a while in changing the subject of equations. If you remember, this just means solving an equation for a specific variable, and nothing changes (except that there are many "letters"!). Let's see an example.

Make y the subject in the following equation:

$$4x + 5y = 8$$

It's exactly the same idea as if you were solving a "normal equation": you need to solve for y. And we just use our inverse operations technique!

$$\begin{aligned}
4x + 5y &= 8 \\
4x - 4x + 5y &= 8 - 4x && \text{The inverse of } +4x \text{ is } -4x \\
5y &= 8 - 4x \\
\frac{5y}{5} &= \frac{8-4x}{5} && \text{The inverse of } \times 5 \text{ is } \div 5 \\
\frac{\cancel{5}y}{\cancel{5}} &= \frac{8-4x}{5} \\
y &= \frac{8-4x}{5}
\end{aligned}$$

That's it! It's even easier, if you think: you don't have to do calculations! Just be very careful when multiplying and dividing the whole side of the equation. It's very common to forget that the whole side gets multiplied! The next example shows this very clearly:

Make b the subject of the equation

$$\begin{aligned}
a &= \frac{b}{3} - d + 5 \\
a + d &= \frac{b}{3} - d + d + 5 && \text{The inverse of } -d \text{ is } +d \\
a + d &= \frac{b}{3} + 5 \\
a + d - 5 = \frac{b}{3} &= 5 - 5 && \text{The inverse of } +5 \text{ is } -5 \\
a + d - 5 &= \frac{b}{3} \\
3 \times (a + d - 5) &= \frac{b}{3} \times 3 && \text{The inverse of } \div 3 \text{ is } \times 3 \\
3a + 3d - 15 &= \frac{b}{\cancel{3}} \times \cancel{3} && \text{Expand the brackets} \\
3a + 3d - 15 &= b
\end{aligned}$$

As you can see, nothing changes: just apply the inverse operations on both sides and everything will work out.

14.5. Exam hints

When solving equations and changing the subject, just apply the inverse operations on both sides. On the whole side. Do not “send things to the other side” . Everything will be fine!

Summary

- An **algebraic expression** is something like $2x - 5$, and an **equation has an equal sign**, such as $2x - 5 = 5$;
- **To solve an equation** for a variable x means manipulating the equation to obtain another equation that looks like $x =$ expression;
- You can apply the same operation on both sides of an equation. **Inverse operations** are very useful as they allow you to get rid of things;
- **Changing the subject** to a is the same as solving for a.

15. Indices

15.1. Why learn indices

Because we are lazy! Indices are a classic example of mathematicians being lazy. It's annoying to write a number times itself, as you need to write 3 characters:

$$2 \times 2$$

So it's much easier to write

$$2^2$$

We save one character! Nothing much, you say, but imagine writing this 100000 times. You save a lot of pen ink!

Jokes aside, indices were initially just *notation*, a way to write something. They have interesting properties though, and that's what we are going to study.

15.2. Definition

15.2.1. Positive integer indices

The definition of indices for positive integers is just repeated multiplication. Given any number a and a positive integer n, we define a^n as

$$a^n = \underbrace{a \times a \times a \times \ldots \times a}_{n \text{ times}}$$

This means that we multiply a by itself b times. The number a is called the *base* and the number n is called the *index* or the *power*.

Let's calculate some:

$$2^3 = \underbrace{2 \times 2 \times 2}_{3 \text{ times}} = 8$$

We can read 2^3 in many equivalent ways: as "2 to the power of 3", as "2 raised to the 3 power", as "2 to the third", or even "2 to the 3". Another example:

$$1^4 = \underbrace{1 \times 1 \times 1 \times 1}_{4 \text{ times}}$$

A "harder" one:

$$(-3)^3 = \underbrace{-3 \times -3 \times -3}_{3 \text{ times}} = -27$$

There are two powers which have "special" names:

$$5^2 = \underbrace{5 \times 5}_{2 \text{ times}} = 25$$

when raising to the power of 2 we normally say "squared" (as this is the area of a square with side length you are well, squaring). The other "special" power name is for powers of 3:

$$4^3 = \underbrace{4 \times 4 \times 4}_{3 \text{ times}} = 64$$

powers of 3 are normally read as "cubed" (as this is the volume of a cube with side length you are cubing).

15.2.2. Negative integer and fractional indices

We will, using the indices properties, understand what it means to raise a number to a negative integer and to a fraction later.

15.3. Indices properties

There are many indices properties you need to memorise. Some people call these "index laws" or similar, but "law" sounds someone told the indices to behave in this way. They are logical facts that come from the definition we have seen, not a law.

For all of the following, a, b are different from 0 and n and m can be any integer[1] These are the properties:

Property 1 When we have a *product* of powers of the same number, we can *add the powers*:

$$a^n \times a^m = a^{n+m}$$

Property 2 When we have a division of powers of the same number, we can *subtract the powers*:

$$\frac{a^n}{a^m} = a^{n-m}$$

Property 3 When we have a power raised to another power, we can *multiply the powers*:

$$(a^n)^m = a^{n \times m}$$

Property 4 *Any number* to the power of 0 is 1:

$$a^0 = 1$$

[1] n could be any real number, but this has to be shown and does have some problems. The wikipedia page for indices (`https://en.wikipedia.org/wiki/Exponentiation`) cover this very well.

Property 5 Raising fractions to a power is the same as *raising the numerator and the denominator by the power*:

$$\left(\frac{a}{b}\right)^n = \frac{a^n}{b^n}$$

Property 6 Whenever you raise a product to a power *everything gets the power*:

$$(ab)^n = a^n \times b^n$$

Property 7 To raise a number to a negative power is the same as raising the *reciprocal of the number to the positive power*:

$$a^{-n} = \left(\frac{a}{1}\right)^{-n} = \left(\frac{1}{a}\right)^n = \frac{1}{a^n}$$

More generally, if you have a fraction raised to a negative power:

$$\left(\frac{a}{b}\right)^{-n} = \left(\frac{b}{a}\right)^n = \frac{b^n}{a^n}$$

Property 8 To raise a number to a fractional power is the same as taking roots:

$$a^{\frac{n}{m}} = \sqrt[m]{a^n} = (\sqrt[m]{a})^n$$

Notice what is happening:

$$a^{\frac{n}{m}} = (\sqrt[m]{a})^n$$

15.4. Using the properties to solve problems

Let's divide the possible questions into two groups: questions with only numbers when you have a calculator and everything else

15.4.1. Questions with numbers when you can use a calculator

On the unlikely possibility of something like this appearing on your exam:

$$\left(\frac{4}{3}\right)^{\frac{3}{4}}$$

and you are on a calculator paper, you will *simply copy it on your calculator.* There you go. No thinking required.

15.4.2. Interesting questions

Now, for the questions you actually need to know things! Basically, you need to know by heart every property and apply them properly, as you cannot use a calculator or that would not help. Let's see some examples.

1. $x^3 \times x^4$

   In all indices questions, identify what property you can use and then apply it. In this case **Property 1** can be used:

$$x^3 \times x^4 = x^{3+4} = x^7$$

2. $\left(\frac{4}{p}\right)^0$

   One of my favourites. Something to the power of 0. **Property 4** to the rescue: anything to the power of 0 is 1:

$$\left(\frac{4}{p}\right)^0 = 1$$

3. $b^5 \times b^{-6}$

   Again, we can use **Property 1**:

$$b^5 \times b^{-6} = b^{5+-6} = b^{5-6} = b^{-1}$$

   but we are not done! We can use **Property 7:**

$$b^{-1} = \frac{1}{b^1} = \frac{1}{b}$$

   Therefore,

$$b^5 \times b^{-6} = \frac{1}{b}$$

4. $8^{\frac{2}{3}}$

   You have two strategies for questions like this. First is to apply **Property** 8:

$$8^{\frac{2}{3}} = \left(\sqrt[3]{8}\right)^2 = (2)^2 = 4$$

   I prefer doing it differently. Notice that you are raising 8 to a fraction with denominator 3. That is a "hint": you can write 8 as something cubed! In this case it is 2, as $2^3 = 8$. You can rewrite the expression as

$$8^{\frac{2}{3}} = \left(2^3\right)^{\frac{2}{3}}$$

   Now, we can use **Property 3**:

$$\left(2^3\right)^{\frac{2}{3}} = 2^{3\times\frac{2}{3}} = 2^2 = 4$$

   (Just remember: if you can use a calculator, just put $8^{\frac{2}{3}}$ in it during the exam!)

5. $4^{-3/2}$

   Let's begin using **Property 7** to get rid of the negative power:

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

   Now, you can either use **Property 8**:

$$\frac{1}{\left(\sqrt[2]{4}\right)^3} = \frac{1}{2^3} = \frac{1}{8}$$

   or you can use "my approach" again:

$$\frac{1}{(2^2)^{3/2}} = \frac{1}{2^{2\times\frac{3}{2}}} = \frac{1}{2^3} = \frac{1}{8}$$

6. $2a \times 4a^2$

   First, remember that in a multiplication order does not matter, so we can just do 2×4 right off the bat:

$$2a \times 4a^2 = 8 \times a \times a^2$$

   Now, we can just **Property 1** to simplify $a \times a^2$

$$2a \times 4a^2 = 8 \times a^{1+2} = 8a^3$$

7. $x^3y^2 \times x^2y^3$

   Same thing as in the example above, you can "solve the x" part before and then join it with the "y part":

$$x^3 \times y^2 \times x^2 \times y^3 = x^3 \times x^2 \times y^2 \times y^3 = x^{3+2} \times y^{2+3} = x^5 \times y^5 = x^5y^5$$

8. $(4p^2q^{-3})^{-2}$

   First, let's use **Property 6** to expand the brackets (remember that everything gets the power):

$$(4p^2q^{-3})^{-2} = 4^{-2}\left(p^2\right)^{-2}\left(q^{-3}\right)^{-2}$$

   Let's use **Property 7** to simplify 4^{-2}, and **Property 3** to get rid of the brackets:

$$\frac{1}{4^2}p^{2\times -2}q^{-3\times -2} = \frac{1}{16}p^{-4}q^6$$

   Finally, let's just use **Property 7** to simplify p^{-4}:

$$\frac{1}{16} \times \frac{1}{p^4} \times q^6$$

   Let's just put everything in one fraction so that it looks pretty:

$$(4p^2q^{-3})^{-2} = \frac{q^6}{16p^4}$$

9. $\frac{16b^{-4}}{2b \times 4b^{-2}}$

Let's start by using **Property 1** on the denominator:

$$\frac{16b^{-4}}{2b \times 4b^{-2}} = \frac{16b^{-4}}{8b^{1+-2}} = \frac{16b^{-4}}{8b^{-1}}$$

We can now divide 16 by 8 and use **Property 2**:

$$\frac{16b^{-4}}{8b^{-1}} = 2b^{-4--1} = 2b^{-4+1} = 2b^{-3}$$

Finally, we can use **Property 7** to convert the expression to a fraction:

$$2b^{-3} = \frac{2}{b^3}$$

Therefore

$$\frac{16b^{-4}}{2b \times 4b^{-2}} = \frac{2}{b^3}$$

10. $\sqrt{289d^8e^{14}}$

There are two paths here.

The first one is to use **Property 8** straight away:

$$\sqrt[2]{289d^8e^{14}} = \left(289d^8e^{14}\right)^{\frac{1}{2}}$$

We now use **Property 6**:

$$(289d^8e^{14})^{\frac{1}{2}} = 289^{\frac{1}{2}} \times (d^8)^{\frac{1}{2}} \times (e^{14})^{\frac{1}{2}}$$

You could put $289^{\frac{1}{2}}$ on your calculator, or use the fact that $289 = 17^2$, or even use **Property 8** back to change the power of $\frac{1}{2}$ to a square root and reach that $289^{\frac{1}{2}} = 17$. Now, we can use **Property 3** on the other two terms:

$$289^{\frac{1}{2}} \times (d^8)^{\frac{1}{2}} \times (e^{14})^{\frac{1}{2}} = 17 \times d^{8\times\frac{1}{2}} \times e^{14\times\frac{1}{2}} = 17 \times d^4 \times e^7$$

We are done:

$$\sqrt{289d^8e^{14}} = 17d^4e^7$$

The second path is remembering that if you have a product inside a root you can split the root:

$$\sqrt{289d^8e^{14}} = \sqrt{289} \times \sqrt{d^8} \times \sqrt{e^{14}}$$

You can use **Property 8** in the two last terms (after calculating $\sqrt{289} = 17$):

$$\sqrt{289} \times \sqrt{d^8} \times \sqrt{e^{14}} = 17 \times (d^8)^{\frac{1}{2}} \times (e^{14})^{\frac{1}{2}}$$

Notice that we have reached exactly the same expression. So, you can use whatever approach you wish.

15.5. Equations involving indices

The hardest indices questions in the IGCSE are equations with indices. They are hard because you have not learned an "easy" way to deal with them[2]. There are techniques, though. Let's see some examples.

1. $2^x = 8$

   There is a general technique to solve these type of equations. The basic idea is that you want both sides of the equation to be *the same number raised to a power*. It doesn't matter what number, but both sides must be that number. In this case, we would want something like

   $$2^x = 2^{\text{something}}$$

   It is somewhat obvious that you cannot do anything to 2^x (in general, if you have a prime number on either side you are going to have to convert the other side to the same prime number), but you need practice in order to identify these quickly.

   Now that we know what we are aiming for, we need to convert the 8 to a power of 2. Given that $2^3 = 8$, we can rewrite the equation as

   $$2^x = 2^3$$

   You have now an equation that says "what power of 2 is equal to 2 to the power of 3"? Well, it isn't 3! Basically, you can just make the powers equal:

   $$x = 3$$

   There you go.

2. $10^{2x} = 100$

   Same idea: we need both sides to be a power of the same number. Notice that $100 = 10^2$, so we can rewrite the equation as

   $$10^{2x} = 10^2$$

   Again, making the powers equal we have

   $$2x = 2$$

   $$\frac{2x}{2} = \frac{2}{2}$$

   $$x = 1$$

[2] Logarithms.

3. $3^y = \frac{1}{27}$

   Again, the same idea: make both sides powers of the same base. In this case, 3 is prime and there is nothing to change, but we can use **Property 7** and notice that

$$3^{-3} = \frac{1}{3^3} = \frac{1}{27}$$

   Rewriting the equation:

$$3^y = 3^{-3}$$

   Finally, making the powers equal:

$$y = -3$$

4. $17^{3x-1} = 289^{2x}$

   Guess, what? Would you have guessed that $289 = 17^2$? You should, as these questions are always like this, full of hints:

$$17^{3x-1} = (17^2)^{2x}$$

   Using **Property 3**:

$$17^{3x-1} = 17^{2\times 2x}$$

$$17^{3x-1} = 17^{4x}$$

   Finally, making the powers equal:

$$3x - 1 = 4x$$

$$-1 = 4x - 3x$$

$$-1 = x$$

$$x = -1$$

5. $8^{4x-1} = 16^{2x+2}$

   The same technique, both sides with a power of the same base. Now, 16 is not a power of 8, so we cannot just change the left side. Notice that both 8 and 16 are powers of 2, though:

$$\left(2^3\right)^{4x-1} = \left(2^4\right)^{2x+2}$$

   Using **Property 3**:

$$2^{3\times(4x-1)} = 2^{4\times(2x+2)}$$

Making the powers equal:

$$3(4x-1) = 4(2x+2)$$
$$12x - 3 = 8x + 8$$
$$12x - 8x = 8 + 3$$
$$4x = 11$$
$$x = \frac{11}{4}$$

15.6. Exam hints

1. If you see indices questions with *only numbers*, just copy the expression on your calculator;
2. Memorise the properties;
3. Be careful with classic mistakes such as
   a) $2^{-2} \neq -2^2$
   b) $(2x)^2 \neq 2x^2$
4. If you see an equation with the unknown on the power, remember to make both sides powers of the same base.

Summary

- To raise a number to **a power**, often called the **index**, and denoted as

$$a^n = \underbrace{a \times a \times a \times \ldots \times a}_{n \text{ times}}$$

  in which a is called the **base**;

- The properties of indices you need to know are:
  1. $a^n \times a^m = a^{n+m}$
  2. $\frac{a^n}{a^m} = a^{n-m}$
  3. $(a^n)^m = a^{n \times m}$
  4. $a^0 = 1$
  5. $\left(\frac{a}{b}\right)^n = \frac{a^n}{b^n}$
  6. $(ab)^n = a^n b^n$
  7. $a^{-n} = \frac{1}{a^n}$, or the more general $\left(\frac{a}{b}\right)^{-n} = \left(\frac{b}{a}\right)^n$

8. $a^{n/m} = \sqrt[m]{a^n} = (\sqrt[m]{a})^n$

- To solve equations with the unknown on the indices, make the left-hand side and the right-hand side be a power of the same base.

Formality after taste

Proof of the indices properties

We will show the properties when the powers are integers. For any real numbers the properties are also true, but the proofs are different.

Property 1 We want to show that

$$a^n \times a^m = a^{n+m}$$

We know that

$$a^n = \underbrace{a \times a \times a \times \ldots \times a}_{n \text{ times}}$$

and that

$$a^m = \underbrace{a \times a \times a \times \ldots \times a}_{m \text{ times}}$$

(Notice that even though the expansions "look the same", they have different numbers of a: the first has n and the second m "hidden"; the number of a in the $\ldots$ is different!).

We can just use this expansions now:

$$a^n \times a^m = \underbrace{a \times a \times a \times \ldots \times a}_{n \text{ times}} \times \underbrace{a \times a \times a \times \ldots \times a}_{m \text{ times}}$$

therefore, we have n a's in the first part and m a's in the second. Those add up to $n + m$ a's in total:

$$a^n \times a^m = \underbrace{\underbrace{a \times a \times a \times \ldots \times a}_{n \text{ times}} \times \underbrace{a \times a \times a \times \ldots \times a}_{m \text{ times}}}_{n+m \text{ times}}$$

Thus we have that

$$a^n \times a^m = \underbrace{a \times a \times a \times \ldots \times a}_{n+m \text{ times}}$$

and we know that multiplying a by itself $n + m$ times is the same as raising a to the $n + m$:

$$a^n \times a^m = a^{n+m}$$

Property 2 We want to show that

$$\frac{a^n}{a^m} = a^{n-m}$$

Let's split this in two cases, when $n \geq m$ and when $n < m$. Starting with $n \geq m$ and remembering that we can expand a^n and a^m as in the previous property proof:

$$\frac{a^n}{a^m} = \frac{\overbrace{a \times a \times a \times \ldots \times a}^{n \text{ times}}}{\underbrace{a \times a \times a \times \ldots \times a}_{m \text{ times}}}$$

we know that $n \geq m$, so we can rewrite the above as

$$\frac{a^n}{a^m} = \underbrace{\frac{a}{a} \times \frac{a}{a} \times \frac{a}{a} \times \ldots \times \frac{a}{a}}_{m \text{ times}} \times \underbrace{a \times a \times a \times \ldots \times a}_{n-m \text{ times}}$$

dividing a number by itself gives 1:

$$\frac{a^n}{a^m} = \underbrace{1 \times 1 \times 1 \times \ldots \times 1}_{m \text{ times}} \times \underbrace{a \times a \times a \times \ldots \times a}_{n-m \text{ times}}$$

and multiplying by 1 does nothing:

$$\frac{a^n}{a^m} = \underbrace{a \times a \times a \times \ldots \times a}_{n-m \text{ times}}$$

Now, we know that multiplying a by itself $n - m$ times is the same as raising a to the $n - m$ power:

$$\frac{a^n}{a^m} = \underbrace{a \times a \times a \times \ldots \times a}_{n-m \text{ times}} = a^{n-m}$$

The other case, when $n < m$, has a similar proof. I'll leave that one to you.

Property 3 We want to show that

$$(a^n)^m = a^{n \times m}$$

This one is quite simple to prove, we just need to use the definition of raising to a power and **Property 1**:

$$(a^n)^m = \underbrace{(a^n) \times (a^n) \times (a^n) \times \ldots \times (a^n)}_{m \text{ times}} \quad \text{Definition of raising to the } m \text{ power}$$

$$(a^n)^m = \underbrace{a^n \times a^n \times a^n \times \ldots \times a^n}_{m \text{ times}} \quad \text{Removing the brackets}$$

$$(a^n)^m = a^{\overbrace{n + n + n + \ldots + n}^{m \text{ times}}} \quad \text{Using } \textbf{Property 1}$$

$$(a^n)^m = a^{n \times m} \quad \text{Adding } n \text{ to itself } m \text{ times is the same as } n \times m$$

Property 4 We want to show that

$$a^0 = 1$$

for any number a. Let's first assume that $a \neq 0$ and let's choose any number n. We know that a^n is a number, and that any number divided by itself is 1. Thus

$$\frac{a^n}{a^n} = 1$$

Using **Property 2**, we have that

$$\frac{a^n}{a^n} = a^{n-n} = a^0$$

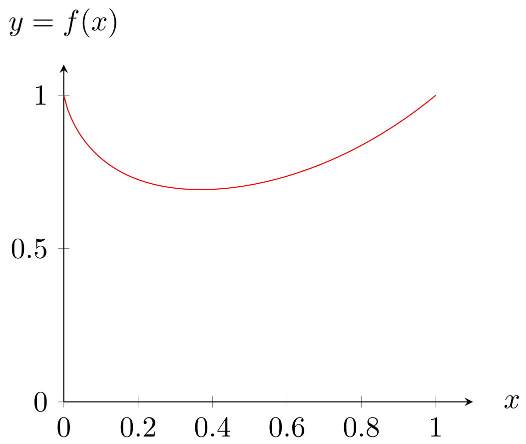
Given that we know that $\frac{a^n}{a^n} = 1$, we have shown that

$$a^0 = 1$$

for any number $a \neq 0$.

Now, when $a = 0$, we have no obvious answer. First, notice that what we did above is not valid, as we $0^n = 0$ and we would be dividing by 0. We could take many possible routes here: we could say $0^0 = 0$, that $0^0 = 1$ or even that 0^0 does not make sense and shouldn't be calculated. I will, however, give you a general intuition to why $0^0 = 1$ is a good answer.

Take the function $f(x) = x^x$ and graph it with x going from 0 to 1. The graph is this:



Do you notice that as x gets closer to 0, the value of x^x *tends to* 1? In that way, it *makes sense* to define 0^0 as 1, as that is the value that we are getting closer and closer to.

Property 5 We want to show that

$$\left(\frac{a}{b}\right)^n = \frac{a^n}{b^n}$$

This is very simple, we just need to use the definition of indices and multiplication of fractions:

$$\left(\frac{a}{b}\right)^n = \underbrace{\frac{a}{b} \times \frac{a}{b} \times \frac{a}{b} \times \ldots \times \frac{a}{b}}_{n \text{ times}} = \frac{\overbrace{a \times a \times a \times \ldots \times a}^{n \text{ times}}}{\underbrace{b \times b \times b \times \ldots \times b}_{n \text{ times}}} = \frac{a^n}{b^n}$$

Property 6 We want to show that

$$(ab)^n = a^n \times b^n$$

Again it is simple, we just use the definition:

$$(ab)^n = \underbrace{(ab) \times (ab) \times (ab) \times \ldots \times (ab)}_{n\text{times}}$$

Removing the brackets:

$$\underbrace{(ab) \times (ab) \times (ab) \times \ldots \times (ab)}_{n \text{ times}} = \underbrace{a \times b \times a \times b \times a \times b \times \ldots \times a \times b}_{n \text{ times}}$$

Remember that multiplication is commutative, that is, the order is not important. Therefore, we can rewrite the above as

$$\underbrace{a \times a \times a \times \ldots \times a}_{n \text{ times}} \times \underbrace{b \times b \times b \times \ldots \times b}_{n \text{ times}}$$

Finally, multiplying a number by itself n times is the same as raising it to the n-th power:

$$\underbrace{a \times a \times a \times \ldots \times a}_{n \text{ times}} \times \underbrace{b \times b \times b \times \ldots \times b}_{n \text{ times}} = a^n \times b^n$$

Property 7 Let's just show the general one. We want to show that

$$\left(\frac{a}{b}\right)^{-n} = \left(\frac{b}{a}\right)^n$$

Notice that

$$\left(\frac{a}{b}\right)^{-n} = \left(\frac{a}{b}\right)^{n-2n}$$

We can now use **Property 2** backwards:

$$\left(\frac{a}{b}\right)^{n-2n} = \frac{\left(\frac{a}{b}\right)^n}{\left(\frac{a}{b}\right)^{2n}}$$

Using the indices definition

$$\frac{\overbrace{\frac{a}{b} \times \frac{a}{b} \times \frac{a}{b} \times \ldots \times \frac{a}{b}}^{n \text{ times}}}{\underbrace{\frac{a}{b} \times \frac{a}{b} \times \frac{a}{b} \times \ldots \times \frac{a}{b}}_{2n \text{ times}}}$$

We can rewrite this expression as

$$\frac{\overbrace{\frac{a}{b} \times \frac{a}{b} \times \frac{a}{b} \times \ldots \times \frac{a}{b}}^{n \text{ times}} \times \overbrace{1 \times 1 \times 1 \times \ldots \times 1}^{n \text{ times}}}{\underbrace{\frac{a}{b} \times \frac{a}{b} \times \frac{a}{b} \times \ldots \times \frac{a}{b}}_{2n \text{ times}}}$$

and now split it

$$\overbrace{\frac{\frac{a}{b}}{\frac{a}{b}} \times \frac{\frac{a}{b}}{\frac{a}{b}} \times \frac{\frac{a}{b}}{\frac{a}{b}} \times \ldots \times \frac{\frac{a}{b}}{\frac{a}{b}}}^{n \text{ times}} \times \overbrace{\frac{1}{\frac{a}{b}} \times \frac{1}{\frac{a}{b}} \times \frac{1}{\frac{a}{b}} \times \ldots \times \frac{1}{\frac{a}{b}}}^{n \text{ times}}$$

Remember that a fraction is a number, so dividing a fraction by itself is 1. Thus, the first part of the expression is just 1 and we can write

$$\overbrace{\frac{1}{\frac{a}{b}} \times \frac{1}{\frac{a}{b}} \times \frac{1}{\frac{a}{b}} \times \ldots \times \frac{1}{\frac{a}{b}}}^{n \text{ times}}$$

We have n times the expression

$$\frac{1}{\frac{a}{b}}$$

This is the same as dividing fractions:

$$\frac{1}{1} \div \frac{a}{b} = \frac{1}{1} \times \frac{b}{a} = \frac{b}{a}$$

Our expression becomes

$$\underbrace{\frac{b}{a} \times \frac{b}{a} \times \frac{b}{a} \times \ldots \times \frac{b}{a}}_{n \text{ times}}$$

finally, this is the same as

$$\underbrace{\frac{b}{a} \times \frac{b}{a} \times \frac{b}{a} \times \ldots \times \frac{b}{a}}_{n \text{ times}} = \left(\frac{b}{a}\right)^n$$

There you have it:

$$\left(\frac{a}{b}\right)^{-n} = \left(\frac{b}{a}\right)^n$$

Property 8 We want to show that

$$a^{\frac{n}{m}} = \sqrt[m]{a^n} = \left(\sqrt[m]{a}\right)^n$$

Let's remember what it means to take a root first. Say we want to find

$$\sqrt[3]{2}$$

that is, we need to find a number that, when we raise to the power of 3 gives 2. Let's call it x. Thus,

$$x = \sqrt[3]{2}$$

Cubing both sides we obtain

$$x^3 = 2$$

You can write that as

$$x^3 = 2^1$$

Now, notice that you can write 1 as

$$1 = \frac{1}{3} + \frac{1}{3} + \frac{1}{3}$$

Therefore

$$x^3 = 2^{\frac{1}{3}+\frac{1}{3}+\frac{1}{3}}$$

We can use **Property 1** to split the right-hand side:

$$x^3 = 2^{\frac{1}{3}} \times 2^{\frac{1}{3}} \times 2^{\frac{1}{3}}$$

Cubing something means multiplying it by itself 3 times, so we have

$$x \times x \times x = 2^{\frac{1}{3}} \times 2^{\frac{1}{3}} \times 2^{\frac{1}{3}}$$

This implies that

$$x = 2^{\frac{1}{3}}$$

Therefore, we have

$$\sqrt[3]{2} = \sqrt[3]{2^1} = 2^{\frac{1}{3}}$$

Let's take a general root now

$$x = \sqrt[n]{a}$$

This means that

$$x^n = a^1 = a^{\overbrace{1/n + 1/n + 1/n + \ldots + 1/n}^{n \text{ times}}} = \underbrace{a^{\frac{1}{n}} \times a^{\frac{1}{n}} \times a^{\frac{1}{n}} \times \ldots \times a^{\frac{1}{n}}}_{n \text{ times}}$$

Again, this implies that

$$\sqrt[n]{a} = x = a^{\frac{1}{n}} \rightarrow \sqrt[n]{a} = a^{\frac{1}{n}}$$

The general

$$a^{\frac{p}{q}} = \sqrt[q]{a^p} = \left(\sqrt[q]{a}\right)^p$$

comes from the fact that

$$a^{\frac{p}{q}} = \left(a^{\frac{1}{q}}\right)^p = \left(\sqrt[q]{a}\right)^p = \sqrt[q]{a^p}$$

16. Factorisation of linear expressions

17. Quadratic expressions and equations

17.1. Why learn quadratics

You have finally reached this point in your education. The great moment when you will be introduced to the power of mathematics. No other subject has the capacity to, with only a single formula, ruin countless generations of people which still think we go to school because it should be 'useful'. The beauty of learning is not its utility, but the empathy building it does: we share this common knowledge, which definitely allows us to pursue our different interests, of course, but this *basic knowledge* allows us to bond with other human beings.

Of course, if you are interested in anything which uses quantitative reasoning learning maths will be useful later, but that is not the main point (forgetting that school needs to cater to every single person with every single possibility of future). You go to school so you are *able* to make choices about your future, and to make wise choices about things you need to have contact with said things.

The maths of school is the basis of much more, but manipulating algebraic expressions and solving equations are fundamental. This is why you have to learn quadratics: they offer a very good introduction to more sophisticated mathematical manipulation.

If, however, you are, sadly, one of those people that think learning a formula is ruining your education and life and love for learning, I am so sorry. You have bigger problems indeed.

17.2. General form of a quadratic expression

An expression is called a quadratic when it has one unknown, and we have a power of two of such unknown. For instance, if we have an expression

$$x^2 + x$$

the unknown is x and it is a quadratic expression, as it has x^2 in it.

A general quadratic expression with unknown x can be written as:

$$ax^2 + bx + c$$

Of course, we can have different unknowns, such as y:

$$ay^2 + by + c$$

Each of the number a, b, and c are called *coefficients*. Thus, a is the coefficient of the unknown which is squared, b is the coefficient with the unknown to the power of one and c is the free value, the one it does not have the unknown.

Let's see some examples of quadratic expressions:

- x^2: we have one single variable, x, which is squared. Here, $a = 1$, $b = 0$, and $c = 0$.
- $-3z^2+5$: we have a variable, z, and z^2 appears in the expression. In this expression, $a = -3$, $b = 0$, and $c = 5$.
- $2y^2 + 5y - 3$: one variable, y, and we have y^2 in it. In this case, $a = 2$, $b = 5$ and $c = -3$.
- $1 + 3x + 2x^2$: we have x squared. Be careful with order, the coefficients are determined by what they multiply, not the order they appear. Thus, we have $a = 2$, $b = 3$, and $c = 1$.
- $-2y + y^2$: here we have y^2. The number multiplying y^2 is 1, so we have $a = 1$, $b = -2$ and $c = 0$.

Identifying coefficients is very important, so be careful when doing it: the order does not matter, and the coefficient is just the number multiplying that part, not the number and the variable.

17.3. Factorising quadratics

Remember that factorising is when we rewrite an expression in a way that when we multiply the new expression we obtain the original. Also known as putting brackets in.

17.3.1. Difference of two squares

In the Expanding brackets chapter we learned some notable products, and was

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

the difference of two squares factorisation is exactly the same, but read from right to left:

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

Notice the pattern: the positive square appears in both brackets positive, and the negative square alternates, once positive, once negative.

It's straightforward to identify when you can use this technique: you have two terms (*two*), they are both square numbers (*squares*) and we are subtracting one from the other (*difference*). If you don't have any of these, you cannot use this.

Let's see examples.

$$a^2 - b^2$$

We have two terms, we are subtracting one from the other, and they are both squares:

$$(a)^2 - (b)^2 \qquad \text{Always good to identify the squares}$$

$$(a+b)(a-b) \qquad \text{Copy the positive and alternate the negative}$$

Another:

$$x^2 - 4y^2$$

As expected, we have two squares, and we are subtracting one from the other:

$$(x)^2 - (2y)^2 \qquad \text{Identify the squares}$$

$$(x+2y)(x-2y) \qquad \text{Copy the positive and alternate the negative}$$

We can have higher degrees:

$$x^6 - y^4$$

however we still have two square numbers, one being subtracted from the other:

$$\left(x^3\right)^2 - \left(y^2\right)^2 \qquad \text{Identify the squares}$$

$$(x^3+y^2)(x^3-y^2) \qquad \text{Copy the positive and alternate the negative}$$

Sometimes you can do it twice:

$$x^4 - y^4$$

$$\left(x^2\right)^2 - \left(y^2\right)^2 \qquad \text{Identify the squares}$$

$$(x^2+y^2)(x^2-y^2) \qquad \text{Copy the positive and alternate the negative}$$

We are not done, as the second bracket is again a difference of two squares!

$$(x^2+y^2)(\underbrace{x^2-y^2}_{\text{Difference of two squares}})$$

$$(x^2+y^2)((x)^2-(y)^2) \qquad \text{Identify the squares}$$

$$(x^2+y^2)(x+y)(x-y) \qquad \text{Copy the positive and alternate the negative}$$

A more annoying case:

$$289b^4 - 144c^2$$

$$\left(17b^2\right)^2 - (12c)^2 \qquad \text{Identify the squares}$$

$$(17b^2-12c)(17b^2+12c) \qquad \text{Copy the positive and alternate the negative}$$

One use of this is to do fast calculations of special cases, such as

$$784 - 729$$

Here we have a difference of two squares:

$28^2 - 27^2$	Identify the squares
$(28+27)(28-27)$	Copy the positive and alternate the negative
$(55)(1) = 55$	

17.3.2. Sum and product

17.3.2.1. $a = 1$

This factorising comes from expansions such as

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

In order to use this technique, we always have to have:

1. Three terms, which normally will be a squared unknown[1], an unknown to the power of 1 and a free term;
2. The coefficient of the squared term **has to be** 1.

Now, let's understand what is going on in the expansion to understand why the technique works. Take

$$(x+2)(x+3) = x^2 + 3x + 2x + 6 = x^2 + 5x + 6$$

First, notice that the free term in the answer, 6, always comes from us multiplying each number in the brackets, in this case 2 and 3. Now, look at the coefficient of x, in this case 5. It appeared when we added $2x$ with $3x$. Notice that the numbers 2 and 3 are responsible for each term, except the x^2 one, which is always going to be x^2: we do $2 \times 3 = 6$ to get the free term and $2 + 3 = 5$ to obtain the the x coefficient.

Thus, the technique is basically the following: when factorising expressions of this form, we always look for two numbers that

1. when multiplied together give us the numerical part of our expression; and
2. when added together give us the x coefficient of our expression.

It's surprisingly easy to do this: we just need to find the factors of the numerical part and see if we can combine them by addition to find the x coefficient. Let's see some examples.

If we want to factorise

$$x^2 + 6x + 8$$

[1] It can also be something like $y^4 + 5y^2 + 6$. The general form we can apply this is $x^{2n} + ax^n + b$.

First, notice we can use this because the x^2 coefficient is 1 and we have three terms that follow the rules above. Let's find all the factors of 8:

$$8 = 1 \times 8$$

$$8 = 2 \times 4$$

Remember that factors always appear in pairs, which is very convenient, as we are searching for pairs of numbers that multiply to 8. Now, we try to combine each of the factors pairs using any combination of negatives and see if we can reach the x coefficient, in this case 6. Starting with 1 and 8:

$$1 \text{ and } 8 \rightarrow \quad \begin{array}{c} +1+8=9\neq 6 \\ +1-8=-7\neq 6 \\ -1+8=7\neq 6 \\ -1-8=-9\neq 6 \end{array}$$

Not a single combination yields 6, so these are not the numbers we are looking for. Let's try 2 and 4 now:

$$2 \text{ and } 4 \rightarrow \quad \begin{array}{c} \boxed{+2+4=6=6} \\ +2-4=-2\neq 6 \\ -2+4=2\neq 6 \\ -2-4=-6\neq 6 \end{array}$$

We have found our numbers: +2 and +4. If you were solving this you would have stopped as soon as you found them, but I continued just for you to see the process again.

Now that we have our numbers, we just do this:

The numbers are +2 and +4

$$(x+2)(x+4)$$

and we are done. Let's factorise now

$$x^2 - 5x + 6$$

Again, we start by finding the pair factors of 6:

$$6 = 1 \times 6$$

$$6 = 2 \times 3$$

Now let's combine them to try to find -5, the coefficient of x:

$$1 \text{ and } 6 \rightarrow \quad \begin{array}{c} +1+6=7\neq -5 \\ +1-6=-5=-5 \\ -1+6=5\neq -5 \\ -1-6=-7\neq -5 \end{array}$$

Even though we found a combination that adds to -5, its product is not $+6$, as $1 \times -6 = -6$! We have to be careful and continue with the other pair, 2 and 3:

$$2 \text{ and } 3 \rightarrow \qquad \begin{array}{c} +2+3=5 \neq -5 \\ +2-3=-1 \neq -5 \\ -2+3=1 \neq -5 \\ \boxed{-2-3=-5=-5} \end{array}$$

We have already found our pair: -2 and -3. Lastly, we just set up the factors:

The numbers are -2 and -3

$$(x-2)(x-3)$$

As you can see, nothing changes, because there is a negative somewhere: it is exactly the same idea. No need to memorise sign combinations or whatever. Let's see now:

$$x^2 - 8x + 16$$

We start by finding the factor pairs:

$$16 = 1 \times 16$$

$$16 = 2 \times 8$$

$$16 = 4 \times 4$$

Now we have to combine them to try to find -8, the coefficient of x. With practice you will identify the possible ones very quickly, but let's do this slowly again:

$$1 \text{ and } 16 \rightarrow \qquad \begin{array}{c} +1+16=17 \neq -8 \\ +1-16=-15 \neq -8 \\ -1+16=15 \neq -8 \\ -1-16=-17 \neq -8 \end{array}$$

$$2 \text{ and } 8 \rightarrow \qquad \begin{array}{c} +2+8=10 \neq -8 \\ +2-8=-6 \neq -8 \\ -2+8=6 \neq -8 \\ -2-8=-10 \neq -8 \end{array}$$

$$4 \text{ and } 4 \rightarrow \qquad \begin{array}{c} +4+4=8 \neq -8 \\ +4-4=0 \neq -8 \\ -4+4=0 \neq -8 \\ \boxed{-4-4=-8=-8} \end{array}$$

We have our pair, -4 and -4, let's finish this:

$$\text{The numbers are } -4 \text{ and } -4 \quad\rightarrow\quad (x-4)(x-4)$$

It is very important to remember that no matter what, we always have to start our factorisation problems by finding common factors. For instance,

$$2x^2 - 32x - 34$$

This expression, at first glance, does not satisfy the restriction that the x^2 coefficient is 1, but if we notice that there is a common factor of 2:

$$2x^2 - 32x - 34 = 2(x^2 - 16x - 17)$$

We can now factorise the expression inside the brackets using our new technique. 17 (or any prime number) is very easy, as its only factors are 1 and itself, so we just combine them:

$$1 \text{ and } 17 \rightarrow \quad \begin{array}{c} +1 + 17 = 18 \neq -16 \\ \boxed{+1 - 17 = -16 = -16} \\ -1 + 17 = 16 \neq -16 \\ -1 - 17 = -18 \neq -16 \end{array}$$

our numbers are, thus, $+1$ and -17. Finally:

$$\text{The numbers are } +1 \text{ and } -17 \quad\rightarrow\quad 2(x+1)(x-17)$$

Let's, to finish, do an interesting one:

$$x^6 - 7x^3 + 12$$

we can use exactly the same technique here, because $x^3 \times x^3 = x^6$, and whenever this happens we are in the same problem[2]. We start by finding factors of 12:

$$12 = 1 \times 12$$

$$12 = 2 \times 6$$

$$12 = 3 \times 4$$

[2] You can even do a substitution such as $y = x^3$, and the expression becomes $y^2 - 7y + 12$.

Let's try now the pairs and see which gives us -7:

$$1 \text{ and } 12 \rightarrow \quad \begin{array}{c} +1+12=13 \neq -7 \\ +1-12=-11 \neq -7 \\ -1+12=11 \neq -7 \\ -1-12=-13 \neq -7 \end{array}$$

$$2 \text{ and } 6 \rightarrow \quad \begin{array}{c} +2+6=8 \neq -7 \\ +2-6=-4 \neq -7 \\ -2+6=4 \neq -7 \\ -2-6=-8 \neq -7 \end{array}$$

$$3 \text{ and } 4 \rightarrow \quad \begin{array}{c} +3+4=7 \neq -7 \\ +3-4=-1 \neq -7 \\ -3+4=1 \neq -7 \\ \boxed{-3-4=-7=-7} \end{array}$$

Thus, our numbers are -3 and -4. We finish in the same way:

$$\begin{array}{c} \text{The numbers are } -3 \text{ and } -4 \\ \downarrow \qquad \downarrow \\ (x^3-3)(x^3-4) \end{array}$$

Notice how we have x^3 inside the brackets, and we had x^3 as the second biggest degree of x in our expression (the same happened in the ones above: we had x as the second biggest degree in the expression, and we had x inside the brackets).

17.3.2.2. $a \neq 1$

We now will see how to factorise generic quadratic expressions of the form

$$ax^2 + bx + c$$

in which a can be any number which is not 0 or 1. There are many methods to do this, but I suggest either one of the first two here.

The "slide and divide" method This idea is very nice: in the same way that to add we fractions we "adapt" the problem to have equal denominators, which we know how to add, here we factorise by changing the quadratic to be able to use sum and product.

The steps are very simple to perform, although a little trickier to understand (see the Formality after taste of this chapter).

Let us use

$$10x^2 + 9x + 2$$

as an example. The first step is to "get rid" of the 10 by "sliding" it to the 2, multiplying both:

$$10x^2 + 9x + 2$$

to obtain another expression:

$$x^2 + 9x + 20$$

Notice that this expression is not equal to the first one, so do not write an equal sign between them!

Now, we use sum and product to factorise $x^2 + 9x + 20$. The numbers that multiply to 20 and add up to 9 are 4 and 5, so we obtain:

$$x^2 + 9x + 20 = (x + 4)(x + 5)$$

We then "divide" both 4 and 5 by 10, which we "slided" at the start:

$$(x + 4)(x + 5) \to \left(x + \frac{4}{10}\right)\left(x + \frac{5}{10}\right)$$

(which again is not equal to $(x + 4)(x + 5)$,so no equals!). Simplify the fractions fully:

$$\left(x + \frac{4}{10}\right)\left(x + \frac{5}{10}\right) = \left(x + \frac{2}{5}\right)\left(x + \frac{1}{2}\right)$$

and finally, if you still have a fraction, "mathemagically" send the denominator to the front of each variable and kill it:

$$\left(x + \frac{2}{5}\right)\left(x + \frac{1}{2}\right) \to (5x + 2)(2x + 1)$$

If we expand $(5x + 2)(2x + 1)$ we do get $10x^2 + 9x + 2$, so we are correct.

Another example:

$$8m^2 + 33m - 35$$

We first "slide" the 8 into the -35:

$$8m^2 + 33m - 35 \to m^2 + 33m - 280$$

Now we factorise $m^2 + 33m - 280$ using sum and product. The numbers that multiply to -280 and add up 33 are 40 and -7:

$$m^2 + 33m - 280 = (m + 40)(m - 7)$$

Then we divide both 40 and -7 by 8, which we slided at the start:

$$(m + 40)(m - 7) \to \left(m + \frac{40}{8}\right)\left(m - \frac{7}{8}\right)$$

which we now simplify:

$$\left(m+\frac{40}{8}\right)\left(m-\frac{7}{8}\right)=(m+5)\left(m-\frac{7}{8}\right)$$

To get rid of the 8 on the denominator in the second brackets we magically multiply m by it:

$$(m+5)\left(m-\frac{7}{8}\right)\rightarrow(m+5)(8m-7)$$

and we are done.

This example shows that, sometimes, the multiplication start gives some annoying numbers to work with (factors of 280 are not the most obvious). Hence why I like the method below.

In summary, to use the "slide and divide" method to factorise ax^2+bx+c:

1. "Slide" the a into the c by multiplying them. You will get a new quadratic that looks like: $x^2+bx+ac$.

2. Use sum and product to factorise $x^2+bx+ac$. You will get an answer that looks like $(x+p)(x+q)$.

3. Divide both p and q by a, the number you slided, to get $\left(x+\frac{p}{a}\right)\left(x+\frac{q}{a}\right)$.

4. Simplify the fractions $\frac{p}{a}$ and $\frac{q}{a}$ fully.

5. If you still have a fraction, "send the denominator" to the front of x.

The criss-cross method This method involves a little bit of guessing, but with practice is very fast and easy. Also, it is the most "obvious" to understand.

Our idea when factorising a quadratic expression is to go back to an expression of the form

$$(ax+b)(cx+d)$$

which, when expanded, gives us

$$acx^2+(ad+bc)x+bd$$

Now, the idea for this method is to "guess" the numbers a, b, c, and d so that ac is equal to the x^2 coefficient of our expression, bd is equal to the free coefficient (the one with no x) and that when we do $ad+bc$ we get the x coefficient.

There is a graphical way to do it, and you will then understand the "criss-cross" name. Let us use

$$2x^2+5x+3$$

as an example. You set up the criss-cross as

$$\begin{array}{ccccc} 2x^2 & +5x & +3 & & \\ \square & & \square & \to & \\ \times & & \times & & + \\ \square & & \square & \to & \end{array}$$

Our goal is to fill the boxes, so that:

- The first column, below the 2, has two numbers that multiply to 2
- The second column, below the 3, has two numbers that multiply to 3
- The "criss-cross" condition: when you multiply the numbers in the boxes in cross and add the result, it should be equal to 5.

In this case, there is only one way to obtain 2, 2×1. The same goes for 3, 3×1. So we can fill the boxes:

$$\begin{array}{ccccc} 2x^2 & +5x & +3 & & \\ 1 & & 1 & \to & \\ \times & & \times & & + \\ 2 & & 3 & \to & \end{array}$$

which we now can "cross-multiply":

$$\begin{array}{ccccc} 2x^2 & +5x & +3 & & \\ 1 & & 1 & \to & 2 \\ \times & & \times & & + \\ 2 & & 3 & \to & 3 \end{array}$$

and the last thing you check is if the sum of the last column is equal to 5, which is. So you now have your factorized quadratic in each row we filled: the first row is the first bracket, the second row the second:

$$\begin{array}{cccccc} 2x^2 & +5x & +3 & & = & (1x+1)(2x+3) \\ 1 & & 1 & \to & 2 & \\ \times & & \times & & + & \\ 2 & & 3 & \to & 3 & \end{array}$$

in which you just have to remember to add the x in the first term, as they need to multiply to x^2.

The idea of the method is this simple, but the guessing is not this easy all the time. For instance

$$2x^2 + 7x + 3$$

is very similar to the previous example, but you have to worry about the order you fill the spaces. We have seen already that filling like this:

$$\begin{array}{ccccc} 2x^2 & +7x & +3 & & \\ 1 & & 1 & \to & 2 \\ \times & & \times & & + \\ 2 & & 3 & \to & 3 \end{array}$$

gives us a "criss-cross" sum of 5, not 7, the coefficient of x. However, if we exchange the 3 and the 1 in the second column:

$$\begin{array}{ccccc} 2x^2 & +7x & +3 & & \\ 1 & & 3 & \to & 6 \\ \times & & \times & & + \\ 2 & & 1 & \to & 1 \end{array}$$

we do get 7, hence telling us that

$$\begin{array}{cccccc} 2x^2 & +7x & +3 & & = & (1x+3)(2x+1) \\ 1 & & 3 & \to & 6 & \\ \times & & \times & & + & \\ 2 & & 1 & \to & 1 & \end{array}$$

A more complicated example to finish this technique. Let us factorize

$$5x^2 + 17x - 40$$

The idea is still the same, but we have more options to obtain -40. Notice that one of the number will also have to be negative. With practice, you will be able to find the following numbers:

$$\begin{array}{ccccc} 5x^2 & +17x & -40 & & \\ 1 & & 5 & \to & 25 \\ \times & & \times & & + \\ 5 & & -8 & \to & -8 \end{array}$$

and get the result

$$5x^2 + 17x - 40 = (x+5)(5x-8)$$

Again, this method does require some "guesing", but if you practise, it is very fast. Also, the best one to understand why it works.

Grouping or the 'ac' method[3] The steps we need to take to factorize

$$ax^2 + bx + c$$

1. Calculate $a \times c$;

[3] I do not recommend this method.

2. Find two numbers, p and q, which satisfy

$$pq = ac$$

$$p + q = b$$

3. Split bx into $px + qx$;
4. Find the common factor of the first pair of terms;
5. Copy the sign between the second and third terms;
6. Find the common factor of the second pair of terms;
7. There will be another common factor inside the brackets, put that one as a common factor and we are done.

Let's see an example, as it is not as complicated as it seems. Say we want to factorise

$$2x^2 + 5x + 2$$

Here we have $a = 2$, $b = 5$ and $c = 2$. The first step is to calculate $ac = 2 \times 2 = 4$. Now, we need to find two numbers, p and q, which multiply to 4 and add to b, that is, 5. We again use the factors of 4 strategy to find the pairs:

$$4 = 1 \times 4$$

$$4 = 2 \times 2$$

Combining them:

$$1 \text{ and } 4 \rightarrow \begin{array}{c} \boxed{+1 + 4 = 5 = 5} \\ +1 - 4 = -3 \neq 5 \\ -1 + 4 = 3 \neq 5 \\ -1 - 4 = -5 \neq 5 \end{array}$$

We have our numbers, +1 and +4. We now split $5x$ into $1x + 4x$ (it could be $4x + 1x$ as well, the order is not relevant):

$$2x^2 + 5x + 2$$

$$2x^2 + 1x + 4x + 2 \qquad \text{Splitting } 5x \text{ into } 1x + 4x$$

$$\underbrace{2x^2 + 1x} + \underbrace{4x + 2} \qquad \text{Factorise the first and second pair of terms}$$

$$x(2x + 1) + 2(2x + 1) \qquad \text{Copy the sign}$$

$$x\underbrace{(2x + 1)} + 2\underbrace{(2x + 1)} \qquad \text{A common factor inside the brackets}$$

$$(x + 2)(2x + 1)$$

It is very important to notice that, when you reach the grouping part, *both brackets will have exactly the same thing inside.* If they do not, you have made a mistake.

Another example:

$$12x^2 + 11x - 5$$

We have $a = 12$, $b = 11$ and $c = -5$. We begin by calculating $ac = 12 \times -5 = -60$. Now, we find the factor pairs of 60:

$$60 = 1 \times 60$$

$$60 = 2 \times 30$$

$$60 = 3 \times 20$$

$$60 = 4 \times 15$$

$$60 = 5 \times 12$$

$$60 = 6 \times 10$$

A lot of possibilities! By now, though, I hope you noticed you can only either the sum of the numbers in the pair or their difference (with all the possible signs), so it should not be that hard to see that the only pair that can give us 11 is 4 and 15:

$$4 \text{ and } 15 \rightarrow \begin{array}{c} +4 + 15 = 19 \neq 11 \\ +4 - 15 = -11 \neq 11 \\ \boxed{-4 + 15 = 11 = 11} \\ -4 - 15 = -19 \neq 11 \end{array}$$

Our numbers are, thus, -4 and $+15$. We split $11x$ into $-4x + 15x$ and continue:

$12x^2 + 11x - 5$	
$12x^2 - 4x + 15x - 5$	Splitting $11x$ into $-4x + 15x$
$\underbrace{12x^2 - 4x} + \underbrace{15x - 5}$	Factorise the first and second pair of terms
$4x(3x - 1) + 5(3x - 1)$	Copy the sign
$4x\underbrace{(3x - 1)} + 5\underbrace{(3x - 1)}$	A common factor inside the brackets
$(4x + 5)(3x - 1)$	

A final example:

$$5x^2 - 11x + 2$$

Here, $a = 5$, $b = -11$ and $c = 2$. First, we calculate $ac = 5 \times 2 = 10$. Now we are searching for two numbers which multiply to 10 and add to -11. Let's first find the factor pairs of 10:

$$10 = 1 \times 10$$

$$10 = 2 \times 5$$

The pair that can yield -11 is 1 and 10:

$$1 \text{ and } 10 \rightarrow \begin{array}{c} +1 + 10 = 11 \neq -11 \\ +1 - 10 = -9 \neq -11 \\ -1 + 10 = 9 \neq -11 \\ \boxed{-1 - 10 = -11 = -11} \end{array}$$

thus our numbers are -1 and -10. We now split $-11x$ into $-1x - 10x$:

$$\begin{array}{rr}
5x^2 - 11x + 2 & \\
5x^2 - 1x - 10x + 2 & \text{Splitting } -11x \text{ into } -1x - 10x \\
\underbrace{5x^2 - x} - \underbrace{(10x - 2)} & \text{Factorise the first and second pair of terms: careful with the minus!} \\
x(5x - 1) - 2(5x - 1) & \text{Copy the sign} \\
x\underbrace{(5x - 1)} - 2\underbrace{(5x - 1)} & \text{A common factor inside the brackets} \\
(x - 2)(5x - 1) &
\end{array}$$

This last example shows you how important is to copy the sign between each pair, sometimes you will have to take extra care with the signs of the last two terms. Don't forget: the numbers inside the brackets will *always* be equal, so if your second brackets are different, you have made a mistake.

17.4. Completing the square

17.4.1. $a = 1$

Remember that we had a special 'notable product' of the form:

$$(a + b)^2 = a^2 + 2ab + b^2$$

which was called a 'perfect square'? This is the square we will be completing now.

For instance, if we expand $(x+1)^2$ we obtain

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

We can now subtract 1 on both sides:

$$\begin{aligned} (x+1)^2 &= x^2 + 2x + 1 \\ (x+1)^2 - 1 &= x^2 + 2x + 1 - 1 && \text{Subtracting 1 on both sides} \\ (x+1)^2 - 1 &= x^2 + 2x \end{aligned}$$

We will learn now how to go from the left-hand side of this last line to the right-hand side, which is called completing the square. In this case, we will be starting from $x^2 + 2x$ and arrive at $(x+1)^2 - 1$.

For now we will only be looking at quadratic expressions that have $a = 1$, such as

$$x^2 - 4x$$

$$x^2 + 5x$$

$$x^2 - 6x + 3$$

and so on.

The method is very simple:

1. Identify b in the expression (the x coefficient);
2. Divide b by 2;
3. Open brackets and write x and write the number you calculated in step 2 with its sign and close the brackets;
4. Square the brackets;
5. Subtract the square of the number you obtained in step 2 from the brackets.
6. Simplify the expression, if possible.

For instance, if we have

$$x^2 - 4x$$

let's do the steps

Step 1:	$x^2 \underbrace{-4}_{b} x$	Identify b: here $b = -4$
Step 2:	$\frac{b}{2} = \frac{-4}{2} = -2$	Calculate $\frac{b}{2}$
Step 3:	$(x-2)$	Write x with the $\frac{b}{2}$ with the sign inside brackets
Step 4:	$(x-2)^2$	Square the brackets
Step 5:	$(x-2)^2 - (-2)^2$	Subtract the square of -2
Step 6:	$(x-2)^2 - 4$	Simplify

Another example:

$$x^2 + 8x$$

Step 1:	$x^2 \underbrace{+8}_{b} x$	Identify b: here $b = 8$
Step 2:	$\frac{b}{2} = \frac{8}{2} = +4$	Calculate $\frac{b}{2}$
Step 3:	$(x+4)$	Write x with the $\frac{b}{2}$ with the sign inside brackets
Step 4:	$(x+4)^2$	Square the brackets
Step 5:	$(x+4)^2 - (4)^2$	Subtract the square of 4
Step 6:	$(x+4)^2 - 16$	Simplify

We are only using the b value in the expression, so if we have c (a number with no unknown) we just ignore it until the simplifying part. For instance

$$x^2 - 10x + 8$$

Step 1: $x^2 \underbrace{-10}_{b} x + 8$ — Identify b: here $b = -10$

Step 2: $\frac{b}{2} = \frac{-10}{2} = -5$ — Calculate $\frac{b}{2}$

Step 3: $(x-5)$ — Write x with the $\frac{b}{2}$ with the sign inside brackets

Step 4: $(x-5)^2$ — Square the brackets

Step 5: $(x-5)^2 - (-5)^2 + 8$ — Subtract the square of 4 and don't forget the +8

Step 6: $(x-5)^2 - 25 + 8 = (x-5)^2 - 17$ — Simplify

17.4.1.1. Formula

Given an expression of the form

$$x^2 + bx$$

we can rewrite it as

$$x^2 \pm bx = \left(x \pm \frac{b}{2}\right)^2 - \frac{b^2}{4}$$

which, as can you see, is the method we learned above: we divide the number multiplying x (b) by 2, copy the sign within the brackets and subtract the number we obtained ($\frac{b}{2}$) squared.

Let's see some examples.

$$x^2 + 4x$$

Here we have $b = 4$ and the sign between is a plus. Using the formula:

$$x^2 + bx = \left(x + \frac{b}{2}\right)^2 - \frac{b^2}{4}$$

$$x^2 + 4x = \left(x + \frac{4}{2}\right)^2 - \frac{4^2}{4}$$

We just need to simplify it:

$$x^2 + 4x = \left(x + \frac{4}{2}\right)^2 - \frac{4^2}{4} = (x+2)^2 - \frac{16}{4} = (x+2)^2 - 4$$

A more complex case:

$$x^2 - 5x - 6$$

here $b = 5$. We 'ignore' the -6 until the end:

$$x^2 - 5x - 6 \qquad b = -5$$

$$\left(x - \frac{5}{2}\right)^2 - \frac{5^2}{4} - 6 \qquad \text{Using the formula}$$

$$\left(x - \frac{5}{2}\right)^2 - \frac{25}{4} - 6 \qquad \text{Calculating the square}$$

$$\left(x - \frac{5}{2}\right)^2 - \frac{25}{4} - \frac{24}{4} \qquad \text{Subtracting the fractions}$$

$$\left(x - \frac{5}{2}\right)^2 - \frac{49}{4}$$

17.4.2. $a \neq 1$

For expressions such as

$$2x^2 + 4x$$

we cannot just use the 'divide b by 2' approach. Don't you think it would be great if that 2 were to 'vanish'? Perhaps we could ignore it for a while. The good news is that we can, simply by factorising the 2 out:

$$2x^2 + 4x = 2(x^2 + 2x)$$

Now, the only thing we need to do is to complete the square inside the brackets:

$$x^2 + 2x$$

which has $a = 1$, so we can use our method from above, with a twist at the end. Let's do everything from the beginning

$$2x^2 + 4x$$

Step 0:	$2(x^2 + 4x)$	Factorise a out
Step 1:	$2\left(x^2 \underbrace{+4}_{b} x\right)$	Identify b inside the brackets
Step 2:	$\frac{b}{2} = \frac{4}{2} = +2$	Calculate $\frac{b}{2}$
Step 3:	$2\left((x{+}2)\right)$	Write x with the $\frac{b}{2}$ with the sign inside brackets
Step 4:	$2\left((x+2)^2\right)$	Square the brackets
Step 5:	$2\left((x+2)^2 - (2)^2\right)$	Subtract the square of 1
Step 6:	$2\left((x+2)^2 - 4\right)$	Simplify
Step 7:	$2(x+2)^2 - 8$	The twist: expand the brackets

Sometimes this does get a little bit trickier, for instance, to complete the square

$$3x^2 - 5x$$

$$3x^2 - 5x$$

Step 0: $3(x^2 - \frac{5}{3}x)$ — Factorise a out

Step 1: $3\left(x^2 \underbrace{-\frac{5}{3}}_{b} x\right)$ — Identify b inside the brackets

Step 2: $\frac{b}{2} = \frac{-\frac{5}{3}}{2} = -\frac{5}{6}$ — Calculate $\frac{b}{2}$

Step 3: $3\left((x - \frac{5}{6})\right)$ — Write x with the $\frac{b}{2}$ with the sign inside brackets

Step 4: $3\left(\left(x - \frac{5}{6}\right)^2\right)$ — Square the brackets

Step 5: $3\left(\left(x - \frac{5}{6}\right)^2 - \left(\frac{5}{6}\right)^2\right)$ — Subtract the square of $-\frac{5}{6}$

Step 6: $3\left((x - \frac{5}{6})^2 - \frac{25}{36}\right)$ — Simplify

Step 7: $3\left(x - \frac{5}{6}\right)^2 - \frac{25}{12}$ — The twist: expand the brackets

17.5. Quadratic equations

17.5.1. Factorisation

Let us understand the basic idea first. If we are multiplying two numbers together, and the answer is 0, what can you say about the numbers? The only way for a multiplication to result in 0 is if, at least, one of the numbers is 0 itself. Let's put some algebra into this. Take two numbers, x and y. Their product is 0:

$$xy = 0$$

This implies that either $x = 0$ or $y = 0$[4]. This may sound obvious, but it is really helpful when solving equations.

For instance, given the following equation:

$$x(x - 2) = 0$$

[4]Notice this 'or' is a mathematical or: we can have that $x = 0$, that $y = 0$ or that x and y are both 0.

Here, we have a number x which, when multiplied to another number, $x-2$, gives us 0. This means that either x itself is 0 or that $x-2$ is 0. In algebra:

$$x = 0$$
$$\text{or}$$
$$x - 2 = 0$$

$x = 0$ is already a solution, and we obtain another solution by solving $x - 2 = 0$:

$$\begin{aligned} x - 2 &= 0 \\ x - 2 + 2 &= 0 + 2 && \text{Adding 2 on both sides} \\ x &= 2 \end{aligned}$$

Thus, the solutions of the equation $x(x-2) = 0$ are $x = 0$ or $x = 2$. Here we have two solutions, which is something to remember: for us, an equation will have at most n solutions, where n is its degree. In our example, $x(x-2) = 0$ can be expanded to $x^2 - 2x = 0$, which has degree 2. Equations that have degree 2 are called *quadratic equations.*

Let's see some other examples:

$$(x+3)(x-1) = 0$$

Remember that $x+3$ and $x-1$ are numbers:

$$\underbrace{(x+3)}_{\text{number}} \overbrace{(x-1)}^{\text{other number}} = 0$$

again we have a number multiplied by another with result 0. That means that one of the numbers, at least, has to be 0:

$$x + 3 = 0$$
$$\text{or}$$
$$x - 1 = 0$$

We just have to solve each equation now:

$$\begin{aligned} x + 3 &= 0 \\ x + 3 - 3 &= 0 - 3 && \text{Subtracting 3 on both sides} \\ x &= -3 \end{aligned}$$

$$\begin{aligned} x - 1 &= 0 \\ x - 1 + 1 &= 0 + 1 && \text{Adding 1 on both sides} \\ x &= 1 \end{aligned}$$

The solutions to $(x+3)(x-1)=0$ are, therefore, $x=-3$ or $x=1$.

Sometimes you can have more than 2 factors:

$$x(x+1)(x-2)(2x+3)=0$$

The same principle is true: if we have a product equal to 0, at least one of the numbers has to be equal to 0. We just have to see the possibilities:

$$\begin{gathered} x=0 \\ \text{or} \\ x+1=0 \\ \text{or} \\ x-2=0 \\ \text{or} \\ 2x+3=0 \end{gathered}$$

Again, just solve all equations:

$$\begin{gathered} x=0 \\ \text{or} \\ x+1=0 \rightarrow x=-1 \\ \text{or} \\ x-2=0 \rightarrow x=2 \\ \text{or} \\ 2x+3=0 \rightarrow x=-\tfrac{3}{2} \end{gathered}$$

So we have that the solutions to $x(x+1)(x-2)(2x+3)=0$ are $x=0$ or $x=-1$ or $x=2$ or $x=-\frac{3}{2}$.

Wouldn't it be great if we had all the equations already factorised? We can, however, factorise them ourselves, and that is the main idea to solve quadratic equations (in particular, but higher degrees as well): we make them equal to 0 and factorise them. After that, we just use the technique from above.

Let's see some examples, starting with

$$x^2+3x=0$$

This equation already is equal to 0, so we can just factorise it. There is a common factor, x:

$$x^2+3x=0$$

$$x(x+3)=0$$

after this, we have two numbers, x and $x+3$, which when multiplied give us 0. That means either one of them has to be 0:

$$\begin{gathered} x=0 \\ \text{or} \\ x+3=0 \rightarrow x=-3 \end{gathered}$$

and we are done. The solutions to $x^2 + 3x = 0$ are $x = 0$ or $x = -3$.

In an IGCSE setting, it is more common for the question to be divided into parts: first they will ask you to factorise a quadratic expression, and then solve the equation expression$= 0$ in the next item:

Solved exercise: an 'IGCSE' solving a quadratic by factorisation style question

1. Factorise the expression $12x^3 + 26x^2 - 10x$.

   The first thing in all factorisation questions is to check for common factors. Here every term has a factor $2x$, so we can put that into evidence:

   $$12x^3 + 26x^2 - 10x = 2x(6x^2 + 13x - 5)$$

   Now, we need to check the brackets for different factorisation possibilities. We don't have a difference of two squares, as we have 3 terms, so we have to use the 'ac' method in it. Inside the brackets we have

   $$6x^2 + 13x - 5$$

   therefore, we have $a = 6$, $b = 13$ and $c = -5$. We begin by calculating $ac = 6 \times -5 = -30$. The factor pairs of 30:

   $$30 = 1 \times 30$$
   $$30 = 2 \times 15$$
   $$30 = 3 \times 10$$
   $$30 = 5 \times 6$$

   Unfortunately, we have two possibilities to check for, as they both may result in 13: 2 and 15 is one of them, and 3 and 10 the other. Let's check them and hope to find the pair quickly:

   $$2 \text{ and } 15 \rightarrow \begin{array}{c} +2 + 15 = 17 \neq 13 \\ +2 - 15 = -13 \neq 13 \\ \boxed{-2 + 15 = 13 = 13} \\ -2 - 15 = -17 \neq 13 \end{array}$$

   Luckily -2 and $+15$ are the pair we want, so no need to check 3 and 10. Now we split $13x$ into $-2x + 15x$:

$$6x^2 + 13x - 5$$

$$6x^2 - 2x + 15x - 5 \qquad \text{Splitting } 13x \text{ into } -2x + 15x$$

$$\underbrace{6x^2 - 2x} + \underbrace{15x - 5} \qquad \text{Factorise the first and second pair of terms}$$

$$2x(3x - 1) + 5(3x - 1) \qquad \text{Copy the sign}$$

$$2x\underbrace{(3x - 1)} + 5\underbrace{(3x - 1)} \qquad \text{A common factor inside the brackets}$$

$$(2x + 5)(3x - 1)$$

Remember that this was just our bracket from $2x(6x^2 + 13x - 5)$, so our final answer is

$$12x^3 + 26x^2 - 10x = 2x(2x + 5)(3x - 1)$$

2. Solve the equation $12x^3 + 26x^2 - 10x = 0$

   We already have factorised the expression, so we have

$$12x^3 + 26x^2 - 10x = 0$$

$$2x(2x + 5)(3x - 1) = 0 \qquad \text{Factorised form}$$

   So we have three numbers, $2x$, $(2x+5)$ and $(3x-1)$ which are being multiplied and give 0. This means they can each be 0, so we have to solve the equations:

$$2x = 0 \rightarrow x = 0$$
$$\text{or}$$
$$2x + 5 = 0 \rightarrow x = -\tfrac{5}{2}$$
$$\text{or}$$
$$3x - 1 = 0 \rightarrow x = \tfrac{1}{3}$$

   Thus, our solutions are $x = 0$ or $x = -\frac{5}{2}$ or $x = \frac{1}{3}$.

17.5.2. Quadratic formula

Given a general quadratic equation

$$ax^2 + bx + c = 0$$

We can find both solutions using the formula[5]:

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

It is very important to notice you have a positive solution:

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

and a negative solution:

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

I like labeling each solution, x_+ and x_-, just to be sure of what I am doing.

In order to use the formula, we just have to identify the coefficients correctly, and then input the formula correctly on our calculators (if available). For instance, to solve:

$$3x^2 - 4x - 5 = 0$$

Remember, a is the coefficient of x^2, b is the coefficient of x and c the free term (the one with no unknown), so we have:

$$\underbrace{3}_{a} x^2 \underbrace{-4}_{b} x \underbrace{-5}_{c} = 0$$

Is is very important to remember that the negatives are *parts of the coefficients*, so they have to be used. Let's set up the formula:

$$\underbrace{3}_{a} x^2 \underbrace{-4}_{b} x \underbrace{-5}_{c} = 0$$

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

$$x = \frac{-(\overbrace{-4}^{b}) \pm \sqrt{(\overbrace{-4}^{b})^2 - 4 \times (\overbrace{3}^{a}) \times (\overbrace{-5}^{c})}}{2 \times (\underbrace{3}_{a})} \qquad a = 3, b = -4, c = -5$$

You could just use your calculator now, by copying this formula exactly as it is two times, one for the positive solution and one for the negative solution:

$$x_+ = \frac{-(-4) + \sqrt{(-4)^2 - 4 \times (3) \times (-5)}}{2 \times (3)} = \frac{2 + \sqrt{19}}{3} = 2.12 \text{ (2 d.p.)} \qquad \text{Positive solution}$$

$$x_- = \frac{-(-4) - \sqrt{(-4)^2 - 4 \times (3) \times (-5)}}{2 \times (3)} = \frac{2 - \sqrt{19}}{3} = -0.79 \text{ (2 d.p.)} \qquad \text{Negative solution}$$

[5]You have to memorise this formula if you are taking maths 0580. But worry not, you can always remember this `https://www.youtube.com/watch?v=2lbABbfU6Zc`.

If you identify the coefficients correctly, the only way to get one of these questions wrong is to input the values incorrectly on your calculator. Be very careful with negatives!

Solving a quadratic equation using the formula I

Solve to 3 d.p.:

$$-5 + 2y = -4y^2$$

Solution: Notice that the question says to solve the equation to a certain number of decimal places. This is an indication that we should use the formula.

The first thing when solving a quadratic equation using the formula is identifying the coefficients. Remember that the formula is based on this general expression:

$$ax^2 + bx + c = 0$$

In our case, we have y as the variable, so we want the equation to be like

$$ay^2 + by + c = 0$$

so let's rearrange our equation to put it into that format:

$$\begin{aligned} -5 + 2y &= -4y^2 \\ 4y^2 - 5 + 2y &= -4y^2 + 4y^2 && \text{Adding } 4y^2 \text{ on both sides} \\ 4y^2 + 2y - 5 &= 0 && \text{Rearranging the order of } 2y \text{ and } -5 \end{aligned}$$

Now we have a quadratic equation in the format we like. Let's identify the coefficients:

$$\underbrace{4}_{a} y^2 + \underbrace{2}_{b} y \underbrace{-5}_{c} = 0$$

Substituting into the formula:

$$\begin{aligned} 4y^2 + 2y - 5 &= 0 \\ y &= \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} && \text{The formula} \\ y &= \frac{-(2) \pm \sqrt{(2)^2 - 4 \times (4) \times (-5)}}{2 \times (4)} && a = 4, b = 2, c = -5 \end{aligned}$$

Finally, we can just input the two expressions (the positive and the negative) into the calculator:

$$y_+ = \frac{-2 + \sqrt{(2)^2 - 4 \times (4) \times (-5)}}{2 \times (4)} = \frac{-1 + \sqrt{21}}{4} = 0.896 \text{ (3 d.p.)} \quad \text{Positive solution}$$

$$y_- = \frac{-2 - \sqrt{(2)^2 - 4 \times (4) \times (-5)}}{2 \times (4)} = \frac{-1 - \sqrt{21}}{4} = -1.396 \text{ (3 d.p.)} \quad \text{Negative solution}$$

17.5.3. Completing the square

The basic idea to solve quadratic equations using completing the square comes from the fact that is easy to solve equations such as

$$(x+1)^2 = 16$$

the idea is that we can just square root both sides

$$\begin{aligned} (x+1)^2 &= 16 \\ \sqrt{(x+1)^2} &= \pm\sqrt{16} \\ x+1 &= \pm 4 \end{aligned}$$

and, finally, we solve the two equations that we have, one with $+4$ and one with -4:

$$x+1 = \pm 4 \rightarrow \begin{cases} x+1 = 4 \rightarrow x = 3 \\ x+1 = -4 \rightarrow x = -5 \end{cases}$$

The basic idea, then, is to complete the square in the quadratic equation we're trying to solve and do the above steps. For instance, to solve

$$x^2 - 6x + 8 = 0$$

We start by completing the square in $x^2 - 6x$:

$$x^2 - 6x + 8 = 0$$

$$(x-3)^2 - 3^2 + 8 = 0 \qquad \text{Completing the square}$$

$$(x-3)^2 - 9 + 8 = 0$$

$$(x-3)^2 - 1 = 0$$

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

$$\sqrt{(x-3)^2} = \pm\sqrt{1} \qquad \text{Square rooting both sides}$$

$$x - 3 = \pm 1$$

$$x - 3 = \pm 1 \rightarrow \begin{cases} x - 3 = 1 \rightarrow x = 4 \\ x - 3 = -1 \rightarrow x = 2 \end{cases} \qquad \text{Solving the plus and minus equations}$$

17.6. Exam hints

- In factorisation questions, always try to factorise using this order:
  1. Find all common factors between all terms of the expression;
  2. Try to do grouping (such as $ax + ay + bx + by = a(x + y) + b(x + y) = (a + b)(x + y)$);
  3. Check to see if you have difference of two squares;
  4. Finally do the sum/product if you have quadratic expressions.
- In questions to solve quadratic equations, if the instructions say to find the solutions to a certain number of decimal places, *use the quadratic formula*. Do not try to factorise the expression;
- Be very careful when inputting numbers in your calculator: remember that $-2^2 = -4$ in the calculator, for instance. If you have a fraction button on your calculator, use it, and literally copy the expression you obtained from the formula **and add brackets around negative numbers** just to be sure. If your calculator gives you a 'math error', you are most likely calculating a negative square root, so check your square root (you are probably forgetting a minus or are doing -2^2 without adding brackets);

Summary

- A *quadratic expression* is one that has this general format

$$ax^2 + bx + c$$

a is called the *coefficient* of x^2, b the coefficient of x and c the free term.

- Factorisation techniques:
  - *Difference of two squares*: use it when you have two square numbers being subtracted:
  $$a^2 - b^2 = (a+b)(a-b)$$
  - *Sum and product*: use it when you have an expression of the form $x^2 + bx + c$, and look for two numbers, p and q, which satisfy:
  $$p + q = b$$
  $$pq = c$$
  $$x^2 + bx + c = (x+p)(x+q)$$
  - *grouping*: use it when you have an expression of the form $ax^2 + bx + c$
- *Completing the square* is changing expressions of the form $ax^2 + bx$ into something that looks like $a(x+h)^2 - k^2$
- You can solve quadratic equations by factorisation and then solving each factor to 0
- The quadratic formula is
$$\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Formality after taste

Derivation of the sum-product method when $a = 1$

Take this generic quadratic expression:

$$x^{2n} + bx^n + c$$

We want to show that are two integers, p and q, which satisfy

$$(x^n + p)(x^n + q) = x^{2n} + bx^n + c$$

and they are given by

$$p + q = b$$

$$pq = c$$

Let's do some expansion:

$$(x^n + p)(x^n + q)$$

$$x^n \times x^n + qx^n + px^n + pq$$

$$x^{2n} + x^n(p + q) + pq \qquad x^n \text{ as a common factor}$$

We want this to be equal to $x^{2n} + bx^n + c$

$$x^{2n} + x^n(p + q) + pq = x^{2n} + bx^n + c$$

For this to happen, we need to

$$p + q = b$$

$$pq = c$$

and this shows why we find two numbers which multiply to c and add up to b.

Why "slide and divide" works[6]

Take the general quadratic equation

$$ax^2 + bx + c = 0$$

We first divide both sides by a, and we get:

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

What we have to do, now, is to do a change of variables:

$$x = \frac{u}{a}$$

which gives us:

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

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

$$\frac{u^2}{a^2} + \frac{bu}{a^2} + \frac{c}{a} = 0$$

[6]Kindly researched on this page to find why this witchcraft works: `https://www.cambridgecollege.edu/news/math-matters-slide-and-divide`.

If we now multiply both sides by a^2 we obtain:

$$\frac{u^2}{a^2} + \frac{bu}{a^2} + \frac{c}{a} = 0$$

$$u^2 + bu + ac = 0$$

and we know have the equivalence to the "slide" part. The "divide" part is just to replace x back using $x = \frac{u}{a}$:

$$(ax)^2 + bax + ac = 0$$

$$a^2x^2 + abx + ac = 0$$

that, if we divide by a, gives us the original equation back.

The simplification of the fractions and the "mathemagics" of ignoring the denominator is simply due to

$$x - \frac{p}{q} = 0 \text{ and } qx - p = 0$$

having the same solutions, so we can "cheat".

Derivation of the grouping or 'ac' method

Given a quadratic expression of the form

$$ax^2 + bx + c$$

we want to show that we can factorise it in

$$ax^2 + bx + c = (px + q)(rx + s)$$

Let's expand the brackets:

$$(px + q)(rx + s) = prx^2 + psx + qrx + qs$$

we want this expression to be equal to our original one:

$$ax^2 + bx + c = prx^2 + psx + qrx + qs$$

for this to happen we need

$$a = pr$$

$$b = ps + qr$$

$$c = qs$$

Now, notice that when we calculate ac we obtain

$$ac = pr \times qs = prqs$$

and this can be rewritten

$$ac = prqs = psqr = ps \times qr$$

and these are ps and qr which add to b.

Derivation of the quadratic formula

We want to prove that given a quadratic equation

$$ax^2 + bx + c = 0$$

with $a \neq 0$, we can find the solutions using the formula

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

It's actually very simple to do this, we just need to complete the square in order to make x the subject:

$$ax^2 + bx + c = 0$$

$$a\left(x^2 + \frac{b}{a}x\right) + c = 0 \qquad a \text{ as a `common factor'}$$

$$a\left(\left(x + \frac{b}{2a}\right)^2 - \left(\frac{b}{2a}\right)^2\right) + c = 0 \qquad \text{Completing the square inside the brackets}$$

$$a\left(x + \frac{b}{2a}\right)^2 - a \times \frac{b^2}{4a^2} + c = 0 \qquad \text{Expanding the brackets}$$

$$a\left(x + \frac{b}{2a}\right)^2 - \cancel{a} \times \frac{b^2}{4\cancel{a^2}^{\,a}} + c = 0 \qquad \text{Simplifying the product}$$

$$a\left(x + \frac{b}{2a}\right)^2 - \frac{b^2}{4a} + c = 0$$

$$a\left(x + \frac{b}{2a}\right)^2 = \frac{b^2}{4a} - c \qquad \text{Adding } \frac{b^2}{4a} \text{ and subtracting } c \text{ on both sides}$$

$$a\left(x + \frac{b}{2a}\right)^2 = \frac{b^2 - 4ac}{4a} \qquad \text{Common factor on the RHS}$$

$$\left(x + \frac{b}{2a}\right)^2 = \frac{b^2 - 4ac}{4a^2} \qquad \text{Dividing both sides by } a$$

$$\sqrt{\left(x + \frac{b}{2a}\right)^2} = \pm\sqrt{\frac{b^2 - 4ac}{4a^2}} \qquad \text{Square rooting both sides}$$

$$x + \frac{b}{2a} = \pm\frac{\sqrt{b^2 - 4ac}}{\sqrt{4a^2}} \qquad \text{Splitting the root into two}$$

$$x + \frac{b}{2a} = \pm\frac{\sqrt{b^2 - 4ac}}{2a} \qquad \text{Calculating the root on the denominator}$$

$$x = -\frac{b}{2a} \pm \frac{\sqrt{b^2 - 4ac}}{2a} \qquad \text{Subtracting } \frac{b}{2a} \text{ on both sides}$$

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \qquad \text{Adding the fractions}$$

18. Inequalities

18.1. Why learn inequalities?

Unequal.

If there is a word that describes our world, it is *unequal.* Look around you and you will see difference, the good and the bad. It would be very uncharacteristic of mathematics not to use that annoying part of it, the *useful* part, and not have a way to deal with unequal things. That's where inequalities come.

18.2. Inequality symbols

To represent inequalities, we have four symbols:

Symbol	Name	Example	How to read
$<$	*Smaller than*	$2 < 3$	'2 is smaller than 3'
$\leq$ or $\leqslant$	*Smaller than or equal to*	$-4 \leq -3$	'−4 is smaller than or equal to −3'
$>$	*Bigger than*	$0 > -1$	'0 is bigger than −1'
$\geq$or $\geqslant$	*Bigger than or equal to*	$1 \geq 1$	'1 is bigger than or equal to 1'

It is easy to remember what each sign means: the 'number it points to' is the smaller one.

Now, notice the last example:

$$1 \geq 1$$

which we read as '1 is bigger than or equal to 1'. The important word is **or**. No number can be bigger (or smaller than itself), so these are all false

$$0 > 0$$

$$-3 < -3$$

however, any number can be bigger than **or** equal to itself, such as

$$17 \leq 17$$

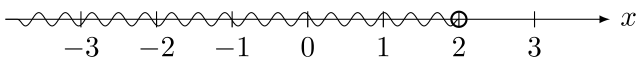
17 is not smaller than 17, but is **equal to** 17. Be careful with this.

18.3. Representing inequalities in a number line

Let's say we have the following inequality:

$$x < 2$$

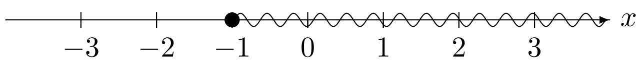
which is read as 'x is smaller than 2'. x stands for a number, any number which satisfies the inequality, that is, any number smaller than 2. There are infinite numbers which you can substitute for x. We can represent this infinity using a number line.



This represents any number smaller than 2. It **does not** include 2 itself, and we show that using that 'open circle' on 2. To represent inequalities such as

$$x \geq -1$$

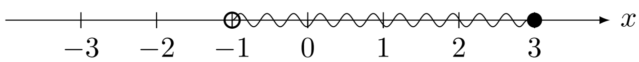
the only difference is that we use a 'filled circle' on the -1:



We can also represent what I like calling 'sandwich' inequalities, such as:

$$-1 < x \leq 3$$

This means that x can be any number between -1, but not including it, and 3, including it. On a line:



Representing inequalities in a number line is useful to see there are an infinite number of solutions to them, but useful to solve a specific type of problem.

Solved exercise: finding integer solutions of an inequality

Find the **integer** solutions to

$$-1 < x \leq 2$$

Solution: First, remember that an integer is a whole number. Now, let's represent the inequality in a number line:

We just need to find all the integers (whole numbers) the squiggly line crosses: $0, 1$, and 2. Notice that -1 is not included, as we have an 'open circle' there.

Thus, the integer solutions are $0, 1$, and 2.

18.4. Solving inequalities

To solve an inequality we use basically the same technique of equations: we do inverse operations on both sides. As you can see, it does work:

$$1 < 2$$

We can add 4 on both sides:

$$1+4 < 2+4$$
$$5 < 6$$

and the inequality still is true. However, we cannot just do *anything* like we did with equations:

$$-2 \leq 1$$

-2 is smaller than or equal to 1, so the inequality is true. Let's square both sides:

$$(-2)^2 \leq 1^2$$
$$4 \leq 1$$

4 is not smaller than or equal to 1. A very important example is the following:

$$-2 < 1$$

We cannot multiply both sides by -1 and do nothing else:

$$-2 < 1$$
$$-2 \times -1 < 1 \times -1$$
$$2 < -1$$

We need to be a bit more careful when solving inequalities, but the basic idea is the same.

Remember that the basic idea when solving an equation was to reach something like

$$\text{unknown} = \text{expression}$$

With inequalities is the same, but we can have any of the $<, >, \leq$ or $\geq$signs.

18.4.1. 'Type 1'

These are inequalities which we don't have to divide or multiply both sides by a negative number. Best way to understand how to solve them is examples.

Let's solve this inequality:

$$x - 3 < 5$$

Let's first understand what this means. We want to find all the numbers, which we are calling x, that when we subtract 3 from them we have a new number which is smaller than 5.

To solve it, we make 'x alone'. Let's get rid of that -3:

$$x - 3 < 5$$

$$x - 3 + 3 < 5 + 3 \qquad \text{Doing } +3 \text{ on both sides}$$

$$x < 8$$

There we go! More examples:

$$\frac{a}{2} \geq 5$$

Same idea:

$$\frac{a}{2} \geq 5$$

$$\frac{a}{2} \times 2 \geq 5 \times 2 \qquad \text{Doing } \times 2 \text{ on both sides}$$

$$a \geq 10$$

Let's try one a bit more complicated:

$$2y - 4 \leq 8$$

Still the same idea:

$$2y - 4 \leq 8$$

$$2y - 4 + 4 \leq 8 + 4 \qquad \text{Doing } +4 \text{ on both sides}$$

$$2y \leq 12$$

$$\frac{2y}{2} \leq \frac{12}{2} \qquad \text{Dividing both sides by } 2$$

$$y \leq 6$$

As you can see, whenever you have one these 'type 1' inequalities, you basically solve them in the same way as an equation. A final example:

$$\begin{aligned}
4x + 3 &> 2x - 5 \\
4x + 3 - 2x &> 2x - 5 - 2x && \text{Doing } -2x \text{ on both sides} \\
2x + 3 &> -5 && \text{Collecting like terms} \\
2x + 3 - 3 &> -5 - 3 && \text{Doing } -3 \text{ on both sides} \\
2x &> -8 && \text{Collecting like terms} \\
\frac{2x}{2} &> \frac{-8}{2} && \text{Dividing both sides by 2} \\
x &> -4
\end{aligned}$$

18.4.2. 'Type 2'

These are the inequalities which we have to either divide or multiply by a negative number.

Let's understand the trick first. Say you have these inequality:

$$-x < 5$$

We cannot just divide or multiply both sides by -1, as that gives us wrong answers:

$$x < -5$$

If we take -6, for instance, and substitute in the original:

$$\begin{aligned}
-x &< 5 \\
--6 &< 5 && \text{Substituting } -5 \\
6 &< 5
\end{aligned}$$

we reach that $6 < 5$, which is not true.

Now, let's try another way: we want x to be positive and alone, right? What if we got rid of $-x$ on the LHS of the inequality? Let's do $+x$ on both sides:

$$\begin{aligned}
-x &< 5 \\
-x + x &< 5 + x && \text{Doing } +x \text{ on both sides} \\
0 &< 5 + x
\end{aligned}$$

Let's just get rid of that annoying 5:

$$0 < 5 + x$$

$$0 - 5 < 5 + x - 5 \qquad \text{Doing } -5 \text{ on both sides}$$

$$-5 < x \qquad \text{Collecting like terms}$$

We have reached

$$-5 < x$$

x is 'alone', so we are done! The solution is any number which -5 is smaller from. Which is weird to say, but you can instead of writing $-5 < x$ write from 'right to left':

$$-5 < x$$

$$x > -5$$

Notice that is the same thing, we just wrote the inequality from right to left. The solution is any number greater than -5. Notice, however, what we started with

$$-x < 5$$

and what we have reached:

$$x > -5$$

We could have got here *quicker* if we had just divided (or multiplied) both sides by -1 and *flipped the inequality sign*:

$$-x < 5$$

$$\frac{-x}{-1} > \frac{5}{-1} \qquad \text{Dividing both sides by } -1 \text{ and } \textit{flipping the sign}$$

$$x < -5$$

That is the trick: when solving inequalities, if we have to either multiply or divide both sides by a negative number, **we have to flip the sign**. Let's see some examples:

$$-2x \geq 4$$

We have to divide both sides by -2, but whenever we divide by a negative number we have to flip the inequality sign:

$$-2x \geq 4$$

$$\frac{-2x}{-2} \leq \frac{4}{-2} \qquad \text{Dividing by } -2 \text{ and flipping the sign}$$

$$x \leq -2$$

To be honest, that is the only difference in the 'method' when solving inequalities from equations. For instance:

$$-\frac{y}{3} < 1$$

$$\frac{y}{-3} < 1 \qquad \text{Shiftting the } - \text{ place}$$

$$-3 \times \frac{y}{-3} > 1 \times -3 \qquad \text{Multiplying both sides by } -3 \text{ and flipping the sign}$$

$$y > -3$$

Another example:

$$3a + 2 < 4a - 3$$

$$3a - 4a + 2 < 4a - 4a - 3 \qquad \text{Doing } -4a \text{ on both sides}$$

$$-a + 2 < -3 \qquad \text{Collecting like terms}$$

$$-a + 2 - 2 < -3 - 2 \qquad \text{Doing } -2 \text{ on both sides}$$

$$-a < -5 \qquad \text{Collecting like terms}$$

$$\frac{-a}{-1} > \frac{-5}{-1} \qquad \text{Dividing both sides by } -1 \text{ and flipping the sign}$$

$$a > 5$$

As you can see, we just have to be careful about the inequality sign! A final one:

$$-3(x-2) \leq 2x - 4$$

$$-3x + 6 \leq 2x - 4 \qquad \text{Expanding the brackets}$$

$$-3x + 6 - 2x \leq 2x - 4 - 2x \qquad \text{Doing } -2x \text{ on both sides}$$

$$-5x + 6 \leq -4 \qquad \text{Collecting like terms}$$

$$-5x + 6 - 6 \leq -4 - 6 \qquad \text{Doing } -6 \text{ on both sides}$$

$$-5x \leq -10 \qquad \text{Collecting like terms}$$

$$\frac{-5x}{-5} \geq \frac{-10}{-5} \qquad \text{Dividing both sides by } -5 \text{ and flipping the sign}$$

$$x \geq 2$$

18.4.3. Double inequalities, a.k.a. sandwiches

These are inequalities like this:

$$-3 < x + 1 < 5$$

They are actually a system of simultaneous inequalities, and you can solve them using a different technique than I'll show you here. I will show you a shortcut. This shortcut only works when we only have the unknown on the 'middle part'.

We can solve these sandwich inequalities by doing the inverse operations on both the 'left' side, the 'midle part' and the 'right' side at the same time. For instance:

$$-3 < x + 1 < 5$$

$$-3-1 < x + 1-1 < 5-1 \qquad \text{Doing } -1 \text{ everywhere}$$

$$-4 < x < 4 \qquad \text{Collecting like terms}$$

The same detail when dividing or multiplying by a negative number, we have to flip all the inequality signs:

$$4 \leq -2x \leq 8$$

$$\frac{4}{-2} \geq \frac{-2x}{-2} \geq \frac{8}{-2} \qquad \text{Dividing everywhere by } -2 \text{ and flipping the signs}$$

$$-2 \geq x \geq -4$$

$$-4 \leq x \leq -2 \qquad \text{Writting from right to left}$$

One a bit longer:

$$1 < -3x - 2 < 4$$

$$1+2 < -3x - 2+2 < 4+2 \qquad \text{Doing } +2 \text{ everywhere}$$

$$3 < -3x < 6 \qquad \text{Collecting like terms}$$

$$\frac{3}{-3} > \frac{-3x}{-3} > \frac{6}{-3} \qquad \text{Dividing everywhere by } -3 \text{ and flipping the signs}$$

$$-1 > x > -2$$

$$-2 < x < -1 \qquad \text{Writting from right to left}$$

What if we have a double inequality which has unknowns everywhere, such as

$$1 + x < 2x \leq x + 3$$

To solve this type, we actually need to solve two inequalities:

$$\begin{cases} 1 + x < 2x & (1) \\ 2x \leq x + 3 & (2) \end{cases}$$

Let's solve inequality (1) first:

$$1 + x < 2x$$

$$1 + x - x < 2x - x$$

$$1 < x$$

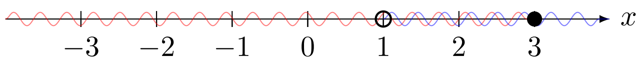
Or $x > 1$ if we read from right to left. Now let's solve inequality (2):

$$2x \leq x + 3$$

$$2x - x \leq x - x + 3$$

$$x \leq 3$$

Now, we need to join these results, and we can do them by representing both inequalities in the same number line:



I represented inequality (1) with a blue line and inequality (2) with a red line. Notice that the place where we have **both** 'snakes' go from 1 to 3, but not including 1. We can write that as an inequality:

$$1 < x \leq 3$$

Integer solutions of an inequality II

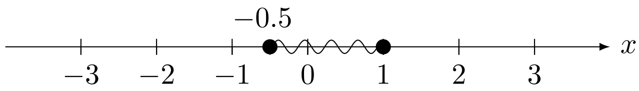
Find the integer solutions to

$$-4 \leq 2x - 3 \leq -1$$

Solution: Let's solve the inequality to begin:

$$\begin{aligned} -4 \leq 2x - 3 \leq -1 & \\ -4{+3} \leq 2x - 3{+3} \leq -1{+3} & \qquad \text{Doing } {+3} \text{ everywhere} \\ -1 \leq 2x \leq 2 & \qquad \text{Collecting like terms} \\ \frac{-1}{2} \leq \frac{2x}{2} \leq \frac{2}{2} & \qquad \text{Dividing everywhere by } 2 \\ -\frac{1}{2} \leq x \leq 1 & \end{aligned}$$

Now, let's represent it in a number line:



We just need to find the integers the squiggly line crosses: 0 and 1. Thus, the integers solutions are 0 and 1.

18.5. Quadratic inequalities

To solve quadratic inequalities, just like quadratic equations, the technique is slightly different. The basic idea is that we are going to sketch the graph of the quadratic, find its roots, and use that to find the solution to the inequality.

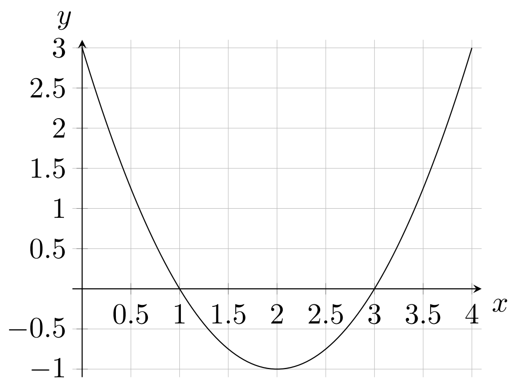
An example: find the values of x for which

$$x^2 - 4x + 3 < 0$$

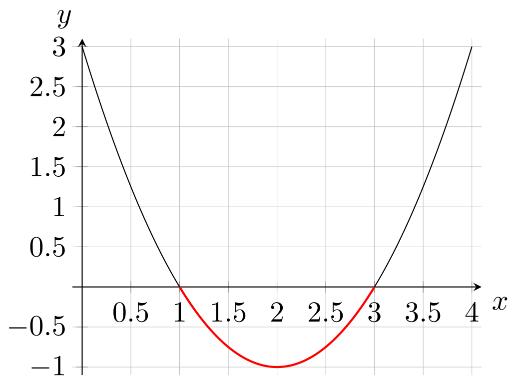
From the quadratics chapter, you know that you can factorise this as

$$(x-1)(x-3) \leq 0$$

We can now focus of the left-hand side of the inequality, $(x-1)(x-3)$. Let's graph it:



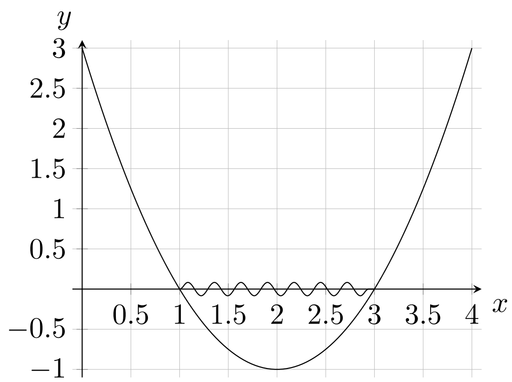
Looking at the graph of it, notice that there is a portion of it which has values which are smaller than or equal to 0:



That part of the parabola in red corresponds to the only possible values that are smaller than or equal to 0, which are precisely what the inequality told us to find

$$x^2 - 4x + 3 \leq 0$$

Thus, we only have to find the x values which determine that region:

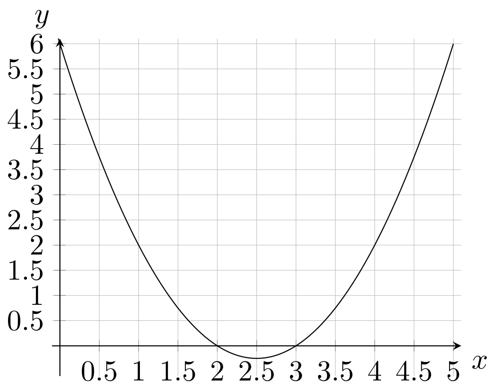


The points which satisfy $1 \leq x \leq 3$ are precisely the ones we want, and the solution to the inequality.

If we wanted to solve

$$x^2 - 5x + 6 > 0$$

we start with the same idea: let us first graph $y = x^2 - 5x + 6$:



We now want to find the pieces of the graph which are greater than 0. Let's colour them:

We have to be careful with the points $(2, 0)$ and $(3, 0)$, though: when $x = 2$, we have that $x^2 - 5x + 6$ is exactly equal to 0, and our inequality this time only wants values which are *greater* than 0. We, therefore, mark those points with an open circle:

Therefore, we want the values of x corresponding to the "snakey" part of the graph:

Finally, the values of x which satisfy

$$x^2 - 5x + 6 < 0$$

are

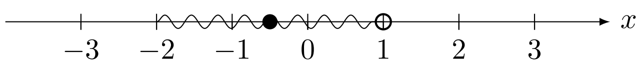
$$x < 2 \text{ and } x > 3$$

18.6. Exam hints

Just be careful when dividing or multiplying by a negative: remember to flip the sign!

Summary

- There are 4 inequality symbols:
  - $<$: smaller than
  - $<$ or $\leqslant$: smaller than or equal to
  - $>$: bigger than
  - $>$ or $\geqslant$: bigger than or equal to
- We can represent inequalities in a number line, remembering that an 'open circle' means we are not including that number and a 'filled circle' means we are. For example, to represent the inequality $-2 \leq x < 1$ we draw this:



- To *solve* an inequality we use the same technique as when solving equations: using inverse operations to make the unknown 'alone'. The only difference is that when we *divide or multiply by a negative number, we have to flip the inequality sign.*
- To solve a quadratic inequality:

- First, rearrange everything just like when solving a quadratic (everything $= 0$) on one side;
- Graph the parabola you obtained;
- Find the parts of the parabola which you want;
- Find the corresponding values of x for those regions.

Formality after taste

Proof of why we "flip" the inequality sign when we divide or multiply an inequality by a negative number

This proof is quite formal, and it depends on some knowledge of the properties of the real numbers. I will just state them here:

- If $a < b$, then $b - a > 0$
- If $c < 0$, then $-c > 0$
- If $a > 0$ and $b > 0$, then $ab > 0$

We can now prove this statement:

Theorem. If $a < b$ and $c < 0$, then $ac > bc$.

Notice that this is saying that if we have a number a, which is smaller than b (or $a < b$), if we multiply both sides of the inequality by a negative number c, we obtain $ac > bc$, in which the inequality sign was "flipped".
The proof goes as follows:

Proof.

$a < b$	Hypothesis
$b - a > 0$	Fact above
$c < 0$	Hypothesis
$-c > 0$	Fact above

We have the two inequalities now:

$$\begin{cases} b - a > 0 & (1) \\ -c > 0 & (2) \end{cases}$$

multiplying both inequalities together:

$$-c(b-a) > 0 \qquad \text{Fact above}$$

$$ac - bc > 0 \qquad \text{Distributive property}$$

$$ac > bc$$

As division does not really exist and it is just a "shortcut" to multiply a number by its reciprocal, this proof is valid for division as well.

□

For more on this, I recommend the book "Elementary Geometry from an Advanced Standpoint" , by Edwin E. Moise.

19. Simultaneous equations

19.1. Why learn simultaneous equations?

To be very honest, most of the applications I know are a bit more advanced than what we can understand right now.

There is one "practical" situation though that can be modeled using simultaneous equations[1].

Let's say we have a company that sells chocolate. The company normally wants to maximise its profits, so it wants to sell as much chocolate as possible at the highest price. The consumer, on the other hand, would like to buy more chocolate if its price is smaller. In all, a simplistic way to see this is: as the price of the chocolate rises, the company wants to sell more chocolate; on the other side, the consumer buys less chocolate when the price is higher.

Call Q the quantity of chocolate sold and P the price of chocolate. Imagine that the company equation for the quantity it wants to sell is the following:

$$Q = 10P + 9$$

It makes sense according to our assumption that as the price increases, the more chocolate the company wants to sell. The consumer wants to buy less chocolate if the price grows, so an equation as

$$Q = 100 - 3P$$

If we want to discover the *equilibrium price and quantity*, that is, the price and quantity where both consumer and company are "happy", we need to find values for P and Q which satisfy *both equations*! When we have more than one equation to solve at the same time, usually with more than one unknown, we have a **system of simultaneous equations**.

We can write this system like this:

$$\begin{cases} Q = 10P + 9 \\ Q = 100 - 3P \end{cases}$$

And our objective is finding values for Q and for P that solve both equations at the same time. For this particular system the solution is $Q = 79$ and $P = 7$. This means the "equilibrium price" for this chocolate is 7, and at the price the company sells 79 units of it.

[1]Hopefully my economics is not as bad as I think it is.

19.2. Systems of linear equations

19.2.1. Graphical solving

I always say "whenever in doubt, make a drawing". Let's solve a system by graphing, also known as drawing.

Let's say we want to find the solution to

$$\begin{cases} x + y = 6 & (1) \\ x - y = 2 & (2) \end{cases}$$

We have two equations of straight lines, which we have already learned how to graph[2]. You could use the "cover-up" method or rearrange the equations to have the format $y = mx + c$ which we like. Let's do both for a quick revision, one for each equation.

Starting with equation (1), remember that the "cover-up" method fills this table:

x	0	
y		0

Making $x = 0$, or "covering" it:

$$x + y = 6$$

$$\square + y = 6$$

$$y = 6$$

So we have that when $x = 0$, $y = 6$. Let's put that in our table

x	0	
y	6	0

We now "cover" y, that is, make $y = 0$:

$$x + y = 6$$

$$x + \square = 6$$

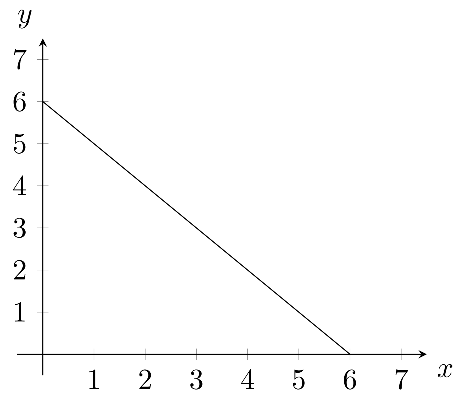
$$x = 6$$

Filling our table:

x	0	6
y	6	0

Thus we know that the line $x + y = 6$ goes through the points $(0, 6)$ and $(6, 0)$. Let's graph it:

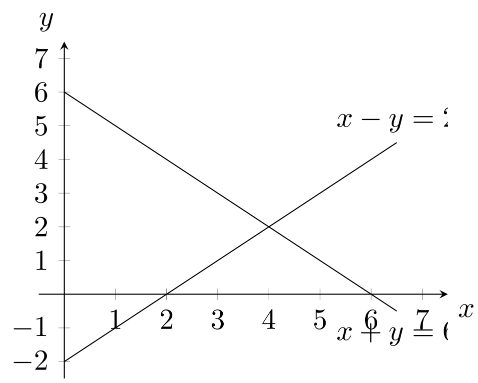
[2]Future reference to chapter on linear graphs



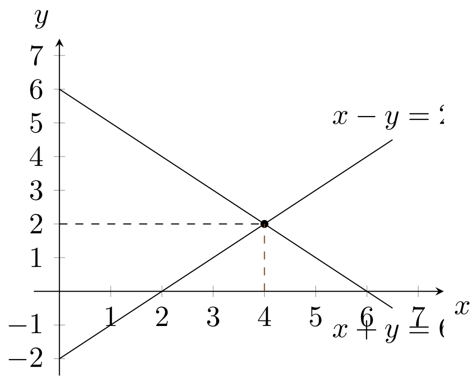
Let's rearrange equation (2):

$$
\begin{aligned}
x - y &= 2 \\
x - x - y &= 2 - x && \text{Doing } -x \text{ on both sides} \\
-y &= 2 - x \\
y &= -2 + x && \text{Dividing both sides by } -1 \\
y &= x - 2
\end{aligned}
$$

So we have that equation (2) goes through $(0, -2)$, as its y-intercept is -2, and has gradient equal to 1. Therefore, it goes through the point $(1, -1)$, as every time it goes 1 unit to the right it goes 1 unit up. Let's add its graph to our previous one:



Notice that both lines intersect at the point $(4, 2)$:



The intersection between the two lines is the only point (x, y) that satisfies *both* equations at the same time. As that happens at the point $(4, 2)$, the solution to our system of equations is $x = 4$ and $y = 2$. Let's check if this actually works:

$$\begin{cases} x + y = 6 & (1) \\ x - y = 2 & (2) \end{cases}$$

Substituting $x = 4$ and $y = 2$:

$$\begin{cases} 4 + 2 = 6 & (1) \\ 4 - 2 = 2 & (2) \end{cases} \to \begin{cases} 6 = 6 & (1) \\ 2 = 2 & (2) \end{cases}$$

There we go. To summarise, in order to solve the system using a graph, we graph all the equations we have and find their intersection.

19.2.2. Substitution

Let's say we have this system of linear equations:

$$\begin{cases} x + y = 6 & (1) \\ y = x - 2 & (2) \end{cases}$$

Notice that in Eq. (2) we have y as the subject. Whenever this happens, we can substitute y for whatever it is equal in the other equation, and this particular unknown won't be there anymore:

$$\begin{cases} x + y = 6 & (1) \\ y = x - 2 & (2) \end{cases}$$

$$x + \underbrace{(x-2)}_{y \text{ value in } (2)} = 6 \qquad \text{Substituting } y \text{ in Eq. (1)}$$

$$x + x - 2 = 6 \qquad \text{Expanding brackets}$$

$$2x - 2 = 6 \qquad \text{Collecting like terms}$$

$$2x - 2 + 2 = 6 + 2 \qquad +2 \text{ on both sides}$$

$$2x = 8$$

$$\frac{2x}{2} = \frac{8}{2}$$

$$x = 4$$

We have found that $x = 4$. Now, to find y, we go back to the Eq. (2), where we had with y as the subject, and substitute $x = 4$:

$$y = x - 2$$

$$y = 4 - 2$$

$$y = 2$$

We are finished, the solution is $x = 4$ and $y = 2$.

Basically, to solve a system of equations by substitution, we make one of the unknowns the subject in one of the equations, and substitute the expression in the other. This will result in an equation with only one unknown, and we are in known ground.

Another example:

$$\begin{cases} 2x + y = 4 & (1) \\ x + y = 3 & (2) \end{cases}$$

Let's make x the subject in Eq. (2):

$$x + y = 3$$

$$x + y - y = 3 - y$$

$$x = 3 - y$$

Now that we have an equation with x as the subject, we can substitute it in the other:

$$\begin{aligned} 2x + y &= 4 && (1) \\ 2(3-y) + y &= 4 && \text{Substituting } x \\ 6 - 2y + y &= 4 && \text{Expanding the brackets} \\ 6 - y &= 4 && \text{Collecting like terms} \\ 6 - 6 - y &= 4 - 6 \\ -y &= -2 \\ -1 \times -y &= -1 \times -2 && \text{Multiplying both sides by } -1 \\ y &= 2 \end{aligned}$$

Thus, $y = 2$. To obtain the value of x, we substitute y into the equation we have that has x as the subject:

$$x = 3 - y$$
$$x = 3 - 2$$
$$x = 1$$

We are done, $x = 1$ and $y = 2$.
In summary, to use the substitution method:

1. Choose an unknown to make the subject in any of the equations;
2. Make this unknown the subject in one of the equations;
3. Substitute the expression you obtained for this unknown in the other equation;
4. Solve the equation you obtained to obtain the value for the other unknown;
5. Substitute the value you obtained in step 4 in the equation you obtained in step 2 and solve the equation and you are done.

19.2.3. Elimination

Before we actually see the elimination method, let's understand the idea. Say we have this super exciting system of equations:

$$\begin{cases} 2 = 2 & (1) \\ 3 = 3 & (2) \end{cases}$$

I know, nothing much to say here. Notice, however, what happens when we *add* both these equations. By adding the two, we mean that we add left-hand side with left-hand side, and add right-hand side with right-hand side:

$$\begin{cases} 2 = 2 & (1) \\ 3 = 3 & (2) \end{cases} \qquad \text{The system}$$

$$\overbrace{2}^{\text{Eq. (1) LHS}} + \underbrace{3}_{\text{Eq. (2) LHS}} = \overbrace{2}^{\text{Eq. (1) RHS}} + \underbrace{3}_{\text{Eq. (2) RHS}} \qquad \text{Adding Eq. (1) with Eq. (2)}$$

$$5 = 5$$

Notice that, in the end, we continue with an equation which is true: $5 = 5$. This shows that if you have two equations and add them like we did, left-hand side with left-hand side, and right-hand side with right-hand side, we obtain a new equation which is still true.

We can also *subtract* one equation from the other:

$$\begin{cases} 2 = 2 & (1) \\ 3 = 3 & (2) \end{cases} \qquad \text{The system}$$

$$\overbrace{2}^{\text{Eq. (1) LHS}} - \underbrace{3}_{\text{Eq. (2) LHS}} = \overbrace{2}^{\text{Eq. (1) RHS}} - \underbrace{3}_{\text{Eq. (2) RHS}} \qquad \text{Eq. (1) - Eq. (2)}$$

$$-1 = -1$$

Again, we obtain a new equation which is still true.

Now, I am sure you are thinking "great, I just went from an obvious fact such as $2 = 2$ and got *another* obvious fact, $5 = 5$, what is the point?". Let's not disappoint you!

19.2.3.1. "Case 1"

What is a "case 1" system? Let's define it.

We'll call a system a "case 1" when at least one of the unknowns, in both equations, has the same coefficient when we ignore the signs. For instance, this is a "case 1" system:

$$\begin{cases} x + y = 6 & (1) \\ x - y = 2 & (2) \end{cases}$$

The coefficients of x, in both Eq. (1) and Eq. (2) is 1. The same happens for y when we ignore the signs: in both equations the coefficient is 1 when we ignore the signs ($+1$ in Eq. (1) and -1 in Eq. (2)). Therefore, this system is a "case 1". Another example:

$$\begin{cases} x + y = 6 & (1) \\ x - 2y = 0 & (2) \end{cases}$$

In both equations, we have that the coefficient of x is 1. Notice, however, that the coefficients of y are different. Finally, this is not a "case 1" system:

$$\begin{cases} 2x + 2y = 12 & (1) \\ 3x - 6y = 0 & (2) \end{cases}$$

Looking at the unknown coefficients, you can see they are not equal. In Eq. (1), the coefficient for x is 2, whereas in Eq. (2) the coefficient for x is 3. For y, in Eq. (1) we have 2, and in Eq. (2) we have 6 (remember the sign is not important).

Let's learn how to solve "case 1" systems now.

Say we want to find the solution to our dear friend from previous sections:

$$\begin{cases} x + y = 6 & (1) \\ x - y = 2 & (2) \end{cases}$$

Let's try our new found technique of *adding* two equations, and do Eq. (1) plus Eq. (2):

$$\begin{cases} x + y = 6 & (1) \\ x - y = 2 & (2) \end{cases} \qquad \text{The system}$$

$$\overbrace{(x+y)}^{\text{Eq. (1) LHS}} + \underbrace{(x-y)}_{\text{Eq. (2) LHS}} = \overbrace{6}^{\text{Eq. (1) LHS}} + \underbrace{2}_{\text{Eq. (2) RHS}} \qquad \text{Eq. (1)+Eq. (2)}$$

$$x + y + x - y = 8 \qquad \text{Removing the brackets}$$

$$\underline{x}\,\boxed{+y}\,\underline{+x}\,\boxed{-y} = 8 \qquad \text{Like terms}$$

$$\underline{x + x}\,\boxed{+y - y} = 8 \qquad \text{Collecting like terms}$$

$$2x + 0 = 8 \qquad y \text{ is gone!}$$

$$2x = 8$$

$$\frac{2x}{2} = \frac{8}{2} \qquad \text{Dividing both sides by 2}$$

$$x = 4$$

By adding Eq. (1) with Eq. (2), we managed to *eliminate* the y unknown and discover x! Now, to discover y, we can just substitute $x = 4$ in either Eq. (1) or Eq. (2). Let's put it into Eq. (1):

$$x + y = 6$$

$$4 + y = 6$$

$$4 + y - 4 = 6 - 4$$

$$y = 2$$

Now we know that $y = 2$, and have finished! The solution is $x = 4$ and $y = 2$.

There is a catch, though: not every time adding two equations helps us. This system, for example:

$$\begin{cases} 2x + y = 4 & (1) \\ x + y = 3 & (2) \end{cases}$$

If we add both equations, we obtain:

$$\overbrace{(2x + y)}^{\text{Eq. (1) LHS}} + \underbrace{(x + y)}_{\text{Eq. (2) LHS}} = \overbrace{4}^{\text{Eq. (1) LHS}} + \underbrace{3}_{\text{Eq. (2) RHS}} \qquad (1) + (2)$$

$$2x + y + x + y = 7 \qquad \text{Removing brackets}$$

$$3x + 2y = 7 \qquad \text{After collecting like terms}$$

As you can see, we still have both unknowns, x and y. Adding both equations didn't help us at all! However, see what happens when we *subtract* Equation (2) from Equation (1):

$$\begin{cases} 2x + y = 4 & (1) \\ x + y = 3 & (2) \end{cases} \qquad \text{The system}$$

$$\overbrace{(2x + y)}^{\text{Eq. (1) LHS}} - \underbrace{(x + y)}_{\text{Eq. (2) LHS}} = \overbrace{4}^{\text{Eq. (1) LHS}} - \underbrace{3}_{\text{Eq. (2) RHS}} \qquad (1) - (2)$$

$$2x + y - x - y = 1 \qquad \text{Removing the brackets}$$

$$\underline{2x}\boxed{+y}\underline{-x}\boxed{-y} = 1 \qquad \text{Identifying like terms}$$

$$\underline{2x - x}\boxed{+y - y} = 1 \qquad \text{Collecting like terms}$$

$$x + 0 = 1 \qquad y \text{ is gone!}$$

$$x = 1$$

Now we managed to eliminate y, and reached an equation with only x. Solving it, we reached $x = 1$. To find y, we substitute $x = 1$ into any of the equations. Equation (2) is simpler, so let's do it there:

$$x + y = 3$$

$$1 + y = 3$$

$$1 + y - 1 = 3 - 1$$

$$y = 2$$

Thus, $y = 2$ and we are finished. The solution is $x = 1$ and $y = 2$.

When we have a "case 1" system, then, we can always eliminate one of the unknowns by either adding or subtracting the two equations. If you like memorising "rules":

- If you want to eliminate two unknowns that have *the same sign*, you **subtract** one equation from the other;
- If you want to eliminate two unknowns that have *different signs*, you **add** both equations together.

Solved exercise: solving a "case 1" system of two linear equations by elimination

Let's solve

$$\begin{cases} 2x - y = -5 & (1) \\ 2x + 3y = -1 & (2) \end{cases}$$

First, notice this is a "case 1" system, as x has coefficient 2 in both equations. Thus, we can eliminate x. We could either add or subtract both equations, but adding $2x$ with $2x$ gives $4x$, which is a bit pointless. Let's subtract them:

$$\overbrace{(2x - y)}^{\text{Eq. (1) LHS}} - \underbrace{(2x + 3y)}_{\text{Eq. (2) LHS}} = \overbrace{-5}^{\text{Eq. (1) RHS}} - \underbrace{(-1)}_{\text{Eq. (2) LSH}} \qquad (1) - (2)$$

$$2x - y - 2x - 3y = -5 + 1 \qquad \text{Expanding the brackets}$$

$$\cancel{2x} - y - \cancel{2x} - 3y = -4 \qquad \text{Collecting like terms}$$

$$-4y = -4$$

$$\frac{-4y}{-4} = \frac{-4}{-4}$$

$$y = 1$$

Thus, $y = 1$. Let's substitute this into Eq. (1):

$$2x - y = -5$$

$$2x - 1 = -5$$

$$2x - 1 + 1 = -5 + 1$$

$$2x = -4$$

$$\frac{2x}{2} = \frac{-4}{2}$$

$$x = -2$$

Done, $x = -2$ and $y = 1$. It's always good to check our answer by substituting:

$$\begin{cases} 2x - y = -5 & (1) \\ 2x + 3y = -1 & (2) \end{cases}$$

$$\begin{cases} 2 \times -2 - 1 = -5 & (1) \\ 2 \times -2 + 3 \times 1 = -1 & (2) \end{cases}$$

$$\begin{cases} -4 - 1 = -5 & (1) \\ -4 + 3 = -1 & (2) \end{cases} \rightarrow \begin{cases} -5 = -5 & (1) \\ -1 = -1 & (2) \end{cases}$$

We're correct.

19.2.3.2. "Case 2"

Very uncreatively, "case 2" systems are all the ones which are not a "case 1".

To solve them, you notice that I have been putting quotes around "case 1" and "case 2" all the time, and that gives you a hint that we'll use the classic mathematician trick of turning one problem into another: we'll convert "case 2" systems into "case 1" systems! That's why I have been putting quotes: they are the same[3].

In order to do this "conversion", we just need to remember that we are still dealing with equations. Thus, we can do everything we know works for equations, including multiplying or dividing both sides of an equation by the same number.

Let's do our magic to this system:

$$\begin{cases} 8x + 5y = 43 & (1) \\ 5x + 3y = 26 & (2) \end{cases}$$

This is not a "case 1" system, as the x coefficients are 8 and 5, and the y coefficients are 5 and 3. Wouldn't it be great if the x coefficients were the same? Then we could just do what already know for "case 1" systems...

Let's try to think to what we know and remember a situation when we had *different numbers* and we wanted them *to be the same*... When adding fractions! To add fractions with unlike denominators we multiplied both numerator and denominator of the fractions so that they had the same denominator. We'll use exactly the same idea here. If we were adding two fractions that had denominators 8 and 5, such as

$$\frac{3}{8} + \frac{1}{5}$$

we would do this:

$$\frac{3}{8} + \frac{1}{5} = \frac{15}{40} + \frac{8}{40} = \frac{23}{40}$$

To obtain the 40, the common denominator, we simply multiplied 8 and 5. Look at our system:

$$\begin{cases} 8x + 5y = 43 & (1) \\ 5x + 3y = 26 & (2) \end{cases}$$

To make x have the same coefficient in both equations, we could multiply the first by

[3] Dramatic music plays.

5 and the second by 8:

$$\begin{cases} 8x + 5y = 43 & (1) \\ 5x + 3y = 26 & (2) \end{cases}$$

$$\begin{cases} 5(8x + 5y) = 5 \times 43 & (1) \times 5 \\ 8(5x + 3y) = 8 \times 26 & (2) \times 8 \end{cases} \quad \text{Multiplying each equation by the } x \text{ coefficient in the other}$$

$$\begin{cases} 40x + 25y = 215 & (1) \\ 40x + 24y = 208 & (2) \end{cases} \quad \text{Expanding the brackets}$$

We have obtained an *equivalent system*:

$$\begin{cases} 40x + 25y = 215 & (1) \\ 40x + 24y = 208 & (2) \end{cases}$$

which is now a "case 1"! Let's solve it. To eliminate x we can do Eq. (1)− Eq. (2):

$$\overbrace{(40x + 25y)}^{\text{Eq. (1) LHS}} - \underbrace{(40x + 24y)}_{\text{Eq. (2) LHS}} = \overbrace{215}^{\text{Eq. (1) RHS}} - \underbrace{208}_{\text{Eq. (2) RHS}} \qquad (1) - (2)$$

$$40x + 25y - 40x - 24y = 7 \qquad \text{Expanding the brackets}$$

$$\cancel{40x} + 25y - \cancel{40x} - 24y = 7 \qquad \text{Cancelling } x$$

$$25y - 24y = 7$$

$$y = 7$$

As usual, we need to substitute $y = 7$ in one of the equations. It's *much* easier to use one of the original equations, as their coefficients are much smaller, so we have smaller number. Let's do $y = 7$ in the *original* Eq. (2):

$$5x + 3y = 26$$

$$5x + 3 \times 7 = 26$$

$$5x + 21 = 26$$

$$5x + 21 - 21 = 26 - 21$$

$$5x = 5$$

$$\frac{5x}{5} = \frac{5}{5}$$

$$x = 1$$

There we have it, the solution is $x = 1$ and $y = 7$.

Just like when adding fractions, I say. And just like when adding fractions, you can go with the "obvious" common coefficient, or you can always use the LCM of the coefficients you want to eliminate. See the next example:

$$\begin{cases} x - 2y = 11 & (1) \\ 2x + 4y = -10 & (2) \end{cases}$$

Let's eliminate y in this system. If we did the "simple" way, we would multiply Eq. (1) by 4, as it is the y coefficient in Eq. (2), and multiply Eq. (2) by 2, as it is the coefficient (without the sign) of y in Eq. (1). That works fine, but gives larger numbers than needed. We want to eliminate y, which has coefficients 2 and 4 (ignoring the signs). The LCM of 2 and 4 is 4, so we just need to obtain an equivalent equation to Eq. (1) in which the y coefficient is 4. We can do that by multiplying it by 2. Then, we don't have to do anything to Eq. (2), as the y coefficient there is already 4. Thus:

$$\begin{cases} x - 2y = 11 & (1) \\ 2x + 4y = -10 & (2) \end{cases}$$

$$\begin{cases} 2x - 4y = 22 & (1) \times 2 \\ 2x + 4y = -10 & (2) \end{cases}$$

We can now add both equations and eliminate y:

$$\begin{aligned} \overbrace{(2x - 4y)}^{\text{Eq. (1) LHS}} + \underbrace{(2x + 4y)}_{\text{Eq. (2) LHS}} = \overbrace{22}^{\text{Eq. (1) RHS}} + \underbrace{-10}_{\text{Eq. (2) RHS}} & \qquad (1)\times 2\text{-}(2) \\ 2x - 4y + 2x + 4y = 22 - 10 & \qquad \text{Simplifying} \\ 4x - \cancel{4y} + \cancel{4y} = 12 & \qquad \text{Collecting like terms} \\ 4x = 12 & \\ \frac{4x}{4} = \frac{12}{4} & \\ x = 3 & \end{aligned}$$

So we have that $x = 3$. Again, we need to substitute this in any of the equations, but it is easier to go the originals, as the numbers are usually smaller. Substituting $x = 3$ in the original Eq. (1):

$$x - 2y = 11$$

$$3 - 2y = 11$$

$$3 - 3 - 2y = 11 - 3$$

$$-2y = 8$$

$$\frac{-2y}{-2} = \frac{8}{-2}$$

$$y = -4$$

There you go, the solution is $x = 3$ and $y = -4$.

Solved exercise: solving a "case 2" system of two linear equations by elimination

$$\begin{cases} 4x - 3y = 7 & (1) \\ 6x + 2y = 4 & (2) \end{cases}$$

First, we notice this is a "case 2" system. So let's change it into a "case 1". Let's eliminate x (you can choose whatever unknown you want). We need to obtain equivalent equations in which x the same coefficient. In Eq. (1) the coefficient is 4 and in Eq. (2) it is 6. The LCM of 4 and 6 is 12, so we need to multiply Eq. (1) by 3 ($4 \times 3 = 12$) and we need to multiply Eq. (2) by 2 ($6 \times 2 = 12$):

$$\begin{cases} 4x - 3y = 7 & (1) \\ 6x + 2y = 4 & (2) \end{cases}$$

$$\begin{cases} 3(4x - 3y) = 3 \times 7 & (1) \times 3 \\ 2(6x + 2y) = 2 \times 4 & (2) \times 2 \end{cases}$$

$$\begin{cases} 12x - 9y = 21 & (1) \\ 12x + 4y = 8 & (2) \end{cases}$$

We have a "case 1" system now, and we can subtract Eq. (2) from Eq. (1) to eliminate x:

$$\overbrace{(12-9y)}^{\text{Eq. (1) LHS}} - \underbrace{(12x+4y)}_{\text{Eq. (2) LHS}} = \overbrace{21}^{\text{Eq. (1) RHS}} - \underbrace{8}_{\text{Eq. (2) RHS}} \qquad (1)-(2)$$

$$12x - 9y - 12x - 4y = 13 \qquad \text{Expanding brackets}$$

$$\cancel{12x} - 13y - \cancel{12x} = 13 \qquad \text{Collecting like terms}$$

$$-13y = 13$$

$$\frac{-13y}{-13} = \frac{13}{-13}$$

$$y = -1$$

We have that $y = -1$. Let's substitute this value in the original Eq. (2):

$$6x + 2y = 4$$

$$6x + 2 \times (-1) = 4$$

$$6x - 2 = 4$$

$$6x - 2 + 2 = 4 + 2$$

$$6x = 6$$

$$\frac{6x}{6} = \frac{6}{6}$$

$$x = 1$$

So $x = 1$. Let's check our answer:

$$\begin{cases} 4x - 3y = 7 & (1) \\ 6x + 2y = 4 & (2) \end{cases}$$

$$\begin{cases} 4 \times 1 - 3 \times -1 = 7 & (1) \\ 6 \times 1 + 2 \times -1 = 4 & (2) \end{cases}$$

$$\begin{cases} 4 + 3 = 7 & (1) \\ 6 - 2 = 4 & (2) \end{cases} \rightarrow \begin{cases} 7 = 7 & (1) \\ 4 = 4 & (2) \end{cases}$$

We are correct.

19.2.4. Formula

As usual, you know my suggestion with formulae: don't use it.

A general system of two linear equations can be written like this:

$$\begin{cases} a_1x + b_1y = c_1 & (1) \\ a_2x + b_2y = c_2 & (2) \end{cases}$$

We have that a_1 and a_2 are the coefficients for x in Eq. (1) and Eq. (2), respectively. The same for b_1 and b_2 for y.

The solution for this system is given by:

$$\boxed{x = \frac{b_1c_2 - b_2c_1}{a_2b_1 - a_1b_2}} \text{ and } \boxed{y = \frac{a_2c_1 - a_1c_2}{a_2b_1 - a_1b_2}}$$

Solved exercise: solving a system of two linear equations with the formula

Let's solve our dear friend using the formula:

$$\begin{cases} x + y = 6 & (1) \\ x - y = 2 & (2) \end{cases}$$

Let's write all the coefficients:

$$\begin{cases} 1x + 1y = 6 & (1) \\ 1x - 1y = 2 & (2) \end{cases}$$

So we have:

$$a_1 = 1, b_1 = 1, c_1 = 6$$

$$a_2 = 1, b_2 = -1, c_2 = 2$$

Substituting this in the formula for x:

$$x = \frac{b_1c_2 - b_2c_1}{a_2b_1 - a_1b_2} = \frac{1 \times 2 - -1 \times 6}{1 \times 1 - 1 \times -1} = \frac{2+6}{1+1} = \frac{8}{2} = 4$$

Same for y:

$$y = \frac{a_2c_1 - a_1c_2}{a_2b_1 - a_1b_2} = \frac{1 \times 6 - 1 \times 2}{1 \times 1 - 1 \times -1} = \frac{6-2}{1+1} = \frac{4}{2} = 2$$

We have, then, $x = 4$ and $y = 2$.

19.3. System of non-linear equations

Until now we have been dealing with systems of *linear equations*, that is, equations in which the unknowns have degree 1. A system of non-linear equations has at least one equation with degree greater than 1. For instance:

$$\begin{cases} x + y = 1 & (1) \\ y = x^2 - 4x + 3 & (2) \end{cases}$$

Equation (2) here has a term x^2, so we are dealing with a non-linear system. In a way, these systems are even easier to solve, as we usually solve them by substitution.

Let's solve the one above:

$$\begin{cases} x + y = 1 & (1) \\ y = x^2 - 4x + 3 & (2) \end{cases}$$

We already have y as the subject in Eq. (2), so let's substitute the expression it is equal to in Eq. (1):

$$x + (\underbrace{x^2 - 4x + 3}_{\text{Subsitute } y}) = 1 \qquad \text{Substituting } y \text{ in (1)}$$

$$x + x^2 - 4x + 3 = 1 \qquad \text{Expanding the brackets}$$

$$x^2 - 3x + 3 = 1 \qquad \text{Collecting like terms}$$

$$x^2 - 3x + 3 - 1 = 1 - 1 \qquad \text{Making the LHS} = 0$$

$$x^2 - 3x + 2 = 0$$

$$(x - 1)(x - 2) = 0 \qquad \text{Factorising}$$

$$\begin{array}{l} x - 1 = 0 \rightarrow x = 1 \\ x - 2 = 0 \rightarrow x = 2 \end{array} \qquad \text{Solving}$$

We have, therefore, two possible values for x: either $x = 1$ or $x = 2$. We now need to substitute **both** of them in Eq. (2) to obtain the possible values for y.

$$y = x^2 - 4x + 3$$

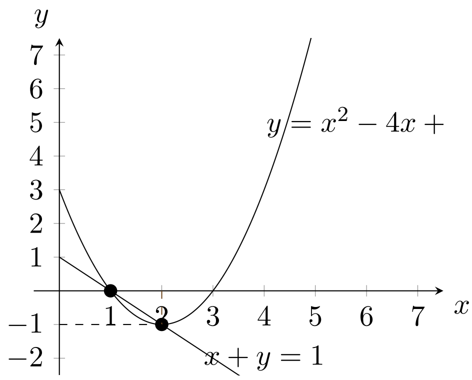
$$y = (1)^2 - 4 \times (1) + 3 \qquad \text{Substituting } x = 1$$

$$y = 1 - 4 + 3 \rightarrow y = 0$$

$$y = (2)^2 - 4 \times (2) + 3 \qquad \text{Substituting } x = 2$$

$$y = 4 - 8 + 3 \rightarrow y = -1$$

Now, we have to be careful how to write the solutions. We have that, when $x = 1$, $y = 0$ as one of the solutions, and we have that when $x = 2$, $y = -1$. A good way to write this is an ordered pair (x, y): the first solution is $(1, 0)$ and the second $(2, -1)$. In this way there is no possible confusion. Remember that we could solve this graphically, and the solutions would be the points themselves:



In all, the only difference from a "standard" substitution is the number of solutions, as you can have more than one.

Solved exercise: solving a system of two non-linear equations

$$\begin{cases} y = x^2 - 3x + 2 & (1) \\ y = 3x^2 - 8x + 4 & (2) \end{cases}$$

Same strategy: let's substitute the expression we have for y in Eq. (1) in Eq. (2):

$$(\underbrace{x^2 - 3x + 2}_{y \text{ in Eq. (1)}}) = 3x^2 - 8x + 4 \qquad \text{Substituting } y \text{ in Eq. (2)}$$

$$x^2 - 3x + 2 = 3x^2 - 8x + 4$$

$$3x^2 - 8x + 4 = x^2 - 3x + 2 \qquad \text{Flipping sides}$$

$$3x^2 - 8x + 4 - x^2 + 3x - 2 = 0 \qquad \text{Making LHS} = 0$$

$$2x^2 - 5x + 2 = 0$$

$$(2x - 1)(x - 2) = 0 \qquad \text{Factorising the LHS}$$

$$\begin{matrix} 2x - 1 = 0 \to x = \frac{1}{2} \\ x - 2 = 0 \to x = 2 \end{matrix} \qquad \text{Solving each possibility}$$

So we have that x can be either 2 or $\frac{1}{2}$. We need to substitute *each* in Eq. (1):

$$y = x^2 - 3x + 2$$

$$y = (2)^2 - 3 \times 2 + 2 \qquad \text{Subsituting } x = 2$$

$$y = 4 - 6 + 2 \to y = 0$$

$$y = \left(\frac{1}{2}\right)^2 - 3 \times \frac{1}{2} + 2 \qquad \text{Substituting } x = \frac{1}{2}$$

$$y = \frac{1}{4} - \frac{3}{2} + 2$$

$$y = \frac{1}{4} - \frac{6}{4} + \frac{8}{4} \to y = \frac{3}{4}$$

The solutions are, thus, $(2, 0)$ and $\left(\frac{1}{2}, \frac{3}{4}\right)$. Let's us check them. Starting with $(2, 0)$:

$$\begin{cases} 0 = (2)^2 - 3 \times (2) + 2 & (1) \\ 0 = 3 \times (2)^2 - 8 \times 2 + 4 & (2) \end{cases} \to \begin{cases} 0 = 4 - 6 + 2 & (1) \\ 0 = 3 \times 4 - 16 + 4 & (2) \end{cases} \to \begin{cases} 0 = 0 & (1) \\ 0 = 0 & (2) \end{cases}$$

This one is correct, let's check $\left(\frac{1}{2}, \frac{3}{4}\right)$ now:

$$\begin{cases} \frac{3}{4} = \left(\frac{1}{2}\right)^2 - 3 \times \left(\frac{1}{2}\right) + 2 & (1) \\ \frac{3}{4} = 3 \times \left(\frac{1}{2}\right)^2 - 8 \times \left(\frac{1}{2}\right) + 4 & (2) \end{cases} \to \begin{cases} \frac{3}{4} = \frac{1}{4} - \frac{3}{2} + 2 & (1) \\ \frac{3}{4} = \frac{3}{4} - 4 + 4 & (2) \end{cases} \to \begin{cases} \frac{3}{4} = \frac{3}{4} & (1) \\ \frac{3}{4} = \frac{3}{4} & (2) \end{cases}$$

Also correct. Therefore, the solutions are $(2, 0)$ and $\left(\frac{1}{2}, \frac{3}{4}\right)$.

19.4. Exam hints

1. If you have a system of non-linear equations, don't overthink it: just use substitution. You'll normally end up in a quadratic equation, then you just work your magic;

2. If you have a system of linear equations, you have to decide which method to use. Usually elimination or substitution. I recommend using substitution only if one of the unknowns is already the subject of one of the equations. If that's not the case, I suggest elimination.

Summary

- A *system of simultaneous equations* is composed by at least two equations which we need to solve *at the same time*, that is, finding values for each of the unknowns that satisfy *all the equations*;

- We can solve systems of simultaneous equations using the *graphical method*, which means we draw each equation graph and find the *intersections*;

- If our system has only *linear equations*, that is, equations with the unknowns to the power of 1, we call the system *linear;*

- We can solve linear systems of equations using either *substitution* or *elimination*;

- If we have a system with at least one equation with degree 2 or above, we have a *non-linear system* of simultaneous equations;

- We usually solve non-linear systems using *substitution.*

Formality after taste

A formula for systems of linear equations with two unknowns

Let's use what is known as *Gaussian elimination*[4] to derive our lovely formula. By the name, you can notice we'll do some kind of elimination.

Remember that a general system of two linear equations can be written as:

$$\begin{cases} a_1x + b_1y = c_1 & (1) \\ a_2x + b_2y = c_2 & (2) \end{cases}$$

and that the solution is given by:

$$\boxed{x = \frac{b_1c_2 - b_2c_1}{a_2b_1 - a_1b_2}} \text{ and } \boxed{y = \frac{a_2c_1 - a_1c_2}{a_2b_1 - a_1b_2}}$$

[4] To see more about it, go to `https://en.wikipedia.org/wiki/Gaussian_elimination`.

The idea to use is the same as the "case 2" systems we have learned and eliminate x: we'll multiply Eq. (2) by a number, and then subtract it from Eq. (1). To discover the number we'll multiply Eq. (2) by, let's think on what we want to happen: we want, when we do Eq. (1)− Eq. (2), to eliminate x. Thus, we need the coefficient of x, in Eq. (2), to be the same as in Eq. (1). So, how can we make a_2 "become" a_1? In equations, we want

$$a_2 \times ? = a_1$$

Solving this, we get

$$? = \frac{a_1}{a_2}$$

Let's multiply Eq. (2) by $\frac{a_1}{a_2}$:

$$\frac{a_1}{a_2}(a_2x + b_2y) = \frac{a_1}{a_2} \times c_2 \qquad \text{Multiplying both sides by } \frac{a_1}{a_2}$$

$$\frac{a_1}{a_2} \times a_2x + \frac{a_1}{a_2} \times b_2y = \frac{a_1c_2}{a_2} \qquad \text{Expanding the brackets}$$

$$\frac{a_1}{\cancel{a_2}} \times \cancel{a_2}x + \frac{a_1b_2}{a_2}y = \frac{a_1c_2}{a_2} \qquad \text{Multiplying the fractions}$$

$$a_1x + \frac{a_1b_2}{a_2}y = \frac{a_1c_2}{a_2} \qquad \text{Eq. (2) after the operation}$$

Now, our system has changed:

$$\begin{cases} a_1x + b_1y = c_1 & (1) \\ a_2x + b_2y = c_2 & (2) \end{cases} \rightarrow \begin{cases} a_1x + b_1y = c_1 & (1) \\ a_1x + \frac{a_1b_2}{a_2}y = \frac{a_1c_2}{a_2} & (2) \end{cases}$$

Let's do (1) − (2):

$$\overbrace{(a_1x + b_1y)}^{\text{Eq. (1) LHS}} - \underbrace{\left(a_1x + \frac{a_1b_2}{a_2}y\right)}_{\text{Eq. (2) LHS}} = \overbrace{c_1}^{\text{Eq. (1) RHS}} - \underbrace{\frac{a_1c_2}{a_2}}_{\text{Eq. (2) RHS}} \qquad (1) - (2)$$

$$a_1x + b_1y - a_1x - \frac{a_1b_2}{a_2}y = c_1 - \frac{a_1c_2}{a_2} \qquad \text{Expanding brackets}$$

$$b_1y - \frac{a_1b_2}{a_2}y = c_1 - \frac{a_1c_2}{a_2} \qquad \text{Eliminating } x$$

$$y\left(b_1 - \frac{a_1b_2}{a_2}\right) = c_1 - \frac{a_1c_2}{a_2} \qquad y \text{ as a common factor}$$

$$y\left(\frac{a_2b_1 - a_1b_2}{a_2}\right) = \frac{a_2c_1 - a_1c_2}{a_2} \qquad \text{Fraction magic}$$

$$y = \frac{a_2c_1 - a_1c_2}{a_2} \div \frac{a_2b_1 - a_1b_2}{a_2} \qquad \text{Dividing both sides by } \frac{a_2b_1 - a_1b_2}{a_2}$$

$$y = \frac{a_2c_1 - a_1c_2}{a_2} \times \frac{a_2}{a_2b_1 - a_1b_2}$$

$$y = \frac{a_2c_1 - a_1c_2}{\cancel{a_2}} \times \frac{\cancel{a_2}}{a_2b_1 - a_1b_2}$$

$$y = \frac{a_2c_1 - a_1c_2}{a_2b_1 - a_1b_2}$$

Finally, a formula for y. Now, let's substitute this pretty value in Eq. (1):

$$a_1x + b_1y = c_1 \quad \text{Eq. (1)}$$

$$a_1x + b_1\left(\frac{a_2c_1 - a_1c_2}{a_2b_1 - a_1b_2}\right) = c_1 \quad \text{Subsituting } y$$

$$a_1x = c_1 - b_1\left(\frac{a_2c_1 - a_1c_2}{a_2b_1 - a_1b_2}\right) \quad \text{Subtracting } b_1\left(\frac{a_2c_1 - a_1c_2}{a_2b_1 - a_1b_2}\right) \text{ from both sides}$$

$$a_1x = c_1 - \frac{a_2b_1c_1 - a_1b_1c_2}{a_2b_1 - a_1b_2} \quad \text{Expanding the bracket}$$

$$a_1x = \frac{a_2b_1c_1 - a_1b_2c_1 - a_2b_1c_1 + a_1b_1c_2}{a_2b_1 - a_1b_2} \quad \text{Fraction magic again!}$$

$$a_1x = \frac{\cancel{a_2b_1c_1} - a_1b_2c_1 - \cancel{a_2b_1c_1} + a_1b_1c_2}{a_2b_1 - a_1b_2} \quad \text{Cancelling equal terms}$$

$$a_1x = \frac{a_1b_1c_2 - a_1b_2c_1}{a_2b_1 - a_1b_2}$$

$$x = \frac{b_1c_2 - b_2c_1}{a_2b_1 - a_1b_2} \quad \text{Dividing both sides by } a_1 \text{and simplifying}$$

There we have it, the formula for x. In summary, we have that the solution for any system of two linear equations:

$$\begin{cases} a_1x + b_1y = c_1 & (1) \\ a_2x + b_2y = c_2 & (2) \end{cases}$$

is given by

$$\boxed{x = \frac{b_1c_2 - b_2c_1}{a_2b_1 - a_1b_2}} \text{ and } \boxed{y = \frac{a_2c_1 - a_1c_2}{a_2b_1 - a_1b_2}}$$

20. Sequences

20.1. Why learn sequences?

Well, an obvious reason is that it is going to be on your test. A less obvious reason is that people tend to confuse being "good" at identifying the general rule of a sequence as a sign of "intelligence", and who does not want to look intelligent?[1].

In a more historical note, sequences have always fascinated people. A classic example is one of Zeno's paradoxes:

> "That which is in locomotion must arrive at the half-way stage before it arrives at the goal."[2]

Translating the Greek, imagine that Achilles will run a track. Zeno argues[3] that, in order to reach the end of the track, Achilles first needs to run *half* the total distance ($\frac{1}{2}$), then *half of what is remaining* ($\frac{1}{2}$ of $\frac{1}{2}$, therefore $\frac{1}{4}$). After, he would need to run *half of the half of the half*, or half of the previous $\frac{1}{4}$, therefore $\frac{1}{8}$, and so on, *ad infinitum*. Therefore, Zeno said, Achilles would never finish the track, as he would always have some distance to cover!

As with everything in life, whenever in doubt, make a drawing:

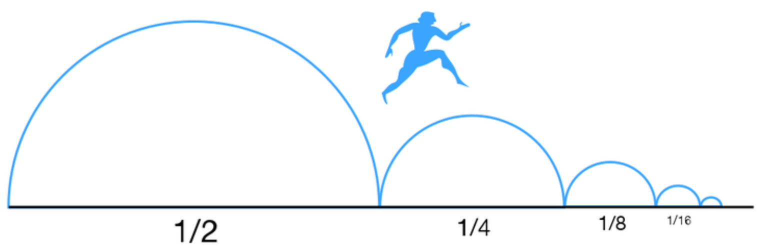


Figure 20.1.: Source: Wikipedia

This, other than telling that Zeno did not go out much, can be written as a *sequence* of the distances that Achilles must run:

[1]Clearly looking intelligent is more important than actually being intelligent.
[2]From https://en.wikipedia.org/wiki/Zeno%27s_paradoxes. Accessed on 06/03/2018.
[3]Argued, he is somewhat dead (he lives in our hearts and minds!)

$$\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \ldots$$

We could say that, by the "end", Achilles would have run:

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \ldots$$

Is this a number? Does Achilles ever finish the race? Interestingly, this was a problem that ancient Greeks could not solve, along with some others. Therefore, they gave up on life and were conquered by the Romans. True story.

Some time later, as Mr Apostol[4] puts:

> "In the 17th and 18th centuries, mathematicians began to realize that it is possible to extend the ideas of ordinary addition from *finite* collections of numbers to *infinite* collections so that sometimes infinitely many positive numbers have a finite 'sum'." [5]

Therefore it took around 2500 years to solve this problem! Infinite sequences are the basis of the discussion, however. They are a very important mathematical entity. That's why it is import to study them.

20.2. Linear sequences

20.2.1. Definition and key terms

The first type of sequence we need to learn are the so called *arithmetic* or *linear* sequences.

Definition. A sequence is called **linear** if the difference between any two consecutive elements is always the same. This is called the **common difference** and is usually denoted by the letter d.

Let's see some examples of them.

Example 1 Let's count:

$$1, 2, 3, 4, 5, 6, \ldots$$

This sequence is a linear sequence, as the difference between any two consecutive terms is always 1:

$$\underbrace{1, \quad 2,}_{+1} \quad \underbrace{\ \ 3,}_{+1} \quad \underbrace{\ \ 4,}_{+1} \ldots$$

[4] A Greek author!

[5] Apostol, T. *Calculus.* Page 374.

Example 2 The following sequence is also linear:

$$4, 2, 0, -2, \ldots$$

The difference between any two terms is, taking the first two, $2 - 4 = -2$:

$$4, \quad 2, \quad 0, \quad -2, \ldots$$
$$-2 \quad -2 \quad -2$$

Example 3 Let's put some fractions into this:

$$1, \frac{3}{2}, 2, \frac{5}{2}, \ldots$$

Again, the difference between any two consecutive terms is always the same $\frac{5}{2} - 2 = \frac{5}{2} - \frac{4}{2} = \frac{1}{2}$:

$$1, \quad \frac{3}{2}, \quad 2, \quad \frac{5}{2}, \ldots$$
$$+\frac{1}{2} \quad +\frac{1}{2} \quad +\frac{1}{2}$$

Solved exercise: first term and common difference of an arithmetic sequence

The following sequence is a linear sequence. Find its:

- first term
- common difference
- next three elements

$$7, 11, 15, 19, 23, \ldots$$

Solution To find the first term, look no further than the first element of the sequence: 7.
To find the common difference, just choose any two consecutive elements and subtract **the one that comes first from the one that comes after**. So, if we choose 7 and 11, we just need to do $11 - 7 = 4$. So the common difference is 4.
Now that we know the common difference, to find the next three elements, just go to the last element and add the common difference, then add the common difference to the answer and so on:

$$7, 11, 15, 19, \quad 23, \quad 27, \quad 31$$
$$+4 \quad +4 \quad +4$$

Therefore, the next three elements are 23, 27 and 31.

20.2.2. Finding the nth term for linear sequences

Finding the nth term of a sequence is simply finding a *rule* which, given a position in the sequence, which we refer by the variable n, gives you the element at that position. It is very important to notice that the position of a number in a sequence, n, **can only be a positive whole number** $(1, 2, 3, 4, \ldots)$.

For example, if we take the sequence $1, 3, 5, 7, \ldots$:

n	1	2	3	4
sequence	1	3	5	7

Looking at the table, you can say that when $n = 1$ the sequence is valued 1, when $n = 3$ the sequence has value 5, and so on. The idea of finding the nth term is finding an expression with n appearing in it that gives you any value of the sequence! For example, the nth term for the sequence $1, 3, 5, 7, \ldots$ is $2n - 1$. You can now calculate any value of the sequence. For example, if $n = 10$ (the 10th element of the sequence), it is going to be $2 \times 10 - 1 = 19$. We'll now learn how to find the nth term of linear sequences.

20.2.2.1. Finding the nth term: method 2 - the "0th" term

A very convenient manner to find the nth term of a linear sequence is to

1. First find the common difference;
2. Discover the element that would appear before the first;
3. Set up the nth term.

This method is excellent to use as it is quite quick and in the IGCSE exam you usually have the sequence laid out from its start.

Say that you have

$$5, 12, 19, 26, \ldots$$

The first thing is to find the common difference, which we can do subtracting any two consecutive terms:

$$5, \quad 12, \quad 19, \quad 26, \quad \ldots$$
$$+7 \quad +7 \quad +7$$

Now we use that to discover the "ghost element" to the left of the first, the 0th element, by subtracting the common difference from the first element:

$$-2, \quad 5, \quad 12, \quad 19, \quad 26, \quad \ldots$$
$$-7 \quad +7 \quad +7 \quad +7$$

After that, we just need to set up the nth term by getting the 0th term and adding the common difference times n. In our case,

$$\underbrace{-2}_{\text{0th term}} + \overbrace{7}^{\text{diff}} n$$

and we are done.

Another example:

$$8, 5, 2, -1, \dots$$

Find the common difference:

$$8, \quad 5, \quad 2, \quad -1, \quad \dots$$
$$-3 \quad -3 \quad -3$$

and then use it to find the term to the left of the first (careful with the minus negative, as we have to add 3 (from the $--3$) to find it):

$$11, \quad 8, \quad 5, \quad 2, \quad -1, \quad \dots$$
$$+3 \quad -3 \quad -3 \quad -3$$

finally obtaining

$$11 - 3n$$

20.2.2.2. Finding the nth term: method 1 - table

Let us say we have the following linear sequence:

$$5, 12, 19, 26, \dots$$

To find the nth term expression for it, you can do the following:

Step 1 Find the common difference:

We just need to subtract consecutive elements:

$$12 - 5 = 7$$

Step 2 Set up this table:

n	1	2	3	4
sequence	5	12	19	26

Step 3 Now, you'll add a third row to the column. You obtain this column by taking the common difference you obtained in **Step 1**, and multiplying it by the corresponding value of n:

n	1	2	3	4
sequence	5	12	19	26
common difference $\times n$	7×1	7×2	7×3	7×4

However, with practice, you'll eventually write the table like this:

n	1	2	3	4
sequence	5	12	19	26
$7n$	7	14	21	28

Step 4 You'll subtract the third row from the second. You'll notice all the results are the same:

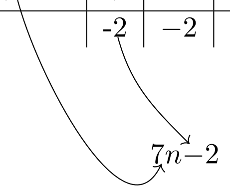
n	1	2	3	4
sequence	5	12	19	26
$7n$	7	14	21	28
sequence $-7n$	$5 - 7 = -2$	$12 - 14 = -2$	$19 - 21 = -2$	$26 - 28 = -2$

Again, with practice, you will just write it like this:

n	1	2	3	4
sequence	5	12	19	26
$7n$	7	14	21	28
	-2	-2	-2	-2

Step 5 Finally, you'll just need to get the correct data from your table:

n	1	2	3	4
sequence	5	12	19	26
7n	7	14	21	28
	-2	-2	-2	-2



And there you go! The nth term of the sequence $5, 12, 19, 26, \ldots$ is $7n - 2$. A very good idea is to check if your answer is correct by trying some substitutions:

$$n = 1 \rightarrow 7 \times 1 - 2 = 5$$

$$n = 2 \rightarrow 7 \times 2 - 2 = 12$$

$$n = 4 \rightarrow 7 \times 4 - 2 = 26$$

As they are all equal to the sequence elements, we are correct! Therefore, the nth term for the sequence $5, 12, 19, 26, \ldots$ is

$$\boxed{7n - 2}$$

Solved exercise: finding the nth term of a linear sequence using the table method

Find the nth term of the following linear sequence:

$$6, 4, 2, 0, -2, \ldots$$

Solution

Step 1 Find the common difference:

$$4 - 6 = -2$$

Step 2 and Step 3 Set up the table:

n	1	2	3	4
sequence	6	4	2	0
$-2n$	-2	-4	-6	-8

Step 4 Subtract the third row from the second (**suggestion**: always put the sign of the result):

n	1	2	3	4
sequence	6	4	2	0
$-2n$	-2	-4	-6	-8
	$+8$	$+8$	$+8$	$+8$

Step 5 Find the nth term by joining the correct parts:

n	1	2	3	4
sequence	6	4	2	0
-2n	-2	-4	-6	-8
	$+8$	$+8$	$+8$	8+

$$-2n+8$$

Let's check to see if we are correct:

$$n = 1- > -2 \times 1 + 8 = -2 + 8 = 6$$

$$n = 2- > -2 \times 2 + 8 = -4 + 8 = 4$$

$$n = 3- > -2 \times 3 = 8 = -6 + 8 = 2$$

They are all equal, so we are correct! Thus, the nth term for the sequence $6, 4, 2, 0, \ldots$ is

$$\boxed{-2n + 8}$$

20.2.2.3. Finding the nth term: method 2 - formula

Let's use the same sequence we used for the table method:

$$5, 12, 19, 26, \ldots$$

Before we deduce the general formula[6], let's think for that particular sequence.
We can calculate the common difference:

$$12 - 5 = 7$$

As it is a linear sequence, we know that each term is the previous added 7. Given that the first element is 5, we can obtain the second one by adding 7 to it:

$$5, 5 + 7$$

We can now obtain the third element by adding 7 to the second:

$$5, 5 + 7, (5 + 7) + 7$$

Repeating the same idea for the fourth:

$$5, 5 + 7, (5 + 7) + 7, ((5 + 7) + 7) + 7 \tag{20.1}$$

We could repeat this as much as we can, always adding 7 to the element that came before it.
Let's rewrite the expressions in (20.1). Again, let's add some brackets to make it easier to see the pattern:

$$5, 5 + 7, 5 + (7 + 7), 5 + (7 + 7 + 7), \ldots$$

We can write this in a more usual way:

$$5, 5 + 7, 5 + 2 \times 7, 5 + 3 \times 7, \ldots$$

Finally, let's just add some things that change nothing:

$$5 + 0 \times 7, 5 + 1 \times 7, 5 + 2 \times 7, 5 + 3 \times 7, \ldots$$

I hope you can see the pattern now! Let's make a table:

n	1	2	3	4
sequence	$5 + 0 \times 7$	$5 + 1 \times 7$	$5 + 2 \times 7$	$5 + 3 \times 7$

Therefore, when $n = 1$, the sequence is equal to $5+0\times 7$, that is, the first term (5) plus zero times the common difference (7). When $n = 2$, the sequence is equal to $5 + 1 \times 7$, that is, the first term plus one time the common difference. When $n = 3$, the sequence

[6] See the formality after taste.

is equal to $5 + 2 \times 7$, that is, the first term (5) plus two times the common difference (7). Finally, for any n, the sequence will be the first term plus $n - 1$ times 7:

$$\underbrace{5}_{\text{1st term}} + \underbrace{7}_{\text{common difference}} \times (n-1)$$

That is actually the formula: for any arithmetic sequence with first term u_1 and common difference d, the nth term is given by the formula[7]:

$$\boxed{u_1 + d(n-1)}$$

Solved exercise: finding the nth term of a linear sequence using the formula

Find the nth term of the following arithmetic sequence:

$$6, 4, 2, 0, -2, \ldots$$

To use the formula we just need the first term and the common difference. The first term is 6. Let's calculate the common difference:

$$4 - 6 = -2$$

Solution Therefore, $u_1 = 6$ and $d = -2$. Let's put them into the formula:

$$u_1 + d(n-1) = 6 + -2(n-1)$$

Let's just expand the brackets and collect like terms:

$6 - 2(n-1)$	Remember that $+ \times - = -$
$6 - 2n + 2$	Careful when distributing the negative!
$8 - 2n$	Collect and go!

Therefore, the nth term for the sequence $6, 4, 2, 0, -2, \ldots$ is:

$$\boxed{8 - 2n = -2n + 8}$$

As you can see, both the table method and the formula method give the same result. However, the formula is much quicker[8], so I suggest using it.

[7] As curiosity, notice that by expanding the formula and adding some brackets for clarity you obtain $(u_1 - d) + dn$, which is the "0th term" method we saw above.

[8] And useful in your further studies!

Solved exercise showing that a number is not part of a linear sequence

Show that the number 501 is not an element of the sequence

$$4, 9, 14, 19, 24, \ldots$$

Solution If 501 were an element of the sequence, it would have a position in the sequence. Therefore, if we knew the nth term of the sequence we could find the corresponding n which would give 502. Let us see if such n exists.
Let's use the formula to find the nth term. The first term is $u_1 = 4$ and the common difference $d = 9 - 4 = 5$. Plugging those into the formula:

$$u_1 + d(n-1) = 4 + 5(n-1) = 4 + 5n - 5 = 5n - 1$$

The nth term is, thus, $5n - 1$. If 501 was an element of the sequence, there would be a certain n that would make $5n - 1$ equal to 501. Let's solve this equation:

$$\begin{aligned} 5n - 1 &= 501 && \text{Adding 1 on both sides} \\ 5n &= 501 + 1 \\ 5n &= 502 && \text{Dividing both sides by 5} \\ n &= \frac{502}{5} \\ n &= 100.4 \end{aligned}$$

As we can see, the n we calculated is not a positive whole number. Therefore, 501 cannot be an element of the sequence $4, 9, 14, 19, \ldots$.

Solved exercise: showing that a number is part of a linear sequence

Show that -13 is an element of the sequence

$$17, 15, 13, 11, \ldots$$

Solution The same idea applies: we calculate the nth term of the sequence first. The first term $u_1 = 17$ and the common difference is $d = 15 - 17 = -2$. Plugging those into the formula:

$$u_1 + d(n-1) = 17 - 2(n-1) = 17 - 2n + 2 = 19 - 2n$$

The nth term is, then, $19 - 2n$.
Now, if -13 is an element of the sequence, that means there is a value of n for which $19 - 2n$ is equal to -13:

$$\begin{aligned}
19 - 2n &= -13 && \text{Subtracting 19 on both sides} \\
-2n &= -13 - 19 \\
-2n &= -32 && \text{Dividing both sides by } -2 \\
n &= \frac{-32}{-2} \\
n &= 16
\end{aligned}$$

Given that $n = 16$ is a positive whole number, -13 is an element of the arithmetic sequence $17, 15, 13, 13, \ldots$.

20.3. Geometric sequences

20.3.1. Definition and key terms

The second type of sequence we'll learn about are the *geometric* sequences.

Definition. A sequence is called **geometric** is the *ratio* between any two consecutive terms is always the same. We'll call this number the **common ratio** and denote it by the letter r.

Remember that ratio is a fancy word for division. Let's see some examples:

Example 1 Let us multiply by 2:

$$1, 2, 4, 8, 16, 32, \ldots$$

This sequence is geometric because if you take any two consecutive terms and divide one by the other the ratio is always 2:

$$1, \quad 2, \quad 4, \quad 8, \quad 16 \ldots$$
$$\times 2 \quad \times 2 \quad \times 2 \quad \times 2$$

Example 2 A very interesting geometric sequence is the following:

$$17, -17, 17, -17, 17, -17, \ldots$$

If you take any two consecutive terms and divide them, you always obtain -1:

$$17, \quad -17, \quad 17, \quad -17, \quad 17 \ldots$$
$$\times -1 \quad \times -1 \quad \times -1 \quad \times -1$$

Whenever the *common ratio* of the sequence is negative, the signs of the elements of the sequence will alternate!

Example 3 The Zeno's paradox sequence:

$$\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \ldots$$

If you divide any two consecutive terms you always obtain $\frac{1}{2}$:

$$\tfrac{1}{2}, \quad \tfrac{1}{4}, \quad \tfrac{1}{8}, \quad \tfrac{1}{16}, \quad \tfrac{1}{32} \ldots$$
$$\times \tfrac{1}{2} \quad \times \tfrac{1}{2} \quad \times \tfrac{1}{2} \quad \times \tfrac{1}{2}$$

Some people prefer saying that you are dividing each term by 2 to obtain the next, but you and I know that division is just a multiplication, so let's stick to our guns.

Solved exercise: first term and common ratio of a geometric sequence
The following sequence is geometric. Find its:

- first term
- common ratio
- next three elements

$$81, -27, 9, -3, 1, -\frac{1}{3}, \dots$$

Solution Again, the first term is the first element of the sequence: 81.
To find the common ratio, just choose any two consecutive terms and divide **the one that comes after by the one that comes first**. If we choose −27 and 9:

$$\frac{9}{-27} = \frac{1}{-3} = -\frac{1}{3}$$

Therefore $r = -\frac{1}{3}$.
To find the next three terms we go to the last term, multiply it by the common ratio, and repeat that two times:

$$81, -27, 9, -3, 1, \quad -\tfrac{1}{3}, \quad \tfrac{1}{9}, \quad -\tfrac{1}{27}$$
$$\times -\tfrac{1}{3} \quad \times -\tfrac{1}{3} \quad \times -\tfrac{1}{3}$$

Thus, the three next terms are $-\frac{1}{3}$, $\frac{1}{9}$ and $-\frac{1}{27}$.

20.3.2. Finding the nth term: formula

To find the nth term of a geometric sequence the best way is to notice the pattern[9].
Let's find the nth term for the sequence:

$$3, 6, 12, 24, 48, \dots$$

First, let's identify it's common ratio:

$$d = \frac{6}{3} = 2$$

That means that we are always multiplying each term in the sequence by 2 to obtain the next.
Therefore, we can write the sequence like this:

[9]Again, check the Formality after taste for a more general example.

$$
\begin{aligned}
3 & \qquad \text{The first term!} \\
6 = 3 \times 2 & \qquad \text{The previous term times 2} \\
12 = 6 \times 2 & \qquad \text{Again, the previous term times 2} \\
24 = 12 \times 2 & \qquad \text{You got the deal!}
\end{aligned}
$$

Now, pay attention to the third term, 12. We obtained 12 by multiplying the previous term, 6 by 2:

$$12 = 6 \times 2$$

Well, we also obtained 6 by multiplying the first term by 2:

$$6 = 3 \times 2$$

Let's substitute this into the 12 equality:

$$
\begin{aligned}
12 &= 6 \times 2 & & \\
12 &= (3 \times 2) \times 2 & & \text{Substituting } 6 = 3 \times 2 \\
12 &= 3 \times 2 \times 2 & & \\
12 &= 3 \times 2^2 & & 2 \times 2 = 2^2
\end{aligned}
$$

We can repeat the same reasoning for 24:

$$
\begin{aligned}
24 &= 12 \times 2 & & \\
24 &= (3 \times 2^2) \times 2 & & \text{Substituting } 12 = 3 \times 2^2 \\
24 &= 3 \times 2^2 \times 2 & & \\
24 &= 3 \times 2^3 & & 2^2 \times 2 = 2^{2+1} = 2^3
\end{aligned}
$$

Now, let's look at those terms regarding their position in the sequence:

$$n = 1 \to 3 = 3 \times 2^0$$

$$n = 2 \to 6 = 3 \times 2^1$$

$$n = 3 \to 12 = 3 \times 2^2$$

$$n = 4 \to 24 = 3 \times 2^3$$

I hope you can see the pattern: the element at position n is obtained by multiplying the first term, in our case 3, by 2^{n-1}, that is, the common ratio to the power of $n-1$!

So, for the sequence $3, 6, 12, 24, 48, \ldots$, as the first term is 3 and the common ratio 2, the nth term is:

$$3 \times 2^{n-1}$$

Finding the nth term of a geometric sequence

Find the nth term of the sequence

$$5, 15, 45, 135, \ldots$$

Solution We just need to find the first term and the common ratio. The first term is, well, the first element: 5. The common ratio we just need to divide consecutive terms:

$$r = \frac{45}{15} = 3$$

We just need to join those two pieces of information, as the nth term is just the first term times the common ratio to the power of $n-1$:

$$5 \times 3^{n-1}$$

Therefore, the nth term of the sequence $5, 15, 45, 135, \ldots$ is:

$$\boxed{5 \times 3^{n-1}}$$

20.4. Quadratic sequences

We now arrive at the second most popular type of sequence in the exam, the quadratic sequences! Let's see an example of quadratic sequence:

$$1, 4, 9, 16, 25, 36, 49, 64, 81, 100, \ldots$$

You have probably seen this numbers before: they are the square numbers[10]!

Looking at the sequence, you can see that the difference between any two consecutive terms is not always equal:

$$\begin{array}{ccccccccccccc} 1, & & 4, & & 9, & & 16, & & 25, & & 36, & & 49, \ldots \\ & +3 & & +5 & & +7 & & +9 & & +11 & & +13 & \end{array}$$

We can also see that the ratio between any two consecutive terms is not constant:

$$\begin{array}{ccccccccccccc} 1, & & 4, & & 9, & & 16, & & 25, & & 36, & & 49, \ldots \\ & \times 4 & & \times\frac{9}{4} & & \times\frac{16}{9} & & \times\frac{25}{16} & & \times\frac{36}{25} & & \times\frac{49}{36} & \end{array}$$

As you can, see quadratic sequences are totally different from arithmetic and geometric sequences! If you remind our study of quadratic expressions, you'll remember that a quadratic expression can always be rearranged to look like this:

$$an^2 + bn + c$$

All the nth term expressions of quadratic sequences will look like this. In particular, the sequence $1, 2, 4, 9, 16, 25, \ldots$ is given by n^2.

20.4.1. Finding the nth term of a quadratic sequence: finding the second difference

There are three methods to find the nth term of a quadratic sequence[11]. Both, however, require you to find the *second difference* of the sequence, so let's learn how to that before actually learning the methods.

Let's use this sequence as an example:

$$1, 4, 11, 22, 37, \ldots$$

The first thing to do is seeing what you are adding to each term from the previous, pretty much in the same way we did for a linear sequence:

$$\begin{array}{ccccccccc} 1, & & 4, & & 11, & & 22, & & 37, \ldots \\ & +3 & & +7 & & +11 & & +15 & \end{array}$$

The numbers $3, 7, 11, 16, \ldots$ are called the *first difference* of the sequence.

The idea with a quadratic sequence is to the same process again, but calculating how the terms in the first difference are changing:

[10] What a surprise!

[11] That I know of.

Sequence	1		4		11		22		37
1st difference		+3		+7		+11		+15	
2nd difference			+4		+4		+4		

Did you notice that the second difference, that is, *how the first differences are changing*, are always the same? That is the characteristic of a quadratic sequence: **the second difference is always the same!** To find the nth term of a quadratic sequence we'll always find this constant second difference.

20.4.2. Finding the nth term of a quadratic sequence method 1: simultaneous equations

First remember that any quadratic sequence has nth term of the form: $an^2 + bn + c$.

As you can imagine, the first step to find the nth term of a quadratic sequence is to calculate its second difference. Let's use this sequence as an example:

$$12, 20, 30, 42, 56, \ldots$$

Let's calculate its second difference:

Sequence	12		20		30		42		56
1st difference		+8		+10		+12		+14	
2nd difference			+2		+2		+2		

The second difference of this sequence is 2. We have already found the a of the nth term! We just need to divide the second difference by 2:

$$a = \frac{2}{2} = 1$$

Therefore, to find the a of the nth term we just need to **divide the second difference by 2!** Until here, then, we know that our sequence is given by

$$1n^2 + bn + c = n^2 + bn + c$$

How could we find b and c? Let's think. We need to find the value of two unknowns, correct? If we could find two equations, we could solve them and done! Guess what: we can find those equations![12]The idea is that we know the values the sequence has at each position, so if we substitute values for n in our incomplete nth term expression, we know what *they should* give. For example, we know that the first term of the sequence is 12. That means that when $n = 1$ the nth term should give us 12:

$$n = 1 \to 1^2 + b \times 1 + c = 12$$

[12]We can actually find infinite equations!

Here is our first equation:

$$1 + b + c = 12$$

$$b + c = 11$$

We can find another by substituting another value for n. Let's do it:

$$n = 2 \to 2^2 + b \times 2 + c = 20$$

If we put $n = 2$, the second term of the sequence is 20, hence our equation. Let's simplify it:

$$4 + 2b + c = 20$$

$$2b + c = 16$$

Here is our second equation! We just need to solve the simultaneous equations system now:

$$\begin{cases} b + c = 11 & (1) \\ 2b + c = 16 & (2) \end{cases}$$

We can subtract one equation from the other to eliminate c:

$$\underbrace{b + c}_{\text{eq.1}} - (\underbrace{2b + c}_{\text{eq.2}}) = \underbrace{11}_{\text{eq.1}} - \underbrace{16}_{\text{eq.2}}$$

Be very careful with the negative!

$$b + c - 2b - c = -5$$

$$-b = -5$$

$$\boxed{b = 5}$$

There you go! We found b! We just need to substitute it into one of our original equations and we're done. Let's do it in the equation $b + c = 11$:

$$5 + c = 11$$

$$c = 11 - 5$$

$$\boxed{c = 6}$$

We're done! The nth term of the sequence $12, 20, 30, 42, 56, \ldots$ is

$$\boxed{n^2 + 5n + 6}$$

Let's check our answer:

$$n = 1 \rightarrow 1^2 + 5 \times 1 + 6 = 1 + 5 + 6 = 12$$

$$n = 2 \rightarrow 2^2 + 5 \times 2 + 6 = 4 + 10 + 6 = 20$$

$$n = 3 \rightarrow 3^2 + 5 \times 3 + 6 = 9 + 15 + 6 = 30$$

$$n = 4 \rightarrow 4^2 + 5 \times 4 + 6 = 16 + 20 + 6 = 42$$

We're correct!

Solved exercise: finding the th term of a quadratic sequence using simultaneous equations

Find the nth term of the sequence

$$-2, 2, 12, 28, 50, \ldots$$

Solution We first need to calculate the second difference of the sequence:

Sequence	−2		2		12		28		50
1st difference		+4		+10		+16		+22	
2nd difference			+6		+6		+6		

Given that the second difference is 6, we know that the n^2 coefficiente in the nth term is given by $\frac{6}{2} = 3$:

$$3n^2 + bn + c$$

Now, we need to set up a system of simulateneous equations to find b and c. Substituting $n = 1$:

$$n = 1 \rightarrow 3 \times 1^2 + b \times 1 + c = -2$$

$$3 + b + c = -2$$

$$b + c = -5$$

To obtain a second equation, let's substitute $n = 2$:

$$n = 2 \rightarrow 3 \times 2^2 + b \times 2 + c = 2$$

$$12 + 2b + c = 2$$

$$2b + c = -10$$

Thus, our system is

$$\begin{cases} b + c = -5 & (1) \\ 2b + c = -10 & (2) \end{cases}$$

Let's calculate equation (2) – equation (1):

$$\underbrace{(2b + c)}_{\text{eq. 2}} - \underbrace{(b + c)}_{\text{eq. 1}} = \underbrace{-10}_{\text{eq. 2}} - \underbrace{-5}_{\text{eq. 1}}$$

$$2b + c - b - c = -10 + 5$$

$$\boxed{b = -5}$$

We have our b. Substituting it back in equation (1):

$$-5 + c = -5$$

$$\boxed{c = 0}$$

Done! The nth term of the sequence $-2, 2, 12, 28, 50, \ldots$ is

$$3n^2 - 5n$$

Let's check our answer:

$$n = 1 \rightarrow 3 \times 1^2 - 5 \times 1 = 3 - 5 = -2$$

$$n = 2 \rightarrow 3 \times 2^2 - 5 \times 2 = 12 - 10 = 2$$

$$n = 3 \rightarrow 3 \times 3^2 - 5 \times 3 = 27 - 15 = 12$$

$$n = 4 \rightarrow 3 \times 4^2 - 5 \times 4 = 48 - 20 = 28$$

That's eveidence enough: we are correct!

Finding the nth term of a quadratic sequence method 2: table

We'll use the same sequence as the example before:

$$12, 20, 30, 42, 56, \ldots$$

The second method starts in exactly the same way: by finding the second difference of the sequence:

Sequence	12		20		30		42		56
1st difference		+8		+10		+12		+14	
2nd difference			+2		+2		+2		

Again, we divide the second difference by 2 to obtain the a in the nth term expression:

$$a = \frac{2}{2} = 1$$

Now we know that the sequence looks like this:

$$n^2 + bn + c$$

Until here exactly the same steps as the previous method! Now they differ. We already know the sequence has a n^2part, correct? We'll put that in a table:

n	1	2	3	4	5
sequence	12	20	30	42	56
n^2	1	4	9	16	25

We continue by subtracting the n^2 line from the "sequence" line:

n	1	2	3	4	5
sequence	12	20	30	42	56
n^2	1	4	9	16	25
sequence$-n^2$	$12 - 1 = 11$	$20 - 4 = 16$	$30 - 9 = 21$	$42 - 16 = 26$	$56 - 25 = 31$

Let's just write this more clearly:

n	1	2	3	4	5
sequence	12	20	30	42	56
n^2	1	4	9	16	25
sequence$-n^2$	11	16	21	26	31

Look at the last line. The numbers there

$$11, 16, 21, 26, 31$$

They form a linear sequence!

$$11, \underset{+5}{\curvearrowright} 16, \underset{+5}{\curvearrowright} 21, \underset{+5}{\curvearrowright} 31$$

We already know how to find the nth term of a linear sequence. Let's use the formula, as it's quicker. We just need to identify the first term, in this case $u_1 = 11$, and the common difference, in this case $d = 5$. We know the formula is given by $u_n + (n-1)d$. Subsituting:

$$11 + (n-1) \times 5$$

$$11 + 5n - 5$$

$$5n + 6$$

We just need to add this to the n^2 part and we're done[13]:

$$\underbrace{n^2}_{\text{2nd diff. part}} + \underbrace{5n+6}_{\text{table part}}$$

Good thing we obtained the same answer, which we have already checked. As you have probably noticed, the table method is faster, so I recommend you using it.

[13] Notice that we had taken away the n^2 from the original sequence, that's why we need to add it back now!

Solved exercise: finding the th term of a quadratic sequence using the table

Find the nth term of the sequence:

$$1, -5, -15, -29, -47, \ldots$$

Solution The first thing is, as usual with a quadratic sequence, find its second difference:

Sequence	1		-5		-15		-29		-47
1st difference		-6		-10		-14		-18	
2nd difference			-4		-4		-4		

So we know our nth term we'll have n^2 coefficient given by $\frac{-4}{2} = -2$. Let's fill our table:

n	1	2	3	4	5
sequence	1	-5	-15	-29	-47
$-2n^2$	-2	-8	-18	-32	-50

We continue by subtracting the n^2 line from the "sequence" line:

n	1	2	3	4	5
sequence	1	-5	-15	-29	-47
$-2n^2$	-2	-8	-18	-32	-50
sequence$-n^2$	$1 - -2 = 3$	$-5 - -8 = 3$	$-15 - -18 = 3$	$-29 - -32 = 3$	$-47 - -50 = 3$

Let's just write this more clearly:

n	1	2	3	4	5
sequence	1	-5	-15	-29	-47
$-2n^2$	-2	-8	-18	-32	-50
sequence$-n^2$	3	3	3	3	3

We need to find the nth term for the sequence in the last row of our table. Well, all the terms are 3, to the sequence is just 3! Let's add it to our $-2n^2$:

$$-2n^2 + 3$$

Let's check our answer:

$$n = 1 \rightarrow -2 \times 1^2 + 3 = 1$$

$$n = 2 \rightarrow -2 \times 2^2 + 3 = -5$$

$$n = 3 \rightarrow -2 \times 3^2 + 3 = -15$$

$$n = 4 \rightarrow -2 \times 4^2 + 3 = -29$$

That's evidence enough, we are correct! Therefore, the nth term of the sequence $1, -5, -15, -29, -47, \ldots$ is

$$\boxed{-2n^2 + 3}$$

Another example of this type. Let us find the nth term of the sequence:

$$4, 13, 26, 43, \ldots$$

Solution We start by finding its second difference (which actually tells us it is a quadratic sequence):

Sequence	4		13		26		43
1st difference		9		13		17	
2nd difference			4		4		

And as we have seen, our nth term we'll have n^2 coefficient given by $\frac{4}{2} = 2$. Table time:

n	1	2	3	4
sequence	4	13	26	43
$2n^2$	2	8	18	32

We continue by subtracting the n^2 line from the "sequence" line:

n	1	2	3	4
sequence	4	13	26	43
$2n^2$	2	8	18	32
sequence$-n^2$	$4-2=2$	$13-8=5$	$26-18=8$	$43-32=11$

Writing it more concisely:

n	1	2	3	4
sequence	4	13	26	43
$2n^2$	2	8	18	32
sequence$-n^2$	2	5	8	11

We now have to find the nth term of the sequence we obtained in the last line:

$$2, 5, 8, 11$$

This is a linear sequence, with common difference 3:

$$2, \quad 5, \quad 8, \quad 11, \ldots$$

$$+3 \quad +3 \quad +3$$

and we can find the "0-th" term by subtracting 3 from the first term:

$$-1, \quad 2, \quad 5, \quad 8, \quad 11, \ldots$$

$$-3 \quad +3 \quad +3 \quad +3$$

and we can obtain our nth term:

$$3n - 1$$

Joining this with the original $2n^2$ we have our full nth term for the sequence:

$$2n^2 + 3n - 1$$

Let's check our answer:

$$n = 1 \rightarrow 2 \times 1^2 + 3 \times 1 - 1 = 2 + 3 - 1 = 4$$

$$n = 2 \rightarrow 2 \times 2^2 + 3 \times 2 - 1 = 8 + 6 - 1 = 13$$

$$n = 3 \rightarrow 2 \times 3^2 + 3 \times 3 - 1 = 18 + 9 - 1 = 26$$

$$n = 4 \rightarrow 2 \times 4^2 + 3 \times 4 - 1 = 32 + 12 - 1 = 43$$

And we are correct! Therefore, the nth term of the sequence $4, 13, 26, 43, \ldots$ is

$$\boxed{2n^2 + 3n - 1}$$

20.4.3. Finding the nth term of a quadratic sequence method 3: formula[14]

There is a formula for finding the nth term of a quadratic sequence. It's a very big and useless formula, I'll just put it here if you really like memorizing formulae.

Given a sequence

$$u_1, u_2, u_3, u_4, u_5, \ldots$$

Which has a constant second difference (is quadratic) given by:

$$d = (u_3 - u_2) - (u_2 - u_1) = u_3 - 2u_2 + u_1$$

The nth term of this sequence is given by

$$\frac{d}{2}n^2 + \frac{-5u_1 + 8u_2 - 3u_3}{2}n + 3u_1 - 3u_2 + u_3$$

Let's see how to apply this. Again, don't recommend it!

For our sequence

$$12, 20, 30, 42, 56, \ldots$$

[14] I highly recommend you to ignore this. It's always better to think than to memorize.

We can calculate d:

$$d = 30 - 2 \times 20 + 12 = 30 - 40 + 12 = 2$$

(This step is calculating the second difference, by the way).
We then just substitute values:

$$\frac{2}{2}n^2 + \frac{-5 \times 12 + 8 \times 20 - 3 \times 30}{2}n + 3 \times 12 - 3 \times 20 + 30$$

$$n^2 + \frac{-60 + 160 - 90}{2}n + 36 - 60 + 30$$

$$n^2 + \frac{10}{2}n + 6$$

$$n^2 + 5n + 6$$

As you can see, it gives the same answer. Again, don't use it.

Solved exercise: finding the nth term of a quadratic sequence using the formula

Find the nth term of the sequence:

$$-1, 0, 7, 20, 39, \dots$$

First we calculate d:

$$d = (u_3 - u_2) - (u_2 - u_1) = (7 - 0) - (0 - -1) = 7 - 1 = 6$$

Now we substitute things:

$$\frac{d}{2}n^2 + \frac{-5u_1 + 8u_2 - 3u_3}{2}n + 3u_1 - 3u_2 + u_3$$

$$\frac{6}{2}n^2 + \frac{-5 \times -1 + 8 \times 0 - 3 \times 7}{2}n + 3 \times -1 - 3 \times 0 + 7$$

$$3n^2 + \frac{5 - 21}{2}n - 3 + 7$$

$$3n^2 + \frac{-16}{2}n + 4$$

$$3n^2 - 8n + 4$$

Let's check our answer:

$$n = 1 \rightarrow 3 \times 1^2 - 8 \times 1 + 4 = 3 - 8 + 4 = -1$$

$$n = 2 \rightarrow 3 \times 2^2 - 8 \times 2 + 4 = 12 - 16 + 4 = 0$$

$$n = 3 \rightarrow 3 \times 3^2 - 8 \times 3 + 4 = 27 - 24 + 4 = 7$$

$$n = 4 \rightarrow 3 \times 4^2 - 8 \times 4 + 4 = 48 - 32 + 4 = 20$$

This is a good indication we are correct. Do not use this formula.

20.5. Other sequences

There are two other types of sequences which may appear, but more in a "identify me" matter and less on calculating anything. These are very rare, but better safe than sorry.

The first one is the cube numbers sequence:

$$1, 8, 27, 64, 125, \dots, n^3, \dots$$

You won't be asked to find a general cubic sequence nth term, but they may give you "simple" derivations of the basic sequence, such as:

$$n^3 \rightarrow 1, 8, 27, 64, 125, \ldots$$

$$n^3 - 2 \rightarrow -1, 6, 25, 63, 123, \ldots$$

$$2n^3 \rightarrow 2, 16, 54, 128, 250, \ldots$$

and so on.

The second type of sequence it may appear is Fibonacci sequences, in which you obtain the next term by adding the previous two. The classic one is

$$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, \ldots$$

But you can start with any two numbers that you want. For instance:

$$3, 4, 7, 11, 18, 29, \ldots$$

Just keep those in your mind!

20.6. Exam hints

When presented a sequences question, a good "roadmap" to follow is:

1. Identify the type of the sequence you are dealing with;
2. Find the nth term of the sequence;
3. Do whatever they ask you to.

Usually you will stop at step 2, but sometimes they may ask you to show that a number is or isn't part of the sequence, or calculate a particular term.

Summary

- A **sequence** is, well, a sequence of numbers. Each number has a **position** in the sequence denoted by the letter n. This position can only be a **positive whole number**;
- To **find the nth** term of a sequence means finding an expression involving n to calculate any term of the sequence;
- An **arithmetic** or **linear** sequence is a sequence in which all consecutive terms have a constant difference between them;

- The nth term of a linear sequence can be found using the table method or the formula

$$u_1 + (n-1)d$$

in which u_1 is the first term of the sequence and d the common difference;

- A **geometric** sequence is a sequence in which all consecutive terms have a constant ratio between them;

- The nth term of a geometric sequence can be found using the formula:

$$u_1 \times r^{n-1}$$

in which u_1 is the first term of the sequence and r the common ratio;

- A **quadratic sequence** is a sequence in which **the second difference** is **constant;**

- The nth term of a quadratic sequence is always an expression of the form $an^2 + bn + c$, and you can find a by **dividing the second difference by 2.** You can find the remaining part of the expression by either using simultaneous equations or the "table method" (there is also the formula, but don't use it);

- It is a good idea to remember the cube numbers sequence $1, 8, 27, 64, 125, \ldots$ and "Fibonacci" type sequences, in which you obtain the next term by adding the previous two.

Formality after taste

Sequences definition

Remember functions? A sequence is nothing but a special function! For example, we could think of the function:

$$f(x) = 3x + 1$$

And, usually, you could take any number you wanted in its domain[15] and substitute x for it, correct? Some examples:

$$f(2) = 3 \times 2 + 1 = 6 + 1 = 7$$

$$f(\frac{1}{2}) = 3 \times \frac{1}{2} + 1 = \frac{3}{2} + 1 = \frac{5}{2}$$

$$f(\pi) = 3\pi + 1$$

[15]For this f any real number!

However, if we are only allowed to choose *natural numbers*[16], and we always start at 1 and go to 2 and then 3 and so on, look what the results are[17]:

$$f(1) = 3 \times 1 + 1 = 4$$

$$f(2) = 3 \times 2 + 1 = 7$$

$$f(3) = 3 \times 3 + 1 = 10$$

$$f(4) = 4 \times 4 + 1 = 14$$

And so on! In the end, what we have is that the $f(x)$ when x is a natural number is a sequence:

x	1	2	3	4
$f(x)$	4	7	10	14

Now, as the symbol representing the natural numbers is $\mathbb{N}$, we prefer not using x for sequences, but n:

n	1	2	3	4
$f(n)$	4	7	10	14

It is exactly the same thing, but with n!

Now, for the very precise definition:

Definition. An **infinite sequence**[18] is a function f whose domain is the natural numbers $1, 2, 3, 4, \ldots$. The value $f(n)$ is called the nth term of the sequence.

Therefore, finding the nth term of a sequence is simply finding the function which generates the sequence.

Deriving the nth term formula for arithmetic sequences

A general arithmetic sequence can be written as:

$$u_1, u_2, u_3, u_4, \ldots$$

in which the difference between consecutive terms is always a constant. Or, in mathematicianese:

[16]The natural numbers are the numbers $1, 2, 3, \ldots$, the ones we use to count!

[17]If you remember one the functions formality after tastes, we are saying that the domain of f is the natural numbers.

[18]It is infinite as the natural numbers you can choose from is somewhat endless.

$$u_{n+1} - u_n = d$$

That is, take any term u_{n+1}. The term that comes before it is u_n. If you subtract this from the one forward, you obtain the common difference d.

Now, we'll basically repeat the process we used for a numeric sequence with variables! We know that each term is just the previous one added d:

$$u_1 = u_1$$
$$u_2 = u_1 + d$$
$$u_3 = u_2 + d$$
$$u_4 = u_3 + d$$

And so on. Do you notice that when calculating u_3 you use u_2? We know that $u_2 = u_1 + d$, so let's substitute that into u_3:

$$u_1 = u_1$$
$$u_2 = u_1 + d$$
$$u_3 = u_2 + d = (u_1 + d) + d = u_1 + d + d = u_1 + 2d$$
$$u_4 = u_3 + 3$$

Now, when calculating u_4 we need u_3. We do know that $u_3 = u_1 + 2d$, so why not plug that into u_4:

$$u_1 = u_1$$
$$u_2 = u_1 + d$$
$$u_3 = u_1 + 2d$$
$$u_4 = u_3 + 3 = (u_1 + 2d) + d = u_1 + 2d + d = u_1 + 3d$$

This pattern would continue forever: you would put u_4 into the expression for $u_5 = u_4 + d$ and obtain $u_5 = u_1 + 4d$ and so on. As you can notice, when you are calculating the term at position n in the sequence, we always reach that it is equal to the first term, u_1, plus the common difference multiplied by $n - 1$.

Thus, the formula for the nth term of a linear sequence $u_1, u_2, u_3, \ldots$ is:

$$\boxed{u_n = u_1 + (n-1)d}$$

It's usually written with the d at the end, but you can also write like this:

$$\boxed{u_n = u_1 + d(n-1)}$$

Deriving the nth term formula for geometric sequences

A general geometric sequence can be written as

$$u_1, u_2, u_3, u_4, \ldots$$

in which the *ratio* of any two consecutive terms is a constant called the common ratio:

$$\frac{u_{n+1}}{u_n} = r$$

This means that, to obtain the next term of the sequence, you simply multiply the previous one by r:

$$u_1 = u_1$$

$$u_2 = u_1 \times r$$

$$u_3 = u_2 \times r$$

$$u_4 = u_3 \times r$$

$$\vdots$$

$$u_n = u_{n-1} \times r$$

and so on. Notice that u_3 uses u_2, that u_4 uses u_3 and so on. In that way, the term u_n always uses the term u_{n-1}:

$$u_1 = u_1 = u_1 \times 1 = u_1 \times r^0$$

$$u_2 = u_1 \times r = u_1 \times r^1$$

$$u_3 = u_2 \times r = (u_1 \times r) \times r = u_1 \times r^2$$

$$u_4 = u_3 \times r = ((u_1 \times r) \times r) \times r = u_1 \times r^3$$

$$\vdots$$

As you can see, to calculate term u_4 we multiply the first term, u_1, by the common ratio to the power of 3; to calculate the term u_3 we multiply u_1 by the common ratio to the power of 2, and so on: if we need to calculate term u_n (the nth term!), we simply multiply the first term by the common ratio to the power of $n-1$:

$$\boxed{u_n = u_1 \times r^{n-1}}$$

21. Variation

21.1. Why learn variation?

Well, this one is easy! Have you noticed things are constantly changing? Annoying, right? It even reminds me of a quote from *Sandman*:

> "Um, what's the name of the word for things not being the same always? You know, I'm sure there is one. Isn't there? There must be a word for it ... the thing that lets you know time is happening. Is there a word?"
> "Change."[1]

Given that math, annoyingly, is supposed to be useful, it would be nice if there was a way to relate two quantities that change together. Search no more, you will find a technique in this chapter.

21.2. What do we mean by "variation"?

Variation is a fancy name for change. We will, however, limit our discussion of the topic to when two quantities are "linked", in a sense that if one of these two changes, the other one will change as well.

For instance, there *seems* to be a relationship between the amount of hours you study and your grades. There are more certain relationships, though:

- The force you need to apply to a coil to compress it by a distance grows as this distance grows;
- The time it takes for a pendulum to do a full oscillation grows with the square root of its length;
- You can convert mass into energy, and the more mass you have more energy you obtain;

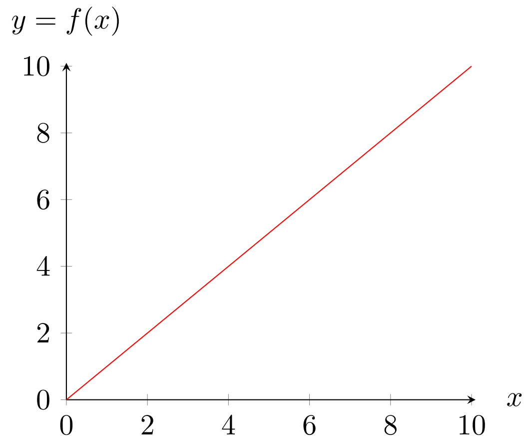
And many more examples exist.

21.3. Direct variation

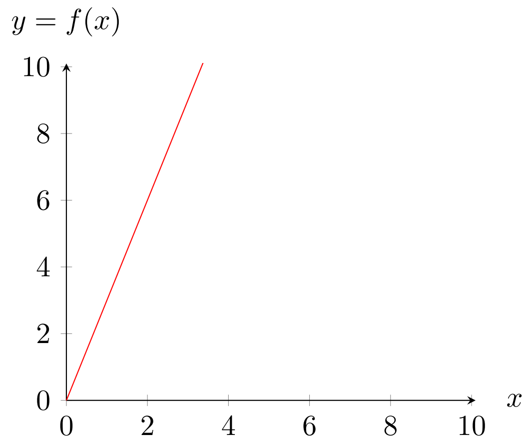
Direct variation means that, given two quantities, if one of them increases the other one also increases, and the converse also happens, if one them decreases, the other one decreases.

[1] Neil Gaiman is a genius.

A way to see this is by looking at some graphs:

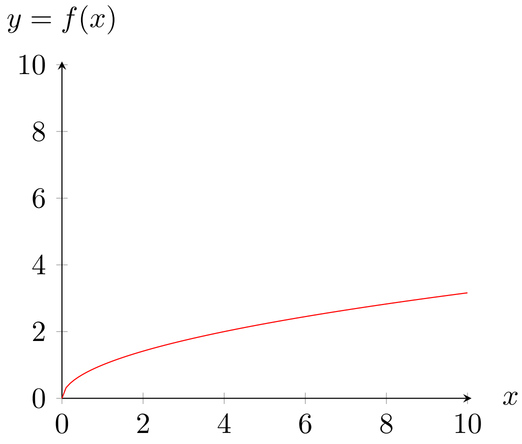


This one you already know: it's linear variation. This is the graph of $y = x$. Here, as x increases, y increases, and vice-versa. Not only that, as x increases by 1, y also increases by 1. We could have a different relationship between x and y:

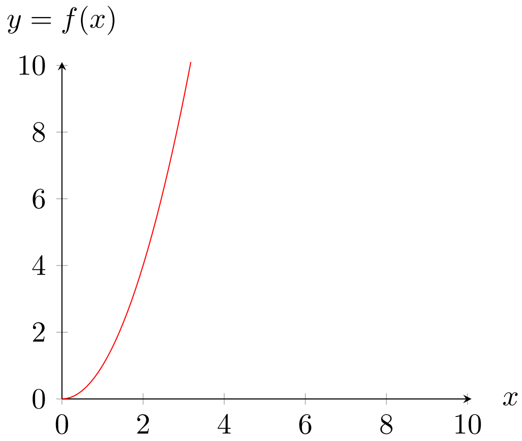


Here we have the graph of $y = 3x$. Again, as x increases y increases, and as x decreases y decreases. However, here as x increases by 1, y increases by 3.

We don't have only linear direct variation:

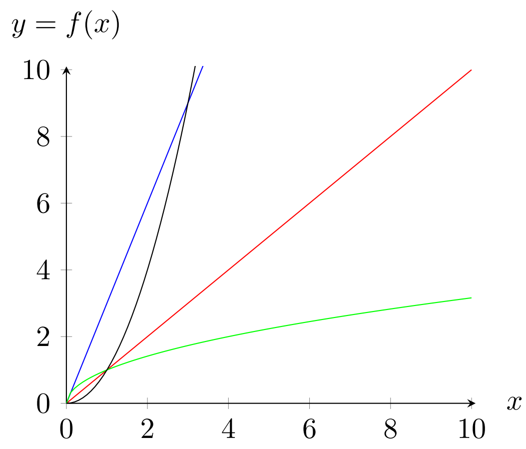


This is the graph of $y = \sqrt{x}$. We still have that as x increases, y increases, and as x decreases, y also decreases. But as x gets bigger, the growth in y gets smaller. We could also have something like:



This is $y = x^2$. Again, as x increases y increases, and as x decreases y decreases. As x gets bigger, though, the growth in y gets bigger.

It's important to notice that, even though the graphs change their shapes, they have similarities. To see that better, let's plot them all in the same axis:



The first thing to notice is that all of the graphs start at the point $(0, 0)$. Remember that we are dealing with variation of a quantity with another, so it there is 0 of one of them, it makes sense for the other to also be 0. The second thing is that all the functions we plotted ($y = x$, $y = 3x$, $y = \sqrt{x}$, $y = x^2$) are of the format

$$y = k \times \text{function of } x$$

in which k is a number and a function of x can be whatever function in x[2]. For instance, in $y = 3x$, we have $k = 3$ and the function of x to be just x. For $y = \sqrt{x}$ we have that $k = 1$ and the function of x is $\sqrt{x}$, and so on.

Thus, we have that, when a quantity y is *directly proportional* (or simply *proportional*) to another quantity x, they are related by an equation of the type:

$$y = k \times \text{function of } x$$

Sometimes, instead of saying that y is directly proportional to a certain function of x, they will give you symbols:

$$y \propto \text{function of } x$$

This is just notation, it means exactly the same: that y varies directly with a function of x.

What we need to figure out in this is the value of k and what kind of function of x we are dealing with. As you will see, in the IGCSE this is very easy!

[2] That's not true, the function has to be increasing, but we don't need to worry about it now.

Solved exercise: finding the equation that relates to proportional quantities

A quantity y is **directly proportional** to a quantity x. Given that when $x = 4$, $y = 2$, find the equation that relates y and x.

Solution: We know that y is directly proportional to x, they are related by an equation of the type

$$y = k \times \text{function of } x$$

Now, the function of x in this case is just x, as nothing else has been said, so we know that

$$y = k \times x$$

To discover k we use the information that when $x = 4$, the value of y is 2. Substituting these:

$$y = kx$$

$$2 = k \times 4$$

$$\frac{2}{4} = \frac{4k}{4}$$

$$\frac{1}{2} = k$$

And we are done! We have found that the formula that relates y and x is

$$y = \frac{1}{2}x$$

Solved exercise: finding the value of a quantity given the other - direct

A quantity y is **directly proportional** to the **square** of x. Given that when $x = 2$, $y = 24$, find the value of y when $x = 3$.

Solution: This is the "type" of exercise that IGCSE really likes. You know that if y is directly proportional to x they are related by an equation of the type

$$y = k \times \text{function of } x$$

The function of x here is x^2, as y is proportional to the *square* of x. This is the only care you have to take in these questions, the remainder is the same:

$$y = k \times x^2$$

Using the fact that when $x = 2$, $y = 24$:

$$24 = k \times 2^2$$

$$24 = k \times 4$$

$$\frac{24}{4} = \frac{4k}{4}$$

$$6 = k$$

The formula that relates y and x is, thus,

$$y = 6x^2$$

We want to find the value of y when x is 3. Substituting again:

$$y = 6x^2$$

$$y = 6 \times 3^2$$

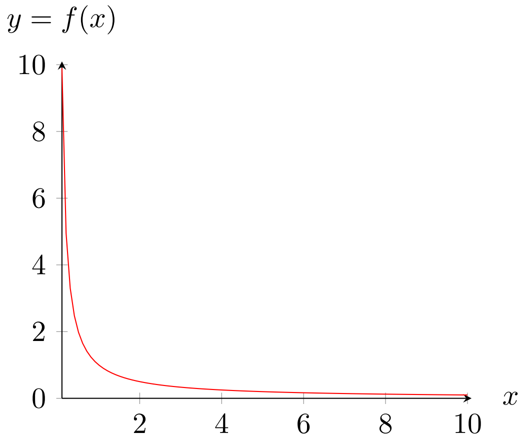
$$y = 6 \times 9$$

$$y = 54$$

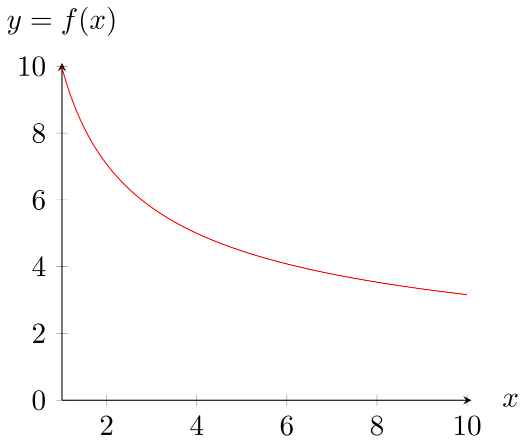
21.4. Inverse variation

Inverse (or indirect) variations is the opposite of direct variation. Given two quantities, if one of them increases, the other one decreases, and vice-versa.

Let's look at some graphs again.



This is the graph of $y = \frac{1}{x}$. Notice that as x increases, the value of y decreases. Also notice that as x gets larger, the change in the value y gets smaller.



This is $y = \frac{10}{\sqrt{x}}$, and again the same behaviour: as x increases, the value of y decreases. Again, as x increases, the y value decreases slower.

In all, whenever we have *inverse variation*, we are going to have an equation relating y and x that looks like

$$y = \frac{k}{\text{function of } x}$$

In the same way as with direct proportion, we can write this as

$$y \propto \frac{1}{\text{function of } x}$$

Solved exercise: finding the equation that relates to inversely proportional quantities

A quantity y is **inversely proportional** to a quantity x. Given that when $x = 2$, $y = 1$, find the equation that relates y and x.

Solution: We are going to use exactly the same idea as when the quantities are proportional, but changing the format of the equation:

$$y = \frac{k}{\text{function of } x}$$

In this case, the function of x is just x, so:

$$y = \frac{k}{x}$$

Substituting $y = 1$ and $x = 2$:

$$1 = \frac{k}{2}$$

$$k = 2$$

Thus, the formula that relates y and x is

$$y = \frac{2}{x}$$

Solved exercise: finding the value of a quantity given the other - inverse

A quantity y is **inversely proportional** to the $\sqrt{x+1}$. Given that when $x = 8$, $y = 3$, find the value of y when $x = 15$.

Solution: Inversely proportional means that

$$y = \frac{k}{\text{function of } x}$$

Very nice of them, in this one they give us the function, $\sqrt{x+1}$:

$$y = \frac{k}{\sqrt{x+1}}$$

Let's use $x = 8$ and $y = 3$:

$$3 = \frac{k}{\sqrt{8+1}}$$

$$3 = \frac{k}{\sqrt{9}}$$

$$3 = \frac{k}{3}$$

$$9 = k$$

So we know that

$$y = \frac{9}{\sqrt{x+1}}$$

Now, to find y when $x = 15$ we just substitute:

$$y = \frac{9}{\sqrt{15+1}}$$

$$y = \frac{9}{\sqrt{16}}$$

$$y = \frac{9}{4}$$

So $y = \frac{9}{4}$ when $x = 15$.

21.5. Exam hints

When these type of questions appear in your exams, search for **the type of variation**, that is, if it is direct or inverse, and what is the function you have to use. After that, set up the problem using the correct structure, substitute the values they give and you're done!

Summary

- *Variation* means that we have two quantities that change together;
- *Direct variation* means that as one quantity increases, the other one also increases, and vice-versa;
- We can denote that y is directly proportional to x using the symbols

$$y \propto x$$

- The "general" equation that relates y and x when are *directly proportional* is

$$y = k \times \text{function of } x$$

- *Inverse* (or *indirect*) variation means that as one quantity increases, the other one decreases, and vice-versa;
- We can denote that y is inversely proportional to x using the symbols

$$y \propto \frac{1}{x}$$

- The "general" equation that relates y and x when are *inversely proportional* is

$$y = \frac{k}{\text{function of } x}$$

22. Functions

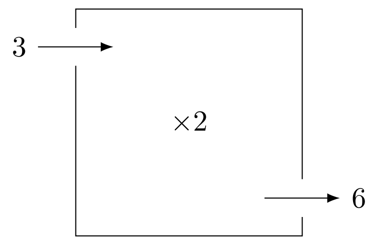
22.1. Why learn functions

The world is full of relationships: how things change depending on other things. There are personal relationships, or how other people affect us; there are physical relationships about 'things', such as how fast something falls depending on the height it is dropped.

Functions are special mathematical relationships. They are maths way to express them, predict them and model them. As you can imagine, given the abundance of relationships it is interesting for us to model and/or predict, you can gather that functions are fundamental. In all, functions are very important in science and engineering, but not only that, they are also one of the most important concepts in maths. Thus, functions are a great topic: not only they are useful in practice, but they are also very important in maths theory.

22.2. Informal definition, key terms and function application

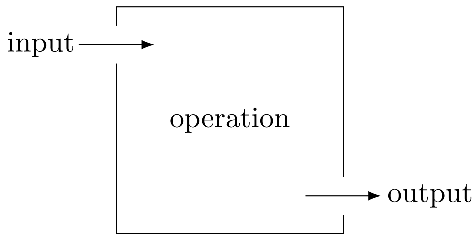
You have probably seen functions before, represented as 'function machines':



'Times 2 machine'

This means that the 'function machine' takes an input, in this example 3, and applies some operations to it, in this example multiplies it by 2, obtaining an output, in this case 6. So, this would be the '×2 function machine', as it is multiplying its input by 2 to obtain its output.

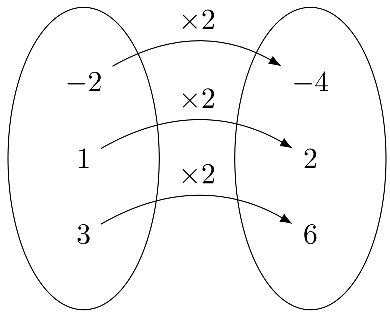
In general, we could draw something such as:



'Machine's name'

A function is simply that: a way to *map* input to output. It cannot be any kind of map, but that's the basic idea: a function is a map between input and output or, to put it differently, a *relation* between input and output.

Another pictorial way to represent functions is something as:



I really like this representation, as it makes clear that a function is nothing more than a *relation between* ***two sets*** of numbers: the input set and the output set. In 'mathematicianese', then, a function is a rule that, given two sets A and B, associates each element of A an element of B. We call the set A (the input of the function) the *domain* of the function, and we call the set of elements of B which were associated by the function its *range* (the output). In the IGCSE exam, both sets are usually the real numbers (every single number you know), so there's no much need for us to dwell in this.

You have to agree that drawing functions would be time consuming, so we have a more compact notation: a faster way to write the same thing.

For instance, the '×2 machine' could be written as:

$$f(x) = 2x$$

Here, we called the function f (very creative!). The input is a number x, and its output is taking this input x and multiplying it by 2, also known as $2x$:

$$\underbrace{f}_{\text{name}}(\overbrace{x}^{\text{input}}) = \underbrace{2x}_{\text{output}}$$

Now, to give the function f an input, such as 3, we write

$$f(3)$$

which is read as 'apply the function f to the value 3' or simply 'f of 3'. To do this, we substitute *every* x on the function output expression (I am going to refer to it as the 'function expression' from now on) by the value given, in this case 3:

$$f(x) = 2x$$

$$f(3) = 2 \times (3) = 6$$

You would get the same result without adding those brackets around 3, but I suggest adding them until you get used to this.

Let's see some more examples.

Take the function

$$g(x) = x^2$$

First, notice that we are not very creative when given names to functions: they are usually called f, g and h. Second, here we have a function called g, that takes as input a number x and gives as output x^2. Let's apply g to some values. To find $g(1)$, we substitute every x on the g expression by 1:

$$g(x) = x^2$$

$$g(1) = (1)^2 = 1$$

Now $g(2)$:

$$g(x) = x^2$$

$$g(2) = (2)^2 = 4$$

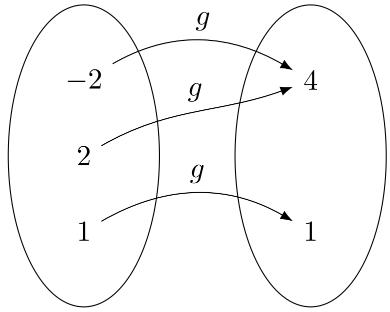
Finally, $g(-2)$:

$$g(x) = x^2$$

$$g(-2) = (-2)^2 = 4$$

This last example of application, $g(-2)$, shows that when you are applying the function to negative values, you **have to use brackets**. It is a classic mistake to forget this and calculate $g(-2) = -2^2$ and put this in your calculator, giving you as answer -4. Be very careful! Again, I suggest you to always add the brackets, at least until you get used to this operation.

Let's also represent the function g using our 'ball drawing':



It is more common for us to label the arrows that show the mapping with the function name, not with the operation (as the operation can be complex).

A more complex example. Given

$$h(x) = x^3 - 2x^2 - x + 3$$

let us calculate some values of h, starting with $h(0)$. I'll be adding a lot of brackets to help visualise:

$$h(x) = x^3 - 2x^2 - x + 3$$

$$h(0) = (0)^3 - 2(0)^2 - (0) + 3$$

$$h(0) = 0 - 2 \times 0 - 0 + 3 = 3$$

So we have that $h(0) = 3$. Let's try $h(1)$:

$$h(x) = x^3 - 2x^2 - x + 3$$

$$h(1) = (1)^3 - 2(1)^2 - (1) + 3$$

$$h(1) = 1 - 2 \times 1 - 1 + 3 = 1$$

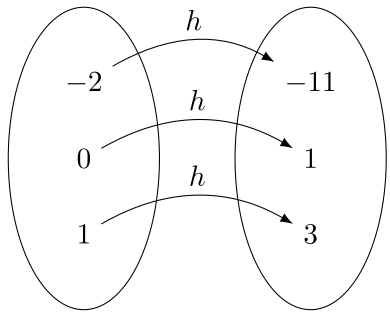
Finally, let's see an example where the brackets are mandatory, $h(-2)$:

$$h(x) = x^3 - 2x^2 - x + 3$$

$$h(-2) = (-2)^3 - 2(-2)^2 - (-2) + 3$$

$$h(-2) = -8 - 2 \times 4 + 2 + 3 = -11$$

Finally, let us represent these values of h we calculated using our 'circle drawings':



By the way, we can use any variable for functions:

$$f(y) = y + 3$$

This is *exactly* the same function as $f(x) = x + 3$, but we are using y instead of x for the input variable.

22.3. Function arithmetic

Let's see now how to do some basic function arithmetic.

To add or subtract functions, we simply *add or subtract their expressions.* For instance, given $f(x) = x^2 - 4$ and $g(x) = -2x + 3$, let's calculate $f(x) + g(x)$:

$$f(x) + g(x)$$

$$\underbrace{(x^2 - 4)}_{f \text{ expression}} + \underbrace{(-2x + 3)}_{g \text{ expression}} \qquad \text{Add } f \text{ and } g \text{ expressions}$$

$$x^2 - 4 - 2x + 3 \qquad \text{Expand the brackets}$$

$$x^2 - 2x - 1 \qquad \text{Collect like terms}$$

We have that $f(x) + g(x) = x^2 - 2x - 1$. Notice you don't need to add the brackets, but I recommend you do it until you get used to this. Now let's calculate $f(x) - g(x)$:

$$f(x) - g(x)$$

$$\underbrace{(x^2 - 4)}_{f \text{ expression}} - \underbrace{(-2x + 3)}_{g \text{ expression}} \qquad \text{Subtract } g \text{ expression from } f \text{ expression}$$

$$x^2 - 4 + 2x - 3 \qquad \text{Expand the brackets}$$

$$x^2 + 2x - 7$$

So $f(x) - g(x) = x^2 + 2x - 7$.

To multiply functions, we just need to multiply the corresponding expressions together, and be very careful with brackets. Let's say we have $f(x) = x-1$ and $g(x) = x^2-2x+3$. To calculate $f(x) \times g(x)$ we do

$$\begin{aligned}
&f(x) \times g(x) \\
&\underbrace{(x-1)}_{f \text{ expression}} \times \underbrace{(x^2-2x+3)}_{g \text{ expression}} && \text{Multiply } f \text{ and } g \text{ expression} \\
&x^3 - 2x^2 + 3x - x^2 + 2x - 3 && \text{Expand the brackets} \\
&x^2 - 3x^2 + 5x - 3 && \text{Collect like terms}
\end{aligned}$$

We can divide functions, and it would just be dividing the functions' expressions, but normally we would just leave them like that, as a fraction. You will learn later how to divide some special types of functions, such as polynomials.

22.4. Equations with functions

A very common type of question is something as: given that $f(x) = 2x - 1$, solve the equation $f(x) = 5$. To do this, you just need to *use the function expression whenever there is the function in the equation.* After it, you will obtain an equation, which you solve. In this case we have

$$\begin{aligned}
f(x) &= 5 \\
\underbrace{(2x-1)}_{f \text{ expression}} &= 5 && \text{Substitute } f(x) \text{ by its expression} \\
2x - 1 &= 5 && \text{Expand the brackets} \\
2x - 1 + 1 &= 5 + 1 && \text{Add 1 on both sides} \\
2x &= 6 && \text{Collect like terms} \\
\frac{2x}{2} &= \frac{6}{2} && \text{Divide both sides by 2} \\
x &= 3
\end{aligned}$$

There we have it.

Sometimes you'll obtain a quadratic. For instance, given that $g(x) = x^2 - 4x$ and

$h(x) = x - 6$, solve $g(x) = h(x)$. Again, substitute the expression and have fun

$$
\begin{aligned}
g(x) &= h(x) && \\
\underbrace{(x^2 - 4x)}_{g \text{ expression}} &= \underbrace{(x - 6)}_{h \text{ expression}} && \text{Substitute } g \text{ and } h \text{ by their expressions} \\
x^2 - 4x &= x - 6 && \text{Expand brackets} \\
x^2 - 4x - x + 6 &= x - x - 6 + 6 && \text{Adding 6 and subtracting } x \text{ from both sides} \\
x^2 - 5x + 6 &= 0 && \text{Collecting like terms} \\
(x - 2)(x - 3) &= 0 && \text{Factorising} \\
x - 2 = 0 \rightarrow x = 2 & && \\
x - 3 = 0 \rightarrow x = 3 & && \text{Solve each bracket} = 0
\end{aligned}
$$

Thus, we have that $x = 2$ or $x = 3$.
In summary, substitute the function name by its expression and work your magic.

Equations with functions

Given

$$f(x) = x^2, \quad g(x) = 2x^2 - 1, \quad h(x) = 3$$

Solve the equation

$$f(x) \times h(x) = g(x) + 2$$

Solution: Just substitute each function by its expression and go crazy on your equation solving:

$$
\begin{aligned}
f(x) \times h(x) &= g(x) + 2 \\
\underbrace{(x^2)}_{f\text{ expression}} \times \underbrace{(3)}_{h\text{ expression}} &= \underbrace{(2x^2 - 1)}_{g\text{ expression}} && \text{Substitute each function name by its expression} \\
3x^2 &= 2x^2 - 1 && \text{Expanding brackets and multiplying} \\
3x^2 - 2x^2 &= 2x^2 - 1 - 2x^2 && \text{Subtracting } 2x^2 \text{ on both sides} \\
x^2 &= 1 && \text{Collecting like terms} \\
\sqrt{x^2} &= \pm\sqrt{1} && \text{Square rooting both sides} \\
x &= \pm 1
\end{aligned}
$$

We have that $x = \pm 1$.

22.5. Composition of functions

22.5.1. Basic composition: only numbers

Let's say we have two functions

$$f(x) = x + 1 \text{ and } g(x) = x^2$$

and that we shall do the following: we are going to calculate $f(2)$, and then apply g to the result of $f(2)$.

Calculating $f(2)$ we have

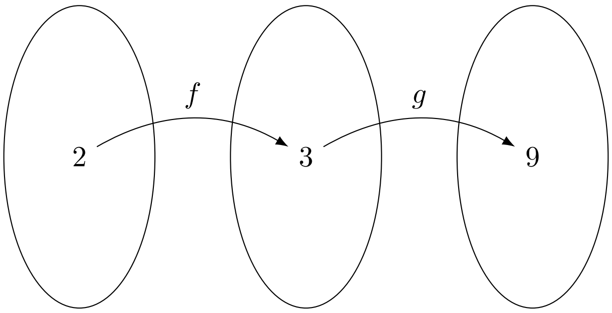
$$
\begin{aligned}
f(x) &= x + 1 \\
f(2) &= (2) + 1 = 3
\end{aligned}
$$

So we have $f(2) = 3$. Now let's take the result of $f(2)$, which is 3, and apply g to it: we'll calculate $g(3)$:

$$
\begin{aligned}
g(x) &= x^2 \\
g(3) &= (3)^2 = 9
\end{aligned}
$$

We have that $g(3) = 9$.

Thinking about our 'circle drawing' for functions, we could draw this situation like this:



Applying one function to the result of another is called *function composition*. What we have just calculated, applying g to the result of $f(3)$ is one example of composition.

The notation for function composition is the following. Let's say we want to apply the function f to the result of $g(2)$ now. We write this as

$$fg(2)$$

We read from right to left here: first we calculate $g(2)$ and then apply f to the result. I prefer this notation, which clearly shows you have to calculate $g(2)$ first:

$$f(g(2))$$

but it is not used in the IGCSE.

Now, remember that $f(x) = x + 1$ and $g(x) = x^2$ and let's calculate $fg(2)$. One way to set this up is like this:

$$g(x) = x^2$$
$$g(2) = (2)^2 = 4$$
$$f(4) = (4) + 1 = 5$$

We have, then, that $fg(2) = 5$. Another way is to add brackets:

$$\underbrace{f(\overbrace{g(2)}^{\text{Calculate this first}})}_{\text{Apply } f \text{ to the result}}$$

$$\begin{aligned} & f((2)^2) && \text{Calculating } g(2) \\ & f(4) && \text{Applying } f \text{ to the result of } g(2) \\ & 4+1 && \text{Calculating } f(4) \\ & 5 && \end{aligned}$$

We obtain the same result. Personally I prefer the second approach, as I think it is better suited for a more complex exercise we'll see in a second.

Basic composition of functions example I

Given the functions

$$f(x) = 2x - 1, \quad g(x) = -x^2$$

Calculate $ff(0)$, $fg(2)$, $gf(2)$, $fg(-1)$ and $gf(-1)$.

Solution: Let's start with $ff(0)$ and use the first set up above:

$$\begin{aligned} f(x) &= 2x - 1 \\ f(0) &= 2(0) - 1 = -1 \\ f(-1) &= 2(-1) - 1 = -2 - 1 = -3 \end{aligned}$$

Then we have it, $ff(0) = -3$. Just for us to see there is no problem on composing a function with itself.

Now $fg(2)$ and use the first set up again:

$$\begin{aligned} g(x) &= -x^2 \\ g(2) &= -(2)^2 = -4 \\ f(-4) &= 2(-4) - 1 = -8 - 1 = -9 \end{aligned}$$

We have $fg(2) = -9$. Now let's try the second set up for $gf(2)$:

$$\underbrace{g(\overbrace{f(2)}^{\text{Calculate this first}})}_{\text{Apply } g \text{ to the result}}$$

$g(2(2)-1) = g(3)$ — Calculating $f(2)$

$g(3)$ — Applying g to the result of $f(2)$

$-(3)^2$ — Calculating $g(3)$

-9

We obtained $gf(2) = -9$. This is a mere coincidence, and does not always happen. Now, let's calculate $fg(-1)$ using the first set up again:

$$g(x) = -x^2$$

$$g(-1) = -(-1)^2 = -(1) = -1$$

$$f(-1) = 2(-1) - 1 = -2 - 1 = -3$$

Thus, $fg(-1) = -3$. Finally, calculating $gf(-1)$:

$$\underbrace{g(\overbrace{f(-1)}^{\text{Calculate this first}})}_{\text{Apply } g \text{ to the result}}$$

$g(2(-1)-1) = g(-3)$ — Calculating $f(-1)$

$g(-3)$ — Applying g to the result of $f(-1)$

$-(-3)^2 = -(9)$ — Calculating $g(3)$

-9

So $gf(-1) = -9$, which is different from $fg(-1)$. You may use whatever method you prefer, as usual.

Basic composition of functions example II

Given

$$f(x) = 2x^2 \text{ and } g(x) = \frac{x}{2}$$

solve the equation

$$f(x) - x - x^2 = gf(1)$$

to **2 decimal places.**

Solution: Let's start calculating $gf(1)$:

$$\underbrace{g(\quad \overbrace{f(1)}^{\text{Calculate this first}} \quad)}_{\text{Apply } g \text{ to the result}}$$

$$g(2(1)^2) = g(2) \qquad \text{Calculating } f(2)$$

$$g(2) \qquad \text{Applying } g \text{ to the result of } f(2)$$

$$g(2) = \frac{2}{2} = 1 \qquad \text{Calculating } g(2)$$

$$1$$

We have $gf(1) = 1$. Let's now substitute the f expression in $f(x)$ and $gf(1) = 1$ and solve the equation using the quadratic formula (remember that whenever a question asks you to solve a quadratic equation to some decimal places is a hint to use the formula):

$$
\begin{aligned}
f(x) - x &= gf(1) \\
2x^2 - x - x^2 &= 1 && \text{Substituting } gf(1) \text{ and } f \text{ expression} \\
x^2 - x &= 1 && \text{Collecting like terms} \\
x^2 - x - 1 &= 1 - 1 && \text{Subtracting 1 on both sides} \\
x^2 - x - 1 &= 0
\end{aligned}
$$

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \to x = \frac{- -1 \pm \sqrt{(-1)^2 - 4 \times 1 \times -1}}{2 \times 1} \qquad a = 1, b = -1, c = -1$$

$$x = \frac{1 \pm \sqrt{1+4}}{2} = \frac{1 \pm \sqrt{5}}{2}$$

$$x = \frac{1 \pm \sqrt{5}}{2} \begin{cases} \xrightarrow{+} x = \frac{1+\sqrt{5}}{2} = 1.62 \\ \xrightarrow{-} x = \frac{1-\sqrt{5}}{2} = -0.62 \end{cases}$$

Thus, $x = 1.62$ or $x = -0.62$.

22.5.2. Algebraic composition of functions

We have seen how to calculate composition of functions given a number on the previous section. It is important to notice, though, that given two functions, their composition is *also a function*. Being it a function, we can find its expression, and that's what we are doing now.

Let's say we have

$$f(x) = 2x - 1 \text{ and } g(x) = x + 3$$

and we want to find $fg(x)$ and $gf(x)$. We don't have any particular number to substitute, we just want to find an expression for two new functions, the compositions.

To do this, we do exactly the same as before: we substitute. However, we now are going to substitute *the whole expression* of the function to the right. Remember that $fg(x)$ is the same as doing $f(g(x))$, and we know how to apply numbers to functions: we substitute each x on the function for it, right? We now substitute the *whole function expression* in each variable:

$$f(x) = 2x - 1$$

$$fg(x) = f(g(x)) = 2 \times (g(x)) - 1 \qquad \text{Each } x \text{ is substituted by } g(x)$$

$$f(x+3) = 2 \times (x+3) - 1 \qquad g(x) = x + 3$$

$$f(x+3) = 2x + 6 - 1 \qquad \text{Expanding the brackets}$$

$$f(x+3) = 2x + 5 \qquad \text{Collecting like terms}$$

We have, then, that $fg(x) = 2x + 5$. Notice that this is a new function, not a number. Let's now calculate $gf(x)$:

$$g(x) = x + 3$$

$$gf(x) = g(f(x)) = (f(x)) + 3 \qquad \text{Each } x \text{ is substituted by } f(x)$$

$$g(2x-1) = (2x-1) + 3 \qquad f(x) = 2x - 1$$

$$g(2x-1) = 2x - 1 + 3 \qquad \text{Expanding the brackets}$$

$$g(2x-1) = 2x + 2 \qquad \text{Collecting like terms}$$

Again, notice that $gf(x)$ is a new function.

Algebraic composition of functions example I

Given $f(x) = x^2$ and $g(x) = 2x - 1$, find $fg(x)$ and $gf(x)$, simplifying your answers fully.

Solution: To calculate $fg(x)$, we need to substitute every x in the expression of f by the *whole expression* of g:

$$f(x) = x^2$$

$$fg(x) = f(g(x)) = (g(x))^2 \qquad \text{Each } x \text{ is substituted by } g(x)$$

$$f(2x-1) = (2x-1)^2 \qquad g(x) = 2x - 1$$

$$f(2x-1) = (2x-1)(2x-1) = 4x^2 - 2x - 2x + 1 \qquad \text{Expanding the brackets}$$

$$f(2x-1) = 4x^2 - 4x + 1 \qquad \text{Collecting like terms}$$

Now, to find $gf(x)$, we have to substitute every x in the expression of g by the *whole expression* of f:

$$g(x) = 2x - 1$$

$$gf(x) = g(f(x)) = 2(f(x)) - 1 \quad \text{Each } x \text{ is substituted by } f(x)$$

$$g(x^2) = 2(x^2) - 1 \quad f(x) = x^2$$

$$g(x^2) = 2x^2 - 1 \quad \text{Expanding the brackets}$$

Algebraic composition of functions example II

Given

$$f(x) = 2x - 1 \text{ and } g(x) = x^2$$

solve the equation

$$gf(x) + 1 = 2g(x) + x$$

Solution: Let's start by calculating $gf(x)$. Remember that to do this, we substitute *every* x in the expression of g by the *whole expression* of f:

$$g(x) = x^2$$

$$gf(x) = g(f(x)) = (f(x))^2 \quad \text{Each } x \text{ is substituted by } f(x)$$

$$g(2x - 1) = (2x - 1)^2 \quad f(x) = 2x - 1$$

$$g(2x - 1) = (2x - 1)(2x - 1) = 4x^2 - 2x - 2x + 1 \quad \text{Expanding the brackets}$$

$$g(2x - 1) = 4x^2 - 4x + 1 \quad \text{Collecting like terms}$$

Thus, we have that $gf(x) = 4x^2 - 4x + 1$. We can now substitute every thing we know in the equation we need to solve:

$$
\begin{aligned}
gf(x) - 1 &= 2g(x) + x \\
4x^2 - 4x + 1 + 1 &= 2(x^2) + x && \text{Substituting } gf(x) \text{ and } g(x) \\
4x^2 - 4x + 2 &= 2x^2 + x && \text{Collecting like terms on the LHS} \\
4x^2 - 4x + 2 - 2x^2 - x &= 2x^2 + x - 2x^2 - x && \text{Subtracting } 2x^2 \text{ and } x \text{ on both sides} \\
2x^2 - 5x + 2 &= 0 && \text{Collecting like terms}
\end{aligned}
$$

We have reached this lovely quadratic equation. Let's solve it using the grouping method. Multiplying the x^2 coefficient (a) by the free variable (c), we have $2 \times 2 = 4$. Now, we are searching for two numbers, p and q, that

$$
\begin{aligned}
p + q &= -5 \\
pq &= 4
\end{aligned}
$$

The numbers are $p = -1$ and $q = -4$. We can now split the $-5x$ into $-1x - 4x$:

$$
\begin{aligned}
2x^2 - 5x + 2 &= 0 \\
2x^2 \underbrace{-1x - 4x}_{-5x} + 2 &= 0 && \text{Splitting } -5x \text{ into } -1x - 4x \\
\underbrace{2x^2 - x} - \underbrace{4x + 2} &= 0 && \text{Factorising each pair} \\
x(2x - 1) - 2(2x - 1) &= 0 && \text{Remember to copy the sign between each pair} \\
x\underbrace{(2x - 1)} - 2\underbrace{(2x - 1)} &= 0 && \text{Factorising again} \\
(2x - 1)(x - 2) &= 0 \\
2x - 1 = 0 \rightarrow x &= \tfrac{1}{2} \\
x - 2 = 0 \rightarrow x &= 2 && \text{Solving each bracket} = 0
\end{aligned}
$$

There we have it, $x = 2$ or $x = \frac{1}{2}$.

22.6. Inverse of a function

22.6.1. Informal definition and an important property

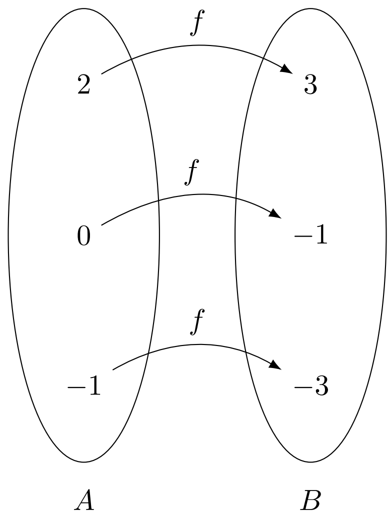
Take the function $f(x) = 2x - 1$. We can calculate some values of f:

$$f(0) = 2 \times 0 - 1 = -1$$

$$f(-1) = 2 \times -1 - 1 = -3$$

$$f(2) = 2 \times 2 - 1 = 3$$

and we can represent this using our 'circle drawing', but with the extra detail of calling the domain set A and the range set B:



This representation shows that the function f takes elements from the set A and maps (associates) each of them with one element of the set B.

There is a 'friend' function to f, called the *inverse of* f, which goes the opposite way: the inverse of f associates each element of the set B with an element of set A. We denote the inverse of function f by f^{-1} (notice that the inverse of a function g would be g^{-1}, the inverse of h would be h^{-1}, and so on). Thus, if we have f associating elements from set A to set B:

$$f : A \to B$$

we have that the inverse of f, f^{-1}, associates elements from set B to set A:

$$f^{-1} : B \to A$$

Now, how does the inverse associates elements of B to set A? What is the rule? It is simple. Let's say we have a number a we can apply f , that is, calculating $f(a)$ and we obtain a number b:

$$f(a) = b$$

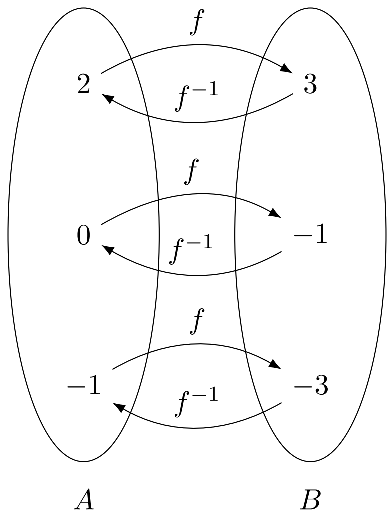
Then, the inverse of f does this

$$f^{-1}(b) = a$$

For instance, in our $f(x) = 2x - 1$, we have $f(0) = -1$. Thus, if we calculate $f^{-1}(-1)$ we obtain 0:

$$f(0) = -1 \text{ implies } f^{-1}(-1) = 0$$

Basically, the inverse of f takes a number back, which is easily seen on our drawing:



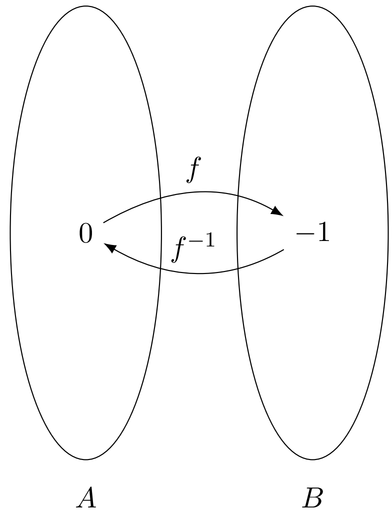
Using this representation is easy to see that

$$f(2) = 3 \text{ then } f^{-1}(3) = 2$$

$$f(0) = -1 \text{ then } f^{-1}(-1) = 0$$

$$f(-1) = -3 \text{ then } f^{-1}(-3) = -1$$

So, remember: the inverse of a function does the 'inverse operation' of the original. Now, let's use our representation of $f(x) = 2x - 1$ focusing on a single input:



What happens if we calculate $f(0)$ and apply f^{-1} to the result of $f(0)$? That is, what is the result of calculating

$$f^{-1}f(0)$$

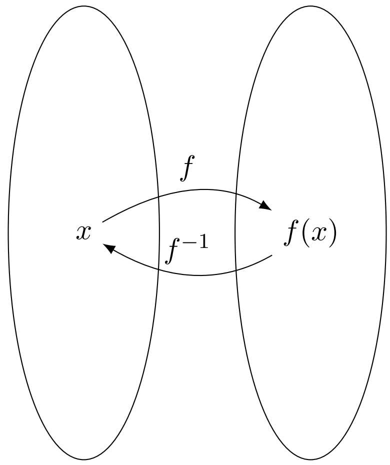
or, in words, what is the result of doing the composition of $f(0)$ with its inverse? We know that $f(0) = -1$, and when we calculate $f^{-1}(-1)$ we obtain 0. Thus, we started with 0, applied f to it, obtained -1, then applied f^{-1} to -1 and obtained 0 again:

$$0 \xrightarrow{f(0)} -1 \xrightarrow{f^{-1}(-1)} 0$$

We started with 0 and went back to 0! Or, using our composition notation:

$$f^{-1}f(0) = 0$$

Notice that this would happen to any other starting input we wanted: if we do the composition of a function f with its inverse, f^{-1}, you always go back to your original number, as f^{-1} 'cancels out' what f does! We can represent this using our picture notation:



which we can write

$$f^{-1}f(x) = x$$

The same thing would happen if you had started with an element on set B, applied the inverse of f to it, then applied f to the result, you would go back to your initial element (you would basically start at the 'right part' of the drawing). Thus, we also have

$$ff^{-1}(x) = x$$

This is a very important result

$$\boxed{ff^{-1}(x) = f^{-1}f(x) = x}$$

that you need to keep on your mind.

22.6.2. Finding the inverse

We are now going to focus on how, given a function, find its inverse expression.

Say we are given a function

$$g(x) = 3x - 2$$

and we would like to find the expression of $g^{-1}(x)$. To do this, we just have to understand that the function g takes an x as input and outputs a value $g(x)$. That is, the function g is taking a number x and changing it somehow. The inverse of g would do the opposite: it would take a value the changed value (that is $g(x)$) and undo what it was done to it, obtaining the original x. Because of that, if we rearranged the equation $g(x) = 3x - 2$ and made x the subject, we would find the inverse of g! Let's make x the subject in $g(x)$:

$$\begin{aligned}
g(x) &= 3x - 2 \\
g(x) + 2 &= 3x - 2 + 2 && \text{Adding 2 on both sides} \\
g(x) + 2 &= 3x && \text{Collecting like terms} \\
\frac{g(x)+2}{3} &= \frac{3x}{3} && \text{Dividing both sides by 3} \\
x &= \frac{g(x)+2}{3} && \text{Swapping sides}
\end{aligned}$$

Although this may seem weird, this expression is merely undoing what g did to x. Remember that $g(x) = 3x - 2$, that is, it multiplies x by 3 and then subtracts 2. This new expression we found does the inverse: it add 2 to the result and divides it by 3, obtaining what we started with.

Looking at our expression, though, it does not look like a function. We were looking for the inverse of g that is, $g^{-1}(x)$, but found an expression for x. Well, we have seen above this expression for x *is* the inverse of g, so we can just rename stuff:

$$\begin{aligned}
x &= \frac{g(x)+2}{3} && \text{Original expression} \\
g^{-1}(x) &= \frac{x+2}{3} && \text{Making } x = g^{-1}(x) \text{ and } g(x) = x
\end{aligned}$$

Much better, we have an expression for $g^{-1}(x)$:

$$g^{-1}(x) = \frac{x+2}{3}$$

We can now check if we are correct by remembering that if we calculate $gg^{-1}(x)$ we have to obtain x:

$$gg^{-1}(x) = g^{-1}g(x) = x$$

Let's try:

$$\begin{aligned}
g(x) &= 3x - 2 \\
gg^{-1}(x) = g(g^{-1}(x)) &= 3(g^{-1}(x)) - 2 && \text{Each } x \text{ is substituted by } g^{-1}(x) \\
g\left(\tfrac{x+2}{3}\right) &= 3\left(\tfrac{x+2}{3}\right) - 2 && g^{-1}(x) = \frac{x+2}{3} \\
g\left(\frac{x+2}{3}\right) = \cancel{3}\left(\frac{x+2}{\cancel{3}}\right) - 2 &= x + 2 - 2 && \text{Expanding the brackets} \\
g\left(\frac{x+2}{3}\right) &= x && \text{Collecting like terms}
\end{aligned}$$

We have obtained x, so we are correct.
Thus, we have now an algorithm to find the inverse of any function f:

1. Identify the variable which is representing the input of the function, usually x;
2. Make x the subject in the $f(x) =$ expression equation;
3. Rename x to $f^{-1}(x)$ and $f(x)$ to x.

Let's see some examples.

Finding the inverse of a function example I: simple function

Find the inverse of $f(x) = x - 4$.

Solution: Just use the algorithm:

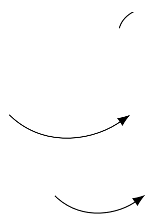
$$\begin{aligned} f(x) &= x - 4 && x \text{ is the variable for input} \\ f(x) + 4 &= x - 4 + 4 && \text{Make } x \text{ the subject by adding 4 on both sides} \\ f(x) + 4 &= x \\ x &= f(x) + 4 && \text{Swapping sides} \\ f^{-1}(x) &= x + 4 && \text{Make } x = f^{-1}(x) \text{ and } f(x) = x \end{aligned}$$

And we are done, the inverse of $f(x) = x - 4$ is $f^{-1}(x) = x + 4$. It does make sense: function f subtracts 4 from x, the inverse of f adds 4 to x.

Finding the inverse of a function example II: rational function

Find the inverse of

$$g(x) = \frac{2x - 3}{x + 1}$$



Solution: Algorithm:

$$g(x) = \frac{2x-3}{x+1} \quad \text{Make } x \text{ the subject}$$

$$g(x) \times (x+1) = \frac{2x-3}{x+1} \times x+1 \quad \text{Multiply both sides by } x+1$$

$$g(x)(x+1) = \frac{2x-3}{\cancel{x+1}} \times \cancel{x+1}$$

$$g(x)x + g(x) = 2x - 3 \quad \text{Expand the brackets on the LHS}$$

$$g(x)x - 2x + g(x) = 2x - 2x - 3 \quad \text{Subtract } 2x \text{ on both sides}$$

$$g(x)x - 2x + g(x) = -3 \quad \text{Collecting like terms}$$

$$g(x)x - 2x + g(x) - g(x) = -3 - g(x) \quad \text{Subtract } g(x) \text{ on both sides}$$

$$g(x)x - 2x = -3 - g(x) \quad \text{Collecting like terms}$$

$$x(g(x) - 2) = -3 - g(x) \quad x \text{ as a common factor on the LHS}$$

$$\frac{x(g(x)-2)}{g(x)-2} = \frac{-3-g(x)}{g(x)-2} \quad \text{Divide both sides by } g(x) - 2$$

$$\frac{x\cancel{(g(x)-2)}}{\cancel{g(x)-2}} = \frac{-3-g(x)}{g(x)-2}$$

$$x = \frac{-3-g(x)}{g(x)-2}$$

$$g^{-1}(x) = \frac{-3-x}{x-2} \quad \text{Make } x = g^{-1}(x) \text{ and } g(x) = x$$

Finding the inverse of a function example III: quadratic function

Find the inverses of $h(x) = x^2 - 4x + 6$.

Solution: We follow the algorithm, but now we need to complete the square:

$$
\begin{aligned}
h(x) &= x^2 - 4x + 6 && \text{We have to make } x \text{ the subject} \\
h(x) &= (x-2)^2 - 2^2 + 6 && \text{Completing the square} \\
h(x) &= (x-2)^2 - 4 + 6 \\
h(x) &= (x-2)^2 + 2 \\
h(x) - 2 &= (x-2)^2 + 2 - 2 && \text{Subtract 2 on both sides} \\
h(x) - 2 &= (x-2)^2 \\
\pm\sqrt{h(x)-2} &= \sqrt{(x-2)^2} && \text{Square root both sides} \\
\pm\sqrt{h(x)-2} &= x - 2 \\
x - 2 + 2 &= 2 \pm \sqrt{h(x)-2} && \text{Add 2 on both sides} \\
x &= 2 \pm \sqrt{h(x)-2} \\
h^{-1}(x) &= 2 \pm \sqrt{x-2} && \text{Make } x = h^{-1}(x) \text{ and } h(x) = x
\end{aligned}
$$

Notice that we have two inverses now, the positive square root one and the negative. Both answers would be correct if nothing else is said, and we do not need to go into further detail now.

22.6.3. The 'swap x with y' algorithm to find inverses

Some people prefer to find the inverse of a function using some extra steps. I'll give you one example. Say we have $f(x) = 3x - 2$ and we want to find $f^{-1}(x)$:

$$
\begin{aligned}
f(x) &= 3x - 2 \\
y &= 3x - 2 && \text{Subsitute } f(x) \text{ by } y \\
x &= 3y - 2 && \text{Make every } x \text{ become } y \text{ and every } y \text{ become } x \\
x + 2 &= 3y - 2 + 2 && \text{Make } y \text{ the subject again} \\
\frac{x+2}{3} &= \frac{3y}{3} \\
y &= \frac{x+2}{3} \\
f^{-1}(x) &= \frac{x+2}{3} && \text{Make } y = f^{-1}(x)
\end{aligned}
$$

Obviously, this gives the same answer, as we are doing exactly the same we did above, but with different variables. use whichever one you prefer.

In all, this algorithm to find the inverse of a function f is:

1. Substitute $f(x)$ by y;
2. Swap y by x, that is, every y becomes x and every x becomes y;
3. Solve for y;
4. Substitute y by $f^{-1}(x)$.

22.7. Exam hints

To be very honest, functions questions are very algorithmic: you just need to be very familiar with the methods of finding inverse and composition of functions, and solving equations.

Summary

- A *function* is a special relationship which associates elements of an input set, the *domain*, with elements of an output set, the *codomain;*
- We write functions as

  $$f(x) = x + 3$$

  This is function f, the input variable is x, and its expression is $x + 3$;

- To apply a function to a value we substitute every single input variable in its expression by the value, eg:

$$h(x) = x^3 - 2x^2 - x + 3$$

$$h(0) = (0)^3 - 2(0)^2 - (0) + 3$$

$$h(0) = 0 - 2 \times 0 - 0 + 3 = 3$$

- To add, subtract or multiply functions we simple add, subtract or multiply their expressions;
- To solve equations with functions we *substitute* each function name by its expression and solve the resulting equation;
- To *compose* functions is the same as applying one function to the result of the other;
- The *inverse* of a function *undoes* what the original function did;
- The algorithm to find the inverse of a function f is:
  1. Identify the variable which is representing the input of the function, usually x;
  2. Make x the subject in the $f(x) =$ expression equation;
  3. Rename x to $f^{-1}(x)$ and $f(x)$ to x.

  If you prefer the 'swap y by x' algorithm, there are its steps:

  In all, this algorithm to find the inverse of a function f is:
  1. Substitute $f(x)$ by y;
  2. Swap y by x, that is, every y becomes x and every x becomes y;
  3. Solve for y;
  4. Substitute y by $f^{-1}(x)$.

Formality after taste

I will basically follow the discussion in Section 1.3 of the great *Calculus*, vol. 1, by Tom Apostol.

First, let's define an *ordered pair*.

Definition. An **ordered pair** is a pair of objects denoted by (a, b) in which a is the *first element* or *member* of the pair and b is the *second element* or *member* of the pair.

You have seen many ordered pairs already when you plot points on the coordinate plane. There, we denote each pair by (x, y), where x is the coordinate of the point on the x-axis and y the coordinate on the y-axis. Thus, you know that order is important: the point $(1, 2)$ is not the same as the point $(2, 1)$.

Let's define ordered pair equality now:

Definition. Two ordered pairs, (a, b) and (c, d), will be said **equal** if and only if we have

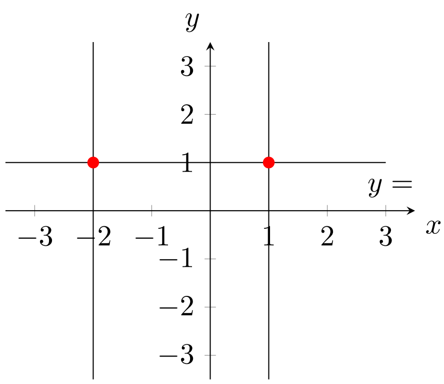
$$a = c \text{ and } b = d$$

Finally we can define a function.

Definition. "A **function** f is a set of ordered pairs (x, y) no two of which have the same first member."[1]

We also have to be very precise what is the set the x values come from, the domain of the f, and what is the set the y values come from, which is called the codomain of f. Notice the difference between codomain and range: range is the set of values that are actually calculated by the function, whereas the codomain can have 'extra elements' which were not mapped by the function.

This definition is very easily visualised by the 'vertical line test' on a graph. If we plot all[2] the (x, y) pairs of a function in the coordinate plane, we can draw vertical lines wherever we want it and those lines will never intersect the function graph more than once (notice that not intersecting is fine). For example, if we plot the function $y = 1$:

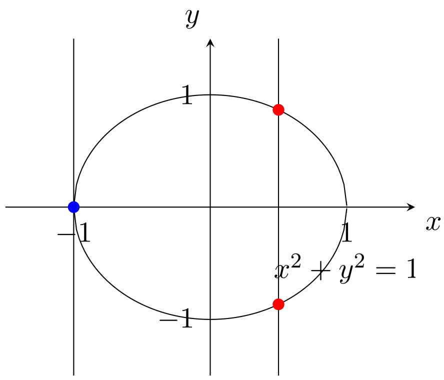


Notice that wherever we were to draw a vertical line (of the form $x = k$), the line would only intersect the line $y = 1$ once.

An example of something which is not a function is the relation $x^2 + y^2 = 1$, which gives this graph:

[1]In the words of our great Apostol.

[2]Plotting 'all' is impossible, we plot a sample. Let's assume 'all' here means 'enough'.



The line $x = -1$ intersects the graph of $x^2 + y^2 = 1$ once, at $(-1, 0)$ (the blue point). However, the line $x = 0.5$ intersects the graph at **two** points. If that happens *once* is enough: the relation is not a function.

Part III.

Functions and graphs

23. The coordinate plane and the basics of coordinate geometry

23.1. Why learn coordinate geometry

> "Cogito, ergo sum"

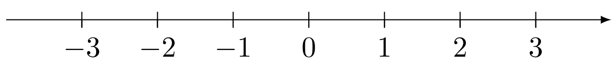
is a famous quote by René Descartes, a French philosopher. In English, it is commonly translated to "I think, therefore I am", and it arises in the process of Cartesian doubt, which I highly recommend you to research (his book *Discourse on Method* is a very fun read).

Descartes was not only a great philosopher, though. He was an outstanding mathematician, and he is one of the inventors of what we now call the *Cartesian coordinate plane*, which we will see now. This development allowed us to work on geometry problems from an algebraic standpoint. Some problems, as we will see, are very easy to solve using the coordinate plane, and they are reasonably complicated from a pure geometrical view.

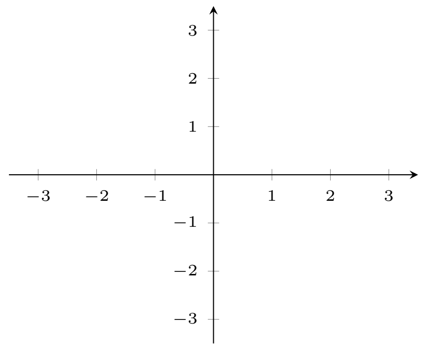
The coordinate plane and the geometry in it is one of the most important ideas in maths, and it is very fun to learn as well.

23.2. The x and y axes and the coordinate plane

The basic problem is: we have the plane, on its infinitude. We want to be able to refer to positions on the plane. What do we do? We think on simpler cases. If we only had one dimension, or a line, we can place points on it by numbering the line:

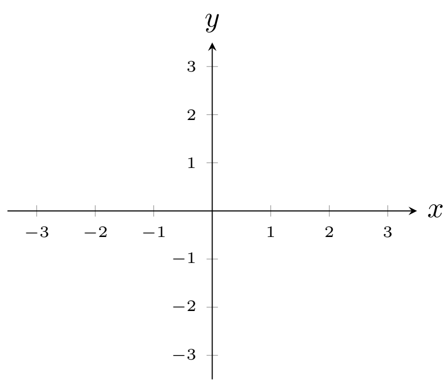


and I am sure you are familiar with the number line. The plane, however, has an extra dimension. So, let us repeat our idea: we can add another number line, but vertical:

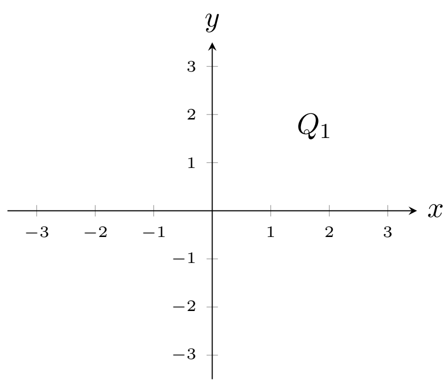


and the brilliance of it is that we just do exactly the same thing for each number line: a point is defined, uniquely, by its position on the horizontal number line **and** by its position on the vertical number line. We need **both**.

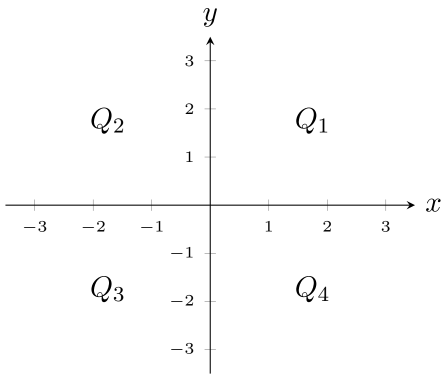
Let us give names to the number lines. The horizontal line is called the x axis, and the vertical line is called the y axis. It is easy to remember without any silly gimmick: the horizontal number line is the one you learned first, and x comes first in the alphabet. Thus, we have two axes, the x and y, which we label to look nice:



As you can see, with the x and y axes we divide the whole plane into 4 regions, each called a *quadrant*. The region where both the values on both x and y are positive is called the *first quadrant*, which we denote as Q_1:



The second quadrant Q_2 is the one to the left of the first, and we proceed counter-clockwise, with the third and fourth quadrants, Q_3 and Q_4:

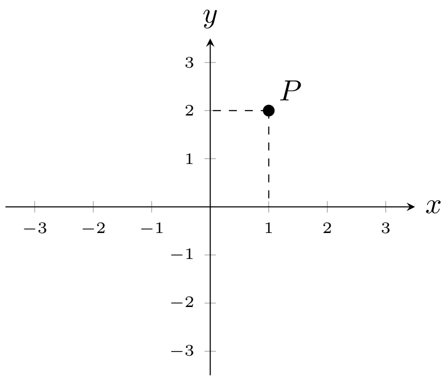


23.3. Points on the coordinate plane

Now that we know the x and y axes and their division of the plane, let us use them to actually do what we wanted: put points on it.

As we saw above, any point is uniquely defined by its position on the horizontal number line, the x axis, and on the vertical number line, the y axis. So, any point needs two numbers to be placed: its *x coordinate*, where it is on the horizontal number line, and its *y coordinate*, where it is on the vertical number line.

Let us place a point there to see how that works. I will call it P:



Focus first on the horizontal number line, the x axis: the position of P on it is 1. Hence, we say that P has x coordinate 1. On the y axis (the vertical one) P has value 2. Hence, its y coordinate is 2. To denote that, we use a pair of brackets:

$$\left(\underbrace{1}_{x \text{ coordinate}}, \underbrace{2}_{y \text{ coordinate}} \right)$$

It is easy to remember that x comes first: it always comes first! It was the "original" number line, so it has priority.

In summary, any point can be uniquely identified by its coordinates x and y:

$$\left(\underbrace{x}_{\substack{\text{horizontal}\\ \text{position}}} , \underbrace{y}_{\substack{\text{vertical}\\ \text{position}}} \right)$$

A very special point is the point $(0, 0)$, which we call the *origin.* It is special because it is where the axes intersect at a right angle.

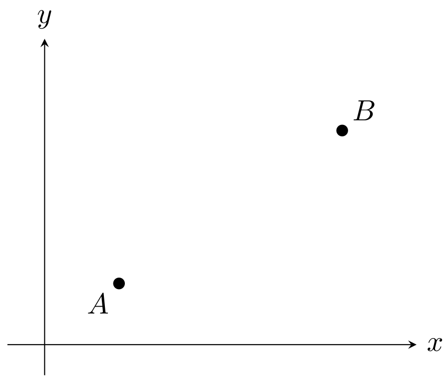
23.4. Midpoint

Now we would like to, given two points in the coordinate plane, to find their *midpoint*, a point which is exactly on the middle of the line connecting our two given points.

I will draw two random points on the first quadrant to reach the formula, but the reasoning would be same for any two points. Let us call the points A and B and give them the coordinates:

$$A = (x_1, y_1)$$

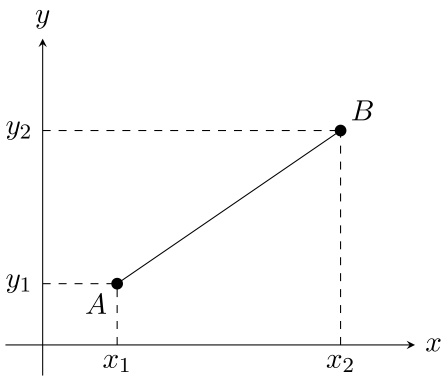
$$B = (x_2, y_2)$$



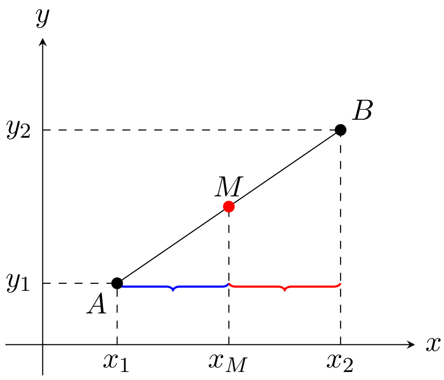
Our goal, again, is to find a point which is exactly on the middle of the line AB, the midpoint. We shall call it M. To find a point implies finding its coordinates, so we need to find the x and y coordinates of M.

To do that, we will break our problem into two parts: we will first find the x coordinate of M, and then its y coordinate. This is a very common strategy when solving problems in coordinate geometry, as the x and y components are independent!

We first start by joining points AB with a straight line, and I will indicate A and B coordinates on the axes:



and now let us think only on the x axis. The midpoint of two points must be exactly on the middle of their x distance:



We want, then, for the blue and red braces to have the same size.

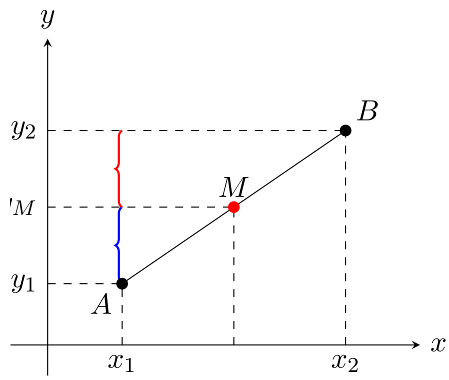
It is much easier to find the midpoint of a horizontal line, as they have the same y coordinates (the same height on the plane). If you were told to find the number right in the middle of 6 and 10, you would add them and divide the result by 2:

$$\frac{6+10}{2} = \frac{16}{2} = 8$$

and that is exactly the same with our x coordinates: we add them and divide the result by 2, which will give the x coordinate of M, x_M

$$x_M = \frac{x_1 + x_2}{2}$$

We now repeat the same idea with the y coordinates. The y coordinate of M, y_M, must be halfway between y_1 and y_2:



and again we would like the red and blue braces to have the same length. To find the midpoint of y_1 and y_2 we can again add and divide their coordinates by 2:

$$y_M = \frac{y_1 + y_2}{2}$$

Which gives us the little formula for the midpoint when we combine the results for x_M and y_M:

Midpoint formula

Given two points (x_1, y_1) and (x_2, y_2), the *midpoint* between them has coordinates

$$\left(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2}\right)$$

Let us say we have the points

$$(2, 3)$$

$$(4, 5)$$

and we want to find their midpoint. The formula just tells us to add and divide each pair of coordinates by 2, so:

$$\left(\underbrace{\frac{2+4}{2}}_{x \text{ coords}}, \underbrace{\frac{3+5}{2}}_{y \text{ coords}}\right) = \left(\frac{6}{2}, \frac{8}{2}\right) = (3, 4)$$

Therefore, the midpoint between $(2, 3)$ and $(4, 5)$ is the point $(3, 4)$.
Another example: say we have the points

$$(-3, -1)$$

$$(-5, -7)$$

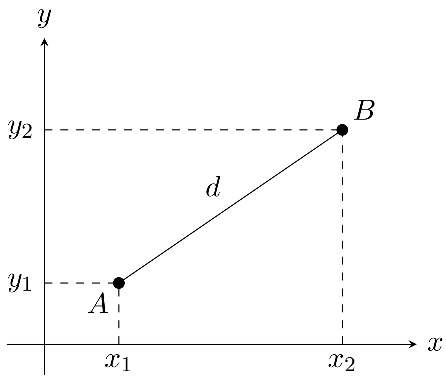
To find their midpoint, we again add each pair of coordinates and divide the result by 2:

$$\left(\underbrace{\frac{-3+-5}{2}}_{x \text{ coords}}, \underbrace{\frac{-1+-7}{2}}_{y \text{ coords}}\right) = \left(\frac{-3-5}{2}, \frac{-1-7}{2}\right) = \left(\frac{-8}{2}, \frac{-8}{2}\right) = (-4, -4)$$

It is important to notice that it does not matter which point you choose to be point 1 (the one that comes first in the formula) as we are just doing addition.

23.5. Distance between two points

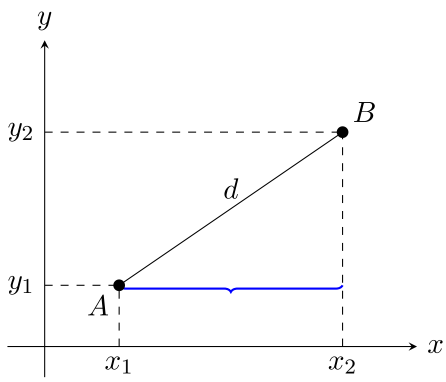
Same setup, different problem: given two points, we want to find their *distance*, or equivalently, the *length of the line which connects them.* In drawing:



in which I called the distance d.

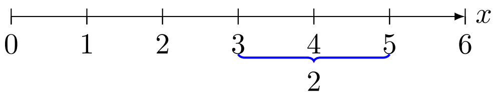
We will again think first on the x axis and then on the y axis.

Let us start by finding the distance between the x coordinates of A and B:

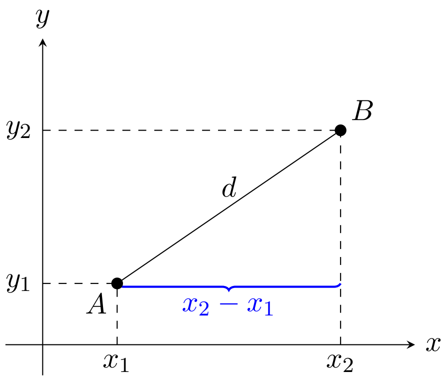


which again is much simpler, as we are only thinking about one dimension.

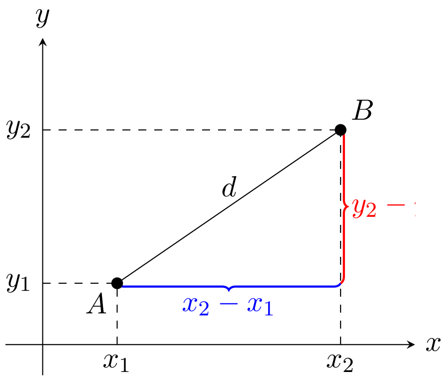
If we were to find the distance between 3 and 5, for instance, we would subtract 3 from 5:



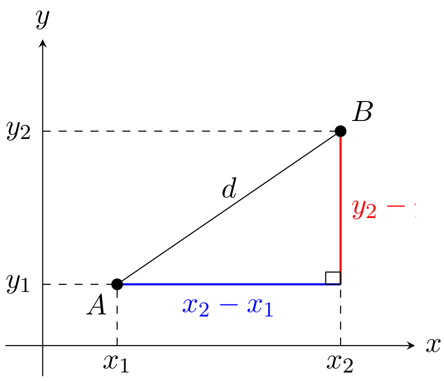
and we will do exactly the same with our x_1 and x_2:



Thinking vertically now, on the y axis, we can do exactly the same with the distance between y_1 and y_2:



We now need to join our independent x and y solutions to find d. Notice that we can draw a right triangle on our graph:



and we can now apply Pythagoras's theorem on it (see Chapter 36). d is the length of our hypotenuse, and the other sides are $x_2 - x_1$ and $y_2 - y_1$. Therefore

$$d^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2$$

which we can square root on both sides to find the formula we want:

Distance between two points formula

Given two points (x_1, y_1) and (x_2, y_2), the *distance*, d, between them is given by:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

For instance, say we have the points $(1, 3)$ and $(4, 7)$. To find the distance between them we can use the formula:

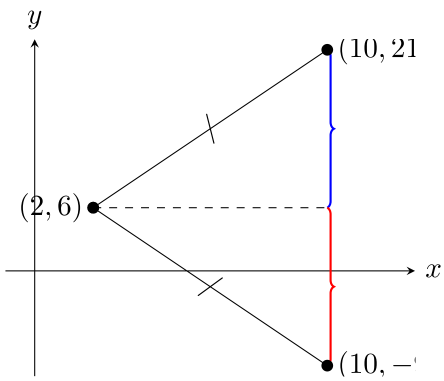
$$d = \sqrt{\underbrace{(1-4)^2}_{\text{dist. between } x} + \underbrace{(3-7)^2}_{\text{dist. between } y}} = \sqrt{(-3)^2 + (-4)^2} = \sqrt{9+16} = \sqrt{25} = 5$$

Notice that you can pick whichever point you want to come before, as long you use the same point for both x and y.

Sometimes, as usual, they can give you the distance between the points and ask you to solve a little equation for a coordinate. For instance, if we have the point $(2, 6)$ and the point $(10, y)$ and we are told the distance between them is 17, we can find y by using the formula:

$$
\begin{aligned}
d &= \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \\
17 &= \sqrt{(2-10)^2 + (6-y)^2} && \text{Subs. into the formula} \\
17^2 &= (2-10)^2 + (6-y)^2 && \text{Squaring both sides} \\
289 &= (-8)^2 + (6-y)^2 \\
289 &= 64 + (6-y)^2 \\
289 - 64 &= (6-y)^2 && \text{Subtracting 64 on both sides} \\
225 &= (6-y)^2 \\
\sqrt{225} &= \sqrt{(6-y)^2} && \text{Square rooting both sides} \\
\pm 15 &= 6 - y && \text{Careful to remember the } \pm \\
-y = -6 \pm 15 \rightarrow &\begin{cases} -y = -6 + 15 & (1) \\ -y = -6 - 15 & (2) \end{cases} && \text{Subtracting 6 on both sides} \\
&\begin{cases} -y = 9 \rightarrow y = -9 & (1) \\ -y = -21 \rightarrow y = 21 & (2) \end{cases}
\end{aligned}
$$

We have two solutions, then, $(10, -9)$ and $(10, 21)$. Geometrically, we have two solutions because the point can be either "above" or "below" $(2, 6)$:



If we were given a restriction, for instance that the point is on the fourth quadrant, we could have a unique solution.

23.6. Exam hints

Be very careful with points: in the heat of the moment it is very easy to confuse the x with y coordinates positions. Remember: x comes first in the alphabet, it is the original horizontal number line, and it is the first coordinate!

When you are using the distance formula, remember that you subtract the coordinates. We add them to find the midpoint. Also, when using the distance formula, just input the whole formula, after substituting the values, into the calculator:

$$\underbrace{\sqrt{\underbrace{(1-4)^2}_{\text{dist. between } x} + \underbrace{(3-7)^2}_{\text{dist. between } y}}}_{\text{copy exactly like this}}$$

Summary

- Any point on the coordinate plane or Cartesian plane or xy-plane or just plane can be uniquely identified by its x and y *coordinates*. The x coordinate represents its position on the *horizontal* number line, and the y coordinate its position on the *vertical* number line;

- We denote a point with coordinates x and y using brackets:

$$(x, y)$$

- The *midpoint* between two points (x_1, y_1) and (x_2, y_2) is the point with coordinates:

$$\left(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2}\right)$$

- The *distance* between two points (x_1, y_1) and (x_2, y_2) (or the *length of the line connecting them*) is given by:

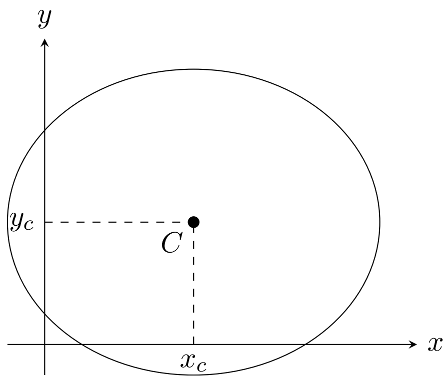
$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Formality after taste

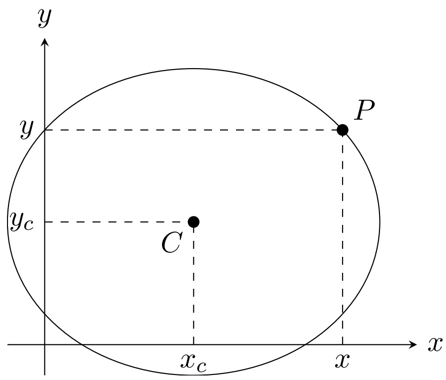
Deriving the equation of a circle in the coordinate plane

Using the distance formula we can derive an equation that represents a circle with a given centre and radius.

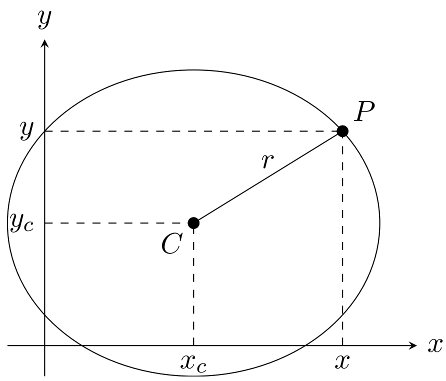
Say we are given the centre C, with coordinates, (x_c, y_c) and the radius r. Let us draw a representation on the plane:



Now let us take an arbitrary point P, with coordinates, (x, y) on the circumference of the circle:



An important feature of circles is that any point on their circumference is equidistant to the centre. We know this distance is the radius, r:



Let us use the distance formula, then! The distance from $C\,(x_c, y_c)$ to $P\,(x, y)$ is

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

$$r = \sqrt{(x - x_c)^2 + (y - y_c)^2} \qquad \text{Substituting coords}$$

$$(x - x_c)^2 + (y - y_c)^2 = r^2 \qquad \text{Squaring both sides}$$

and there we have it: the equation for the circle with centre (x_c, y_c) and radius r. If the circle had centre $(2, 3)$ and radius 3, for instance:

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

And mathematicians's favourite, the unit circle: a circle centered at the origin $(0, 0)$ and with radius equal to the unit, that is, 1:

$$x^2 + y^2 = 1$$

Polar coordinates, a different way to describe points in the plane

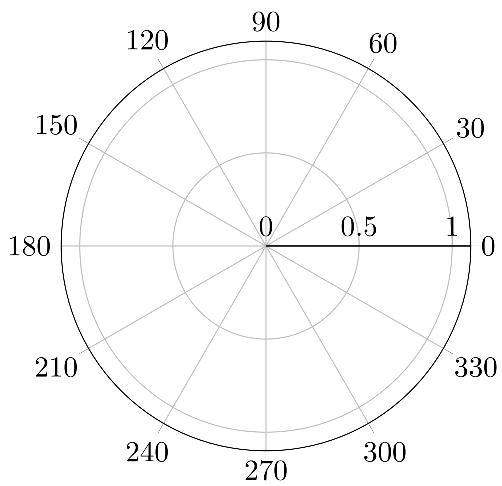
During your school education you will only use the rectangular system of coordinates, in which the x and y coordinates are given in order to place a point. We can, however, use a different system, called *polar*.

In the polar coordinates system, we define a point position with its *distance from the origin* and the *angle* the line that connects the point to the origin makes with the x axis. We need both the angle and the distance because, if just given the distance you have all points on a circumference of distance as radius. If only given the angle, you have all the points on the straight line that makes that angle with the x axis.

Thus, we define points, using polar coordinates, by the pair

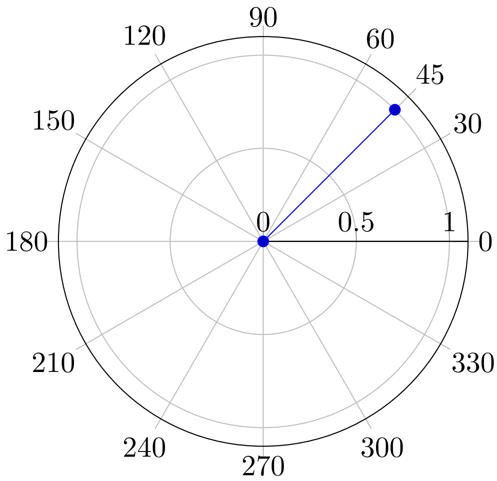
$$(r, \theta)$$

where r is the distance from the origin and θ (Greek letter theta) is the angle. We even draw the "guidelines" of the plane differently:



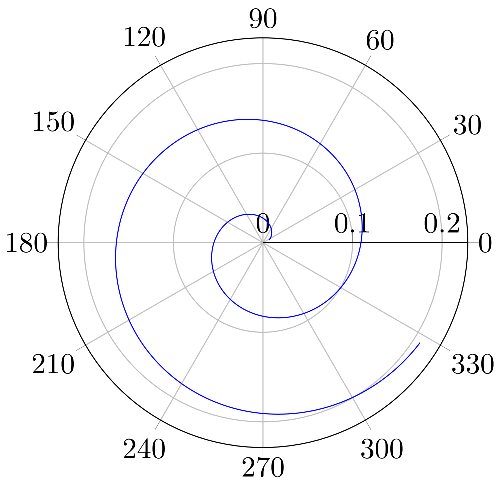
as what matters are the circles centered at the origin, which represent families of points with distance r, and the angles on the circle.

Say, for instance, that we want to plot the point whose distance is 1 from the origin and makes a 45° angle with the x axis:

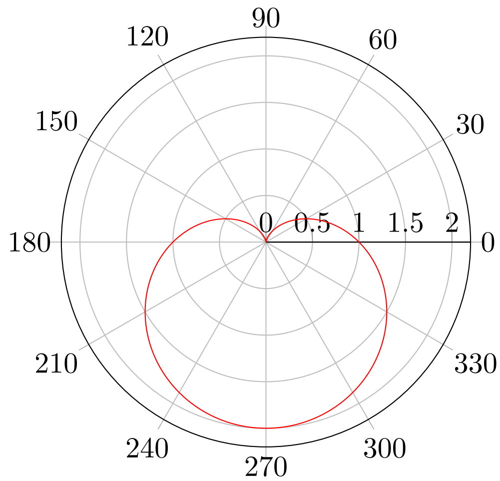


If you already know sine and cosine (ADDREF), it is easy to convert from rectangular to polar coordinates.

In polar coordinates it is simple to define complicated graphs such as spirals:



or a cardioid, which the name says it looks like a heart (card- is heart, -oid is “in the shape of), but it does look like an apple:



I highly recommend you to go to `desmos.com` and play around with polar coordinates.

24. Linear function and its graph

24.1. Why learn about the linear function and its graph

In many situations, having a "visual" representation of a mathematical entity is of great help. Sometimes to help visualize its behaviour, sometimes to help you solve a problem.

With functions, their graphs are not only useful in this regard but also really helpful when solving equations or inequalities. We need to start our study of the graphs of functions with the simplest graph, but the ideas here will be useful later.

24.2. Basic idea

When we want to graph a function $f(x)$, what we actually want is to obtain a set of ordered pairs (points) of the form

$$(x, f(x))$$

which simply means that for every x value, we find the corresponding $f(x)$ and then we plot the point

$$(x, y = f(x))$$

that is, the points have x coordinate that we pick and the corresponding y coordinate is the value of $f(x)$.

This is why graphing the function is sometimes asked of you by "plot the graph of $y = f(x)$".

For all the graphs we will see, this is always the idea: the graph of a function f is a visual representation of the points $(x, f(x))$.

24.3. The equation of a straight line

24.3.1. An exception: $x = k$

Before we look at "usual" straight lines, a special case. Graphs of the form

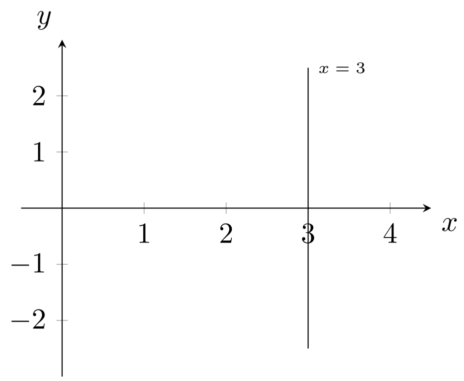
$$x = \text{number}$$

are not graphs of linear functions (you will learn later that $x = k$ is not a function). Despite that, they are straight lines, so we will study them here.

If we are given the equation

$$x = 3$$

and told to plot its graph, the only thing we know about it is that the x coordinate of all points will be equal to 3. As we can have an infinite number of y values with x coordinate equal to 3, the graph of $x = 3$ is a *vertical* straight line:



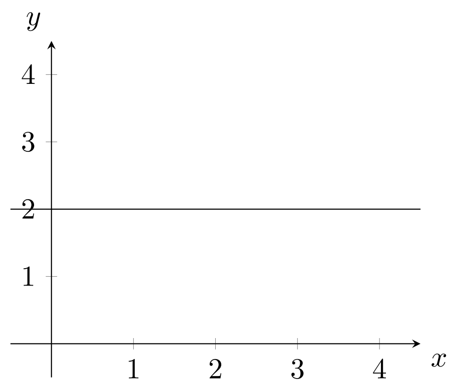
In general, graphs of equations of the form

$$x = \text{number}$$

are vertical straight lines crossing the x axis at the number given.

24.3.2. A simple case: $y = k$

Before we go to the general case, let us look at the "friend" of the vertical line, the horizontal. For instance, say we have the following graph of a line:



we have a horizontal line, and all points in the line have y coordinate equal to 2. Hence, the equation of the horizontal line going through $y = 2$ is

$$y = 2$$

In general, horizontal straight lines have equations of the form

$$y = k$$

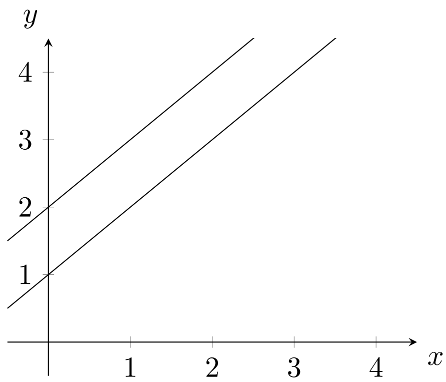
where k is the y coordinate of their points.

24.3.3. The general case

We now want to discuss lines which are "between" vertical and horizontal. We will learn about the two *parameters* that determine a straight line, the y intercept and its gradient.

24.3.3.1. The y intercept, c

Take a look at the two lines on the next graph:



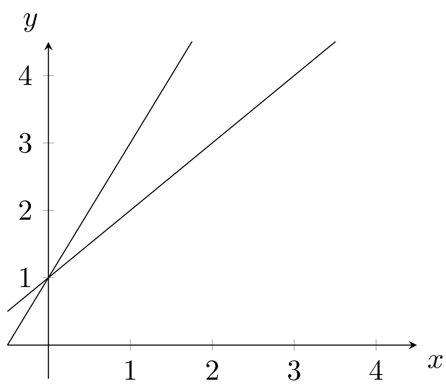
notice that they are similar in the sense they have the same "incline" but they cross the y axis in a different point: one of them at $(0, 1)$ and one at $(0, 2)$. The point in which a line crosses the y axis is called its y *intercept*, the place where the line intercepts the y axis (creative, I know).

The y intercept will be referred by the variable c, and it will be one of the parameters to determine a line.

To find the y intercept of a line from its graph, you just need to see where it touches the y axis. I will show you later (see 24.6) how to find it from the equation.

24.3.3.2. The gradient (a.k.a. slope), m

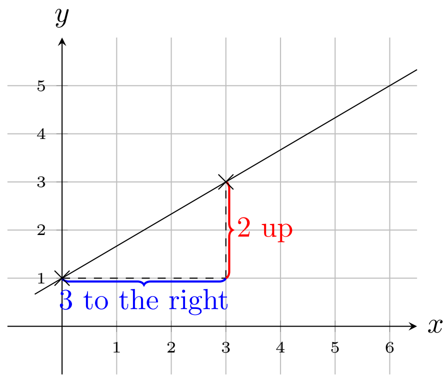
Now look at these two lines:



both of them cross the y axis at the same point, $(0, 1)$. Hence, they have the same y intercepts.

However, notice that they have different inclines: one of them is more steep than the other. The steepness of a line is another way to differentiate them. There are many ways to measure it, but we will use what we call the *gradient* (in the US it is called *slope*).

The gradient of a line is a measure of how steep it is, and it is very simple: we see how much the line has gone up (or down) relatively to how much it has gone to the right (or the left). Let me give you an example:



I chose two points on the line $(0, 1)$ and $(3, 3)$. To see how steep the line is, its gradient, I can see how much it went "up", in this case 2, and divide it by how much it went to the right, in this case 3:

$$\text{gradient} = \frac{\text{how much it went up}}{\text{how much it went right}} = \frac{2}{3}$$

(if you already know trigonometry, particularly the tangent ration, the gradient is the tangent of the right triangle in the graph above).

So, the gradient is a *ratio* (a division) of the *vertical distance* between two points by the *horizontal distance* between the same points. That is why it is commonly said that the gradient is "rise over run", where the rise is the vertical distance and the run the horizontal distance.

Hence,

$$\text{gradient} = \frac{\text{vertical distance between two points}}{\text{horizontal distance between the same points}}$$

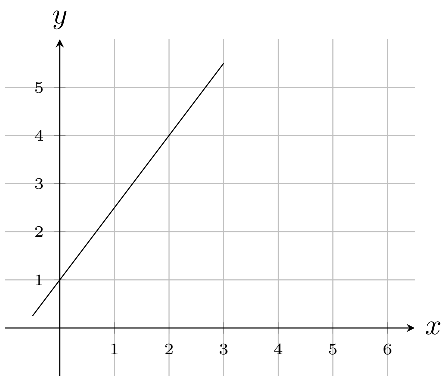
$$\text{gradient} = \frac{\text{rise}}{\text{run}}$$

We refer to the gradient by the variable m, so we can write:

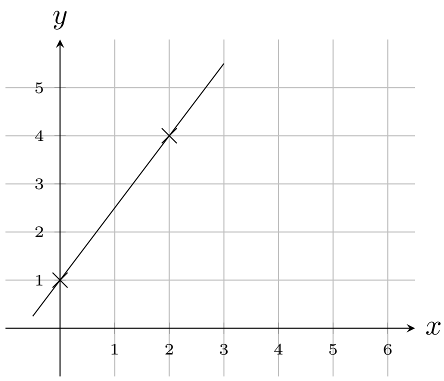
$$m = \frac{\text{vertical distance between two points}}{\text{horizontal distance between the same points}} = \frac{\text{rise}}{\text{run}}$$

There are basically two ways to find the gradient of a line: graphically or using a formula. Let us find it graphically first.

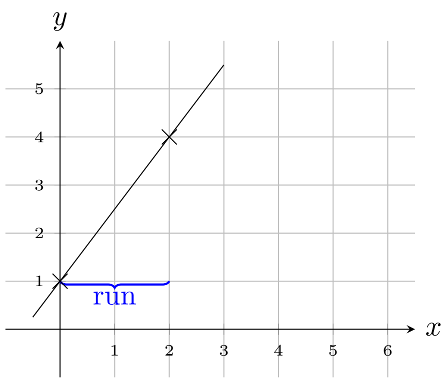
Finding the gradient of a line from its graph Say we are given this line and told to find its gradient:



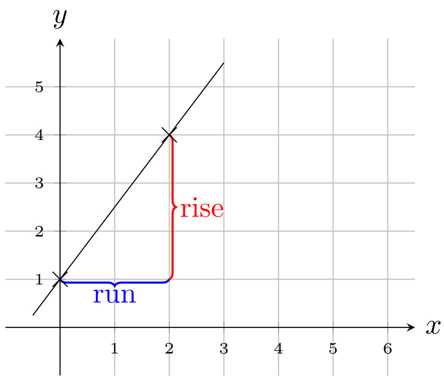
The first thing we do is choose two points on the line. Of course, try to pick points which have whole number coordinates. Let us choose $(0, 1)$ and $(2, 4)$:



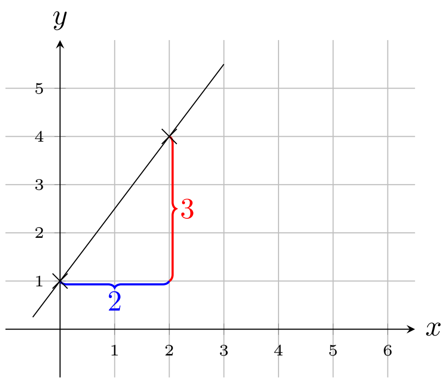
Now, we need to see the horizontal and vertical distance between the points. Let us start with the horizontal distance, the "run":



Now, for the vertical distance, we can just go up from where we stopped. So, for the "rise":



We now can simply count those distances. In our case, the horizontal distance, the "run" is 2. The vertical distance, the "rise", is 3:



and now we just divide them:

$$\text{gradient} = m = \frac{\text{vertical distance}}{\text{horizontal distance}} = \frac{\text{rise}}{\text{run}} = \frac{3}{2}$$

Thus, the gradient of the line is $\frac{3}{2}$.

Finding the gradient of a line algebraically

24.3.4. The equation of a straight line, $y = mx + c$

Now that we know that a line is identified by its gradient and y intercept, we can write any line equation by combining them:

$$y = \underbrace{m}_{\text{gradient}} x + \underbrace{c}_{y-\text{intercept}}$$

Here, m is the gradient and c is the y intercept. y and x are just coordinates of any point on the line.

For instance, say that we have a line

$$y = 2x - 1$$

in this, the gradient is 2 (the number multiplying x) and the y intercept is -1.

24.4. Plotting linear graphs

24.4.1. Using a substitution table

When graphing any function, you can always construct a table of values relating x and y values, which will give you as many points as you need.

For straight lines, only 2 points are needed, but many teachers say you should find 3 to "be sure". I suggest you to do whatever you prefer.

Let us say we want to plot the graph of

$$y = 2x - 1$$

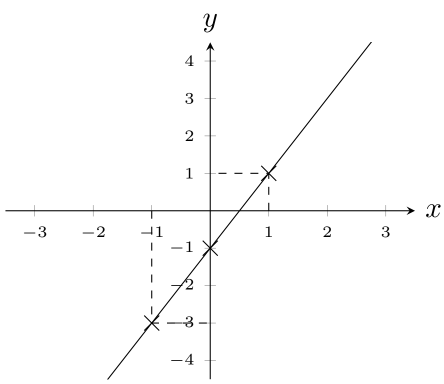
We can *choose* some values of x (any values we want) to substitute into the equation. I recommend values close to 0:

x	$y = 2x - 1$
-1	$y = 2 \times -1 - 1 = -2 - 1 = -3$
0	$y = 2 \times 0 - 1 = 0 - 1 = -1$
1	$y = 2 \times 1 - 1 = 2 - 1 = 1$

and each row will give you a point to plot:

x	$y = 2x - 1$	(x, y)
-1	-3	$(-1, -3)$
0	-1	$(0, -1)$
1	1	$(1, 1)$

which we now plot and connect:



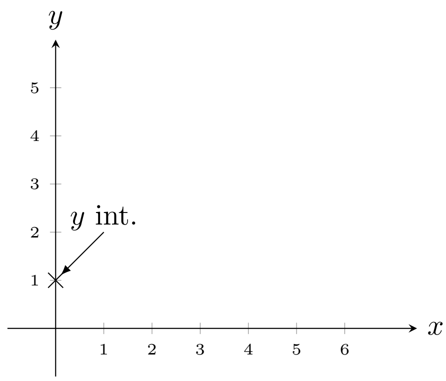
I will now show you two different techniques to plot linear graphs, depending on how the equation is given. Of course, you can go from one to another by rearranging the equation, but each method is very convenient and easy.

24.4.2. Given $y = mx + c$, the "slope-intercept" form

Let us say we would like to draw the graph of

$$y = \frac{2}{3}x + 1$$

The first thing we do is remember that the c, the y intercept, is a point that we have for free: we can always, when given the equation like this, plot the point $(0, c)$. In our case, we have $(0, 1)$:



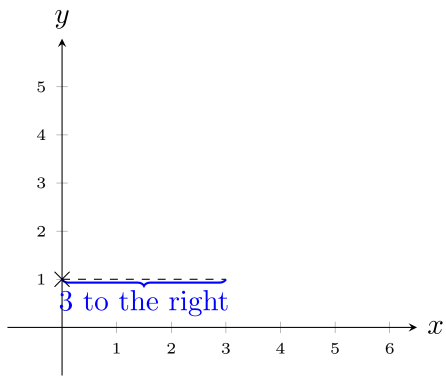
To find another point, we need to interpret the gradient, m. Always keep in mind that we define the gradient as "rise over run", that is, how much we move up or down divided by how much we move to the right. In our case, the gradient is $\frac{2}{3}$, hence we have

$$\frac{2}{3} = \frac{\text{rise}}{\text{run}}$$

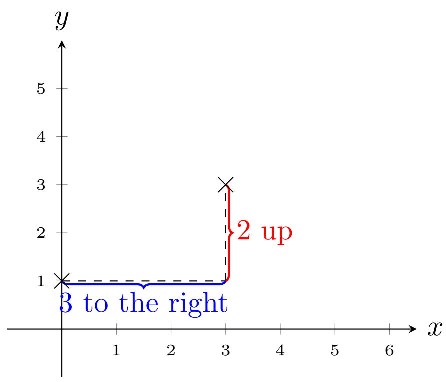
which means that if we move 2 units up, as the numerator says, we move 3 units to the right:

$$\frac{2}{3} = \frac{\text{how much we move up}}{\text{how much we move right}}$$

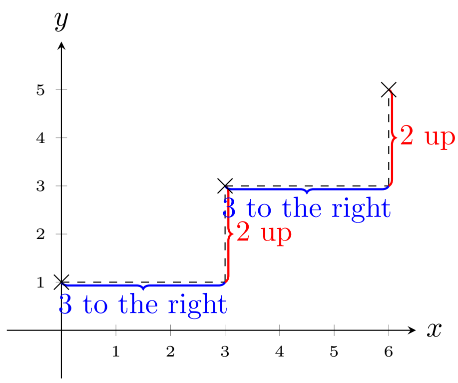
So, from any point in the line, if we move 3 units to the right, we move 2 units up. But we already have a point, the y intercept, plotted: so we just go 3 units to the right from it:



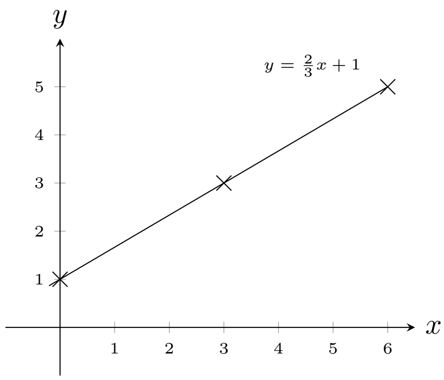
and from where we are, we move 2 units up (the "run" part, as the numerator tells us), and we have our point:



Let us repeat the process to get another point just to practice. From our newly found point, $(3, 3)$, we go another 3 units right and 2 units up:



which we simply connect to the other 2 points to obtain our line:



We do not need 3 points, but some people like the confirmation, as I mentioned above.

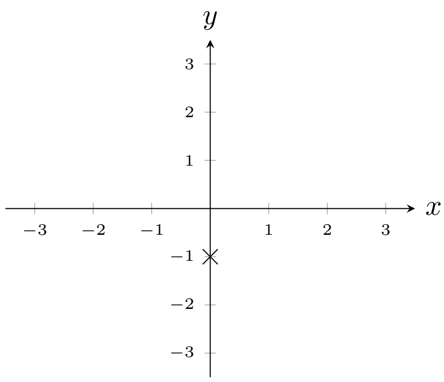
When the gradient is not a fraction, we simply write the number with 1 in the denominator. Say we wanted to plot

$$y = 3x - 1$$

This is the same as

$$y = \frac{3}{1}x - 1$$

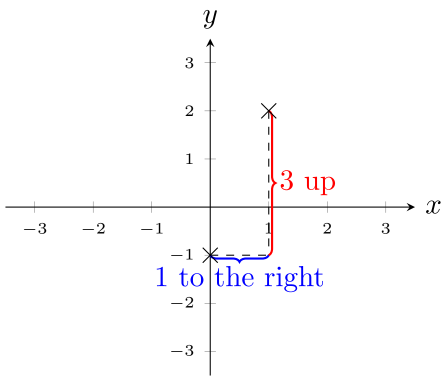
and we can use the same technique. We start with the y intercept, $(0, -1)$:



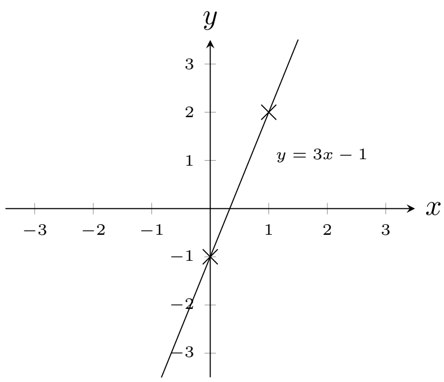
and again look at the gradient:

$$3 = \frac{3}{1} = \frac{\text{how much we move up}}{\text{how much we move right}}$$

Thus, when we go 1 unit right, we move 3 units up, which gives us another point:



Now we have two points, so we can use them to draw the line:



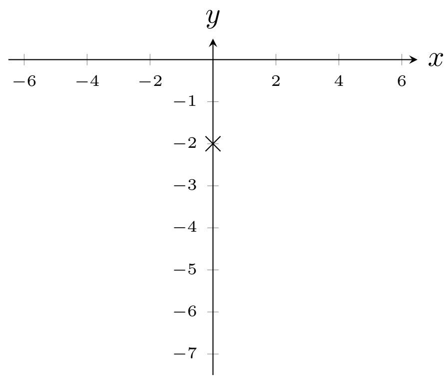
If the gradient is negative, either you memorize (which I never recommend) that the line is decreasing, or we can simply think of "rise over run" again, but this time "negative rise". For instance, if we want to plot

$$y = -\frac{3}{5}x - 2$$

we can remember, from Chapter 3, that

$$-\frac{3}{5} = \frac{-3}{5}$$

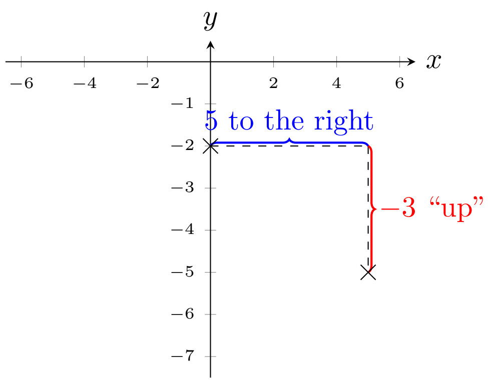
and now we do exactly the same. We start with the y intercept point, $(0, -2)$:



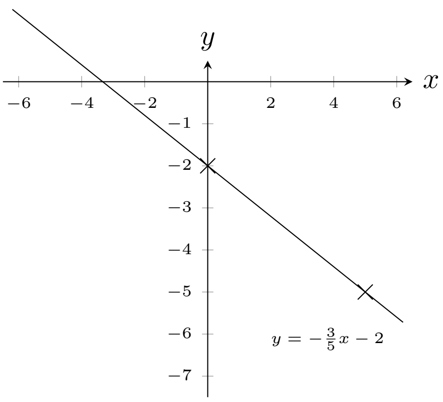
and now use the gradient

$$\frac{-3}{5} = \frac{\text{how much we move ``up''}}{\text{how much we go to the right}}$$

When we move 5 units to the right, we go 3 units down, hence the quotes around up. We obtain our point:



which gives us our line:



24.4.3. Given $ax + bx = d$, the standard form

When the equation is given in standard form, we can use what most people know as the "cover-up method".

Say we want to plot

$$2x + 5y = 10$$

If we make $x = 0$ into it, the whole $2x$ terms vanishes:

$$2 \times 0 + 5y = 10$$

$$0 + 5y = 10$$

$$5y = 10$$

and, as you can see, we now have a simple equation with only y, which we can solve dividing both sides by 5:

$$5y = 10$$

$$y = \frac{10}{5} = 2$$

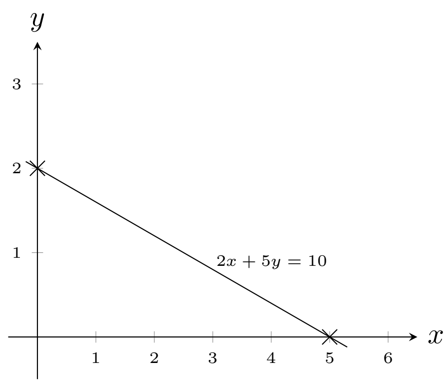
What this means is that, when $x = 0$, the line goes through $y = 2$. Thus, we have the point $(0, 2)$. Now, we repeat the same process but making $y = 0$:

$$2x + 5 \times 0 = 10$$

$$2x = 10$$

$$x = \frac{10}{2} = 5$$

which tells us that when $y = 0$, the line goes through $x = 5$. Hence, we have the point $(5, 0)$. As we now have two points, $(5, 0)$ and $(0, 2)$, we can plot the line:



A quick way to find these two points is by "covering up" the $2x$ term in the equation:

$$2x + 5y = 10$$

$$\underbrace{\square}_{\text{cover up}} \; 5y = 10$$

$$5y = 10$$

$$y = \frac{10}{5} = 2$$

and "covering up" means we made $x = 0$, so we obtain the equation $5y = 10$, which we solve to get $y = 2$. Thus, we got the point $(0, 2)$.

We now "cover up" the $5y$ term:

$$2x + 5y = 10$$

$$2x \; \underbrace{\square}_{\text{cover up}} = 10$$

$$2x = 10$$

$$x = \frac{10}{2} = 5$$

which means that when $y = 0$, $x = 5$. Therefore, we obtain the point $(5, 0)$.

Another example. Say we would like to graph

$$3x - 2y = -6$$

We can "cover up" the $3x$ term:

$$3x - 2y = -6$$

$$\square - 2y = -6$$

$$-2y = -6$$

$$y = \frac{-6}{-2} = 3$$

which tells us that when $x = 0$, $y = 3$. Hence we have the point $(0, 3)$. We now "cover up" the $-2y$ term:

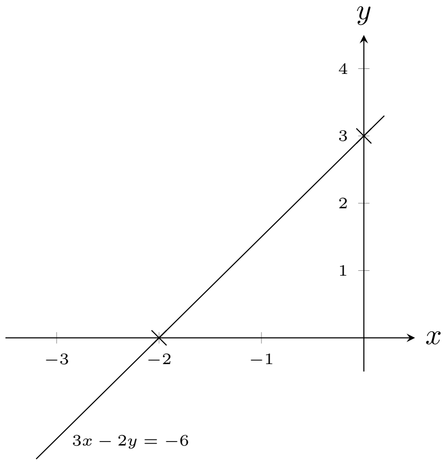
$$3x - 2y = -6$$

$$3x\square = -6$$

$$3x = -6$$

$$x = \frac{-6}{3} = -2$$

which gives us that when $y = 0$, $x = -2$ and the point $(-2, 0)$. We can now plot the line:



24.5. Checking if a point is on a line

Now that we know how the graph of a straight line is, we can answer if a given point is on the line or not.

Say we have the line

$$f(x) = y = 4x - 1$$

(from now on I will just write $y = 2x - 1$, but do remember that what we are actually doing is graphing the function $f(x) = 2x - 1$ using the idea of $(x, f(x))$)

To check if the point $(2, 3)$ is on the line, we substitute the x and y coordinates of the point in $y = 4x - 1$:

$$(2, 3)$$

$$y = 4 \times x + 3$$

and we obtain:

$$3 \stackrel{?}{=} 4 \times 2 + 3$$

I am using the $\stackrel{?}{=}$ as we are not sure if the left side is equal to the right side. However, if they are equal, the point will be on the line; if not, the point is not on the line. In our case:

$$3 \stackrel{?}{=} 8 + 3$$

$$3 \neq 11$$

as 3 is not equal to 11, the point $(2, 3)$ is *not* on the line $y = 4x - 1$.

Another example: let us check if the point $(2, 5)$ is on the line $y = -2x + 9$. Again, we substitute the point into the line equation:

$$(2, 5)$$

$$y = -2 \times x + 9$$

which gives us:

$$5 \stackrel{?}{=} -2 \times 2 + 9$$

$$5 = 5$$

Thus, the point $(2, 5)$ is on the line $y = -2x + 9$.

So, in summary: to find out if a point (x, y) is on a line $y = mx + c$, substitute its x and y coordinates on the line equation. If the sides are equal, the point is on the line; if not, the point is not.

24.6. Finding the y and x intercepts from the equation of a line

24.7. Parallel and perpendicular lines

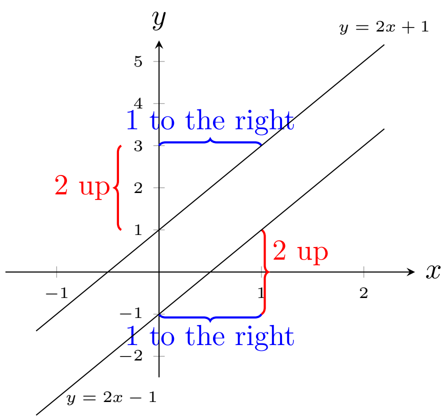
24.7.1. Parallel lines

Parallel lines have *the same gradient*. This makes sense, as parallel lines "never meet". If two lines have the same gradient (and pass through two different points), they will never be able to meet. For instance, the lines

$$y = 2x - 1$$

$$y = 2x + 1$$

have the same gradient, 2, but different y intercepts. If you graph them:



and you can see that the distance between the lines will remain constant and they will never meet.

Hence, do memorize the fact:

Parallel lines have *equal gradients.*

24.7.2. Perpendicular lines

Perpendicular lines have an interesting relation on their gradients: if you multiply one by the other, the answer is -1. This means that one of the gradients has to be the *negative reciprocal* of the other: it's "flipped" and with the opposite sign.

For instance, if the gradient of a line is $\frac{2}{3}$, to find the gradient of any perpendicular line to it we "flip and change the sign":

$$\underbrace{\frac{2}{3}}_{\text{original}} \xrightarrow{\text{negative}} -\frac{2}{3} \xrightarrow{\text{"flip"}} \underbrace{-\frac{3}{2}}_{\text{perpendiular}}$$

I like doing both the "flipping" and changing the sign in one step and remember "perpendicular lines are neg rec" (from negative reciprocals):

$$-\frac{3}{4} \xrightarrow[\text{rec}]{\text{neg}} \frac{4}{3}$$

Do remember that whole numbers are also fractions with 1 in the denominator:

$$5 = \frac{5}{1} \xrightarrow[\text{rec}]{\text{neg}} -\frac{1}{5}$$

Perpendicular lines have *negative and reciprocal gradients.*

24.8. Finding the equation of a line

In this type of question, we need to find the equation of the line

$$y = mx + c$$

given some information.

No matter what form of this problem comes, I always suggest:

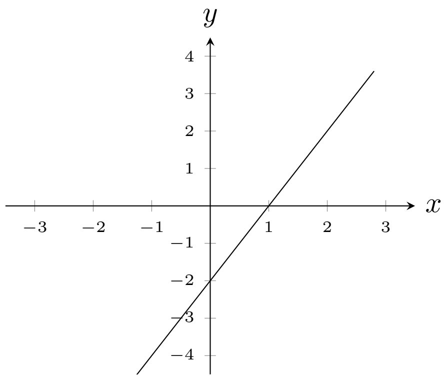
1. Find the gradient; then
2. find the y intercept

This order will also work when working with derivatives (see Chapter ADDREF), so it is very consistent and easy to apply.

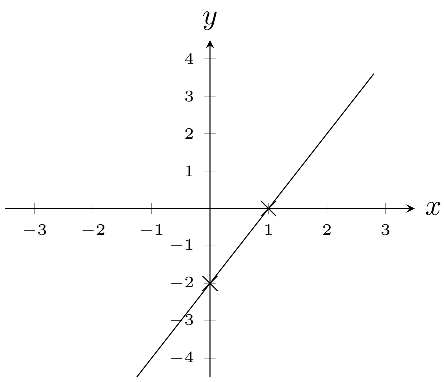
24.8.1. From its graph

If we are given the graph of a straight line, first we find the gradient of the line. We can do that by choosing two points and then finding "rise over run", that is, the difference in height (the y axis) and divide it by the distance in length (the x axis).

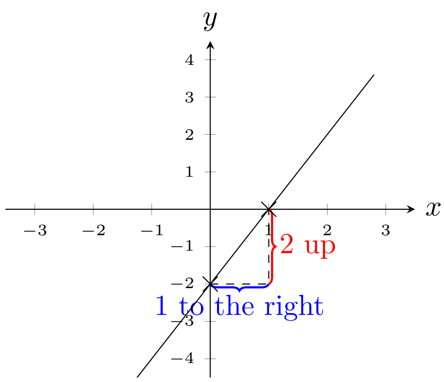
Let us say we have this graph:



First we choose any two points on the line. I will choose $(0,-2)$ and $(1,0)$:



and now we draw the "rise over run" triangle to help us:



and we have our “rise over run”, which allows us to find the gradient, m:

$$m = \frac{\text{rise}}{\text{run}} = \frac{2 \text{ up}}{1 \text{ right}} = \frac{2}{1} = 2$$

Thus, $m = 2$. So far, then, we know our line equation looks like:

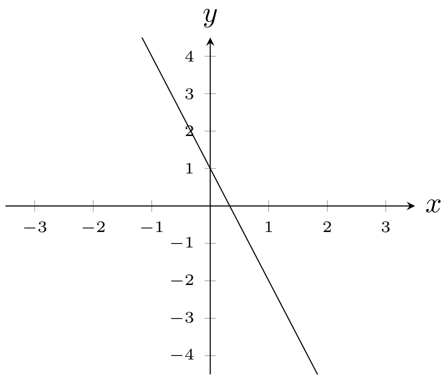
$$y = 2x + c$$

We now need to find the y intercept. In this case, we already have it: it is -2, as the line intercepts the y axis at $(0, -2)$. We can substitute -2 in c:

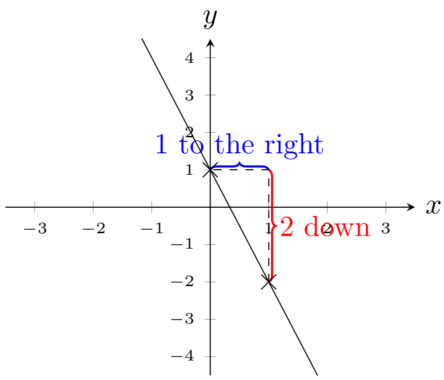
$$y = 2x - 2$$

and we have our line equation.

You have to be careful when the line is decreasing, that is, it has a negative gradient. For instance:



we can choose the points $(0, 1)$ and $(1, -2)$ and draw the triangle:



as you can see, the triangle is “different”: the line has decreased, to the difference in height (the y axis) is *negative*, as we went *down*. So, the gradient becomes:

$$m = \frac{\text{rise}}{\text{run}} = \frac{2 \text{ down}}{1 \text{ right}} = \frac{-2}{1} = -2$$

and the the gradient is negative, as we expected (remember that decreasing lines have negative gradient). So far, then, our line has equation

$$y = -2x + c$$

The y intercept is given again, as the line crosses the y axis at the point $(0, 1)$. So, $c = 1$:

$$y = -2x + 1$$

Some people prefer to "ignore" the direction of the rise, and just memorize that

- Increasing lines have positive gradient;
- Decreasing lines have negative gradient.

24.8.2. From two points

If, instead of giving the graph, the question says the line goes through two points, first we find the gradient using the gradient formula we learned above (see 24.3.3.2). Then, we use one of the points to find the y intercept.

For instance, say we are given that a line goes through $(1, 2)$ and $(3, 6)$. We first find the gradient using

$$m = \frac{\text{rise } (y)}{\text{run } (x)} = \frac{y_2 - y_1}{x_2 - x_1}$$

To do that, we need to decide which point we will use as "point 1" and which as "point 2". The order makes no difference at all, so I usually choose the first point in the question as "point 1":

$$\text{Point } 1 = \left(\underbrace{1}_{x_1}, \underbrace{2}_{y_1}\right) \text{ and Point } 2 = \left(\underbrace{3}_{x_2}, \underbrace{6}_{y_2}\right)$$

which we substitute now into the formula for m:

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{6 - 2}{3 - 1} = \frac{4}{2} = 2$$

So far, then, our line equation looks like this:

$$y = 2x + c$$

To find c, we can substitute any point's coordinates into the equation, and solve it for c. Say we choose "point 1" :

$$y = 2x + c$$

$$\underbrace{2}_{y_1} = 2 \times \underbrace{1}_{x_1} + c$$

$$2 = 2 + c$$

$$0 = c$$

Thus, our line has equation

$$y = 2x + 0$$

or simply

$$y = 2x$$

Do be careful when the points given have negative coordinates, as a very common mistake is to ignore the "minus minus" that appears. For instance, say we want to find the equation of a line which goes through the points $(-2, -4)$ and $(-3, 0)$. Using the first point as point 1 again and plugging the coordinates into the gradient formula gives us

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{0 - -4}{3 - -2} = \frac{4}{5}$$

We know the line has equation, thus far

$$y = \frac{4}{5}x + c$$

and we substitute a point in it, say $(-3, 0)$ to find c:

$$0 = \frac{4}{5} \times -3 + c$$

$$0 = -\frac{12}{5} + c$$

$$c = \frac{12}{5}$$

Therefore, the line has equation

$$y = \frac{4}{5}x + \frac{12}{5}$$

24.8.3. From a point and another piece of information

If instead of two points they tell you that the line is parallel or perpendicular to another line, just use the relations above (see Section 24.7) to find the gradient first and then the idea is the same.

For instance, to find the equation of the line which is parallel to $y = 3x - 1$ and goes through the point $(1, 4)$, first we remember that parallel lines have equal gradients. Thus, the gradient of our line is also 3:

$$y = 3x + c$$

To find c we substitute the point $(1, 4)$, as usual:

$$4 = 3 \times 1 + c$$

$$4 = 3 + c$$

$$c = 1$$

and we have our line equation:

$$y = 3x + 1$$

Be careful if they tell you that the line is parallel to $4x + 2y = 8$, as you first need to rearrange it to get into the "slope-intercept" form, by making y the subject:

$$4x + 2y = 8$$

$$2y = -4x + 8$$

$$y = \frac{-4x}{2} + \frac{8}{2}$$

$$y = -2x + 4$$

Now you can safely say the gradient of the line is -2 and continue.

The same idea applies if the line goes through $(1, 2)$ and is perpendicular to $y = \frac{1}{2}x + 1$. Perpendicular lines have negative reciprocal gradients, so:

$$\frac{1}{2} \xrightarrow{\text{neg}} -\frac{1}{2} \xrightarrow{\text{rec}} -\frac{2}{1} = -2$$

our line, then, has gradient -2 and looks like:

$$y = -2x + c$$

and to find c we substitute the point $(1, 2)$:

$$2 = -2 \times 1 + c$$

$$2 = -2 + c$$

$$c = 4$$

Finally, we have our line:

$$y = -2x + 4$$

24.8.4. Using $y - y_0 = m\left(x - x_0\right)$

The formula you can fondly read as "yo yo mi xo xo" can be used to find the equation of a line if you know its gradient and a point it goes through.

For instance, say that you want to find the line that has gradient 4 and goes through the point $(2, 3)$. We substitute $m = 4$ and $x_0 = 2$ and $y_0 = 3$:

$$(2, 3), m = 4$$

$$y - y_0 = m\left(x - x_0\right)$$

$$y - 3 = 4\left(x - 2\right)$$

and we know simplify the equation we reached:

$$y - 3 = 4\,(x - 2)$$

$$y - 3 = 4x - 8$$

$$y = 4x - 8 + 3$$

$$y = 4x - 5$$

As long as you have the gradient and a point, you can always use "yo yo mi xo xo". For example, say we want to find the line that goes through the points

$$(-2, 1) \text{ and } (1, -5)$$

We first find the gradient, as usual:

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{1 - -5}{-2 - 1} = \frac{6}{-3} = -2$$

and we have two points the line goes through. Let us choose $(-2, 1)$ and substitute:

$$(-2, 1)\,, m = -2$$

$$y - y_0 = m\,(x - x_0)$$

$$y - 1 = -2\,(x - -2)$$

which we simplify again:

$$y - 1 = -2(x - -2)$$

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

$$y - 1 = -2x - 4$$

$$y = -2x - 3$$

24.9. Exam hints

Summary

Formality after taste

Why perpendicular lines have negative reciprocal gradients

25. Quadratic function and its graph

26. Cubic function and its graph

27. Reciprocal function and its graph

28. Exponential function and its graph

29. Absolute value and absolute value of linear functions

30. Tangents and introduction to differentiation

31. Linear programming

31.1. Why learn linear programming

First, a disclaimer: linear programming has nothing to do with computer programming. It has to do with *optimization*, which means finding a solution to a problem which is the *best* depending on the restrictions of the problem. As you can imagine, this is very useful in practice, and optimization itself is probably one of the most important mathematical tools in the real world. Wikipedia, for example, says:

> Linear programming can be applied to various fields of study. It is widely used in mathematics, and to a lesser extent in business, economics, and for some engineering problems. Industries that use linear programming models include transportation, energy, telecommunications, and manufacturing. It has proven useful in modeling diverse types of problems in planning, routing, scheduling, assignment, and design.[1]

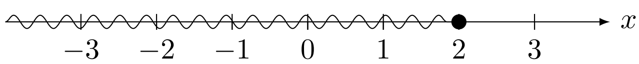
Now, to be very honest with you, we are going to see a very 'watered down' version of linear programming, which basically is an extension of our inequality knowledge. We will be doing some optimization, but it is very limited. However, the beginning of something is always the most important, as it we build on top of it!

31.2. Representing inequalities in the coordinate plane, a.k.a. shading

We know how to represent inequalities with one unknown in a number line. For instance,

$$x \leq 2$$

is represented as



But what about inequalities with *two* unknowns, such as:

$$x + y \leq 3$$

[1] `https://en.wikipedia.org/wiki/Linear_programming`. Accessed on: 25/06/2018.

For this, we need to find a *region* of the coordinate plane. This region will contain every single point (x, y) that satisfies our inequality. In the same way that there was an infinite number of solutions to $x \leq 2$, we have an infinite amount of points that satisfy $x + y \leq 3$, thus the region.

Let's find some points that satisfy $x + y \leq 3$. The point $(0, 0)$ does:

$$x + y \leq 3$$

$$0 + 0 \leq 3$$

$$0 \leq 3$$

0 is indeed smaller than 3, so we know that the point $(0, 0)$ is in the region we want. Now, the point $(1, 3)$ does not satisfy our inequality:

$$x + y \leq 3$$

$$1 + 3 \leq 3$$

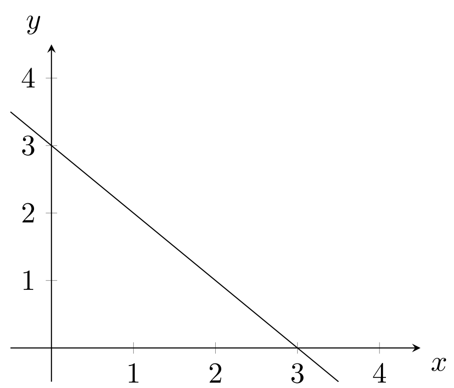
$$4 \leq 3$$

As you can see, the inequality is not true.

To represent this infinitude of points, we will shade **the region which has the points *that do not satisfy the inequality***. I know, it's weird to shade 'what we do not want', but it helps to visualise the answers to the exercises later on. The first thing is to draw the line

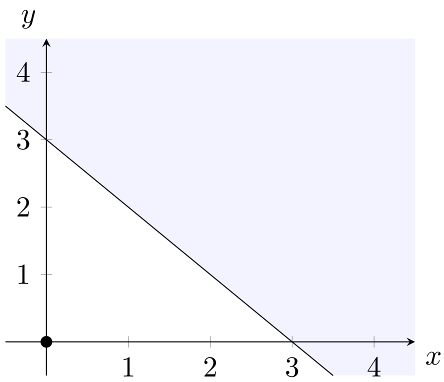
$$x + y = 3$$

as it is going to be the boundary for our region:



Notice that I have drawn the line 'filled'. Remember that we had the 'open circle' for inequalities such as $x < 2$? We will have to draw a dashed line for inequalities such as $x + y < 3$ and 'filled' lines for $x + y \leq 3$.

Now that we have the line, we need to shade the region **we do not want**. Remember that the point $(0, 0)$ was in the region we wanted? We are going to shade the 'half' of the coordinate plane the line $x + y = 3$ creates which has not the point $(0, 0)$:



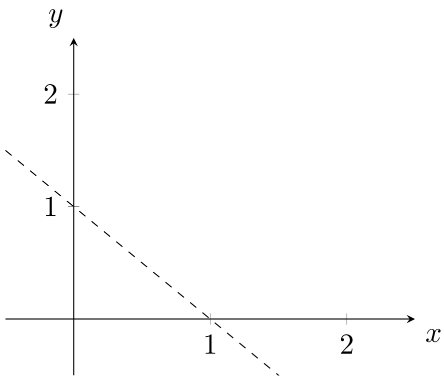
Notice that I have plotted the point $(0,0)$ for you to see that we shade the *region we do not want.*

For inequalities such as

$$y > 1 - x$$

we proceed exactly the same. First we draw the line

$$y = 1 - x$$



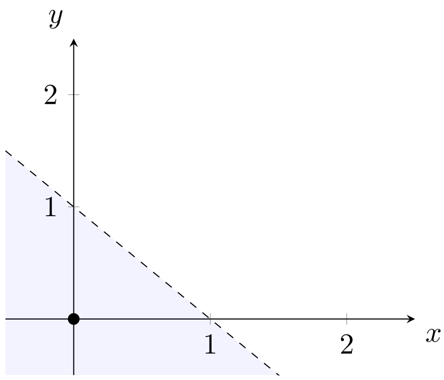
but notice we draw it *dashed*, as we are not including the line in the region (the points in the line are the ones where $y = 1-x$, but we don't want y to be *equal* to $1-x$ we want it to be *bigger).* Now, let's use the same trick to figure out what to shade (remember we shade the region *we don't want*). Substituting the point $(0,0)$:

$$y > 1 - x$$

$$0 > 1 - 0$$

$$0 > 1$$

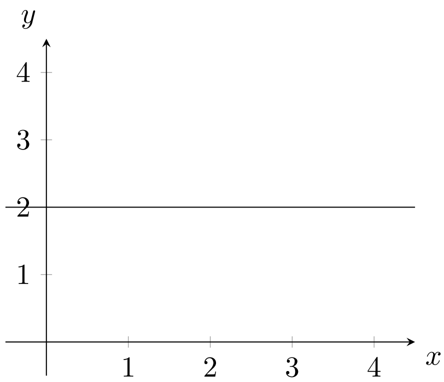
we see that $(0,0)$ does not satisfy the inequality, hence we shade the region that contains $(0,0)$ *because we shade what we do not want*, as you can see by the point plotted:



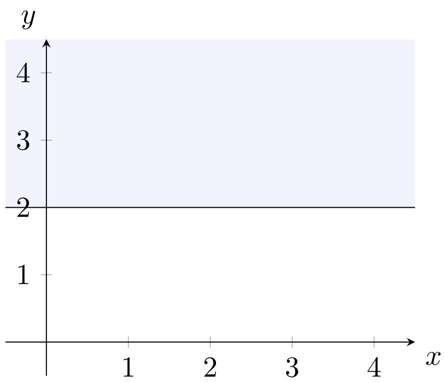
Nothing changes when we have only one unknown, such as

$$y \leq 2$$

To begin, draw the line $y = 2$ *filled* (notice that it is smaller than or *equal to*):



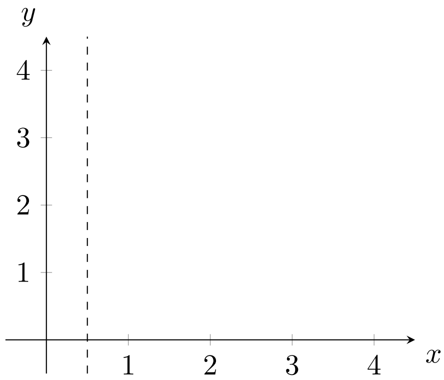
Now, any point which has y coordinate smaller than or equal to 2 is in the region we want, so we shade everything *above* the line:



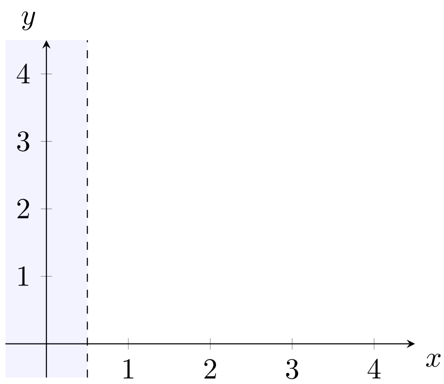
Same thing for

$$x > 0.5$$

First, draw the line $x = -0.5$



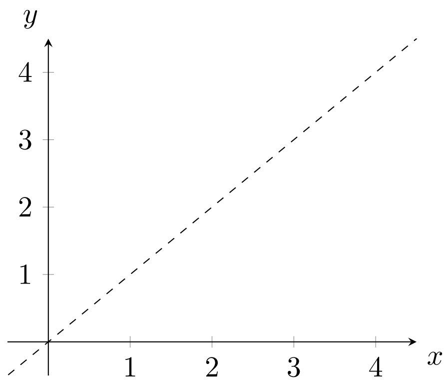
Any point with x coordinate bigger than 0.5 is in the region we want, so we shade everything *to the left of the line*, as we don't want that region:



A final example is when we cannot use the point $(0, 0)$ to decide what region to shade, such as

$$y < x$$

The idea is still the same. draw the line $y = x$ *dashed* (smaller than!):

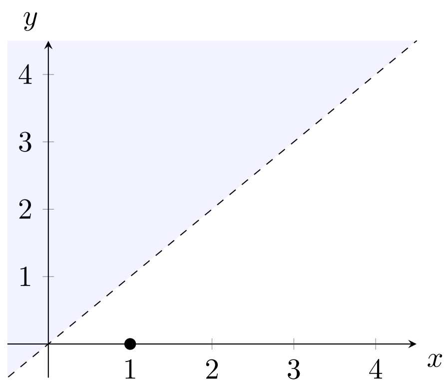


Now just use a point such as $(1, 0)$ to decide:

$$y < x$$

$$0 < 1$$

Thus $(1, 0)$ satisfies the inequality and is the region we want. We shade what we do not want, though:



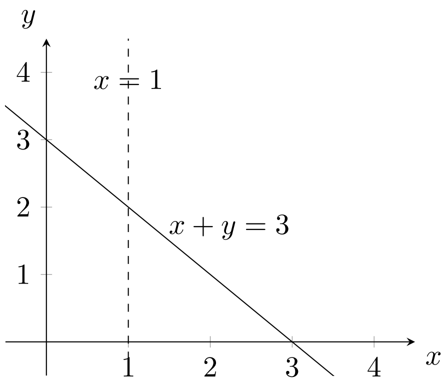
31.3. Representing more than one inequality in the coordinate plane

Now we will focus on representing more than one inequality at the same time. For instance:

$$\begin{cases} x > 1 & (1) \\ x + y \leq 3 & (2) \end{cases}$$

We start by drawing the lines that correspond to the inequalities:

$$\begin{cases} x = 1 & (1) \\ x + y = 3 & (2) \end{cases}$$

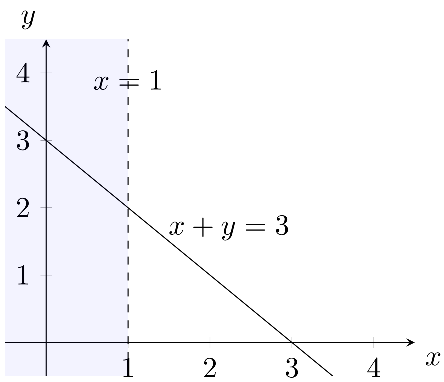


Notice that $x = 1$ is dashed, as the corresponding inequality is $x > 1$. On the other hand, $x + y = 3$ is filled.

Now, we need to shade the *regions we do not want.* Let's start with

$$x > 1$$

We *want* any point with x coordinate bigger than 1, so we shade everything **to the left** of $x = 1$:



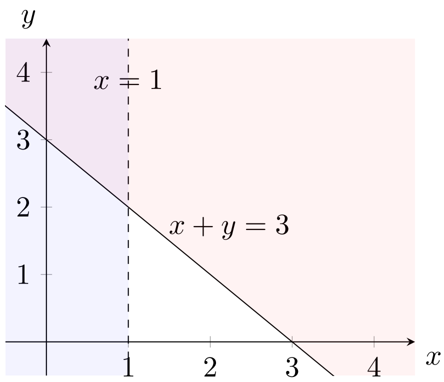
Finally, to shade the region for $x + y = 3$ let's use the $(0, 0)$ test:

$$x + y \leq 3$$

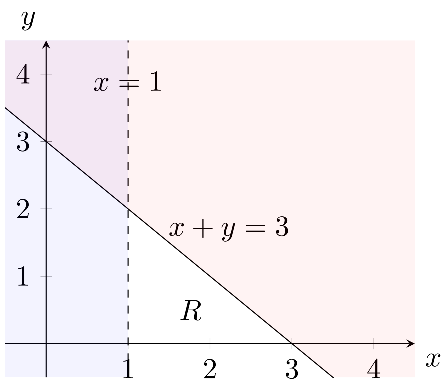
$$0 + 0 \leq 3$$

$$0 \leq 3$$

The point $(0, 0)$ satisfies our inequality, so we shade the region *which does not contain* $(0, 0)$ (everything above the line. Do you notice a pattern with the $<$ and $>$ symbols and what we shade?):



Now we can finally understand why we shade the regions *we do not want*: the region that satisfy all inequalities is the one we didn't shade anything! Let's label it R:

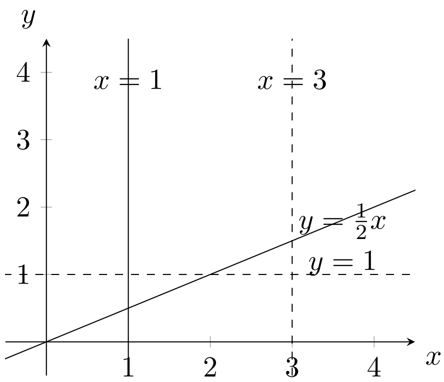


Another example. Let's find the region determined by the following inequalities:

$$\begin{cases} y > 1 & (1) \\ x \geq 1 & (2) \\ x < 3 & (3) \\ y \geq \frac{1}{2}x & (4) \end{cases}$$

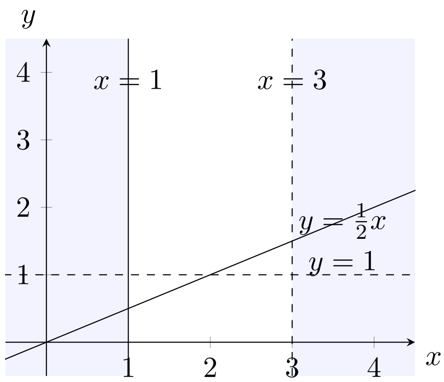
First let's draw the corresponding lines:

$$\begin{cases} y = 1 & (1) \\ x = 1 & (2) \\ x = 3 & (3) \\ y = \frac{1}{2}x & (4) \end{cases}$$

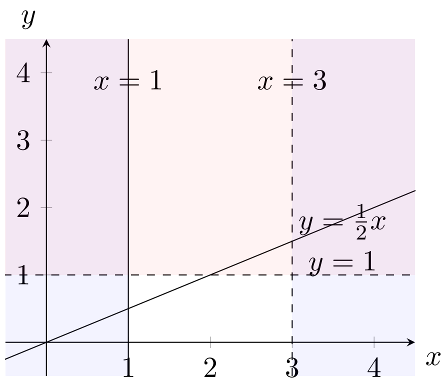


Notice that both $y = 1$ and $x = 3$ are dashed due to their corresponding inequalities having no equal part. Let's now shade the regions we do not want.

Starting with $x \geq 1$, any point with x coordinate greater than 1 is *what we want*, so we shade everything to the left. For $x < 3$, anything with x coordinate smaller than 3 is *what we want*, so we shade everything to the right:



Same reasoning for $y > 1$, any point with y coordinate greater than 1 is what we want, so we shade everything below the line:



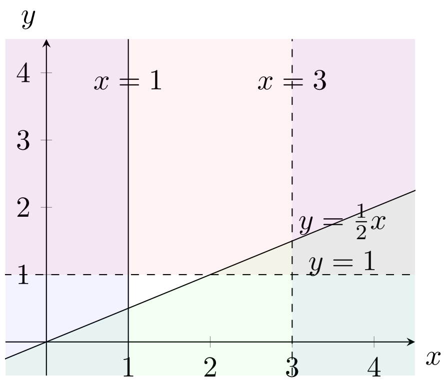
Finally, for $y \geq \frac{1}{2}x$, let's use the point $(2, 0)$:

$$y \geq \frac{1}{2}x$$

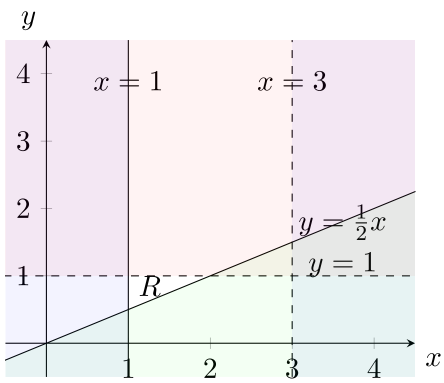
$$0 \geq \frac{1}{2} \times 2$$

$$0 \geq 1$$

The point $(2, 0)$ does not satisfy the inequality, so we shade the region which does contains $(2, 0)$, as we do not want it:



Now, the result is the region which has not been shaded by anything. Let's mark it with a R:



31.4. Linear programming

A certain teacher is going to adopt x cats and y dogs.

He will adopt at least 1 dog, more cats than dogs, more than 1 cat and no more than 6 animals in total.

Not only he's adopting many animals, he wants to discover all the possible number of dogs and cats he can adopt which satisfy the above conditions. He is not well.

Anyway, to do this, we'll first represent the above conditions using inequalities.

He's adopting **at least** 1 dog, which means any number of dogs (y) greater than or equal to 1. In inequality form:

$$y \geq 1$$

More than 1 cat (x) means any number greater than 1:

$$x > 1$$

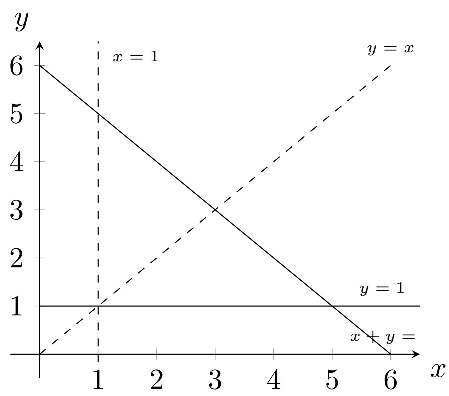
More cats than dogs just means that the number of cats (x) has to be bigger than the number of dogs (y):

$$x > y$$

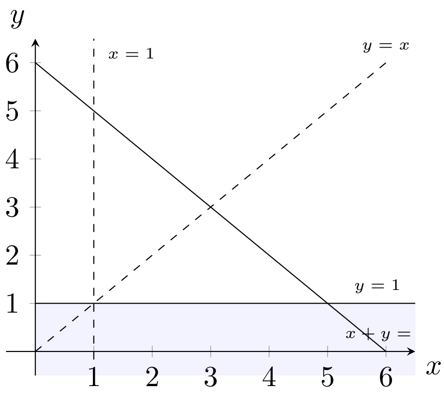
Finally, no more than 6 animals **in total**. This means that adding cats (x) with dogs (y) must be smaller than or equal to 6 (no more than includes 6):

$$x + y \leq 6$$

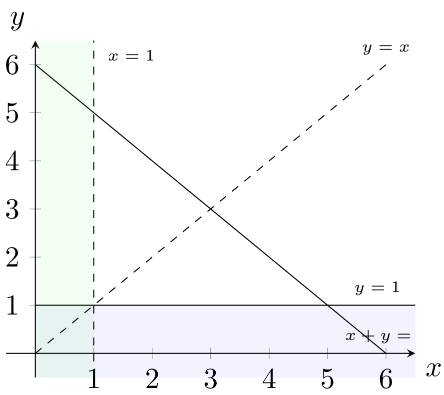
So we have the inequalities we need. In the next graph I already plotted the corresponding lines, and used dashed lines to represent the lines which don't have the 'or equal' part already:



Now comes the shading. Starting with $y \geq 1$, we *want* any point with y coordinate greater than or equal to 1, so let's shade the region which has points with y coordinates smaller than 1 (in blue):



For $x > 1$ we use the same reasoning: any point with x coordinate greater than 1 is what *we want*, so let's shade the region that has points with x coordinates smaller than 1 (in green):

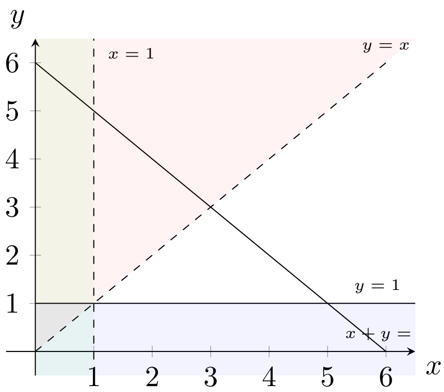


Now, to shade the region for $x > y$ let's use the point $(1, 0)$:

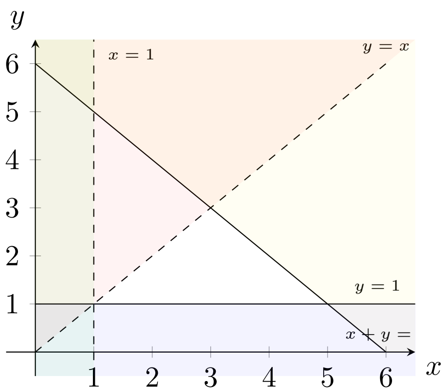
$$x > y$$

$$1 > 0$$

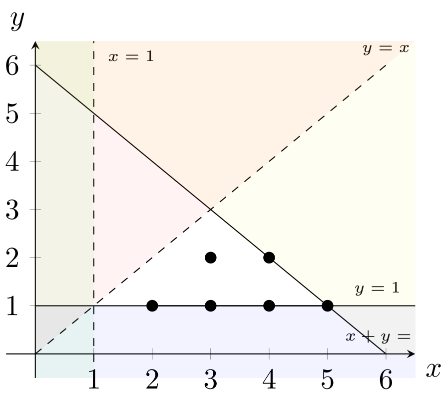
The point $(1, 0)$ satisfies the inequality, so it's in the region we *want*. Let's shade the other side then (in red):



Finally, the region determined by $x+y \leq 6$, we can use the point $(0,0)$, which satisfies the inequality $0+0 \leq 6 \to 0 \leq 6$. Thus, the origin is the region we want, so let's shade the other 'half' of the plane (in yellow):



What this lovely unicorn vomit shows us is the following: any region which is shaded does not satisfy our requirements. The only points which satisfy them are the ones in that triangle in the middle which remains unshaded. We have many points in that triangle that have integer coordinates (it would be hard to adopt $\frac{3}{4}$ of an animal):



To finish our adoption, the teacher needs to decide which of those possibility-point is the *best*. He knows that each cat he adopts will give him 2 happiness and each dog 1 happiness. Clearly he wants to maximise his happiness, that is, get as much of it as he can. If he adopts x cats and y dogs, he gets $2x + y$ happiness. Let's calculate how much happiness he gets in each of those points:

Point	Happiness
(x, y)	$2x + y$
$(2, 1)$	$2 \times 2 + 1 = 5$
$(3, 1)$	$2 \times 3 + 1 = 7$
$(4, 1)$	$2 \times 4 + 1 = 9$
$(5, 1)$	$2 \times 5 + 1 = 11$
$(3, 2)$	$2 \times 3 + 2 = 8$
$(4, 2)$	$2 \times 4 + 2 = 10$

Notice that the point which gives most happiness is $(5, 1)$. Therefore, the teacher should adopt 5 cats and 1 dog.

This thought exercise is very important because it shows us how to solve linear programming questions:

1. Write all the inequalities that represent the constraints of the problem;
2. Plot the line equations corresponding to those inequalities, remembering to draw dashed lines for inequalities with no 'or equal' part;
3. Shade the regions *we do not want* which those inequalities determine;
4. Identify the region which has not been shaded: it satisfies our restrictions;
5. If we have to maximise or minimise the value of an expression, we can try all the possible pairs inside the region, but we don't need to: we just need to try the points near the corners of the region.

In Step 5 you only need to try the points near the corners because if you are in a point inside the region and you can move another unit either up, down, right or left you should, as it would improve your situation. You can repeat this until you reach a corner.

31.5. Exam hints

We have already seen the two types of questions that can appear of this topic. If a questions asks you to identify the inequalities that determine a region, first obtain the line equations that are dividing the plane, then using points (such as $(0,0)$) discover if they are of $<$ or $>$. Remember that dashed lines do not have the 'or equal' part.

Now, for linear programming questions (which usually appear in Paper 4), such as the one in the last section, follow the steps:

1. Write all the inequalities that represent the constraints of the problem;
2. Plot the line equations corresponding to those inequalities, remembering to draw dashed lines for inequalities with no 'or equal' part;
3. Shade the regions *we do not want* which those inequalities determine;
4. Identify the region which has not been shaded;
5. If we have to maximise or minimise the value of an expression try the points near the corners of the unshaded region.

Summary

- To represent inequalities with 2 unknowns, we first plot the straight line they represent, remembering that if the inequality has the 'or equal' part $(\leq, \geq)$ we plot the line *filled* and if does not $(<, >)$ we plot the line *dashed*. To discover what part we want to shade, we substitute any point (x, y) in the inequality: if the inequality is *true*, that point is in the *region we want*; if *not*, the point is in the region *we do not want*. We *always shade the region we do not want*;
- In linear programming questions, follow these steps:
  1. Write all the inequalities that represent the constraints of the problem;
  2. Plot the line equations corresponding to those inequalities, remembering to draw dashed lines for inequalities with no 'or equal' part;
  3. Shade the regions *we do not want* which those inequalities determine;
  4. Identify the region which has not been shaded;
- To maximise or minimise a quantity, try the points near the *corners* of the region you did not shade.

Part IV.

Geometry

32. Angles

33. Symmetry

34. Polygons

35. Circles

36. Pythagoras's theorem

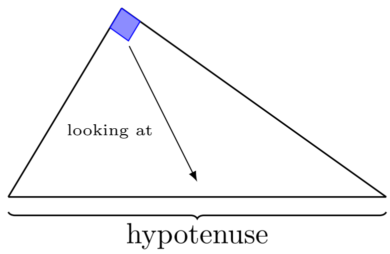
36.1. Why learn Pythagoras's theorem

Pythagoras's theorem is one of the oldest mathematical facts humans have, in many places and independently, discovered. Ian Stewart, in his book *"17 equations that changed the world",* starts his book with Pythagoras's theorem. In all, we have known this result for a long while, and besides practical uses (geometry lends itself to many practical uses), it is probably a fact that the vast majority of educated people know in the world. In that way, Pythagoras's theorem is, arguably, a piece of knowledge that bridges differences and builds empathy: human kind knows the Pythagoras's theorem. In a few minutes, so will you.

By the way, Pythagoras's was not the first to notice the pattern that the theorem describes. Wikipedia says Pythagoras was the first person to record a proof of it.[1]

36.2. Some names

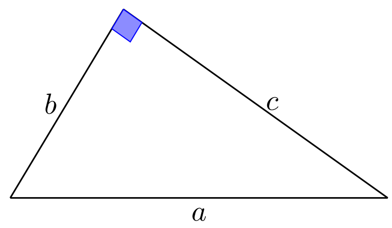
Before we start, let us remind ourselves of a very important name: in a right triangle, the largest side is called the *hypotenuse*. The name comes from Greek, and the most important part of the word is the *hypo*, which means 'under'. Under what, you ask? Under the right angle! The hypotenuse is the side 'under' the right angle, the one the angle is looking at:



36.3. The theorem

The Pythagoras's theorem is very simple to enunciate: given a right triangle, the square of the hypotenuse is equal to the sum of the squares of the other two sides. As always, a drawing helps:

[1] `https://en.wikipedia.org/wiki/Pythagorean_theorem`. Accessed on 15/04/2018.



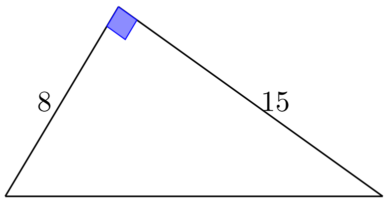
Given that this is a right triangle, we know that

$$\boxed{a^2 = b^2 + c^2}$$

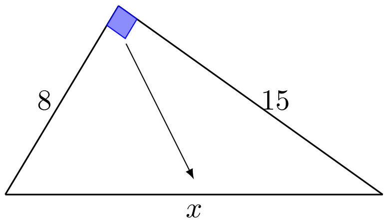
Let me give one of my usual rants: please don't say that the Pythagoras's theorem is "a squared equals b squared plus c squared". These are a bunch of letters (which you can change if you change the variables you are using!). Remember the theorem: the square of the biggest side in a right triangle (the hypotenuse) is equal to the sum of the other two sides squared.

Solved exercise: finding the hypotenuse

Find the length of the hypotenuse in the following right triangle:



Solution: The only difficulty is finding the hypotenuse, but you just have to remember that is always the side the right angle is facing. Let's call its length x on the drawing:



Let's apply Pythagoras's:

$$\text{hypotenuse}^2 = \text{sum of squares of the other sides}$$

$$x^2 = 8^2 + 15^2$$

$$x^2 = 64 + 225$$

$$x^2 = 289$$

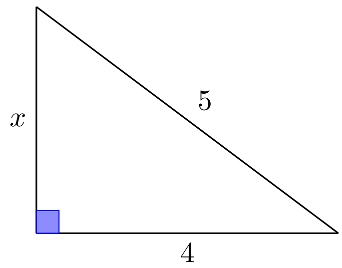
$$\sqrt{x^2} = \sqrt{289}$$

$$x = 17$$

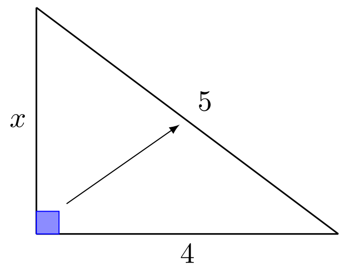
The hypotenuse measures 17. Notice that the negative solution of the square root does not make sense as we cannot have negative side lengths.

Solved exercise: finding other sides

Find x in the right triangle:



Solution: Again, let's identify the hypotenuse:



We can now apply Pythagoras's:

$$5^2 = x^2 + 4^2$$

$$25 = x^2 + 16$$

$$25 - 16 = x^2 + 16 - 16$$

$$9 = x^2$$

$$\sqrt{9} = \sqrt{x^2}$$

$$3 = x$$

The side measures 3. As you can see, there are no "3 versions" of Pythagoras's depending on which side you want so solve for, you just solve equations.

36.3.1. It goes back!

Normally, we say that in a right triangle, the hypotenuse squared is equal to the sum of the other two sides squared. However, the converse is also true: if you have a triangle and the biggest side squared is equal to the sum of the other two sides squared, then the triangle is a right triangle! You can write that like this:

Right triangle $\Longleftrightarrow$ square of the biggest side is equal to the sum of the squares of the other sides

As you can see, you can go from left to right, which is saying that if you have a right triangle, then the square of the hypotenuse (the biggest side) is equal to the sum of the squares of the two other sides. But you can also go from right to left, which is saying that if you have a triangle in which the square of the biggest side is equal to the sum of the squares of the other two sides, that means this triangle is a right triangle.

Solved exercise: showing that a triangle is a right triangle

Show that a triangle with sides 5, $\sqrt{24}$ and 7 is a right triangle.

Solution To show that a triangle is a right triangle, we just need to show that the square of the biggest side is equal to the sum of the squares of the two other sides. Therefore, we just need to identify the biggest side:

$$5 = 5$$

$$\sqrt{24} = 4.899$$

$$7 = 7$$

Therefore, the biggest side measures 7. Let's see if the square of 7 is equal to the sum of the squares of the other sides:

$$7^2 = 49$$

Adding the squares of the other sides:

$$5^2 + \left(\sqrt{24}\right)^2 = 25 + 24 = 49$$

They are equal! Therefore, the triangle with sides 5, $\sqrt{24}$ and 7 is a right triangle.

Solved exercise: showing that a triangle is a right triangle

Given two integers m and n which satisfy $m > n > 0$, show that a triangle with sides $m^2 - n^2$, $2mn$ and $m^2 + n^2$ is a right triangle.

Solution Again, we need to show that the square of the biggest side is equal to the sum of the other two sides squared. Therefore, we need to find the biggest side. We don't know the values for m and n, so we need to think a bit harder this time.

First, you have to agree that $m^2 + n^2$ has to be bigger than $m^2 - n^2$. Both m^2 and n^2 are positive numbers, and adding two positive numbers must be bigger than their difference. Now, who is bigger, $m^2 + n^2$ or $2mn$? You could do some trial and error (such as $m = 3$ and $n = 2$), but that's boring. Let's try to prove that $m^2 + n^2 > 2mn$. As we know that $m > n$, we know that $m - n > 0$. Let's square both sides of the inequality:

$$(m - n)^2 > 0^2$$

$$m^2 - 2mn + n^2 > 0$$

$$m^2 + n^2 > 2mn$$

There you go: $m^2 + n^2$ is the biggest side. Let's now see if when we square it, we get the sum of the other two sides squared.

$$(m^2 + n^2)^2 = m^4 + 2m^2n^2 + n^4$$

The other sides squared:

$$(m^2 - n^2)^2 + (2mn)^2 = m^4 - 2m^2n^2 + n^4 + 4m^2n^2 = m^4 + 2m^2n^2 + n^4$$

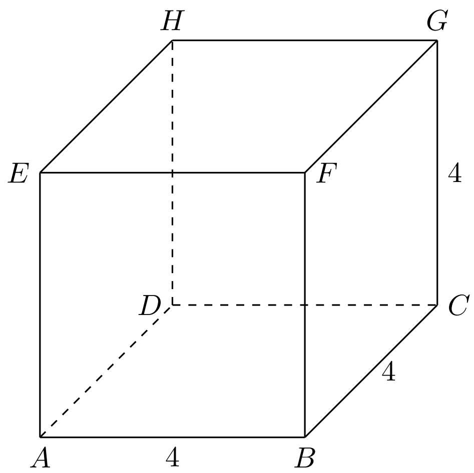
They are equal! Therefore the triangle with sides $m^2 - n^2$, $2mn$ and $m^2 + n^2$ is a right triangle[a].

[a] By the way, this is one possibility of generating right angle triangle sides: pick m and n and calculate the sides.

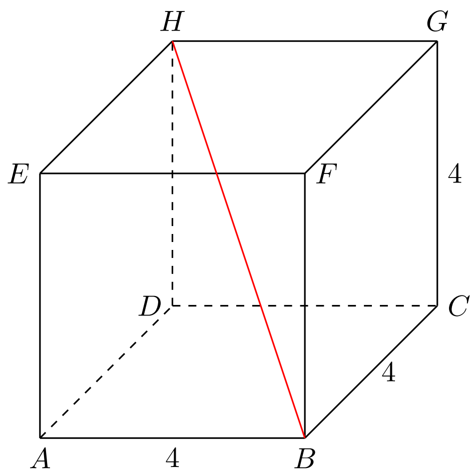
36.4. Pythagoras's applications in 3D geometry

A very interesting use of Pythagoras's theorem is to calculate distances between points in three dimensional shapes.

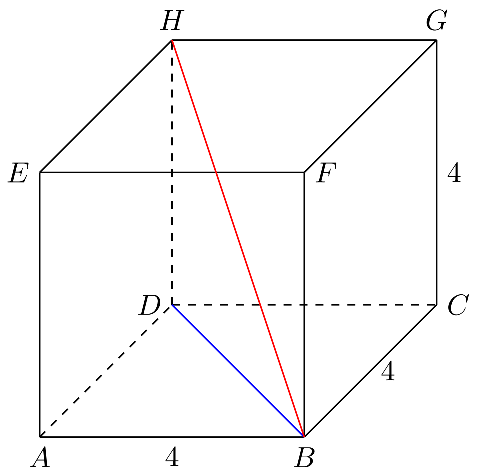
As an example, let's say we have the following cube $ABCDEFGH$:



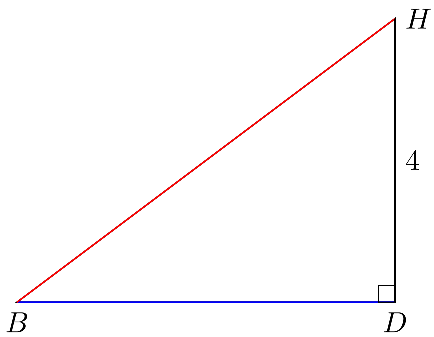
and we want to find the length of the segment BH:



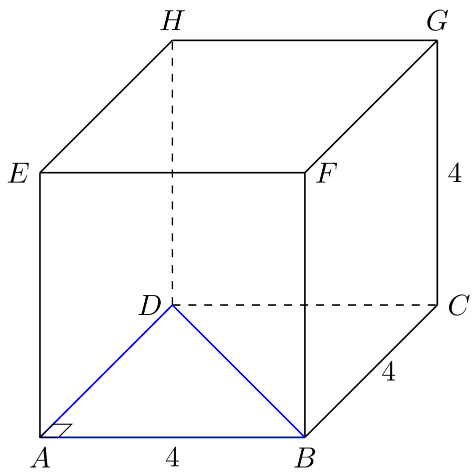
We can find it applying Pythagoras's twice. Before we do that, however, it is important to understand what is the *projection* of the segment BH on the base $ABCD$ of the cube. You can think of it as 'the shadow' of BH on the base if light was shining the cube from above:



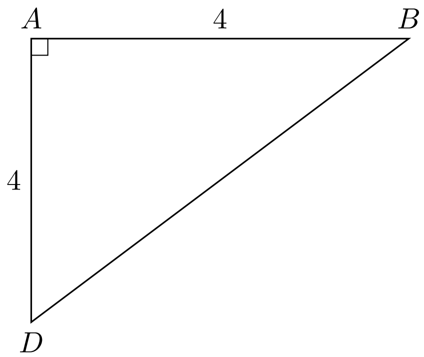
The blue segment, BD, would be that shadow. Therefore we have the right triangle BDH ‘inside’ the cube:



BDH is a right triangle as all angles in a cube are right. We already know that $DH = 4$, so if we knew the length of BD we could find BH by applying Pythagoras’s on BDH. We can, however, find BD by applying Pythagoras’s in another triangle, ABD:



Given that $ABCD$ is a square, all its angles have to be right. Therefore, we have the following triangle:



We can now apply Pythagoras's in ABD[2]:

$$BD^2 = AD^2 + AB^2$$

$$BD^2 = 4^2 + 4^2$$

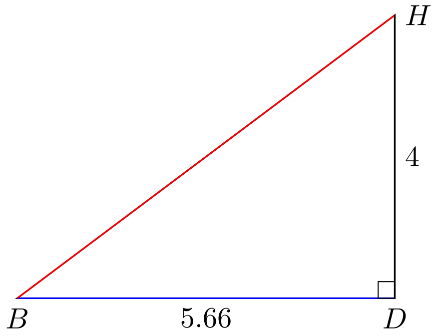
$$BD^2 = 16 + 16$$

$$BD^2 = 32$$

$$BD = 5.6568$$

Now that we know the length of BD, we can use it in our triangle BDH:

[2]To be honest, you don't need BD itself, only its square, as we are using it later in another Pythagoras's calculation.



Again applying Pythagoras's:

$$BH^2 = AD^2 + BD^2$$

$$BH^2 = 4^2 + 5.6568^2$$

$$BH^2 = 16 + 32$$

$$BH^2 = 48$$

$$BH = 6.93 \text{ (3 s.f.)}$$

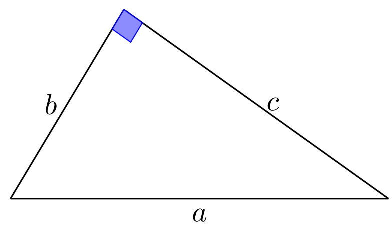
As you can see, the question was about finding right triangles 'inside' the cube and using Pythagoras's normally.

36.5. Exam hints

Let's assume you see a right angle triangle during your exam: it is very likely that you will be using Pythagoras's there (or the trigonometric relations)! My only hint is: identify the hypotenuse and let the theorem magic work itself.

Summary

- The **hypotenuse** of a right triangle is the biggest side, the one in front of the right angle;
- You can apply **Pythagoras's theorem** only to **right triangles**. The theorem states that **the square of the hypotenuse is equal to the sum of the squares of the other sides.** A drawing is worth 17 words:



$$\boxed{a^2 = b^2 + c^2}$$

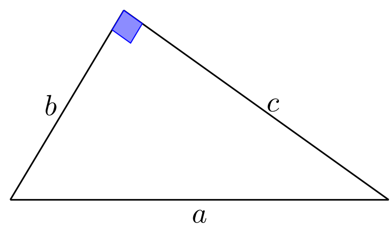
- In some cases it is important to find right angle triangles 'inside' 3D shapes. After you 'isolate' them, just apply Pythagoras's normally.

Formality after taste

A proof of the Pythagoras's theorem

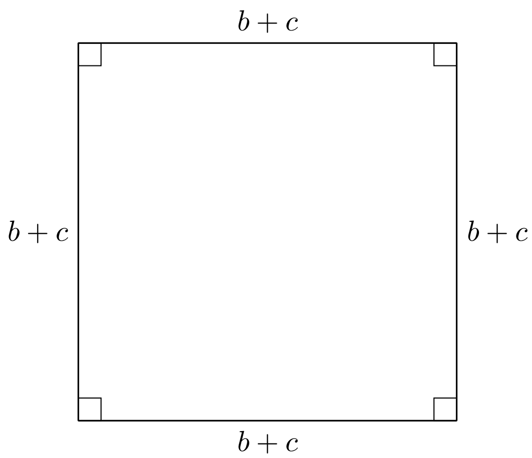
There are many different proofs of the Pythagoras's theorem[3]. I will show you my favourite, which is very simple.

We want to prove that, given a right angle triangle with hypotenuse a and other sides b and c, that $a^2 = b^2 + c^2$. Let us draw the triangle again to help:

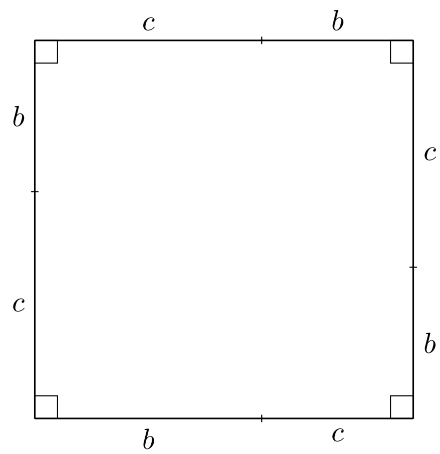


The idea is, as usual when in doubt, to make a drawing! Let's draw a square with side length $b + c$:

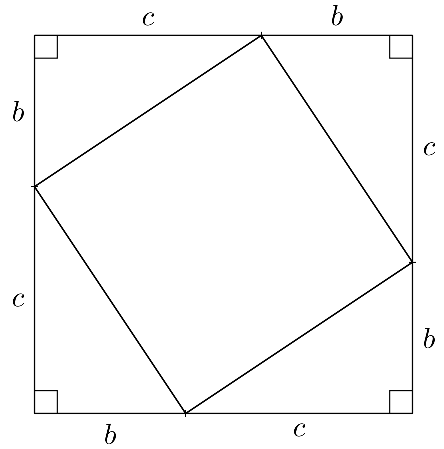
[3] There is a book called *The Pythagorean Proposition*, by Loomis, which has 367 different proofs of it!



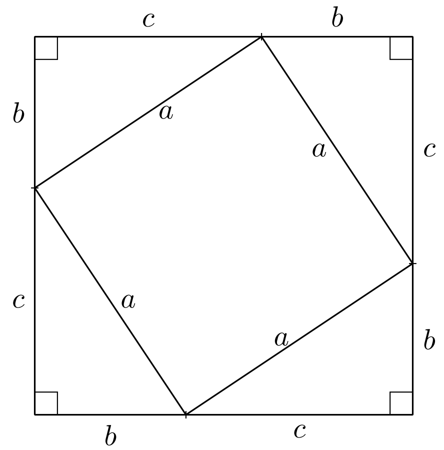
Now, let us split the sides into two parts each: one with length b and one with length c:



Now, let us join the points that are dividing the sides:



If you apply Pythagoras's to any of those triangles with sides b and c, you will find that their hypotenuse must be a:



The set up is complete. Let's calculate some areas now.

First, the area of the big square (the one that has side $b+c$). Let's call this area A_{big}:

$$A_{\text{big}} = (b+c)^2$$

The area of each our triangles is given by the usual base times height. Let's call this area A_t:

$$A_t = \frac{b \times c}{2}$$

Finally, the area of the small square inside. Let's call this area A_{small}:

$$A_{\text{small}} = a^2$$

Now, do you agree that the area of the big square, A_{big}, has to be equal to the area of the small square, A_{small} plus the area of the four triangles, A_t? In mathematicianese:

$$A_{\text{big}} = A_{\text{small}} + 4A_t$$

Let's do some equation solving now:

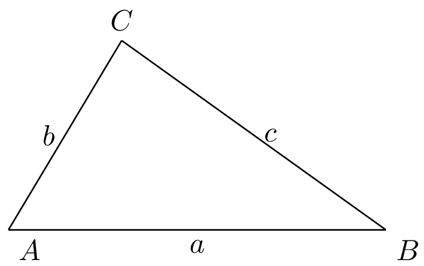
$$\begin{aligned}
A_{\text{big}} &= A_{\text{small}} + 4A_t \\
(b+c)^2 &= a^2 + 4 \times \left(\frac{b \times c}{2}\right) && \text{Substituting our formulas} \\
(b+c)(b+c) &= a^2 + \cancelto{2}{4} \times \left(\frac{b \times c}{\cancelto{1}{2}}\right) && \text{Expanding the brackets} \\
b^2 + cb + bc + c^2 &= a^2 + 2bc \\
b^2 + 2bc + c^2 &= a^2 + 2bc \\
b^2 + 2bc + c^2 - 2bc &= a^2 + 2bc - 2bc && \text{Subtracting } 2bc \text{ on both sides} \\
b^2 + \cancel{2bc} + c^2 - \cancel{2bc} &= a^2 + \cancel{2bc} - \cancel{2bc} && \text{The usual killing} \\
b^2 + c^2 &= a^2
\end{aligned}$$

There you go! We just showed that

$$\boxed{a^2 = b^2 + c^2}$$

A proof of Pythagoras's theorem converse

This time, we want to show that if, in a triangle, the square of the largest side squared is equal to the sum of the squares of the other two sides, then that triangle must be a right one. That is to say, in a triangle:



if we have that

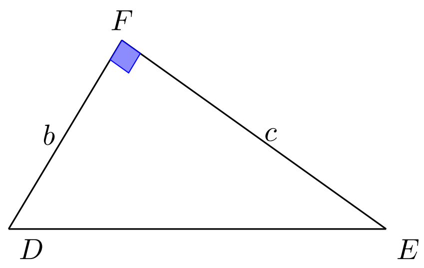
$$a^2 = b^2 + c^2$$

then angle ACB must be a right one.

The proof is quite simple. We start by drawing a right triangle DEF, in which

$$DF = b$$

$$EF = c$$



The triangle DEF is a right triangle, so we can apply Pythagoras's to it:

$$DE^2 = DF^2 + EF^2$$

$$DE^2 = b^2 + c^2$$

But we know that $b^2 + c^2$ is equal to a^2 by hypothesis:

$$DE^2 = b^2 + c^2$$

$$DE^2 = a^2$$

$$DE = a$$

Thus, $DE = a = AB$. Given that all sides of the triangles ABC and DEF are equal, they are congruent. Hence their angles must be equal to, implying that angle ACB must be 90°.

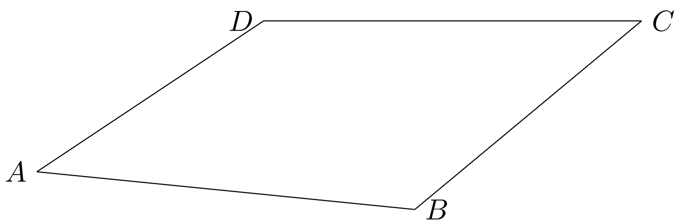
37. Similarity and congruence

38. Prisms

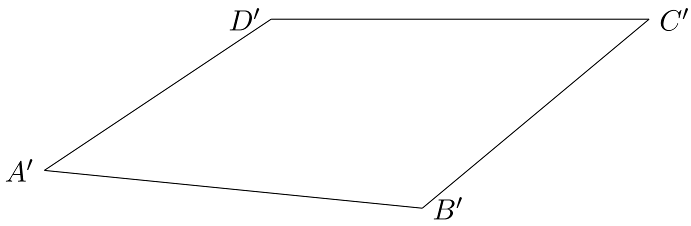
38.1. Why learn about prisms

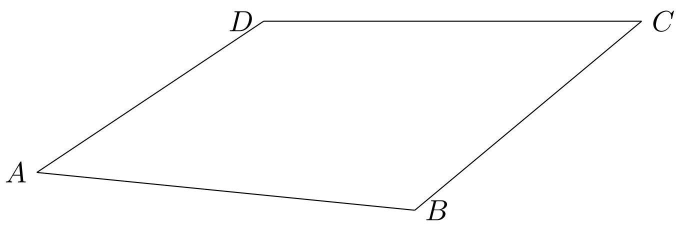
38.2. What is a prism

Let us start with a random two dimensional shape:

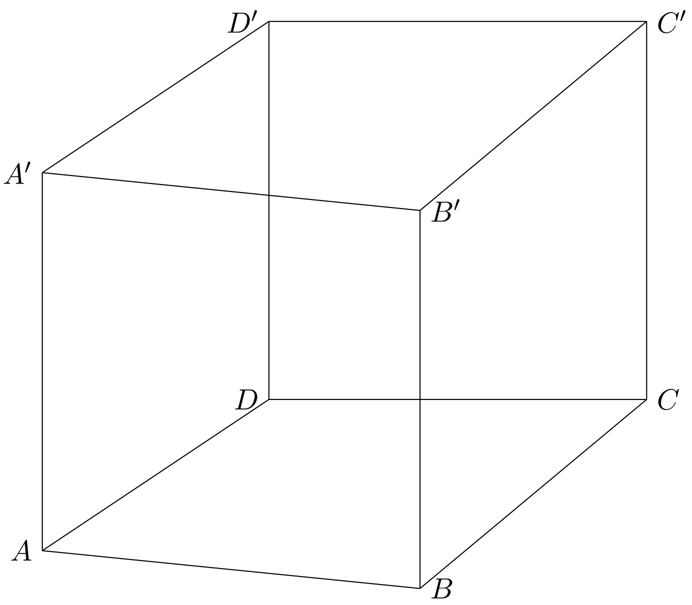


and we know make a copy of it:

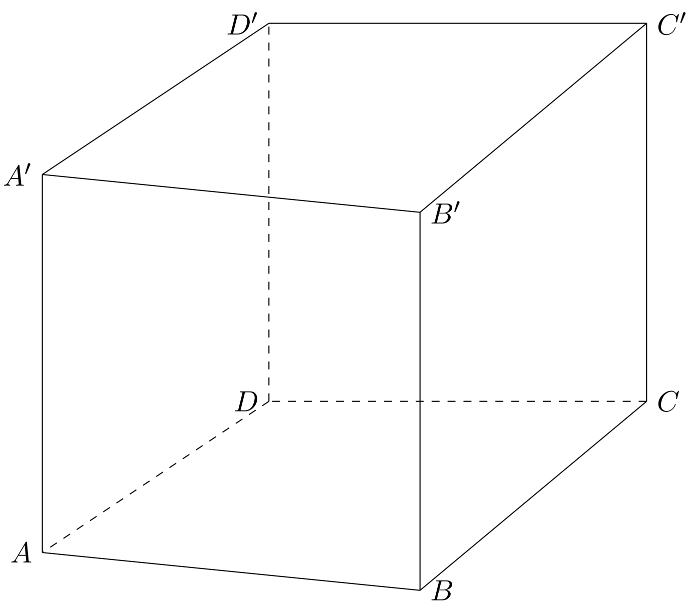




and now we connect the vertices that correspond to each other (A to A', B to B' and so on) in each shape:



and we build a three dimensional shape. Let us dot the lines behind faces to see it better:



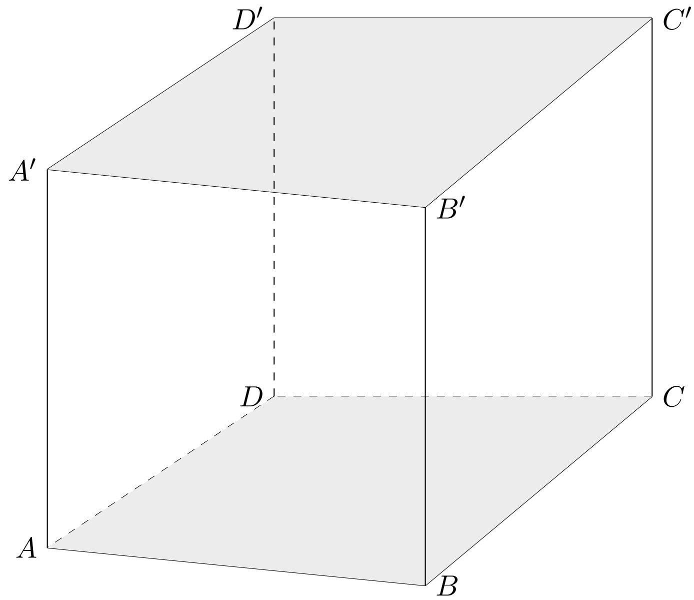
and we obtain a prism!

Formally, a prism is a solid composed of a polygonal face which we *translate* (that is, copy) to have a second base which is exactly equal to the first; we join the corresponding points in each base to compose a prism.

In less formal terms, a prism is when we take a shape with some sides and give it some "depth", that is, we drag it somewhere else.

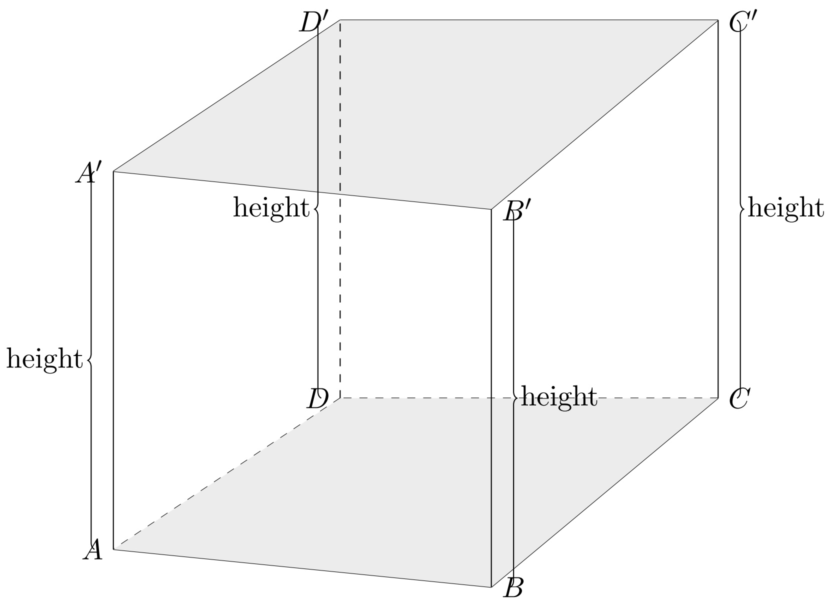
38.3. Key terms: base, height and cross-section

In a prism, we have the two bases, which I shaded here:

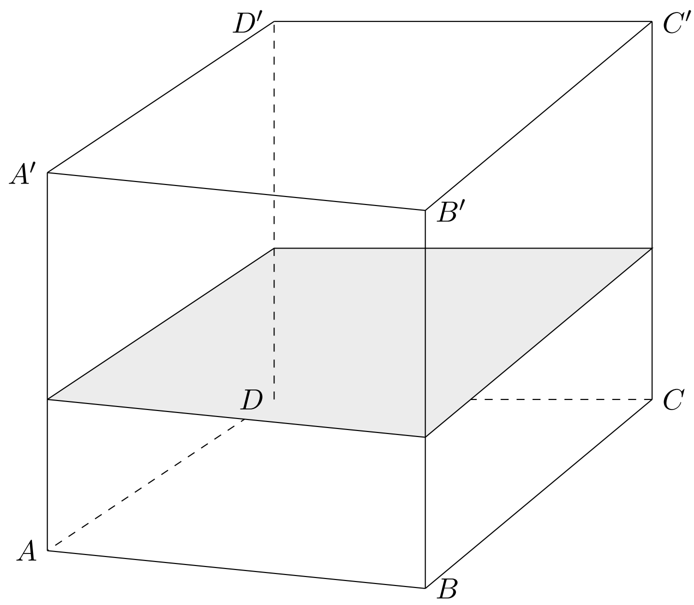


the bases of a prism are always parallel to one another.

The distance between the two bases is called the *height* of the prism, and it is easier to see the height by looking at the length of the edges that connect the bases:



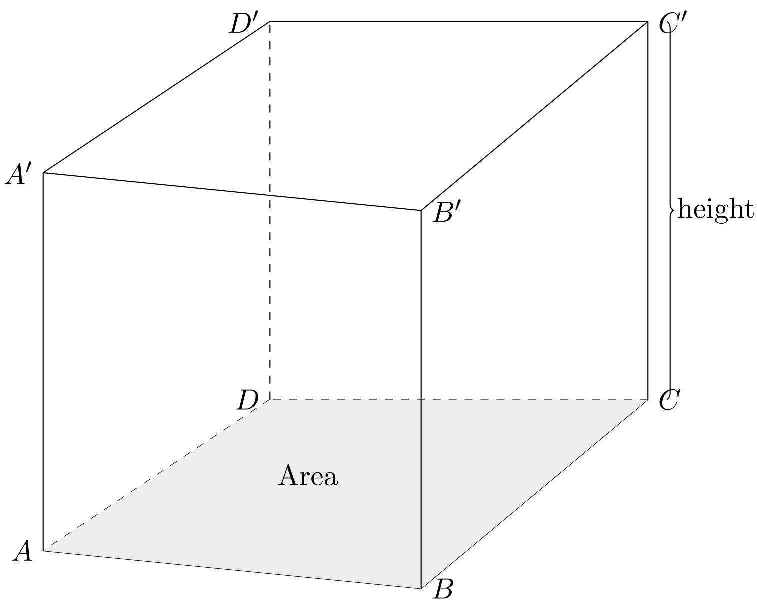
Finally, if you cut the prism anywhere to make a "fake face" inside the prism, such as:



we define what is called a *cross-section* of the prism. Notice that all cross-sections of a prism are identical copies of the base faces.

38.4. Volume of prisms

To find the volume of a prism (any prism), we simply have to multiply the area of their base (or of a cross-section) by the height of the prism:



In formula:

$$\text{Volume} = \text{area of base} \times \text{height}$$

or, more concisely

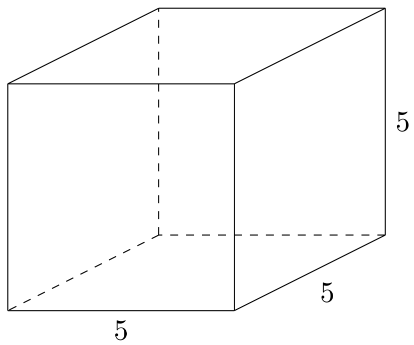
$$V = B \times H$$

where V is the volume, B the area of the base and H the height of the prism.

Let us do some classic ones first.

38.4.1. Volume of cubes

A *cube* is a prism which has a square base and height of the same length as the side of the square: that is a nice "square box". For instance, one that has edges of length 5:



and we can use the formula for volume of prisms (area of base times height):

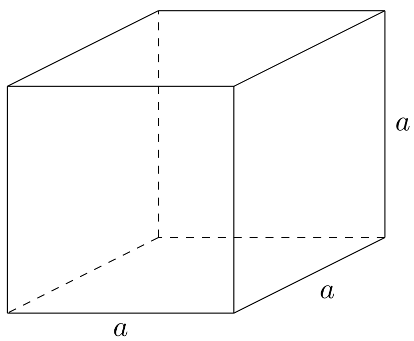
$$V = B \times H$$

The base is a square of side 5, so its area is 5^2, and the height is also 5, thus

$$V = 5^2 \times 5 = 5^3 = 125$$

Which is nice to confirm: the volume of a *cube* with side 5 is equal to 5 *cubed*.

In general, a cube with side a has volume given by a^3:



as we have a base as a square with side a, which has area a^2 and height a:

$$V = B \times H$$

$$V = a^2 \times a$$

$$V = a^3$$

Of course, you can always do the inverse. Say we have a cube with volume 216 and we want to find its side length, which we shall call x:

$$V = B \times H$$

$$216 = \underbrace{x^2}_{\text{area of square}} \times \underbrace{x}_{\text{height}}$$

$$216 = x^3$$

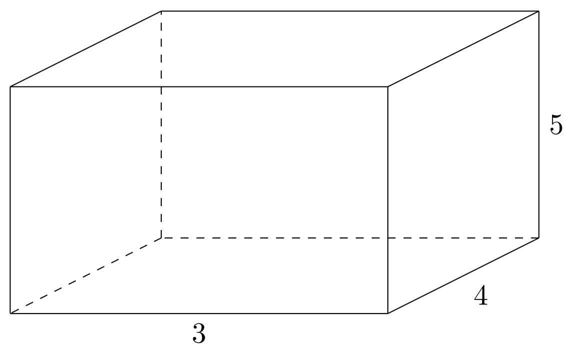
$$\sqrt[3]{216} = \sqrt[3]{x^3}$$

$$6 = x$$

38.4.2. Cuboids

A *cuboid* is a prism that *looks like* a cube, but it has different side lengths. In reality, it is any cube has a rectangle as a base (of which the cube is a special case).

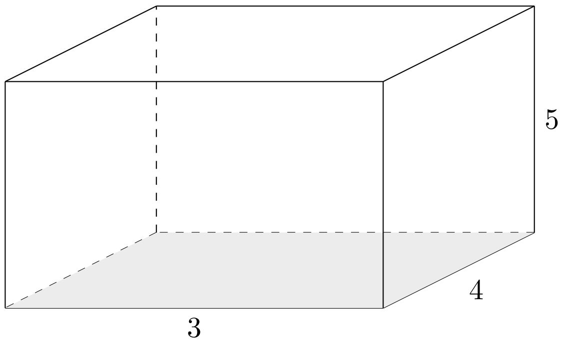
For instance, we can have a cuboid with side lengths 3, 4 and 5:



and we can find its volume by, again, multiplying the area of the base by the distance between bases, the height:

$$V = B \times H$$

Using the rectangle with sides 3 and 4 as a base:



we have the distance between it and the "top" is 5, hence:

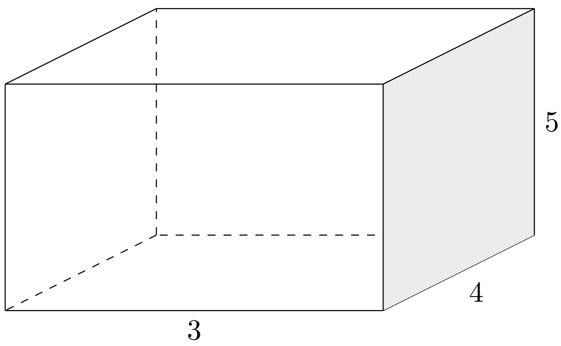
$$V = B \times H$$

$$V = \underbrace{3 \times 4}_{\text{rectangle 3 by 4}} \times 5$$

$$V = 12 \times 5$$

$$V = 60$$

A very interesting thing in cuboids is that you can choose any pair of opposite sides as bases. For instance, if we choose the face in the right as the base:



and now the the base is a 4 by 5 rectangle and the distance between the right and left faces is 3:

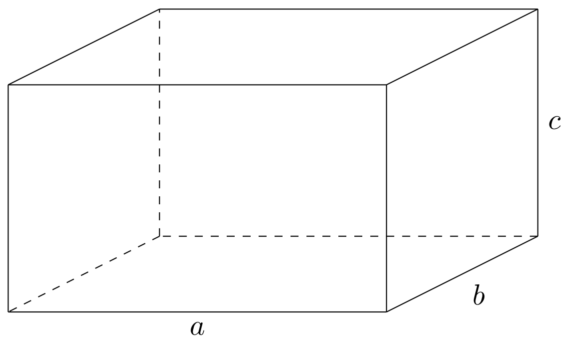
$$V = B \times H$$

$$V = \underbrace{4 \times 5}_{\text{rectangle 4 by 5}} \times 3$$

$$V = 20 \times 3$$

$$V = 60$$

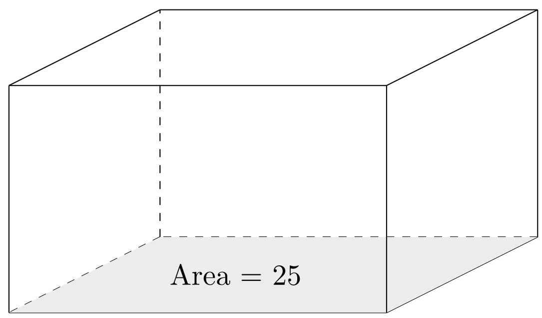
In general, if we have a cuboid of sides with length a, b and c:



its volume is given by

$$V = abc$$

Of course, you can always do the reverse. Say we have the cuboid below:



and we are told its volume is 100 and to find its height. As we know the area of any prism is given by the area of the base times the height, we have

$$V = B \times H$$

$$100 = 25H$$

$$\frac{100}{25} = \frac{25H}{25}$$

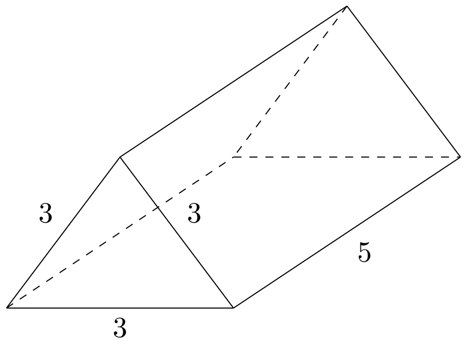
$$4 = H$$

So its height is 4.

38.4.3. Triangular prisms

Triangular prisms are the ones that have triangles as their base. Fairly common in exams, so excellent to know about.

For instance, here a prism that has an equilateral triangle as its base:



notice that I have drawn the prisms with its bases facing you: the rectangle "supporting" the prism is *not* a base! This is a very common mistake, so do be careful.

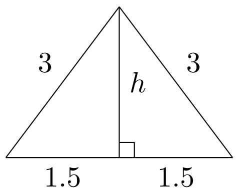
As usual, we can find the volume of this prism multiplying the area of its base, in this case the equilateral triangle with side 3, by its height, 5:

$$V = B \times H$$

$$V = \text{Area of triangle} \times 5$$

and here we reach the annoying part of volume calculations: sometimes it is a little harder to find the area of the base.

There are many ways to find the area of an equilateral triangle (even a formula, as usual. See 38.6). Here let us find it by finding its height using Pythagoras:



It is very important to remember that in isosceles triangles (which the equilateral is a special case) the height divides the base into two equal parts. Hence the two parts of 1.5. Now, let us use Pythagoras's (see Chapter 36) on the right half of the triangle:

$$\text{hypotenuse}^2 = \text{side}^2 + \text{other side}^2$$

$$3^2 = h^2 + 1.5^2 \qquad \text{Subsituting}$$

$$h^2 = 3^2 - 1.5^2 \qquad \text{Subtracting } 1.5^2 \text{ on both sides}$$

$$h = \sqrt{3^2 - 1.5^2} \qquad \text{Square root on both sides}$$

$$h = \frac{3\sqrt{3}}{2} = 2.598... \qquad \text{Calculator}$$

Now that we have the height of the triangle, we can find its area. Here, be very careful: we will use the base and height of the *triangle*, and this triangle is the base of the *prism*, which has a different height still! The terms are the same but referring to different objects.

$$\text{Area of triangle} = \frac{\text{base} \times \text{height}}{2}$$

$$A = \frac{3 \times 2.598}{2} = \frac{9\sqrt{3}}{4} = 3.897...$$

Finally, we can find the volume of the original prism:

$$V = B \times H$$

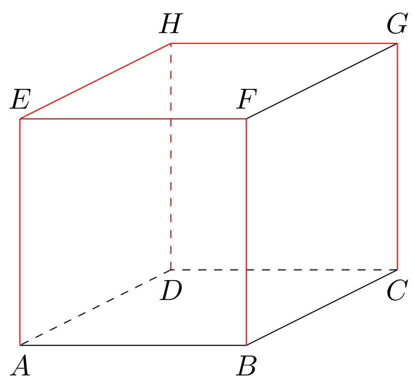
$$V = \underbrace{3.897}_{\text{area of triangular base}} \times 5$$

$$V = \frac{45\sqrt{3}}{4} = 19.5 \text{ (3s.f.)}$$

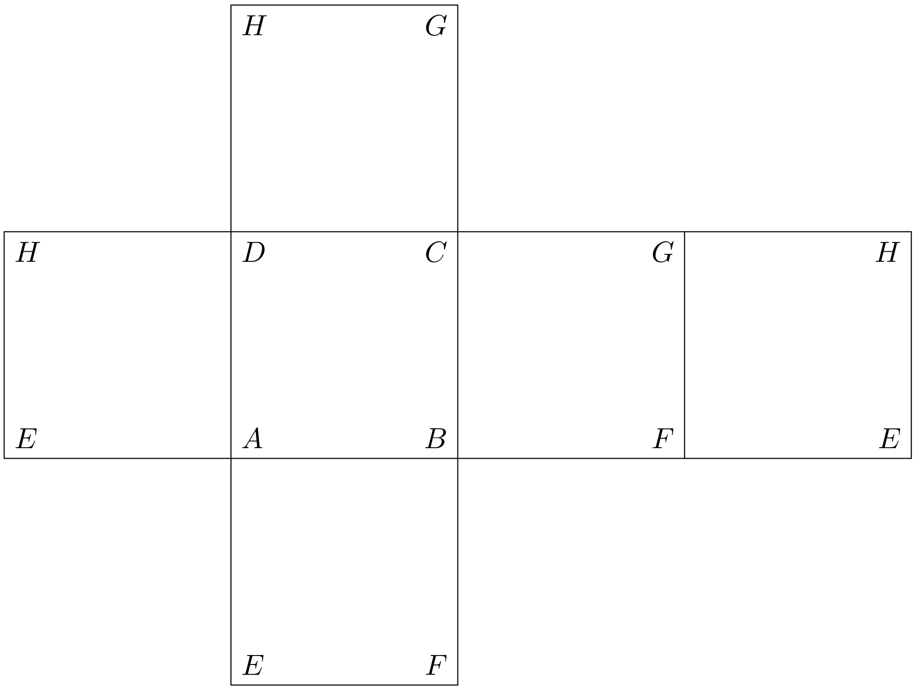
38.5. Nets and surface area of prisms

38.5.1. Nets

A *net* of a solid (a 3d shape) is when we "open it up" and draw its faces on the plane. For instance, we can start with a cube and "cut through" the edges I have coloured in red:



and we can "put the faces" on a plane like this, for instance:



You can picture this as the face $ABCD$ staying glued to a table, and if we joined the vertices with same name we would go back to the original cube.

38.5.2. Surface area

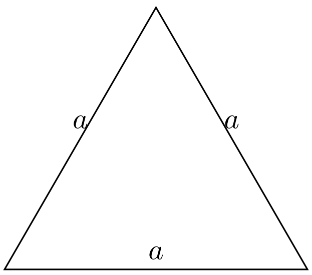
38.6. Exam hints

Summary

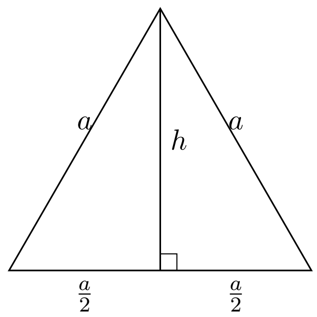
Formality after taste

A formula for the area of an equilateral triangle

Say that we have an equilateral triangle of side a:



Let us do this without trigonometry. First, let us find an expression for the height of the triangle. Remember that the height of an isosceles triangle divides the base into two equal parts, and that equilateral triangles are also isosceles. Hence, we have:



Let us use our favourite Pythagoras's on the half to the right:

$$a^2 = h^2 + \left(\frac{a}{2}\right)^2$$

$$h^2 = a^2 - \frac{a^2}{4}$$

$$h^2 = \frac{3a^2}{4}$$

$$h = \sqrt{\frac{3a^2}{4}}$$

$$h = \frac{a\sqrt{3}}{2}$$

Thus, the height of an equilateral triangle with side a is given by the side times root 3 divided by 2. I have always liked this formula for some strange reason.

We can now find a formula for the area of the triangle:

$$A = \frac{\text{base} \times \text{height}}{2}$$

$$A = \frac{a \times \frac{a\sqrt{3}}{2}}{2}$$

$$A = \frac{a^2\sqrt{3}}{2} \times \frac{1}{2}$$

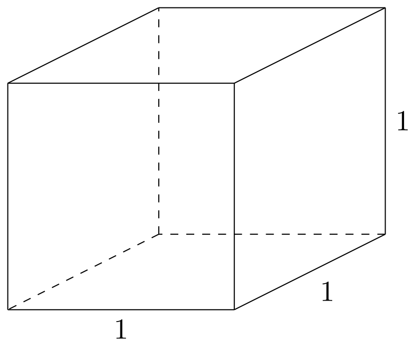
$$A = \frac{a^2\sqrt{3}}{4}$$

and there we have it: just plug a into the formula and find the area of an equilateral triangle!

An intuition on why the volume of a cuboid of sides a, b and c is abc

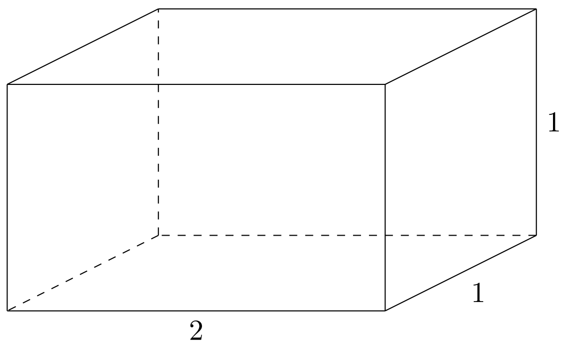
(All this discussion is based on the one in “A Matemática do Ensino Médio”, Volume 2, by Lima, Carvalho, Wagner and Morgado.)

We start from the basic unity of volume, the cube with side 1, which has volume equal to 1:

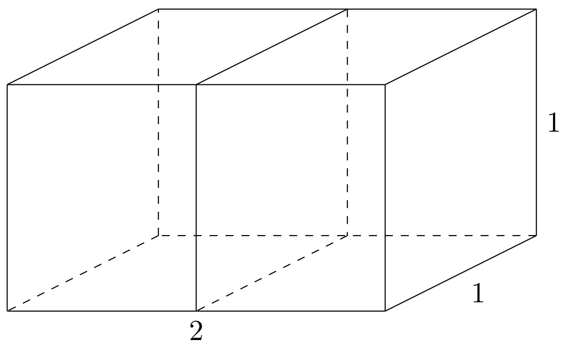


A cuboid of sides a, b and c is obtained by multiplying each side of the cube by a, b and c.

Now, intuitively, if we start with the cube of side 1 and just change *one* of its sides, say to 2, we obtain a cuboid:



which we can split in two equal cubes of side 1:



and, as the volume of each cube is 1, the volume of the cuboid must be $1 + 1 = 2$. This is true if we keep any two sides constant and multiply the other by a value: if we multiply one dimension of the cube by a and keep the others 1, the volume of the cuboid is now a. If we now keep one side as a, one as 1 and multiply the other by b, the new cuboid will have volume ab; finally, keeping a and b constant and multiplying the other side by c we have our final cuboid with volume abc.

If the rectangle with sides a and b is the base of the prism, then c is the distance between bases, the height. Thus, we have:

$$V = abc$$

$$V = (ab)\,c$$

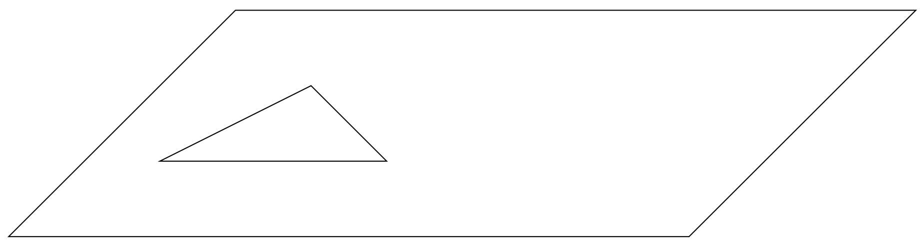
$$V = \text{area of the base} \times \text{height}$$

A proof that the volume of a prism is area of the base multiplied by height

We will use the very strong Cavalieri's principle to show (very easily) the result. You do need to be well versed in the content of Appendix VII.

The only we will assume the "intuitively shown" fact that the cuboid has volume equal to the product of its base area by its height.

Take any prism with a polygonal face of area A and put it on a plane:



39. Pyramids

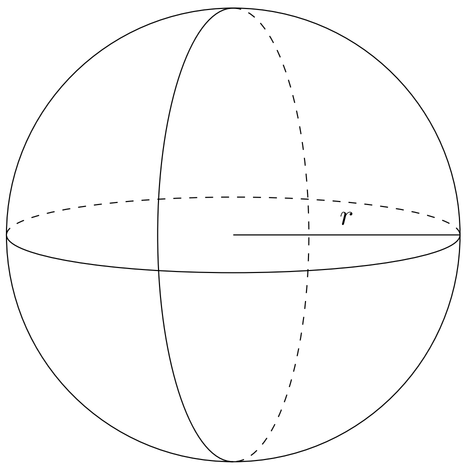
40. Spheres

40.1. Why learn about spheres

Well, a sphere is the shape our world resembles (as a curiosity, the Earth has an ellipsoid shape).

40.2. The radius of a sphere

Much like a circle, a sphere only needs its radius to be determined. And also like a circle, the radius is the distance from the centre of the sphere to any point in its surface:



40.3. The surface area of a sphere

To find the surface area of a sphere with radius r we use the formula:

$$\text{Surface Area} = 4\pi r^2$$

So, if we have a sphere with radius 3, its surface area is given by:

$$\text{Surface Area} = 4\pi r^2$$

$$\text{SA} = 4 \times \pi \times 3^2 = 4 \times \pi \times 9 = 36\pi = 113 \text{ (3s.f.)}$$

Sometimes, they give you the area and want you to find the radius of the sphere. For instance, say that a sphere has surface area 100π. To find its radius we solve an equation:

$$
\begin{aligned}
\text{SA} &= 4\pi r^2 \\
100\pi &= 4\pi r^2 && \text{Substituting} \\
100\cancel{\pi} &= 4\cancel{\pi} r^2 && \text{Cancelling } \pi \text{ on both sides} \\
100 &= 4r^2 \\
\frac{100}{4} &= \frac{4r^2}{r} && \text{Dividing by 4} \\
25 &= r^2 \\
\sqrt{25} &= \sqrt{r^2} && \text{Square root on both sides} \\
5 &= r
\end{aligned}
$$

40.4. The volume of a sphere

To find the volume of a sphere with radius r we use the formula

$$V = \frac{4}{3}\pi r^3$$

For instance, say that a sphere has radius 2. We can use the formula to find its volume:

$$V = \frac{4}{3}\pi r^3$$

$$V = \frac{4}{3} \times \pi \times 2^3$$

$$V = \frac{4}{3} \times \pi \times 8 = \frac{32}{3}\pi = 33.5(3\text{s.f.})$$

Of course, they can also give you the volume and ask you to find the radius. If a sphere has volume 36, we again can set up an equation to find its radius:

$$
\begin{aligned}
V &= \frac{4}{3}\pi r^3 \\
36 &= \frac{4}{3}\pi r^3 && \text{Substituting the volume} \\
36 \times 3 &= \frac{4}{3}\pi r^3 \times 3 && \text{Multiplying both sides by 3} \\
108 &= 4\pi r^3 \\
\frac{108}{4\pi} &= \frac{4\pi r^3}{4\pi} && \text{Dividing both sides by } 4\pi \\
8.59436 &= r^3 \\
\sqrt[3]{8.59436} &= \sqrt[3]{r^3} && \text{Cube rooting both sides} \\
2.05 &= r
\end{aligned}
$$

40.5. Exam hints

In both IGCSE 0580 and 0607 you do not need to memorize the formulas for spheres, they are given to you.

Summary

- A sphere only needs its *radius* to be determined (the radius is the distance from the centre to any point on the surface of the sphere);
- The surface area of a sphere of radius r is given by the formula

$$\text{SA} = 4\pi r^2$$

- The volume of a sphere of radius r is given by the formula

$$V = \frac{4}{3}\pi r^3$$

Formality after taste

As usual when there are formulas involved, I would like to show you a derivation of them. This whole discussion is the same as the one in "A Matemática do Ensino Médio", Volume 2, by Lima, Carvalho, Wagner and Morgado.

A "derivation" of the area formula

At the level we are, there is no proper way to show you how to derive the surface area of the sphere. However, if you allow me to be a bit imprecise, we can have a "demonstration" of it.

Say you take the "shell" of the sphere and divide it with a "grid", which will make a number of tiny circular divisions. Say that we have n divisions, all of them very very tiny, with areas and perimeters being very small.

Now, each of these tiny divisions is a cone with the base on the "shell" of the sphere and vertex at the centre of the sphere. The height of each of these cones is very close to R, the radius of the sphere. The more divisions we have on the "shell", the closest to R the height of these cones will be.

Now, say that the surface area of the sphere is given by A, and we have n divisions of it, all with area A_i: $A_1, A_2, A_3, \ldots, A_n$. So, we have that

$$A = A_1 + A_2 + A_3 + \ldots + A_n$$

We also know that the volume of the sphere with radius R is $\frac{4}{3}\pi R^3$, and that the volume of the sphere is equal to the volume of all the n cones we have added together. Each tiny cone has volume $\frac{1}{3}A_i R$, for $i = 1$ to n. So, we have that

$$\begin{aligned}
\frac{1}{3}A_1R + \frac{1}{3}A_2R + \ldots + \frac{1}{3}A_nR &= \frac{4}{3}\pi R^3 && \text{Sum of volumes of cones = Volume of sphere} \\
\frac{1}{3}R\left(A_1 + A_2 + \ldots + A_n\right) &= \frac{4}{3}\pi R^3 && \frac{1}{3}R \text{ is a common factor} \\
\frac{1}{3}RA &= \frac{4}{3}\pi R^3 && \text{Sum of small areas equal total area} \\
\frac{1}{\cancel{3}}\cancel{R}A &= \frac{4}{\cancel{3}}\pi R^{\cancel{3}\,2} && \text{Cancelling } \frac{1}{3} \text{ and one } R \\
A &= 4\pi R^2
\end{aligned}$$

Now, this is not a proof as have no idea how to formalize this properly since for that we need Calculus. However, the intuition is very solid, in my opinion, and an interesting "consolation" for now.

Derivation of the volume formula

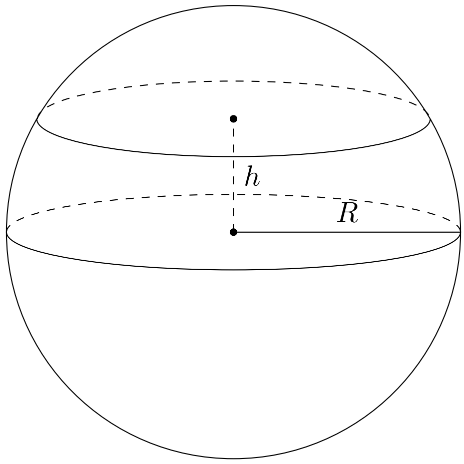
The volume formula we can actually prove if we use Cavalieri's principle (do see Appendix VII).

As a reminder, Cavalieri's principle states that:

> Given two solids and a plane P, if all parallel planes to P determine section of equal areas in each solid, then the solids have the same volume.

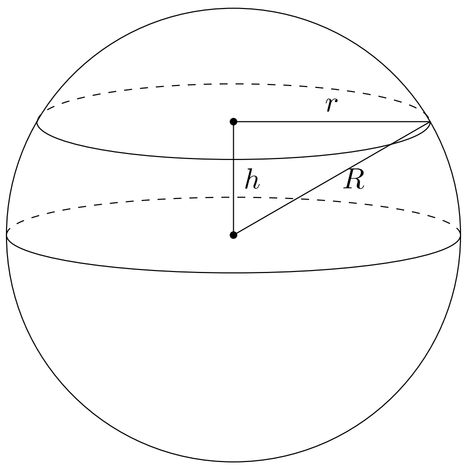
To use it, we will need two solids, then. Of course, one of them will be a sphere. As we will be sectioning the sphere with planes, it is useful to find an expression for the area of each section.

Given a sphere with radius R, we make a cross-section in it which is h away from the plane that contains the centre of the sphere:



Let us find an expression for the area of any cross-section h away from the centre of the sphere.

We can do that by drawing a triangle inside the sphere and the radius of the cross-section above the central one and call the radius of the cross-section r:



and that triangle is square. So we can use Pythagoras's:

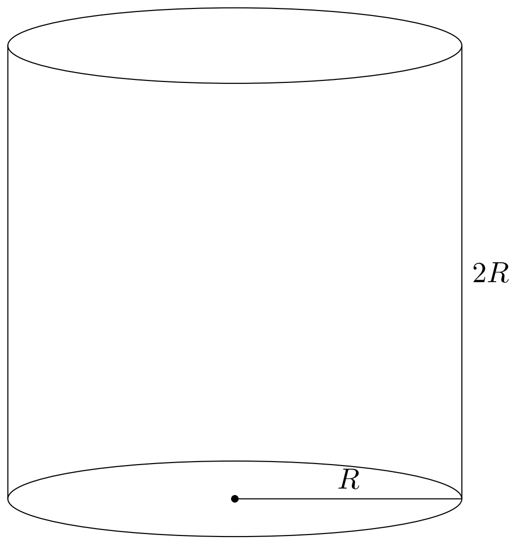
$$R^2 = h^2 + r^2$$

$$r^2 = R^2 - h^2$$

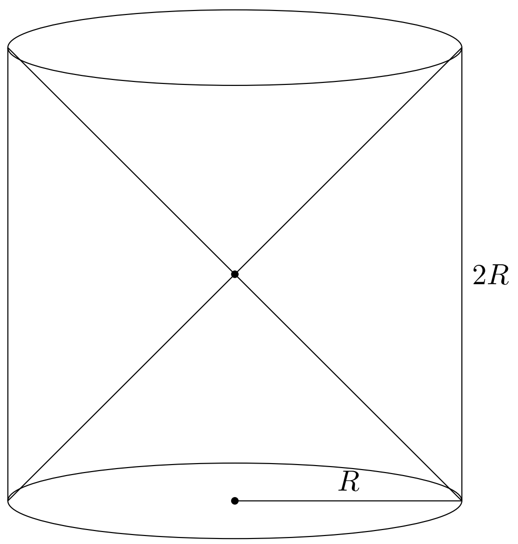
Now that we know the radius squared of each cross-section, we know that their areas must be

$$\text{area of cross-section} = \pi \left(R^2 - h^2\right)$$

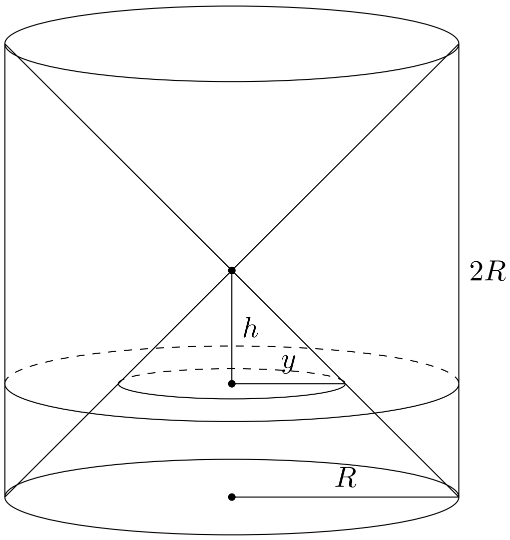
For the second solid (remember we need two to use Cavalieri's principle), we will start with an equilateral cylinder with diameter and height $2R$:



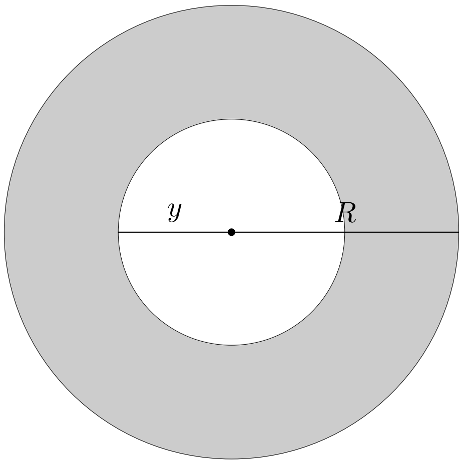
From this cylinder, we remove two cones with apex at the centre of the cylinder:



In this solid, any cross-section h away from its centre (the same as the cylinder, by the way) is a circle with a central piece removed:



the cross-sections in this shape look like this from above:



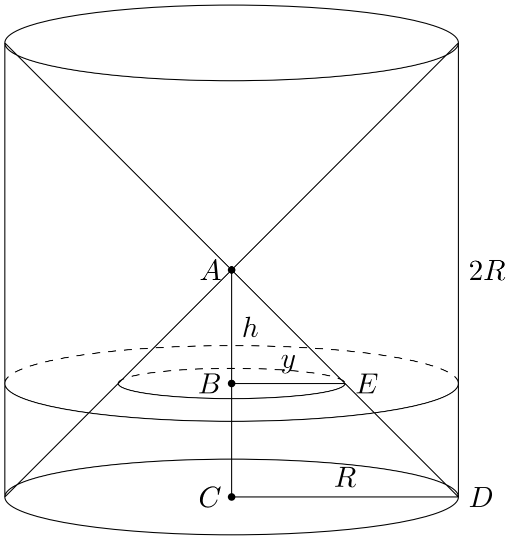
and they have area equal to

$$\text{Area of big circle} - \text{area of small circle}$$

$$\pi R^2 - \pi y^2$$

$$\pi \left(R^2 - y^2\right)$$

To find y, we can focus on two triangles inside our solid:



Triangles ABE and ACD are similar, so we have that

$$\frac{\overline{BE}}{\overline{CD}} = \frac{\overline{AB}}{\overline{AC}}$$

$$\frac{y}{R} = \frac{h}{R}$$

$$y = h$$

and we have that our y is actually equal to h. Hence, the areas of the cross-sections h away from the centre of our weird solid are given by

$$\text{area of cross-sections} = \pi\left(R^2 - h^2\right)$$

Of course, these are the same areas as the cross-sections of the sphere of radius R! So, if we put both solids on the same plane, all their cross-sections would have the same area, and by Cavalieri's principle, they have the same volume:

$$\text{Volume of sphere} = \text{Volume of strange solid}$$

To find the volume of the strange solid we subtract the volume of the two cones from the cylinder:

$$\text{Volume of strange solid} = \text{Volume of cylinder} - 2\text{Volume of cone}$$

The volume of the cylinder is easy to find:

$$\text{Volume of cylinder} = \pi R^2 \times 2R = 2\pi R^3$$

and the volume of a single cone is given by

$$\text{Volume of cone} = \frac{1}{3}\pi R^2 \times R = \frac{\pi R^3}{3}$$

Finally, we are able to find the volume of the solid:

$$2\pi R^3 - 2 \times \frac{\pi R^3}{3} = \frac{6\pi R^3}{3} - \frac{2\pi R^3}{3} = \frac{4}{3}\pi R^3$$

and as the sphere has the same volume, we found that

$$\text{Volume of sphere} = \frac{4}{3}\pi R^3$$

41. Circle theorems

41.1. Why learn circle theorems

I honestly don't know how to justify learning this in a 'practical' sense. It is fun to solve these kind of problems, and is definitely useful to prove the theorems, but I have no justification here. Sorry.

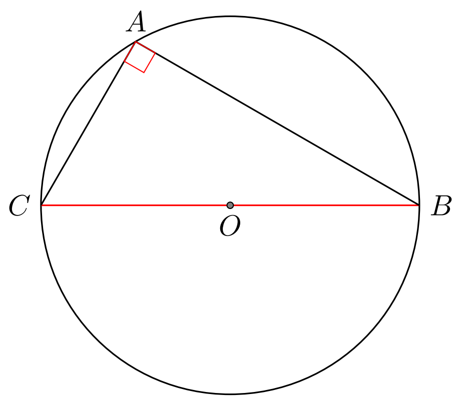
Given that proving the theorems is the only thing outside memorising them (you *do have* to memorise them!), I highly recommend reading the 'Formality after taste' session to see how each theorem is proved.

41.2. The theorems

41.2.1. Inscribed triangles with a side on a diameter

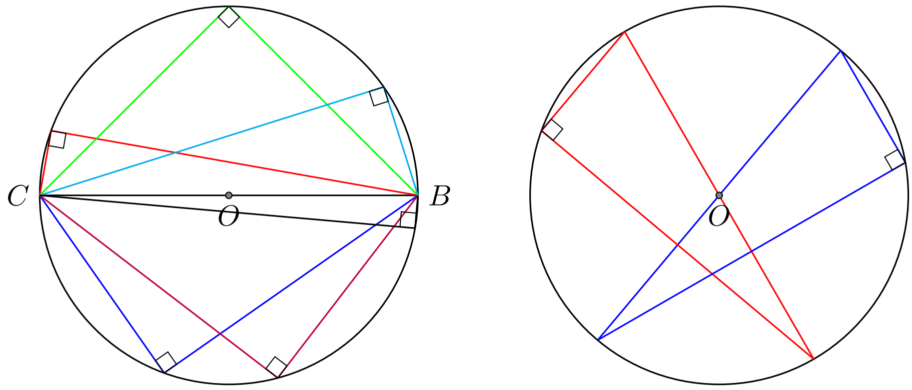
Theorem. Every inscribed triangle in which the largest side is a diameter of a circle is a right triangle.

For instance, in the triangle ABC below, where BC is a diameter of the circle with centre O, the angle $BAC = 90°$:



I have colored BC red to show you the important part: the largest side *has to be* a diameter of the circle.

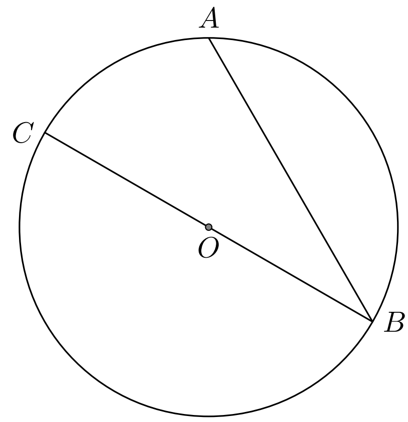
Notice that it does not matter *where* the vertex nor in the diameter is:



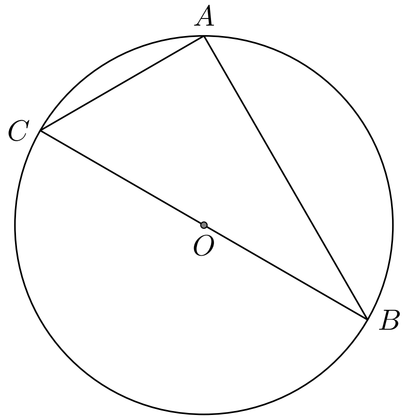
If you see a triangle inscribed in a circle and one of its sides is on a diameter, use this very useful result: it is a right triangle.

Inscribed triangle with side on a diameter example

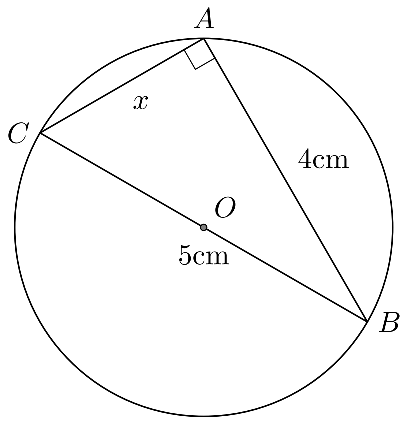
On the image below, the radius of the circle is equal to 2.5cm. The length of the segment AB is 4cm and BC is a diameter of the circle. Find the length of the segment AC.



Solution: Let's draw the segment AC:



We have that the triangle ABC has one side on the diameter BC, therefore it is a right triangle and $\angle BAC = 90°$. We know that the segment $AB = 4$cm and, given that the radius of the circle is 2.5cm, the diameter is equal to $2 \times 2.5 = 5$cm. Adding this information to the drawing and calling the length of the segment AC x:



Being a right triangle we can use our beloved Pythagoras's theorem:

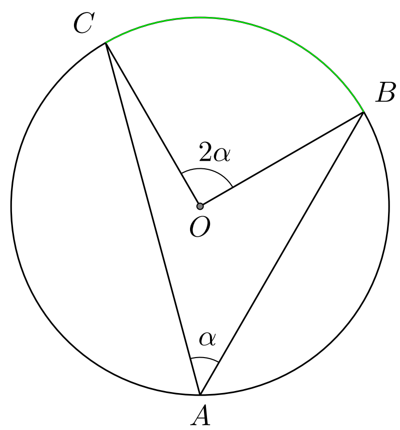
$$
\begin{aligned}
AB^2 + AC^2 &= BC^2 && \text{Pythagoras's} \\
4^2 + x^2 &= 5^2 \\
16 + x^2 &= 25 \\
16 - 16 + x^2 &= 25 - 16 && \text{Subtracting 16 on both sides} \\
x^2 &= 9 \\
\sqrt{x^2} &= \sqrt{9} && \text{Square rooting both sides} \\
x &= 3
\end{aligned}
$$

Thus, $x = AC = 3$cm.

41.2.2. Inscribed angles

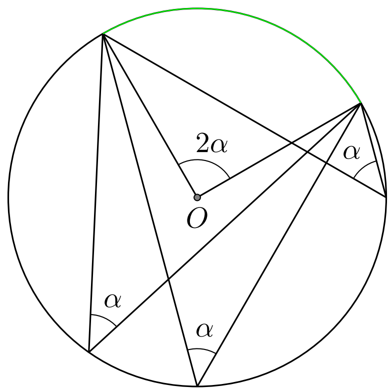
Theorem. Every inscribed angle is equal to half the size of the central angle that determines the same arc.

A picture is worth 18 words:



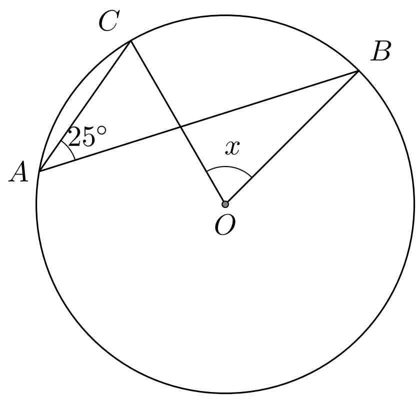
I like identifying these angles by finding the arcs on the circle each of them 'look': I have painted the arc in green in this case. Notice that both the central angle BOC and the inscribed angle BAC 'look at' the same arc.

Again, it does not matter *where* the inscribed angle is: if it is 'looking' at the same arc as a central angle, it measures half the central:

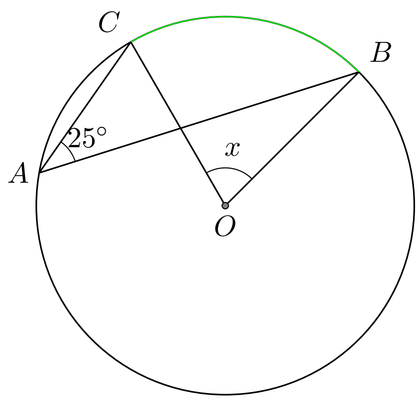


Inscribed angle with central angle example I

In the following circle, find the value of x.



Solution: Notice that both angles BAC and BOC 'look' at the same arc:



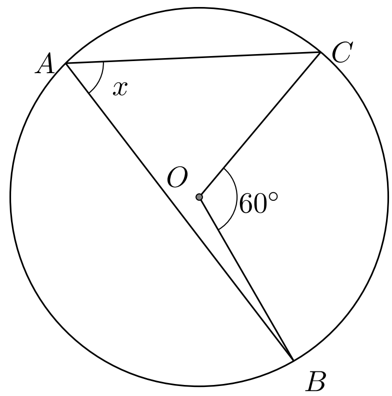
Therefore, we can use the inscribed angle theorem: we know that the central angles is twice the size of the inscribed angle:

$$\begin{aligned} \text{central} &= 2 \times \text{inscribed} && \text{The theorem} \\ x &= 2 \times 25 \\ x &= 50 \end{aligned}$$

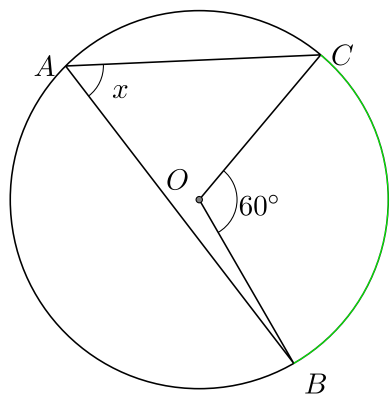
So we have that $x = 50°$.

Inscribed angle with central angle example II

In the following circle, find the value of x.



Solution: Again, notice that both angle BAC and angle OBC 'look' at the same arc:



We can, then, use the inscribed angle theorem:

$$\begin{aligned} \text{central} &= 2 \times \text{inscribed} && \text{The theorem} \\ 60 &= 2 \times x \\ \frac{60}{2} &= \frac{2x}{2} && \text{Dividing both sides by 2} \\ 30 &= x \end{aligned}$$

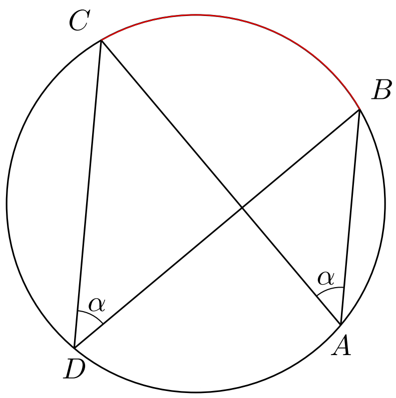
Thus, $x = 30°$.

41.2.2.1. Corollary: angles in the same arc, a.k.a. the 'bow-tie property'

Remember that a corollary is a result that follows easily from another. Here, it follows easily from the inscribed angle theorem the 'bow-tie property':

Corollary. Inscribed angles that are on the same arc are equal.

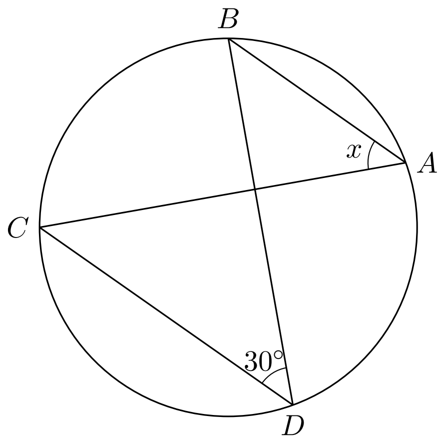
Let's see why it is called the 'bow-tie property':



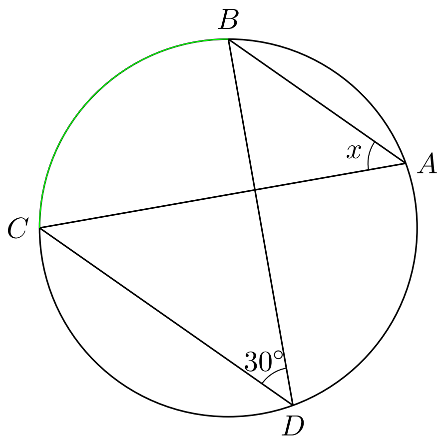
Notice that both angle BAC and angle BDC 'look' at the arc BC and are equal.

'Bow-tie property' example

Find the value of x.



Solution: I like marking the arc both angles are looking to be sure I can use the 'bow-tie':

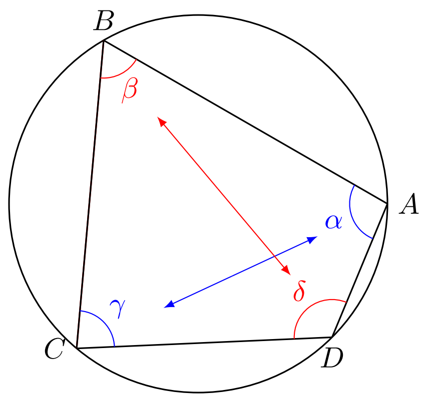


Thus, we know that $x = \angle BAC = \angle CDB$. Given that $\angle CDB = 30°$, we have that $x = 30°$.

41.2.3. Inscribed quadrilaterals (cyclic quadrilaterals)

Theorem. Opposite angles in inscribed quadrilaterals add up to 180°.

As usual, our drawing:



In this the quadrilateral $ABCD$ is inscribed (or cyclical[1]), and the angles ABC and ADC are opposite, so we have:

$$\angle ABC + \angle ADC = 180^\circ \qquad \text{Theorem: inscribed quadrilaterals}$$

$$\beta + \delta = 180^\circ$$

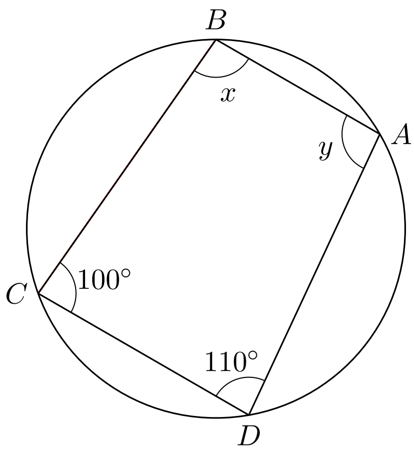
Not only, the angles BAD and BCD are opposite, thus:

$$\angle BAD + \angle BCD = 180^\circ \qquad \text{Theorem: inscribed quadrilaterals}$$

$$\alpha + \gamma = 180^\circ$$

Inscribed quadrilateral example

Find the values of x and y.



[1] I have to say I *hate* the 'cyclic quadrilateral' name. Every quadrilateral is 'cyclic', in the sense you can start walking in a vertex and reach it back.

Solution: We have an inscribed quadrilateral, so we know that opposite angles add up to 180°. Angles BAD and BCD are opposite:

$$\angle BAD + \angle BCD = 180^\circ \qquad \text{Inscribed quadrilateral}$$

$$y + 100 = 180$$

$$y + 100 - 100 = 180 - 100 \qquad \text{Subtrating 100 on both sides}$$

$$y = 80$$

The same thing for angles ABD and ADC:

$$\angle ABD + \angle ADC = 180^\circ \qquad \text{Inscribed quadrilateral}$$

$$x + 110 = 180$$

$$x + 110 - 110 = 180 - 110 \qquad \text{Subtracting 110 on both sides}$$

$$x = 70$$

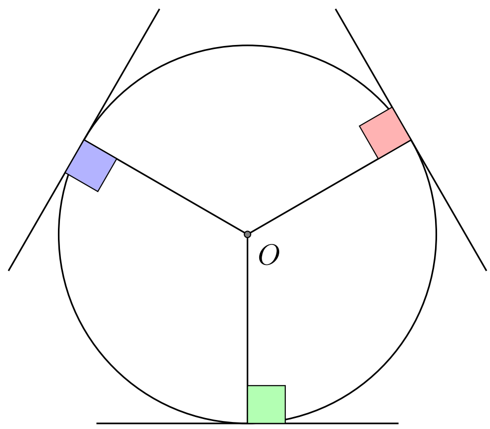
So we have that $x = 70^\circ$ and $y = 80^\circ$.

41.3. Tangent properties

These are useful to know[2].

41.3.1. The angle between a tangent to a circle and a radius is always 90°

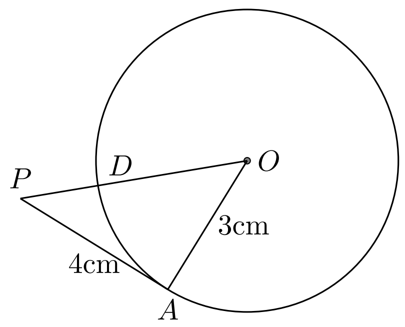
Remember that a tangent is a straight line that only touches the circle **in one point**. Drawing version:



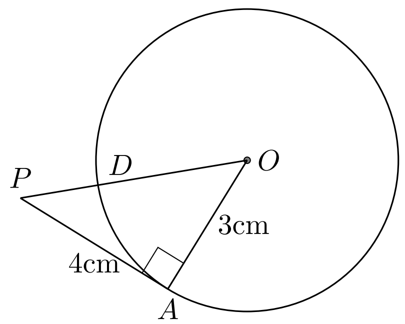
[2] I'll leave the proofs for these properties to you as exercises.

Tangents are perpendicular to radius example

Find the length of the segment PD, given that OA is a radius and has length 3cm.



Solution: We know that tangents are perpendicular to radiuses, thus we know that angle OAP is 90°:



We can now apply Pythagoras's in triangle OAP:

$$OP^2 = OA^2 + AP^2 \quad \text{Pythagoras's}$$

$$OP^2 = 3^2 + 4^2$$

$$OP^2 = 9 + 16$$

$$OP^2 = 25$$

$$\sqrt{OP^2} = \sqrt{25} \quad \text{Squaring both sides}$$

$$OP = 5$$

Now, to find PB, notice that the segment OD is also a radius and measures 3. We have, then:

$$PD = OP - OD$$

$$PD = 5 - 3$$

$$PD = 2$$

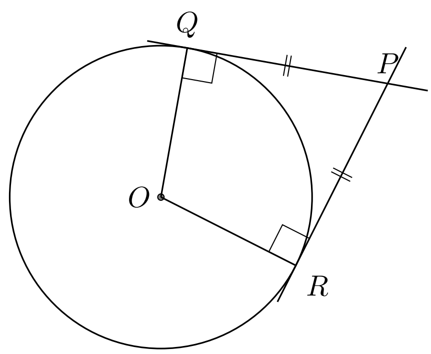
Thus, $PD = 2\text{cm}$.

41.3.2. Tangents from the same point have equal length

Given a point P outside the circle, and the two (why two?) points Q and R defined by the intersection of the tangents and the circle, we have that

$$PQ = PR$$

Our drawing:



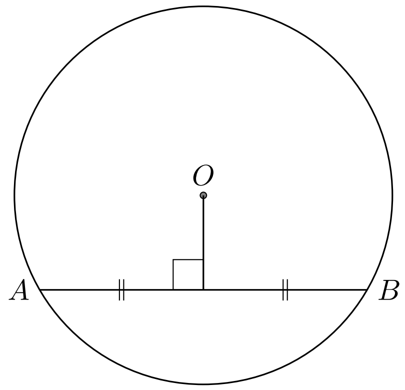
41.4. Chords properties

Also useful to know[3].

41.4.1. Chords are bisected by the perpendicular line from the centre

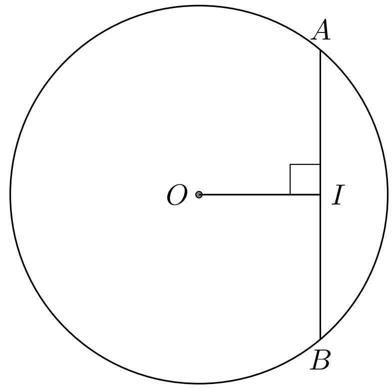
This one is very useful. It says that if you have a chord in a circle, and a radius crosses this chord in a right angle, then the radius bisects the chord. Drawing version:

[3] Also leaving the proofs to you.

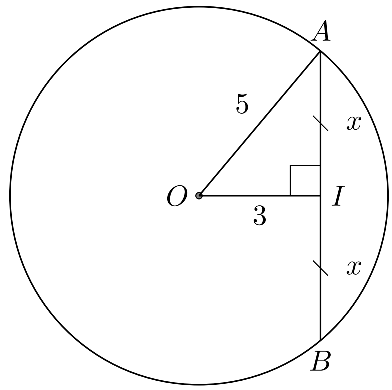


Perpendicular radius bisects chords example

Find the length of AB, given that the radius of the circle is 5cm and the segment OI measures 3cm.



Solution: This is a classic. Let's draw the radius OA and make $AI = x$ and remember that if the radius meets the chord at a right angle, it bisects it:



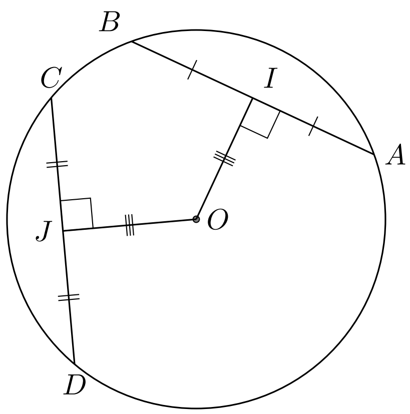
Let's use our favourite theorem, Pythagoras's, in triangle AOI:

$$
\begin{aligned}
OA^2 &= AI^2 + OI^2 && \text{Pythagoras's} \\
5^2 &= x^2 + 3^2 \\
25 &= x^2 + 9 \\
25 - 9 &= x^2 + 9 - 9 && \text{Subtracting 9 on both sides} \\
16 &= x^2 \\
\sqrt{16} &= \sqrt{x^2} && \text{Square rooting both sides} \\
4 &= x
\end{aligned}
$$

We know that $x = AI = 4$, but we want the length of the full chord AB. Now, given that AI is half the full chord, we have that $AB = 2x = 8$cm.

41.4.2. Chords of equal length are at the same distance from the centre

This one may sound complicated, but it is very simple. I'll give the drawing first this time.



Here, we have that if $AB = CD$, then $OI = OJ$. That's it.

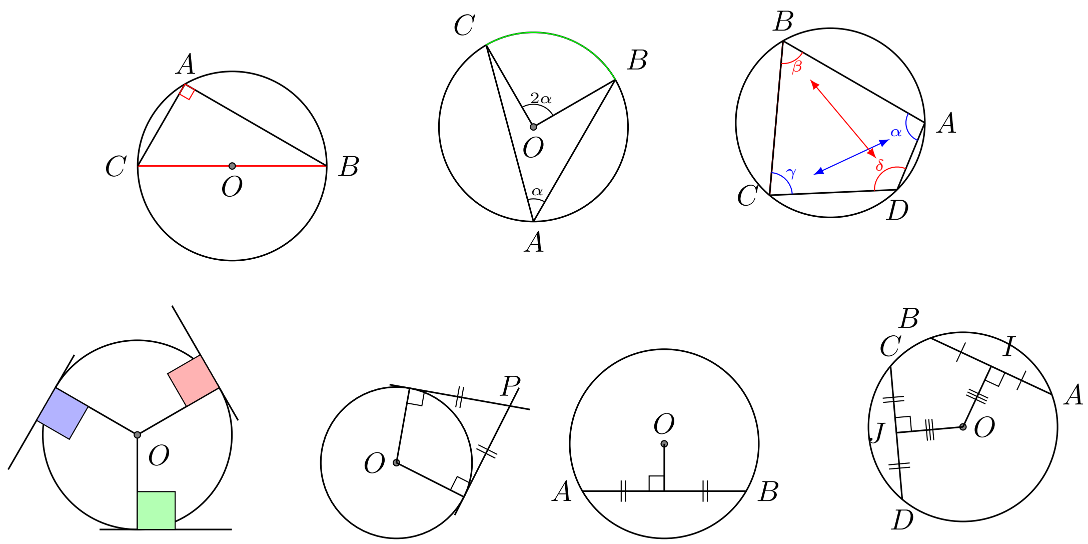
41.5. Exam hints

When you are presented with a circles theorem question, you have to be selectively ignorant: you almost always will be using many of the theorems and properties, and it is important to practice a lot in order to visualise which parts to focus and which parts to ignore.

Be aware of the classic mistakes:

- A quadrilateral has to have **all 4** vertices on the circle to be an inscribed (cyclic) quadrilateral. If any is not on the circle, you cannot use the theorem that opposite vertices add up to 180°;
- The inscribed angle theorem only applies to **central angles**, if you have an angle which is not central it does not follow that it is twice the size of the inscribed angle;
- To be a right triangle, the largest side **has to be a diameter**.

Summary

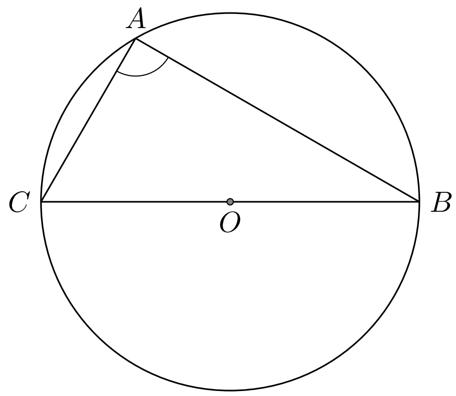


Formality after taste

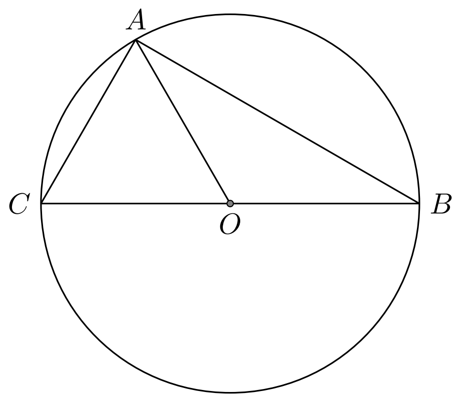
Proof of the right inscribed triangles

Theorem. Every inscribed triangle in which the largest side is a diameter of a circle is a right triangle.

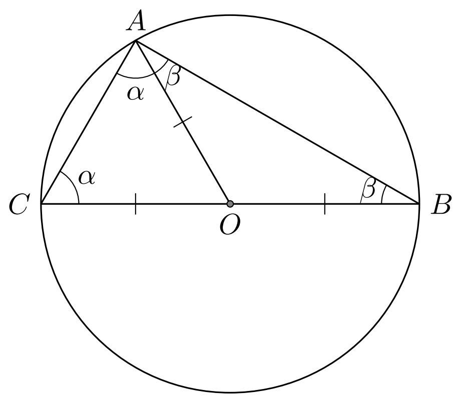
Proof. Let's take our triangle ABC with side BC on a diameter:



We want to prove that the angle CAB is a right angle. Let's start by dividing ABC into two triangles using the radius OA:



Notice that $OC = OA = OB$ as they all radiuses of the circle. Now, because of this, the triangles OAC and OAB are isosceles and, therefore, we know that $\angle OCA = \angle OAC$ and that $\angle OAB = \angle OBA$. Let's say that $\angle OCA = \angle OAC = \alpha$ and $\angle OAB = \angle OBA = \beta$:



Notice that the angle CAB is equal to $\alpha+\beta$. Now, adding all the angles in the triangle ABC we should get 180°:

$$\begin{aligned}
\angle OAC + \angle OAC + \angle OAB + \angle OBA &= 180^\circ && \text{Sum of angles in } \triangle ABC \\
\alpha + \alpha + \beta + \beta &= 180 \\
2\alpha + 2\beta &= 180 \\
2(\alpha + \beta) &= 180 && \text{Factorising} \\
\frac{2(\alpha+\beta)}{2} &= \frac{180}{2} \\
\alpha + \beta &= 90
\end{aligned}$$

Thus, we have shown that $\alpha + \beta = 90°$, and given that angle $CAB = \alpha + \beta$, it follows that the angle CAB is 90° and that the triangle ABC is a right triangle.

□

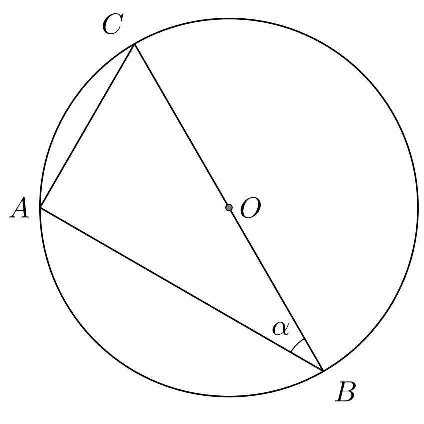
There is a way to prove this theorem using the inscribed angle theorem which is very nice. Give it a try!

Proof of the inscribed angle theorem

Theorem. Every inscribed angle is equal to half the size of the central angle that determines the same arc.[4]

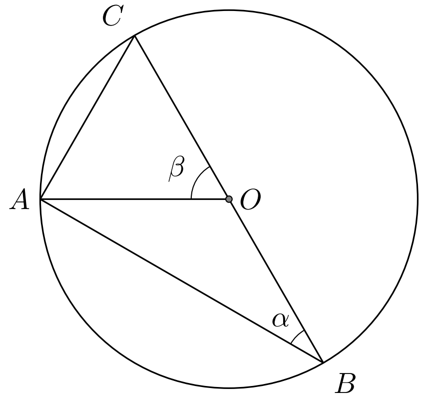
Proof. Let's first use a special case of the theorem. This means we'll start with a simpler version of what we want to prove, and later use this 'weaker' result to help us prove the 'full' one we are actually interested in.

The special case is when one of the chords that determine the arc is a diameter of the circle:

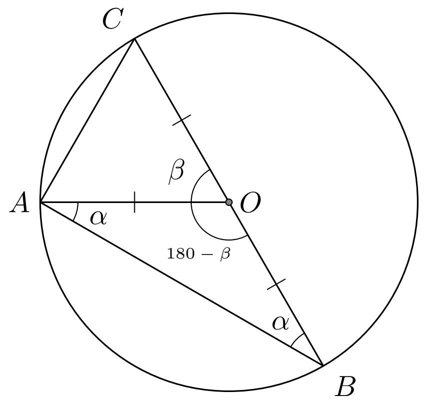


[4]Khan Academy has a great video on this proof: `https://www.khanacademy.org/math/geometry/hs-geo-circles/modal/v/inscribed-and-central-angles`

Our aim here is to prove that the angle ABC is equal to half the size of the angle AOC. Let's call $\angle AOC = \beta$:



Here $OA = OB = OC$ as they are all radiuses of the circle. Thus, we have that $\angle OAB = \angle ABO = \alpha$ and that $\angle AOB = 180 - \beta$, as the angles AOC and AOB are angles on a line:



Now, if we add all the angles in triangle OAB we obtain 180°:

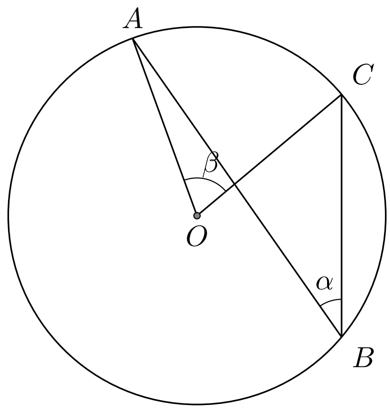
$$\angle OAB + \angle AOB + \angle OBA^{\circ} = 180 \qquad \text{Sum of angles in a triangle}$$

$$\alpha + 180 - \beta + \alpha = 180$$

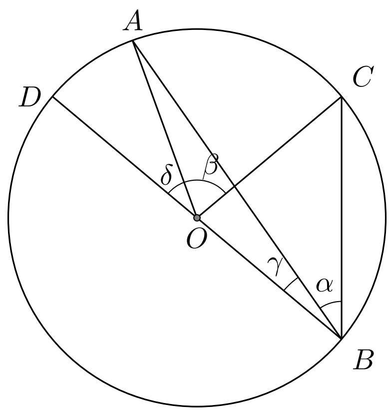
$$2\alpha - \beta + \cancel{180} = \cancel{180}$$

$$2\alpha = \beta$$

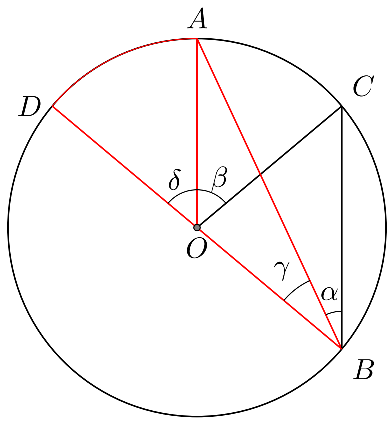
There we go: we have proved that the central angle, β, is twice as big as the inscribed angle α, in the particular case one of the legs of the angle is a diameter. Let's try a general case now, such as:



We know how to solve this problem when we have a diameter with one ends on the vertex of our inscribed angle, as we have just proved it. So let's draw a diameter and label some angles:



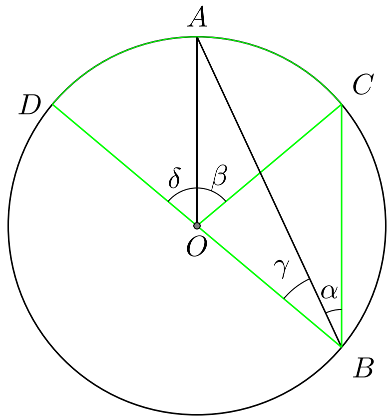
Now let's use abstraction in the sense of 'ignoring' and focus on the coloured parts:



Notice that we have both angle ABD and angle AOD are 'looking' at the same red arc, and that one of the legs of the angle ABD is the diameter BD. Therefore, we can use the result we just proved about inscribed angles with one leg being a diameter:

$$\begin{aligned} \text{central} &= 2 \times \text{inscribed} && \text{In the special case we just proved} \\ \delta &= 2\gamma \end{aligned}$$

Now focus on the green parts:



Here we have that both angle CBO and angle COD are 'looking' at the same green arc. Not only that, one of the legs of angle CBD is the diameter BD, so we can use the result we just proved again:

$$\begin{aligned} \text{central} &= 2 \times \text{inscribed} && \text{Still the special case} \\ \delta + \beta &= 2 \times (\alpha + \gamma) \\ \delta + \beta &= 2\alpha + 2\gamma \\ 2\gamma + \beta &= 2\alpha + 2\gamma && \delta = 2\gamma \\ \cancel{2\gamma} + \beta &= 2\alpha + \cancel{2\gamma} \\ \beta &= 2\alpha \end{aligned}$$

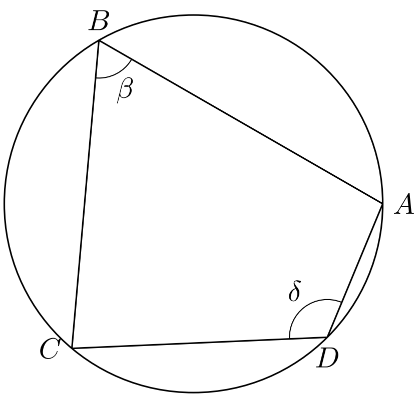
And there we go, we have proved that $\beta = 2\alpha$.

□

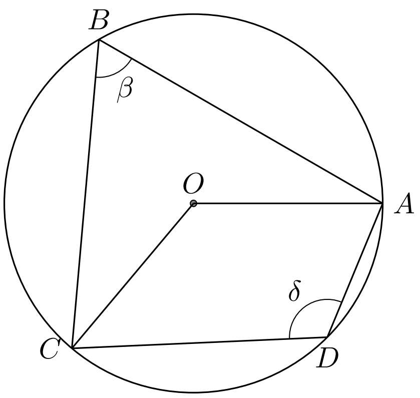
Proof of the inscribed quadrilateral theorem

Theorem. Opposite angles in inscribed quadrilaterals add up to 180°.

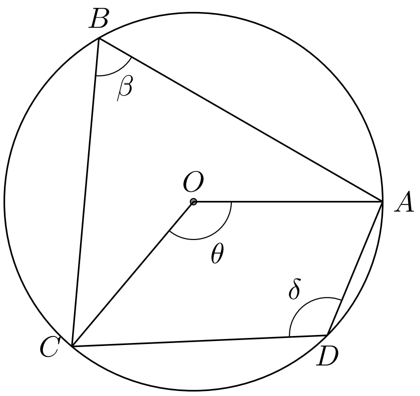
Proof. We only need to prove this for one pair of opposite angles, without loss of generality. Thus, we want to prove that $\beta + \delta = 180°$:



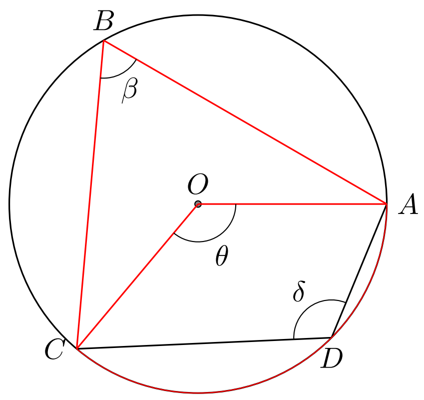
Let's draw the centre of the circle O and draw two radiuses, OA and OB:



Now, notice that we have this new central angle called θ:



Now, let's focus on the red parts:

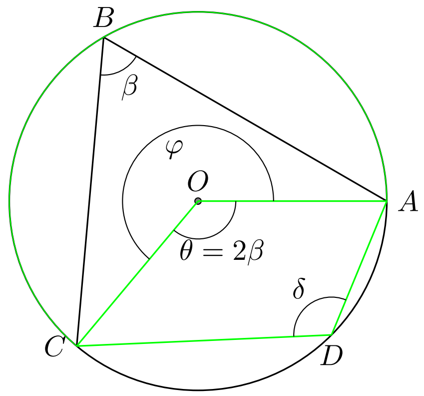


Notice that both angles ABC and AOC are both 'looking' at the arc ADC! Therefore, we can use the inscribed angle theorem:

$$\text{central} = 2 \times \text{inscribed} \qquad \text{Inscribed angle theorem}$$

$$\theta = 2\beta$$

Look now at the reflex angle AOC, called φ and at the green parts:

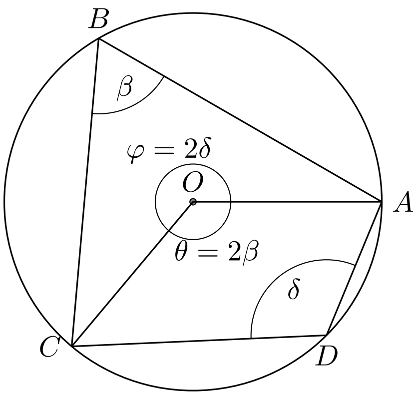


Notice that both angles ADC and AOC are 'looking' at the same arc ABC. Again, we can use the inscribed angle theorem:

$$\text{central} = 2 \times \text{inscribed} \qquad \text{Inscribed angle theorem}$$

$$\varphi = 2\delta$$

Thus, we have



Around the centre O we need to add to 360°, so we have

$$\varphi + \theta = 360^\circ \qquad \text{Angles around a point}$$

$$2\delta + 2\beta = 360$$

$$2(\delta + \beta) = 360 \qquad \text{Factorising}$$

$$\frac{2(\delta + \beta)}{2} = \frac{360}{2}$$

$$\delta + \beta = 180$$

There we have it, $\beta + \delta = 180^\circ$.

□

42. Vectors

42.1. Why learn vectors

Vectors are a "natural" entity as well. Think about force. When you apply a force to something, say you push a door, it is different to push the door to the side than straight forward. If you want to drag something, it's betters to pull it horizontally, parallel to the ground. However, when you apply a force, you also control *how much* force you apply. With velocity is the same: when you are moving, you are moving *somewhere*, you have *direction*, and you also have "how fast" you are moving.

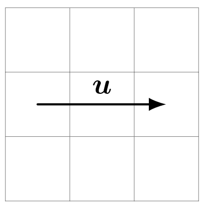
Whenever a quantity has not only "size" but also a *direction*, it is a **vectorial** measure. We have many of those, therefore vectors are very useful in science and engineering. They are very useful in math itself, being a basic entity in a field called linear algebra. In all, a very useful topic.

42.2. Notation and definitions

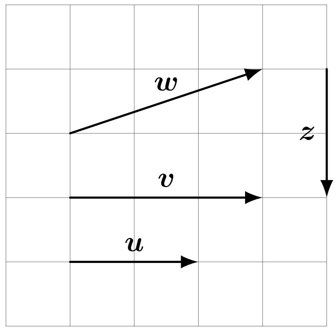
First, on the following discussion, a **bold faced letter** will represent a vector. So $\boldsymbol{u}, \boldsymbol{v}$and $\boldsymbol{w}$ are all vectors. I myself prefer an arrow on top to represent vectors, such as $\overrightarrow{u}$, $\overrightarrow{v}$ and $\overrightarrow{w}$, but the notation used in the IGCSE is the bold face.

42.2.1. Graphical

Remember that a vector is a quantity that has not only "size" but also direction. Let's start using the proper terms: the "size" of a vector is called its **magnitude**. Let's represent a vector graphically:



Here we have vector $\boldsymbol{u}$ in all its majesty. You can see it has *direction*, as it is pointing to the right. Its *magnitude* is denoted by how long the vector is. The next image has 4 vectors:



Notice that vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ have the same direction: they are pointing to the right. Vector $\boldsymbol{z}$ is "pointing" down. Vector $\boldsymbol{w}$ has a "weird" direction, as he is pointing to a "mix" or right and up.

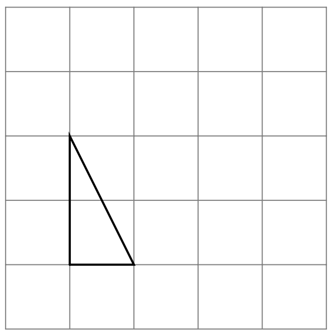
In summary, to represent vectors graphically, we use arrows: the direction they point is the direction of the vector, and the size of the arrow represents its magnitude.

42.2.2. Column vector

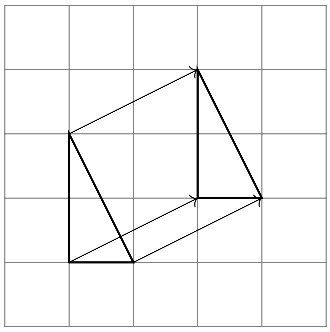
You have already seen column vectors before, when we studied translations. Remember that a translation was denoted like this

$$\begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

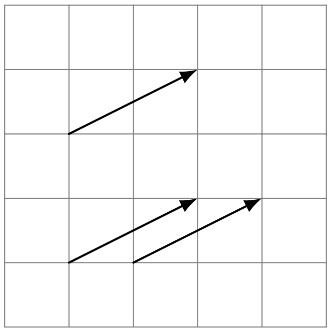
which meant the shape "moved" 2 units in the x direction and 3 units in the y direction (also known as "2 to the right and 3 up"). That is called a translation vector, as it was just defining a translation. Let's remember that with an example. We have the triangle below:



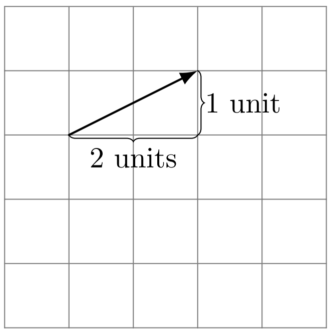
Let's translate this triangle by the vector $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$. To do that, we take each vertex and move it 2 units to the right and 1 unit up:



Now, let's remove the triangle from the image:



Now, let's remove two the vectors from the drawing, and notice that the vector goes 2 units to the right and 1 unit up:



So, just like when we had a translation vector, we can represent *any* vector using the notation

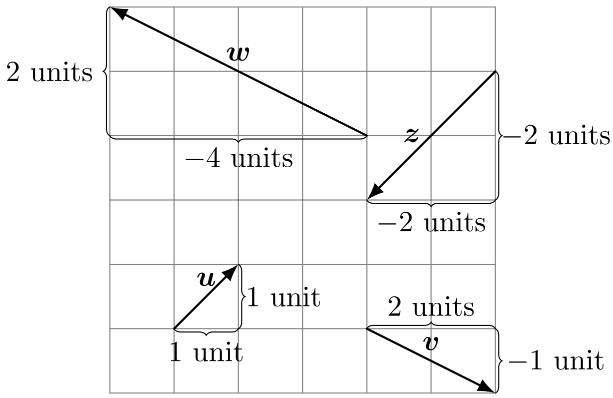
$$\boldsymbol{u} = \begin{pmatrix} x \\ y \end{pmatrix}$$

which means that vector $\boldsymbol{u}$ is the vector that "points" x units on the x-axis and y units on the y-axis. We call the x values the x *component* of the vector, and we call the y value the y *component* of the vector.

For instance, the vectors

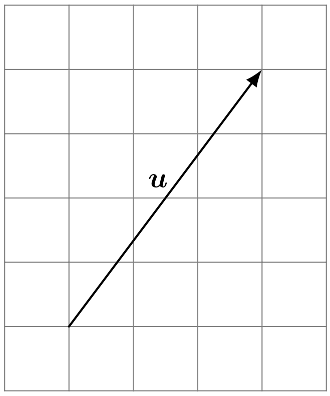
$$\boldsymbol{u} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \boldsymbol{v} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \boldsymbol{w} = \begin{pmatrix} -4 \\ 2 \end{pmatrix}, \boldsymbol{z} = \begin{pmatrix} -2 \\ -2 \end{pmatrix}$$

can be seen graphically as:



42.2.3. Magnitude

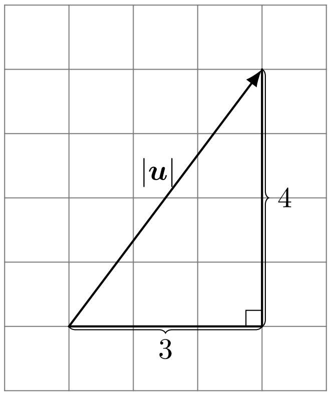
How do we calculate the magnitude of a vector? Let's say we have this vector:



Remember that the magnitude of a vector is its size, so we want to find the *length* of the vector $\boldsymbol{u}$. We denote the *magnitude of a vector* by putting vertical bars around it:

$$|\boldsymbol{u}|$$

Thus, we want to find $|\boldsymbol{u}|$. This can be done using our old dead friend, Pythagoras:



Here, we have a right triangle with $|\boldsymbol{u}|$ as hypotenuse, and sides 3 and 4. Applying Pythagoras's:

$$|\boldsymbol{u}|^2 = 3^2 + 4^2$$

$$|\boldsymbol{u}|^2 = 9 + 16$$

$$|\boldsymbol{u}|^2 = 25$$

$$\sqrt{|\boldsymbol{u}|^2} = \sqrt{25}$$

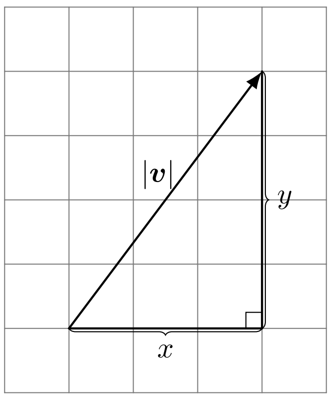
$$|\boldsymbol{u}| = 5$$

There we have it, the magnitude of $|\boldsymbol{u}|$ is 5.

In general, given a vector

$$\boldsymbol{v} = \begin{pmatrix} x \\ y \end{pmatrix}$$

which represents "going x units on the x-axis and y units on the y-axis", we can calculate $|\boldsymbol{v}|$ using Pythagoras's, as the following image shows us:



Again, we have a right triangle with $|\boldsymbol{v}|$ as hypotenuse, and x and y as sides. Applying

Pythagoras's:

$$|\boldsymbol{v}|^2 = x^2 + y^2$$

$$\sqrt{|\boldsymbol{v}|^2} = \sqrt{x^2 + y^2}$$

$$|\boldsymbol{v}| = \sqrt{x^2 + y^2}$$

There we go. Given a vector

$$\boldsymbol{v} = \begin{pmatrix} x \\ y \end{pmatrix}$$

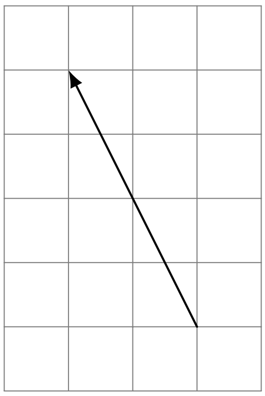
its magnitude is given by

$$\boxed{|\boldsymbol{v}| = \sqrt{x^2 + y^2}}$$

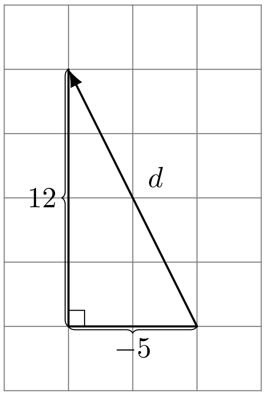
Solved exercise: magnitude of vectors

1. Calculate the magnitude of $\begin{pmatrix} -5 \\ 12 \end{pmatrix}$.

   Solution: You can use the formula straight away, but I always like making a sketch of the vector and applying Pythagoras's. In this case, we have a vector which "points" 5 units to the left (-5 in the x component) and 12 units up. Something like this suffices as a drawing:

   We can complete the drawing with the right triangle, and let's call the hypotenuse d:



We can now apply Pythagoras's:

$$d^2 = 12^2 + (-5)^2$$

$$d^2 = 144 + 25$$

$$d^2 = 169$$

$$\sqrt{d^2} = \sqrt{169}$$

$$d = 13$$

You can use the formula straight away as well:

$$\sqrt{x^2 + y^2}$$

$$\sqrt{(-5)^2 + 12^2}$$

$$\sqrt{25 + 144}$$

$$\sqrt{169} = 13$$

2. A vector $\begin{pmatrix} a \\ a-2 \end{pmatrix}$ has magnitude $\sqrt{2}$. Find the value of a.

Let's use the formula:

$$\sqrt{2} = \sqrt{a^2 + (a-2)^2}$$

$$(\sqrt{2})^2 = \left(\sqrt{a^2 + (a-2)^2}\right)^2$$

$$2 = a^2 + (a-2)^2$$

$$2 = a^2 + a^2 - 4a + 4$$

$$2a^2 - 4a + 4 = 2$$

$$2a^2 - 4a + 4 - 2 = 2 - 2$$

$$2a^2 - 4a + 2 = 0$$

$$2(a^2 - 2a + 1) = 0$$

$$2(a-1)^2 = 0$$

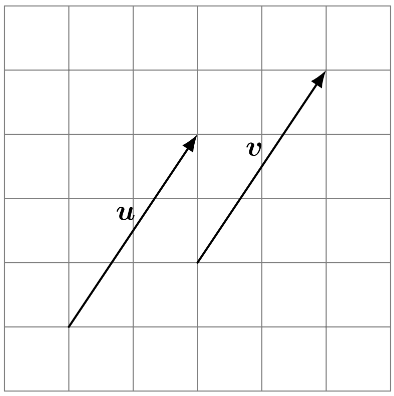
$$(a-1)^2 = 0$$

$$a = 1$$

So we have that $a = 1$.

42.3. Vector equality

I wonder if you have noticed that I have been drawing the vectors wherever I want. Do you think it matters? For instance, take this next figure with two vectors, $\boldsymbol{u}$ and $\boldsymbol{v}$:



Both these vectors have the same direction and magnitude, and we could represent them using our column notation:

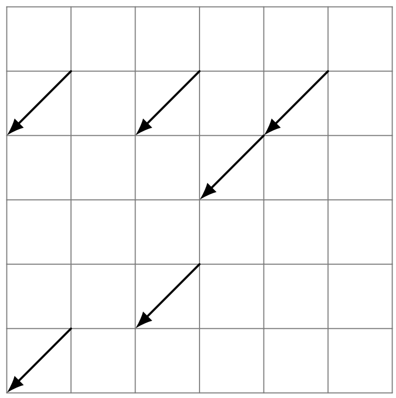
$$\begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

Do you agree that it does not matter *where* we draw the vectors, they both represent the same entity, as they both are "pointing" at the same direction and have the same size? If you imagined they each represented someone's velocities, each person would be moving at the same speed, as both vectors have the same magnitude, and they would be moving parallel to each other, as both of them would be "moving 2 to the right and 3 up".

In all, it does not matter *where* we draw the vectors, what matters is their magnitude and their direction.

Thus, to be *equal*, two vectors must have **same magnitude and same direction**.

To exemplify this further, all the vectors in the next figure are equal:



42.4. Vector arithmetic

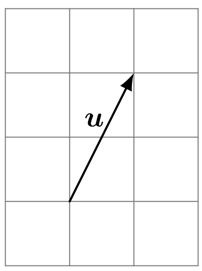
42.4.1. Multiplying a vector by a scalar

Multiplying a vector by a scalar, also know an as number, is the same as multiplying a matrix by a number: you simply multiply both components by the number outside.

Let's see an example. Given

$$\boldsymbol{u} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

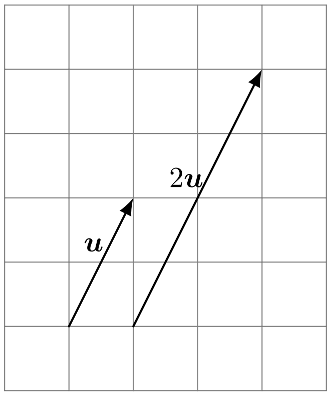
which can be seen graphically as



we can calculate $2\boldsymbol{u}$, which is the same as

$$2\boldsymbol{u} = 2 \times \boldsymbol{u} = 2 \times \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$$

Let's draw $2\boldsymbol{u}$ as well



Notice that $2\boldsymbol{u}$ is twice as big as $\boldsymbol{u}$.
In general, given a vector

$$\boldsymbol{v} = \begin{pmatrix} x \\ y \end{pmatrix}$$

and a scalar (number) k, the vector $k\boldsymbol{v}$ is given by

$$k\boldsymbol{v} = k \times \boldsymbol{v} = k \times \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} kx \\ ky \end{pmatrix}$$

Let's show that the magnitude of $k\boldsymbol{v}$ is equal to k times the magnitude of $\boldsymbol{v}$:

$$|k\boldsymbol{v}| = \sqrt{(kx)^2 + (ky)^2} \qquad \text{Definition of } |k\boldsymbol{v}|$$

$$|k\boldsymbol{v}| = \sqrt{k^2x^2 + k^2y^2}$$

$$|k\boldsymbol{v}| = \sqrt{k^2(x^2 + y^2)} \qquad k^2 \text{ is a common factor}$$

$$|k\boldsymbol{v}| = \sqrt{k^2} \times \sqrt{x^2 + y^2} \qquad \text{We can separate products in roots}$$

$$|k\boldsymbol{v}| = k \times |\boldsymbol{v}| \qquad |\boldsymbol{v}| = \sqrt{x^2 + y^2}$$

$$|k\boldsymbol{v}| = k|\boldsymbol{v}|$$

It is important to remember this: multiplying a vector by a number also multiplies its magnitude by the same number. For instance, if we have the vector

$$\boldsymbol{v} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

we can calculate its magnitude:

$$|\boldsymbol{v}| = \sqrt{3^2 + 4^2} = \sqrt{25} = 5$$

Now, if we multiply $\boldsymbol{v}$ by 2:

$$2\boldsymbol{v} = \begin{pmatrix} 6 \\ 8 \end{pmatrix}$$

and calculate its magnitude:

$$2\boldsymbol{v} = \sqrt{6^2 + 8^2} = \sqrt{100} = 10$$

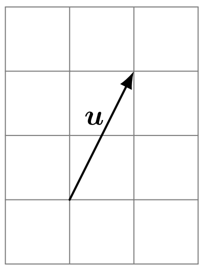
which is twich as big as $\boldsymbol{v}$. So, remember, if you multiply a vector by a scalar its magnitude also gets multiplied.

42.4.2. Opposite vectors

Let's say we have this vector $\boldsymbol{u}$:

$$\boldsymbol{u} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

We can draw it



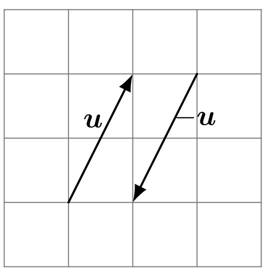
Now, let's calculate $-\boldsymbol{u}$, which is the same as

$$-\boldsymbol{u} = -1 \times \boldsymbol{u}$$

We know that to multiply a vector by a scalar is the same as multiplying both its components by the number, so we have

$$-\boldsymbol{u} = -1 \times \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} -1 \\ -2 \end{pmatrix}$$

We can also draw $-\boldsymbol{u}$:



Notice that $-\boldsymbol{u}$ has **opposite direction** of $\boldsymbol{u}$: they are on parallel lines, but $\boldsymbol{u}$ and $-\boldsymbol{u}$ are each pointing to one end of the line.

Also notice they both have the same magnitude:

$$|\boldsymbol{u}| = \sqrt{1^2 + 2^2} = \sqrt{5}$$

$$|-\boldsymbol{u}| = \sqrt{(-1)^2 + (-2)^2} = \sqrt{5}$$

What we have, then, is that given a vector

$$\boldsymbol{v} = \begin{pmatrix} x \\ y \end{pmatrix}$$

its *opposite vector*, written $-\boldsymbol{v}$, is given by

$$-\boldsymbol{v} = \begin{pmatrix} -x \\ -y \end{pmatrix}$$

has the same magnitude as $\boldsymbol{v}$, but has opposite direction to it.

In summary, "minusing a vector" is the same as putting the arrowhead on the other side!

42.4.3. Adding vectors

To understand vector addition, let's define it first using our column vector notation. Given two vectors

$$\boldsymbol{u} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \boldsymbol{v} = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$$

we can find a new vector $\boldsymbol{u} + \boldsymbol{v}$ by doing

$$\boldsymbol{u} + \boldsymbol{v} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \end{pmatrix}$$

Thus, to add two vectors is exactly as adding two matrices: you add corresponding components.

Now, with an example, let's see what happens graphically.

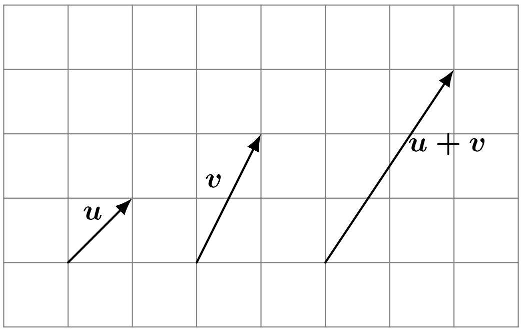
Take two vectors

$$\boldsymbol{u} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \boldsymbol{v} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

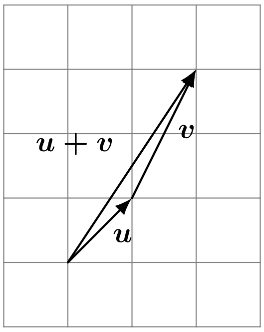
We can calculate $\boldsymbol{u} + \boldsymbol{v}$ using the above definition:

$$\boldsymbol{u} + \boldsymbol{v} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1+1 \\ 1+2 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

Let's draw all these vectors together:



Remember that where we draw a vector does not matter? Let's draw the same vectors, but in different places:

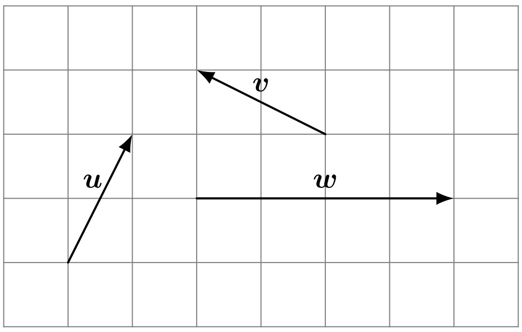


Notice that you could have drawn this by doing the following:

1. Draw $\boldsymbol{u}$ anywhere;
2. Draw $\boldsymbol{v}$ starting *where* $\boldsymbol{u}$ *ends*;
3. Draw another vector that begins where $\boldsymbol{u}$ started and ends where $\boldsymbol{v}$ ends.

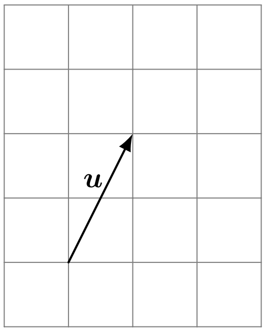
I like remembering that you do vector addition by thinking about *paths*: adding vectors is the same as going from the start to the end, but by some tortuous path, made of different parts; the result of adding all those vectors is a *shortcut, the smallest path between the beginning and the end.*

Let's see another example and add all the following vectors:

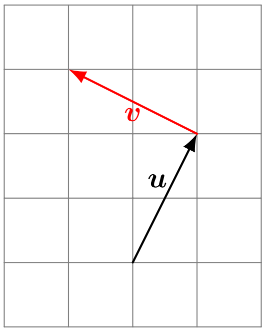


Again, we just need to choose a vector to start, and draw the other vectors at the end of each one. Then, we just join the beginning of everything with the end. Let's do step by step.

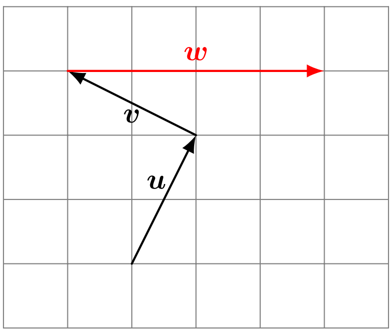
Begin with $\boldsymbol{u}$, as it comes first in the alphabet:



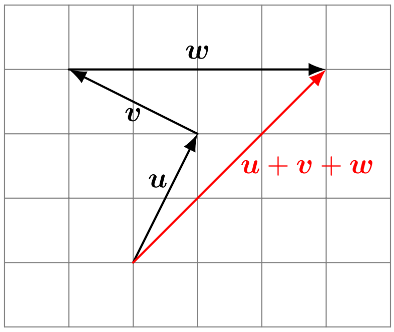
Now, at the end of $\boldsymbol{u}$, let's draw $\boldsymbol{v}$:



Then, we take $\boldsymbol{w}$ and draw it where $\boldsymbol{v}$ ends:



To finish, we draw the vector that begins where $\boldsymbol{u}$ begun and finishes where $\boldsymbol{w}$ finishes:



There we have it, the red vector is $\boldsymbol{u}+\boldsymbol{v}+\boldsymbol{w}$.

Let's check our answer by writing the same operation using column notation. Notice that $\boldsymbol{u}$ goes 1 to the right and 2 up, so it is written as $\begin{pmatrix}1\\2\end{pmatrix}$. $\boldsymbol{v}$ goes 2 to the left and 1 up, so it is written as $\begin{pmatrix}-2\\1\end{pmatrix}$. Finally, $\boldsymbol{w}$ goes 4 to the right and nothing up or down, so it is written as $\begin{pmatrix}4\\0\end{pmatrix}$. Let's find their sum:

$$\boldsymbol{u}+\boldsymbol{v}+\boldsymbol{w}$$

$$\begin{pmatrix}1\\2\end{pmatrix}+\begin{pmatrix}-2\\1\end{pmatrix}+\begin{pmatrix}4\\0\end{pmatrix}$$

$$\begin{pmatrix}1+-2+4\\2+1+0\end{pmatrix}$$

$$\begin{pmatrix}3\\3\end{pmatrix}$$

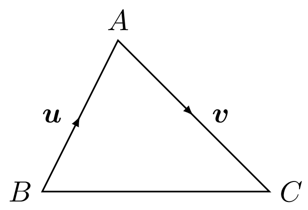
Look at our drawing: the red vector, which is the sum, goes 3 units to the right and 3 units up, which can be written as $\begin{pmatrix}3\\3\end{pmatrix}$.

I recommend you to do two things: first, change the order you draw the vectors to see that it gives the same result, as a good addition should. Second, remember: to add vectors, find the shortcut, the smaller straight path from the beginning to the end.

42.5. Vector paths

We know the basics now. The questions that we will actually be solving build on those, and they can be summarised by: *find a path.* Poetic. Let's see an example.

Given the following triangle ABC and the vectors $\boldsymbol{u}$ and $\boldsymbol{v}$:



We need to find the vector $\overrightarrow{BC}$.

Remember how adding vectors is like finding the shortcut? We are going to do the reverse now: they gave us the shortcut, $\overrightarrow{BC}$, we need to find a path from B to C which uses the vectors we have.

So let's see, starting from B, we can go to A, and from A go to C. Notice that the path starts at B and ends at C: these questions are always like this. If they want you to find the vector that represents BC, you find a path (any path) that starts at B (the first vertex!) and finishes at C (the second vertex!). Writing this as vectors:

$$\overrightarrow{BC} = \overrightarrow{BA} + \overrightarrow{AC} \qquad \text{Find the path}$$

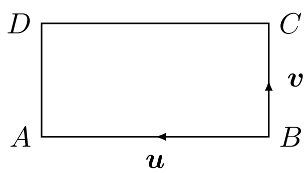
$$\overrightarrow{BC} = \boldsymbol{u} + \boldsymbol{v} \qquad \text{Write the vectors}$$

We are done! Have I mentioned that you just need to find a path? So poetic.

Let us see many examples now.

Finding vectors: negative vectors

Given a rectangle $ABCD$ and vectors $\boldsymbol{u}$ and $\boldsymbol{v}$:



1. Find $\overrightarrow{AB}$

   Remember, we need to find a path that starts at A and ends at B. We have a vector that goes the other way: starts at B and ends at A. Now, remembering that a negative of a vector is a vector which has the same magnitude but points to the opposite direction (or as I said, has the arrow pointing to the other way), we know that $\overrightarrow{AB}$ is equal to $-\overrightarrow{BA}$. So we have

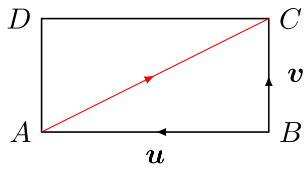
$$\overrightarrow{AB} = -\overrightarrow{BA} \qquad \text{The path is simple this time}$$

$$\overrightarrow{AB} = -\boldsymbol{u} \qquad \overrightarrow{BA} = \boldsymbol{u}$$

So we have that $\overrightarrow{AB} = -\boldsymbol{u}$.

2. Find $\overrightarrow{AC}$

   First, let's draw the vector they want us to find to help us visualise the path:



   We need to find a path that starts at A and ends at C. Notice that we can go from A to B and from B to C, so our path is

$$\overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC}$$

   Good thing they have given us $\overrightarrow{BC} = \boldsymbol{v}$. We also have just found $\overrightarrow{AB} = -\boldsymbol{u}$. We just have to substitute:

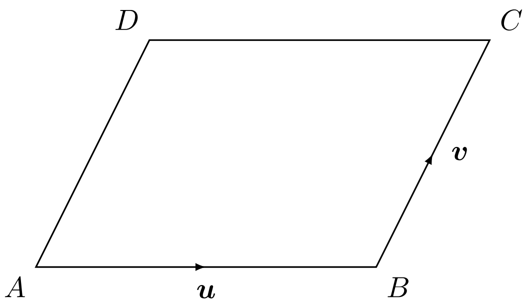
$$\overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC}$$

$$\overrightarrow{AC} = -\boldsymbol{u} + \boldsymbol{v}$$

$$\overrightarrow{AC} = \boldsymbol{v} - \boldsymbol{u}$$

Finding vectors: parallel vectors

Given a parallelogram $ABCD$ and vectors $\boldsymbol{u}$ and $\boldsymbol{v}$.



1. Find $\overrightarrow{AD}$

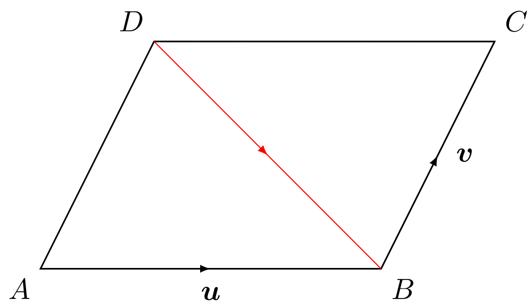
   Remember that a vector is equal to another if they have same direction and same magnitude? A parallelogram has two pairs of parallel and of equal length sides. So we have that the side AD is parallel to the side BC, and they have the same length. That means the vector $\overrightarrow{BC}$ is equal to the vector $\overrightarrow{AD}$, as they have same direction (parallel) and same magnitude (same length). Thus, we have that

$$\overrightarrow{AD} = \overrightarrow{BC}$$

$$\overrightarrow{AD} = \boldsymbol{v}$$

2. Find $\overrightarrow{DB}$

   Again, let's draw the vector we need to find to help us find our path:



We have to find a path starting at D and finishing at B. Let's do this in two ways. First, notice we could go from D to C and from C to B, which in vector notation is

$$\overrightarrow{DB} = \overrightarrow{DC} + \overrightarrow{CB}$$

As we are in a parallelogram, we again have that the length of DC is equal to the length of AB, and they are parallel. That implies that $\overrightarrow{DC} = \overrightarrow{AB}$. Thus, $\overrightarrow{DC} = \boldsymbol{u}$. $\overrightarrow{CB}$, on the other hand, has the same magnitude of $\overrightarrow{BC}$, but has opposite direction, so we have that $\overrightarrow{CB} = -\overrightarrow{BC}$. Given that $\overrightarrow{BC} = \boldsymbol{v}$, we have $\overrightarrow{CB} = -\boldsymbol{v}$. Substituting in our path equation:

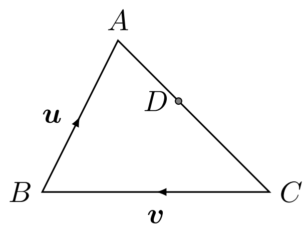
$$\overrightarrow{DB} = \overrightarrow{DC} + \overrightarrow{CB}$$

$$\overrightarrow{DB} = \boldsymbol{u} + -\boldsymbol{v}$$

$$\overrightarrow{DB} = \boldsymbol{u} - \boldsymbol{v}$$

Finding vectors: ratios

Given triangle ABC, vectors $\boldsymbol{u}$ and $\boldsymbol{v}$, and that $AD : DC = 2 : 3$



1. Find $\overrightarrow{AC}$

   As always, let's find our path starting at A and ending at C. We can go from A to B and from B to C, which in vector notation:

   $$\overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC}$$

   We know that $\overrightarrow{BA} = \boldsymbol{u}$, so we have that $\overrightarrow{AB} = -\overrightarrow{BA} = -\boldsymbol{u}$. By the same reasoning, we can find that $\overrightarrow{BC} = -\overrightarrow{CB} = -\boldsymbol{v}$. Substituting:

   $$\overrightarrow{AC} = -\boldsymbol{u} + -\boldsymbol{v} = -\boldsymbol{u} - \boldsymbol{v}$$

2. Find $\overrightarrow{AD}$

   The vector $\overrightarrow{AD}$ is a "small" version of $\overrightarrow{AC}$, do you agree? It has the same direction, as they are both on top of the same side of the triangle, it just has a different magnitude. Now, to calculate how much smaller, we need to use the ratio they gave us $AD : DC = 2 : 3$. This means that whatever the distance between AC is, if we divide it in 5 $(2+3)$, the distance from A to D is 2 parts of those 5. Thus, we have that

   $$AD = \frac{2}{5}AC$$

   So, we have that

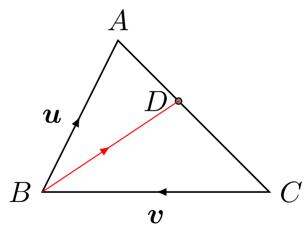
   $$\overrightarrow{AD} = \frac{2}{5}\overrightarrow{AC}$$

   $$\overrightarrow{AD} = \frac{2}{5}(-\boldsymbol{u} - \boldsymbol{v})$$

   $$\overrightarrow{AD} = -\frac{2}{5}\boldsymbol{u} - \frac{2}{5}\boldsymbol{v}$$

3. Find $\overrightarrow{BD}$

   Drawing $\overrightarrow{BD}$ to help us see the path:



Let's find our path. We start at B, go to A, and from there to D. So we have the vector equation:

$$\overrightarrow{BD} = \overrightarrow{BA} + \overrightarrow{AD}$$

We already have $\overrightarrow{BA} = \boldsymbol{u}$ and $\overrightarrow{AD} = -\frac{2}{5}\boldsymbol{u} - \frac{2}{5}\boldsymbol{v}$, we just need to substitute and collect like terms:

$$\overrightarrow{BD} = \overrightarrow{BA} + \overrightarrow{AD}$$

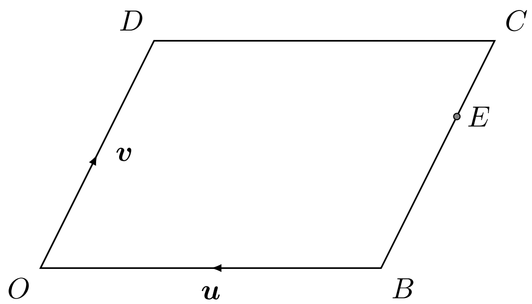
$$\overrightarrow{BD} = \boldsymbol{u} + \left(-\frac{2}{5}\boldsymbol{u} - \frac{2}{5}\boldsymbol{v}\right)$$

$$\overrightarrow{BD} = \boldsymbol{u} - \frac{2}{5}\boldsymbol{u} - \frac{2}{5}\boldsymbol{v}$$

$$\overrightarrow{BD} = \frac{3}{5}\boldsymbol{u} - \frac{2}{5}\boldsymbol{v}$$

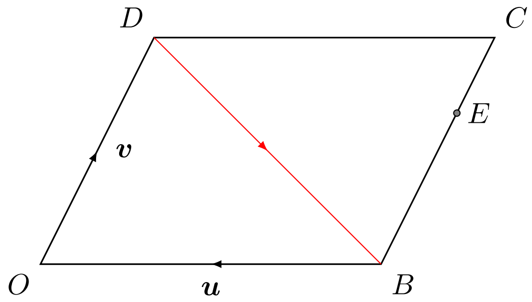
Finding vectors: position vector

Given a parallelogram $OBCD$, vectors $\boldsymbol{u}$ and $\boldsymbol{v}$, and knowing that the point E divided the side BC in the ratio $BE : EC = 2 : 1$.



1. Find $\overrightarrow{DB}$.

   Let's start by drawing $\overrightarrow{DB}$ to help us find our path:



Our path starts at D, then can go to O, and from O to B. In vector form:

$$\overrightarrow{DB} = \overrightarrow{DO} + \overrightarrow{OB}$$

Remember that $\overrightarrow{DO} = -\overrightarrow{OD}$ and that $\overrightarrow{OB} = -\overrightarrow{BO}$. Substituting:

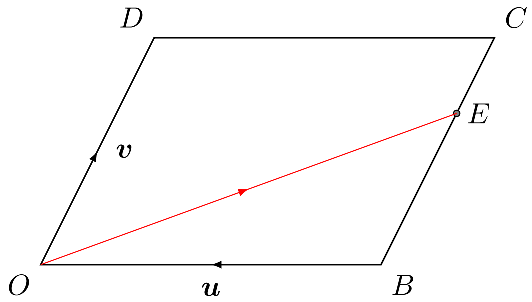
$$\overrightarrow{DB} = \overrightarrow{DO} + \overrightarrow{OB}$$

$$\overrightarrow{DB} = -\overrightarrow{OD} + -\overrightarrow{BO}$$

$$\overrightarrow{DB} = -\boldsymbol{v} - \boldsymbol{u}$$

2. Find the position vector of point E.

A position vector of a point is just the vector that goes from O (the origin) to the point itself. In this case, the position vector of point E is the vector $\overrightarrow{OE}$. Let's draw $\overrightarrow{OE}$:



Starting at O, we could go to B, and from B to E. Thus, our vector equation is

$$\overrightarrow{OE} = \overrightarrow{OB} + \overrightarrow{BE}$$

We already know that $\overrightarrow{OB} = -\overrightarrow{BO} = -\boldsymbol{u}$. We need to find $\overrightarrow{BE}$. Notice that $\overrightarrow{BE}$ is a "tiny" $\overrightarrow{BC}$, as they both same direction, but $\overrightarrow{BE}$ has a smaller magnitude.

Let's use the same strategy: if we divide the segment BC into 3 parts $(2+1)$, we know BE takes 2 of those 3, so we have

$$BE = \frac{2}{3}BC$$

In vector form:

$$\overrightarrow{BE} = \frac{2}{3}\overrightarrow{BC}$$

We need to find $\overrightarrow{BC}$. But we already have it! Remember that it does not matter where a vector is: if it has the same magnitude and the same direction it is the same. Given that $OBCD$ is a parallelogram, the sides OD and BC have the same length and are parallel. Therefore, $\overrightarrow{BC} = \overrightarrow{OD} = \boldsymbol{v}$. Now we know that

$$\overrightarrow{BE} = \frac{2}{3}\overrightarrow{BC} = \frac{2}{3}\boldsymbol{v}$$

Substituting in our equation for $\overrightarrow{OE}$:

$$\overrightarrow{OE} = \overrightarrow{OB} + \overrightarrow{BE}$$

$$\overrightarrow{OE} = -\overrightarrow{BO} + \frac{2}{3}\overrightarrow{BC}$$

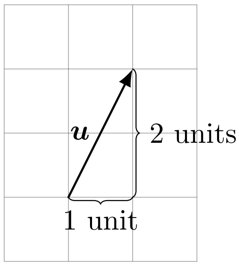
$$\overrightarrow{OE} = -\boldsymbol{u} + \frac{2}{3}\boldsymbol{v}$$

42.6. Exam hints

This may sound repetitive: find your path and write it just like we have been doing. From there everything solves itself.

Summary

- A *vector* is used to express a quantity that has *magnitude*, its "size", and *direction;*
- We can represent vectors both using drawings and *column notation.* For instance, the vector $\boldsymbol{u} = \begin{pmatrix}1\\2\end{pmatrix}$ has x-component 1 and y-component 2, which means it goes "1 to the right" and "2 up". Graphically:



- The *magnitude* of a vector $\boldsymbol{u} = \begin{pmatrix} x \\ y \end{pmatrix}$ is denoted by $|\boldsymbol{u}|$ and we can calculate it using a formula:

$$|\boldsymbol{u}| = \sqrt{x^2 + y^2}$$

- Two vectors are *equal* if the have *the same magnitude and direction*;
- When you *multiply a vector by a scalar*, you multiply each component of the vector by the scalar;
- Given a vector $\boldsymbol{u}$, the *opposite vector* $-\boldsymbol{u}$ is a vector that has *the same magnitude*, but *opposite directions*;
- To *add vectors* graphically, you draw the first vector anywhere and keep drawing the next vector where the previous ended. When you're done, you join the beginning of the first vector with the end of the last. To add vectors in column form, you just add components, like a matrix;
- To find *vectors* that represent a line, *find a path* that starts at the first "letter" and finished at the second.

43. Transformations

Part V.

Trigonometry

44. Trigonometry in the right triangle

45. Trigonometric functions and their graphs

46. Trigonometric equations

47. Trigonometry in the non-right triangle

48. Application of trigonometry: bearings

Part VI.

Probability and statistics

49. Probability

49.1. Why learn probability

We are rarely sure of anything in life. We are not sure if it will rain or not; we are not sure we, the sun, or the universe will be here tomorrow. We are not sure which side a dice will land on, nor which face of a coin will be showing when it lands. *Randomness* is all around us, every day, all the time.

Mathematics, being very intrusive, has a tool to deal with these uncertainties: probability. Given that life is filled with uncertainty, I believe that learning probability is quite important.

49.2. What is a probability and basic facts

Probability is a measure of the likelihood of something random happening. Basically, a probability is a *number* that measures how likely an outcome of a random process is. Probability theory is very complicated, and it is something you can study in much further depth should you choose to become a mathematician or related (highly recommended!).

Let us introduce some important vocabulary. An *event* is the result of anything random that may happen. Say you roll a regular six sided dice: one possible event is rolling a 2. If you throw a coin, there are two possible events: heads or tails (please do not be the person that takes the "side" into consideration).

I mentioned that probabilities are numbers we assign to events, the outcomes of random processes. There are two very important rules: we assign the number 0 to any event that has no possibility of happening and 1 to any event that is sure to happen. Let us some examples:

- You roll a regular six sided dice. The event "rolling a 7" is impossible to happen, so the probability we assign to it is 0;
- You have only black socks in a drawer. You take one sock at random. The event "picking a black sock" is sure to happen, so the probability of it is 1.

Both these events are very extreme, however. Usually, events have *some* probability of happening. Hence, most events will have probabilities between 0 and 1. This is another important fact: probabilities are *always numbers between* 0 *and* 1.

It is kinda annoying to keep writing "the probability of event X", though. So let us introduce a notation. We will write

$$P(X)$$

to refer to the "probability of event X happening", or simply the "probability of X". In our previous examples, then, we could have written:

- You roll a regular six sided dice, and consider the event "rolling a 7". We can write

$$P(7) = 0$$

  to mean that the probability of rolling a 7 is 0.

- In the drawer that only has black socks, we take one sock. The probability of taking a black sock is written as

$$P(\text{black}) = 1$$

Do remember, though, that in general, probabilities are not 0 or 1, but something in between. In general, we can write that the probability of an event X is

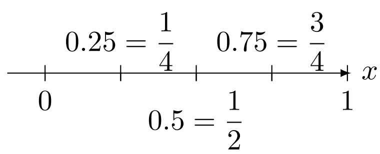
$$0 \leq P(X) \leq 1$$

Knowing this you can always at least know if your answer to probability questions make sense: if you get a negative number or something greater than 1, you are definitely wrong!

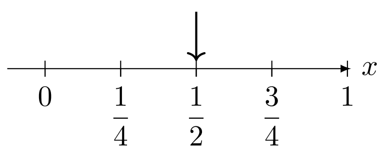
A final, and very important fact about probabilities is that if we add the probabilities of all possible outcomes, we always need to obtain 1 as a result. This makes sense as if we consider all possibilities and add their probabilities, we must reach 100%, or 1.

49.3. Probability scales

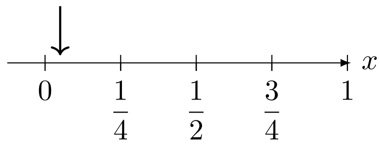
One way to represent probabilities graphically is using probability scales. These are just a number line from 0 to 1:



Each number from 0 to 1 can represent a probability (either in decimal, fraction or percentage format). We can then point at particular places representing probabilities. For instance, if we throw a coin, the probability of it landing on heads is $\frac{1}{2}$:



We can also represent estimates of probabilities. Say we believe there is small probability of raining tomorrow, close to 0. We can draw an arrow "somewhere" around 0 then:



49.4. Relative frequency as estimates of probabilities

Let us now consider a way to estimate probabilities from real world experiments.

A classic example is the following. We want to estimate the probability of throwing a coin and it landing on heads. What we can do is throw the coin a number of times and tabulate the results. Say we throw the coin 100 times and obtain the following:

Face	Heads	Tails
Frequency	63	37

What we can do is use the relative frequency of Heads as an approximation of the probability of the coin landing on Heads. The *relative frequency* of an event is the number of times the event happens divided by the total number of times we tried the experiment (which we normally call a *trial*). In this case, we threw the coin 100 times and we got 63 Heads, so we have

$$P\left(\text{Heads}\right) = \frac{\text{number of heads}}{\text{total number of trials}} = \frac{63}{100}$$

We can also do the same for Tails:

$$P\left(\text{Tails}\right) = \frac{\text{number of tails}}{\text{total number of trials}} = \frac{37}{100}$$

It is important to notice that the relative frequencies follow the probability rules we saw above: they are both positive and add up to 1:

$$P\left(\text{Heads}\right) + P\left(\text{Tails}\right) = \frac{63}{100} + \frac{37}{100} = \frac{100}{100} = 1$$

An important warning: you may have wondered that it is weird to estimate the probability of a coin flip not being 0.5 or 50%. Do remember that when we say that the probability (not the chance!) of a coin landing on heads is 50% we are assuming that the coin is *unbiased*, that is, that both heads and tails have the same probability of landing. Before actually throwing a coin we are not sure if its unbiased or not (and to actually check that you would need some statistics).

49.5. Expected frequency (experimental probability)

We can use relative frequencies to estimate the number of times an event will happen in a fixed number of trials. Notice that this is different than finding probabilities: we now want to find "number of times" something will happen.

Take the coin example in the last section. We estimated, from experiment, that

$$P(H) = \frac{63}{100} \text{ and } P(T) = \frac{37}{100}$$

Let us say we will now throw the coin 2000 times. We want to estimate how many times it will land on heads. We can do this by multiplying the probability of heads by the number of times we will throw the coin:

$$\text{number of throws} \times P(H) = 2000 \times \frac{63}{100} = 1260$$

In general, to estimate the number of times an event will happen, we use the formula:

$$\text{Expected number of occurrences} = \text{number of trials} \times \text{probability of event happening}$$

More concisely, let us define $E(X)$ as the expected number of occurrences of event X, n as the number of times we will repeat the experiment and, as we say above, $P(X)$ as the probability of X occurring. Hence, we have:

$$E(X) = nP(X)$$

Let us see an example of a possible exercise with the results from this and the last section.

I roll a 4 sided dice 100 times, and record the number each face lands facing down. The results are in the following table:

Value	1	2	3	4
Frequency	27	13	38	22

Let us use relative frequencies to estimate the probability of obtaining each face. Remember that

$$P(X) = \frac{\text{number of times } X \text{ happens}}{\text{total number of trials}}$$

Hence, for the number 1:

$$P(1) = \frac{\text{number of 1s}}{\text{total number of rolls}} = \frac{27}{100}$$

and the same reason goes for all other values:

$$P(2) = \frac{\text{number of 2s}}{\text{total number of rolls}} = \frac{13}{100}$$

$$P(3) = \frac{\text{number of 3s}}{\text{total number of rolls}} = \frac{38}{100}$$

$$P(4) = \frac{\text{number of 4s}}{\text{total number of rolls}} = \frac{22}{100}$$

Now, if we roll the same dice 125 times, how many times we expect to roll each value? We can use $E(X) = nP(X)$. For 1:

$$E(1) = \underbrace{125}_{\text{number of rolls}} \times \underbrace{\frac{27}{100}}_{P(1)} = 33.75$$

and for the other values:

$$E(2) = 125 \times \frac{13}{100} = 16.25$$
$$E(3) = 125 \times \frac{38}{100} = 47.5$$
$$E(4) = 125 \times \frac{22}{100} = 27.5$$

It is important to notice that $E(1) + E(2) + E(3) + E(4) = 125$. Also, it is fine for the expected number of times we roll a value to be a decimal number: this is just an estimate, it does not mean we will actually roll 47.5 number 3s if we roll the dice 125 times.

49.6. Theoretical probabilities

It would be quite annoying if we had to do a very large number of experiments to determine probabilities by relative frequencies, though. And being smart lazy people, we can worry about *theoretical probabilities*. Remember that probability is a measure of how likely a random experiment has a certain outcome. What if we counted the number of all possible outcomes the experiment has, and also counted the number of ways the outcome we are interested in happens, and simply divided them one by another? Say we throw a regular 6 sided dice and we want to find the probability of rolling a 4. There are 6 possible outcomes: $1, 2, 3, 4, 5$ and 6. We are interested in one outcome of those 6. Hence, the probability of rolling a 4 is given by:

$$P(4) = \frac{\text{number of ways we can roll a 4}}{\text{total number of outcomes}} = \frac{1}{6}$$

I like remembering this as "the probability that X happens is equal to the number of times X happens divided by the total number of possible outcomes" or "probability of something happening is the number of times it can happen divided by the total number of possibilites":

$$P(X) = \frac{\text{number of ways } X \text{ happens}}{\text{total number of outcomes}}$$

If we have a spinner with the numbers $1, 2, 2, 2, 3, 4$ and we want to find the probability of spinning a 2, then:

$$P(2) = \frac{\text{number of ways we can roll a 2}}{\text{total number of possible rolls}} = \frac{3}{6}$$

49.7. Important results

Before we discuss the next results, we need to understand a concept which is very important, called *independence*. Two events are called independent when the fact of one of the happening does not change the probability of the other one happening. An example makes this very clear: if you throw a coin twice, and the first time it lands of tails, it does not change the probability of the *next* throw landing on tails. When two events are not independent, we call them dependent. For instance, if we remove a card from a standard 52 card deck[1] and burn it, it does affect the probability of the next card being a King.

As a warning, independence is very different from being mutually exclusive: this means that if one event happens, the other one cannot happen. Say you throw a coin and it lands on heads. This makes it impossible for the coin to land on tails, hence heads and tails are mutually exclusive events. Independent events, on the other hand, can happen at the same time.

49.7.1. The "and" rule

If two events are independent, the probability of them happening *together* is equal to the product of their individual probabilities. Using our notation:

$$P(A \text{ and } B) = P(A)P(B)$$

To make some intuitive sense of this, let us think of an absurd example. One of our events will be you being hit by lightning, and the second one you to learn how to fly. Considering these events separately, they are very unlikely. You have to agree that is even *less likely* for you to both be hit by lightning *and* to learn how to fly at the same time! As probabilities are numbers between 0 and 1, when we multiply them the result is a smaller number. So, when we multiply the probabilities of the individual events, the resulting probability is smaller.

[1] A standard deck of cards has 52 cards, divided in 4 suits: hearts ♡, diamonds ♢, spades ♠ and clubs ♣, each with 13 cards. Hearts and diamonds cards are red, whereas clubs and spades cards are black. Finally, each suit has the following cards: A, 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q and K.

49.7.2. The "or" rule

If two events are mutually exclusive, the probability of one or[2] the other happening is equal to the sum of the individual probabilities. In mathematicianese:

$$P(A \text{ or } B) = P(A) + P(B)$$

Again, this makes intuitive sense: we are increasing the possibilities we are interested in, hence the probability of at least one of the events happening is bigger.

As an example, let us consider the event that you graduate your maths university with distinction and the other event that you graduate with merit. These events are mutually exclusive, so the probability of one or the other happening is the sum of their probabilities.

49.8. Sample space diagrams: representing random outcomes

Let us use all this nice foundation to find some probabilities. We are going to use sample space diagrams to represent all possible outcomes of our experiments, and use them to find the probabilities we need.

49.8.1. Tables of outcomes

A nice way to represent outcomes when two random things happen is a table of outcomes.

Say we roll a 6 sided dice and throw a coin. We can represent all possible outcomes in a table:

Dice / Coin	1	2	3	4	5	6
Heads (H)	$H,1$	$H,2$	$H,3$	$H,4$	$H,5$	$H,6$
Tails (T)	$T,1$	$T,2$	$T,3$	$T,4$	$T,5$	$T,6$

This table has all 12 possible outcomes we have when we combine a coin toss with a dice roll (the 12 comes from having 6 possibilities for the dice and 2 for the coin: $6 \times 2 = 12$). If you focus on this cell, for instance:

Dice / Coin	1	2	3	4	5	6
Heads (H)	$H,1$	$H,2$	$H,3$	$H,4$	$H,5$	$H,6$
Tails (T)	$T,1$	$T,2$	$T,3$	$T,4$	$T,5$	$T,6$

It represents the outcome throwing a Heads and rolling a 2 on the dice.

[2] This or is *exclusive*, which means that both events cannot happen at the same time (as they are mutually exclusive). We just have to be careful with our language.

We can now use this table to calculate probabilities. For instance, let us find the probability of rolling a Tails **and** a prime number[3]. Remember that probability can be seen as "what you want over what you have", so let us find what we want in the table:

Coin \ Dice	1	2	3	4	5	6
Heads (H)	$H, 1$	$H, 2$	$H, 3$	$H, 4$	$H, 5$	$H, 6$
Tails (T)	$T, 1$	$T, 2$	$T, 3$	$T, 4$	$T, 5$	$T, 6$

The blue cells are the ones in which the coin landed on Tails and the dice rolled a prime number. We have 3 of them ("what we want") and we have 12 total outcomes ("what we have"), so:

$$P\left(\text{Tails and a prime number}\right) = \frac{3}{12}$$

Let us now ask the question: what is the probability of rolling a power of 2[4] **or** Heads. Again, let us colour the cells which have the outcomes which satisfy this:

Coin \ Dice	1	2	3	4	5	6
Heads (H)	$H, 1$	$H, 2$	$H, 3$	$H, 4$	$H, 5$	$H, 6$
Tails (T)	$T, 1$	$T, 2$	$T, 3$	$T, 4$	$T, 5$	$T, 6$

Notice that we all outcomes with Heads count, as we want Heads **or** a power of 2. The outcomes with Tails that have a power of 2 are also part of our event. Hence, we have 9 outcomes we are interested in ("what we want") and we can find the probability of it by dividing by 12 ("what we have"):

$$P\left(\text{power of 2 or Heads}\right) = \frac{9}{12}$$

A classic example is asking questions when we roll two dice and **add** their results. Let us represent that in a table again:

Dice 1 \ Dice 2	1	2	3	4	5	6
1	$1+1=2$	$1+2=3$	4	5	6	7
2	$2+1=3$	$2+2=4$	5	6	7	8
3	$3+1=4$	$3+2=5$	6	7	8	9
4	$4+1=5$	$4+2=6$	7	8	9	10
5	$5+1=6$	$5+2=7$	8	9	10	11
6	$6+1=7$	$6+2=8$	9	10	11	12

[3]Remember that prime numbers are numbers which have exactly two distinct factors. From 1 to 6, the prime numbers are 2,3 and 5. See Chapter 2 for a reminder.

[4]The powers of 2 in a 6 sided dice are 1, 2 and 4. See Chapter 15 to refresh your memory.

We have 36 possibilities (6 for each dice), and each cell in the table has the result of adding the roll represented in the row with the roll represented by the column. Just from this table we have some interesting information: the minimum sum we have is 2 (blue cell), the maximum sum is 12 (green cell) and the most common sum is 7 (red cells):

Dice 1 \ Dice 2	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

We can, then, answer some cute questions, such as: what is the probability of the most common result when adding the values of the two dice? We know the most common result is 7, and there 6 of them in our table. Hence, we have 6 ways to get "what we want" and a total of 36 outcomes (the "what we have"):

$$P\,(\text{most common value}) = P\,(7) = \frac{6}{36}$$

Another nice question is: what is the probability of rolling a total sum greater than 8? We can colour the cells with values larger than 8:

Dice 1 \ Dice 2	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

We have, then, 10 outcomes larger than 8 (notice we do not include 8, as we just want *larger* results). Using our usual "what we want" divided by "what we have":

$$P\,(\text{result larger than } 8) = \frac{10}{36}$$

A final example: what is the probability of rolling a total sum of **at most** 3? Here we want sums of 2 or 3, as at most 3 includes 3 itself:

Dice 1 \ Dice 2	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

So, we have 3 ways of obtaining 2 or 3, hence we have that

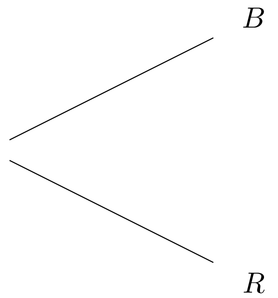
$$P\left(\text{sum of at most 3}\right) = \frac{3}{36}$$

49.8.2. Tree diagrams

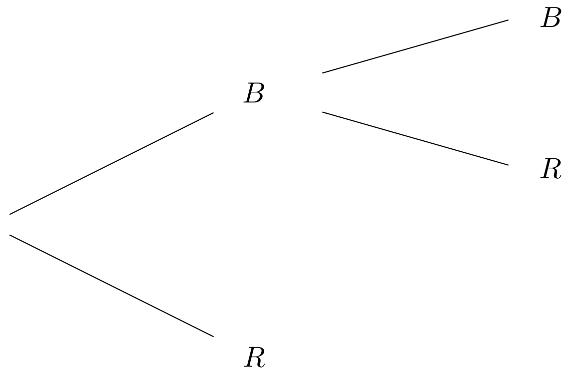
Tree diagrams are very useful to represent events that happen successfully.

A classic example is removing balls from a bag. Say we have a bag with 8 red balls (R) and 6 blue balls (B). We will remove one ball from the bag, note what colour it was, *put it back in the bag*, and remove another ball to look at its colour. Let us think about all the possible outcomes we can have.

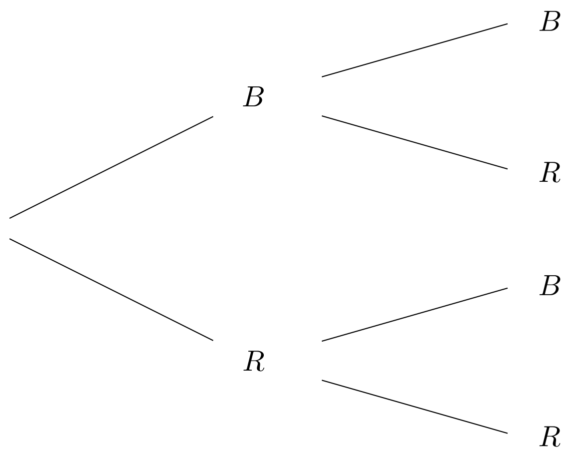
The first ball may be R or B, which we can represent by a "double path":



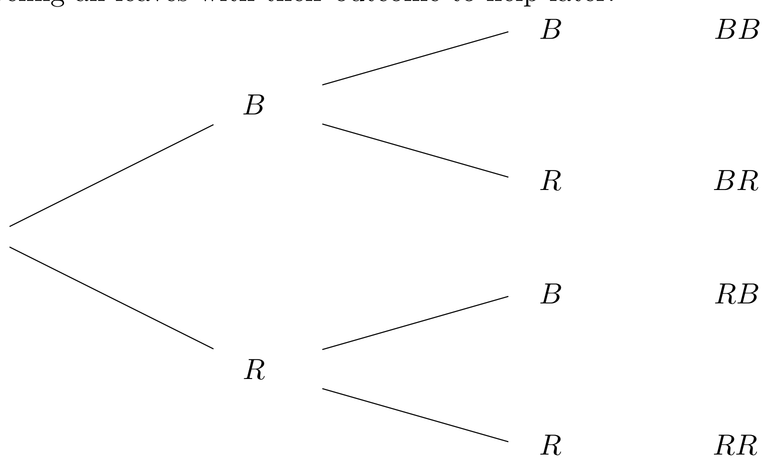
Now, if the first ball was blue (the top branch), we put it back in the bag and take another ball, so we have another split:



and we have the same situation if the first ball was red (the bottom branch), as we put it back and draw another ball:



You can see, then, that we have 4 possibilities how to draw a ball, put it back, and draw another: each possible outcome is one of the "leaves" of the tree. For instance, if the first ball was blue, and the second was also blue, we would be in the top outcome. I like labeling all leaves with their outcome to help later:



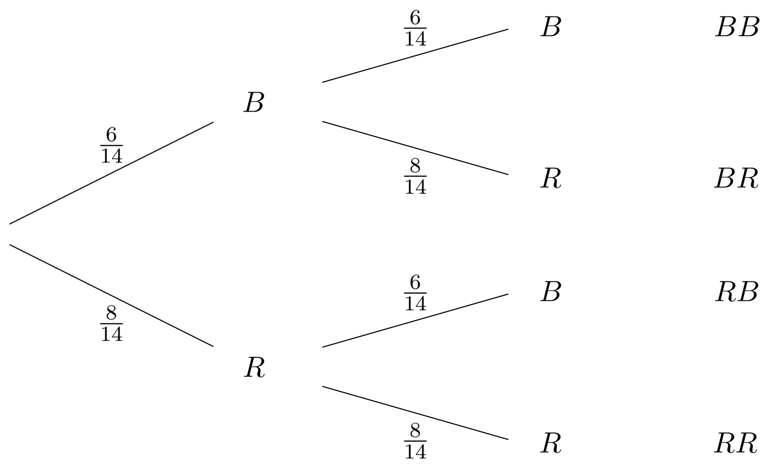
Now, we just need to add the probabilities to our tree. To do this, let us think about our bag: it has 8 red balls and 6 blue balls. In total, then, it has 14 balls. We can find the probability of drawing a red ball by using our "what we want divided by what we have" strategy:

$$P(R) = \frac{\text{number of } R \text{ in bag}}{\text{total number of balls}} = \frac{8}{14}$$

and the same for B:

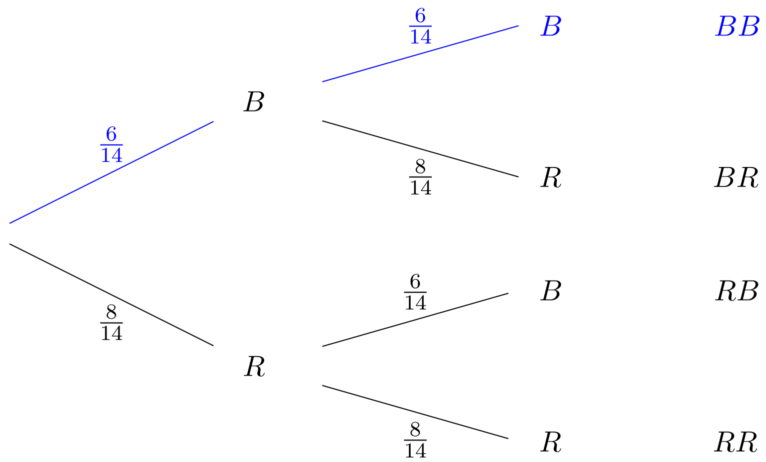
$$P(B) = \frac{\text{number of } B \text{ in bag}}{\text{total number of balls}} = \frac{6}{14}$$

As we *replace* the first ball after we draw it, the second draw has exactly the same probabilities of the first. We can now write the probabilities on our tree as well:



We can now, using the and rule, find the probability of reaching any leaf of the tree (remember that a leaf is any point to the right where the tree ends). As each draw from the bag is independent from the other (as we are replacing the ball we remove), we can just multiply the probabilities in a path to find the overall probability.

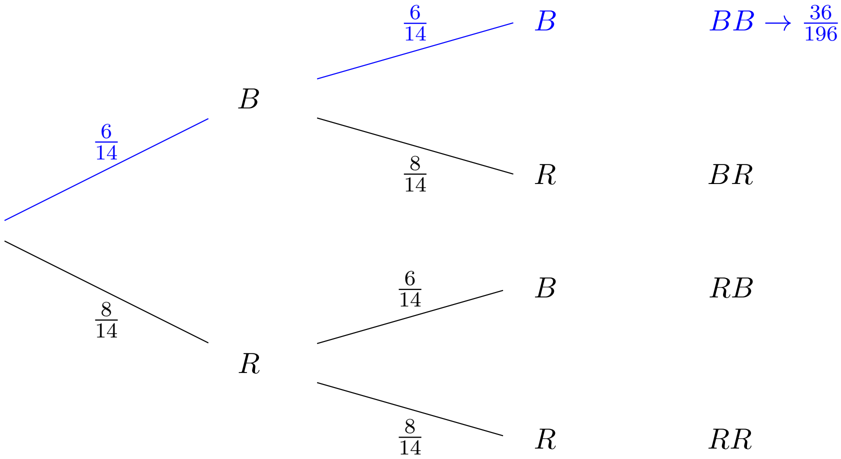
Take, for instance, the path BB, which means we drew a blue ball both times:



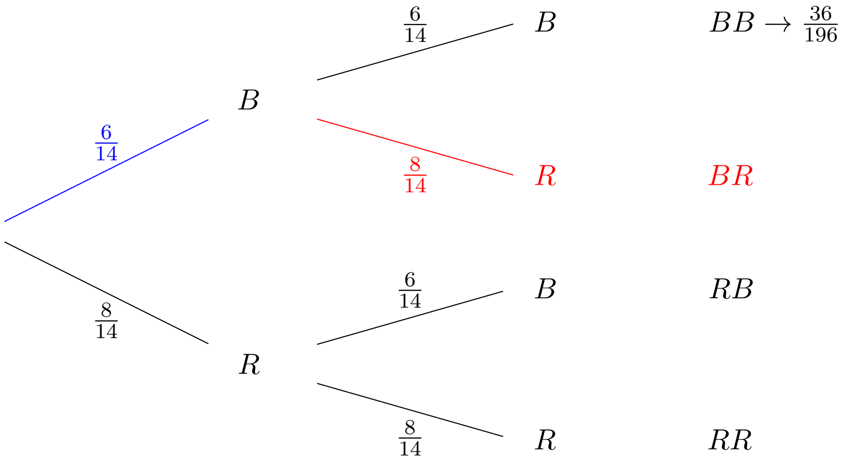
in order to find the probability of getting two blue balls, we multiply the probabilities in the path (we are using the and rule):

$$P(BB) = \underbrace{\frac{6}{14}}_{\text{1st } B \text{ prob}} \times \underbrace{\frac{6}{14}}_{\text{2nd } B \text{ prob}} = \frac{36}{196}$$

which I also like adding to the right of the outcome:



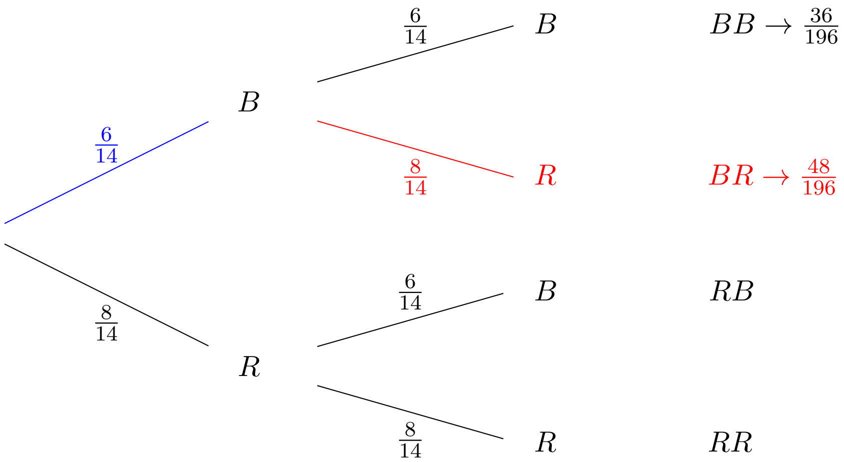
For the case we first draw a blue ball, followed by a red ball:



we can find the probability of this happening in the same way: we multiply the probabilities in the path (but remember, we are just using the and rule):

$$P(BR) = \underbrace{\frac{6}{14}}_{B \text{ prob}} \times \underbrace{\frac{8}{14}}_{R \text{ prob}} = \frac{48}{196}$$

which we can add to the tree:



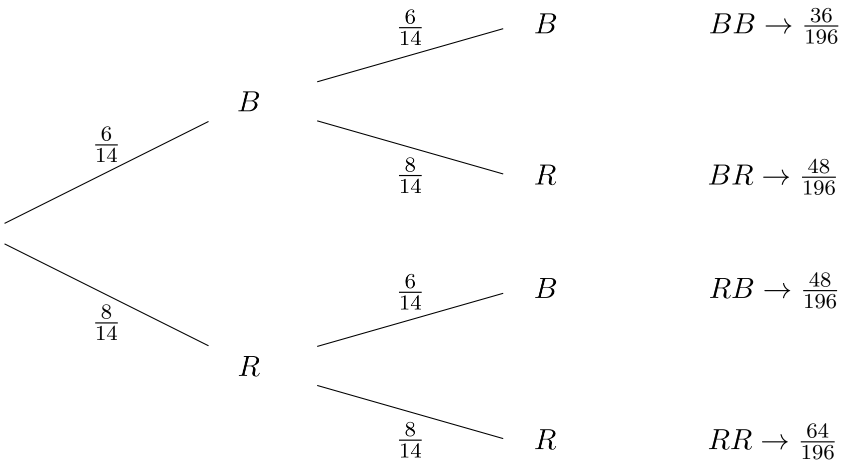
The branches below follow the same idea. For the RB outcome, we have:

$$P(RB) = \underbrace{\frac{8}{14}}_{R} \times \underbrace{\frac{6}{14}}_{B} = \frac{48}{196}$$

and the same for the RR outcome:

$$P(RR) = \underbrace{\frac{8}{14}}_{R} \times \underbrace{\frac{8}{14}}_{R} = \frac{64}{196}$$

Let us add all these probabilities to our tree:



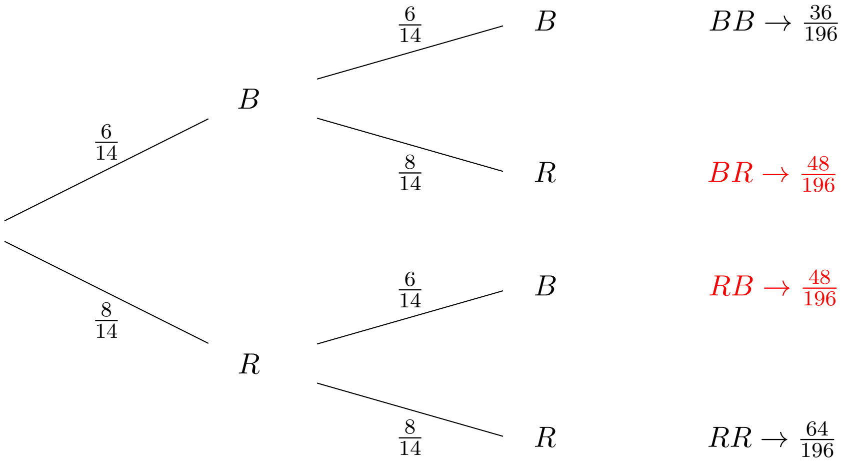
Before we continue, notice that if we add the probabilities of all 4 outcomes we do obtain 1, as expected:

$$P(BB) + P(BR) + P(RB) + P(RR)$$

$$\frac{36}{196} + \frac{48}{196} + \frac{48}{196} + \frac{64}{196} = \frac{196}{196} = 1$$

Also, notice that each time we branch the tree, for instance when we take the first ball, those probabilities also add up to 1. This is really useful to check if you did not do any calculation mistake.

Finally, after completing the tree, we can use it to solve harder problems. For instance, let us find the probability of drawing exactly one red ball. We can find those outcomes in our tree:



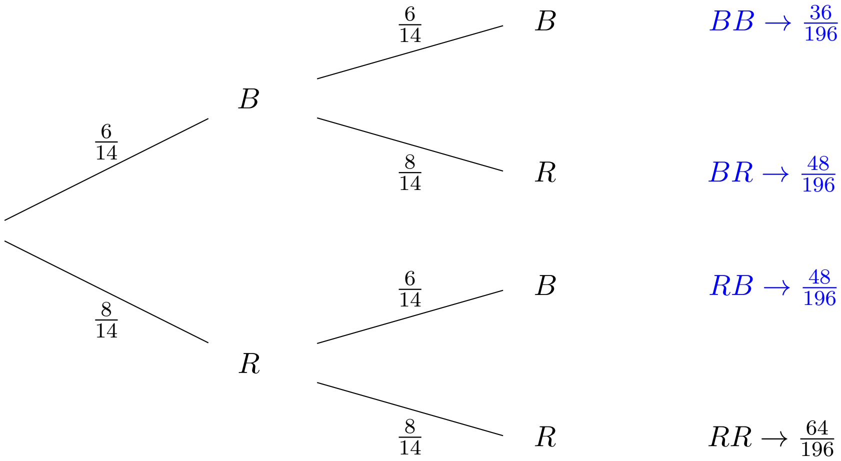
By our lovely colouring, we can see that to get exactly one red ball, we need either the RB or the BR outcomes. Hence, we want

$$P(RB \text{ or } BR)$$

and we do know that when events are mutually exclusive (both cannot happen at the same time), the or rule says we can simply add the probability of each outcome:

$$P(RB \text{ or } BR) = P(RB) + P(BR) = \frac{48}{196} + \frac{48}{196} = \frac{96}{196}$$

Finally, let us find the probability of drawing *at least one* blue ball. "At least one" means either one or two blue balls, which we can show in our tree:



We can again use the or rule and add those probabilities:

$$P(\text{at least one } B) = P(BB) + P(BR) + P(RB)$$

$$P(\text{at least one } B) = \frac{36}{196} + \frac{48}{196} + \frac{48}{196} = \frac{132}{196}$$

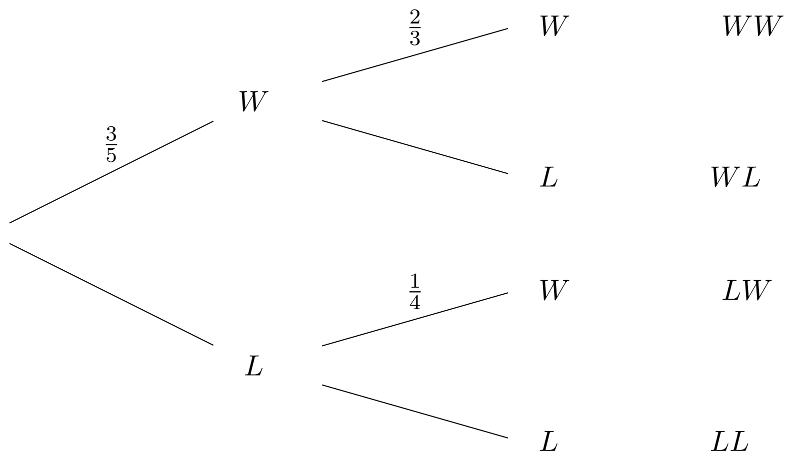
There is a faster way to find this, though. The probability of drawing at least one blue ball is also the total probability *minus* the probability of only drawing red balls:

$$P(\text{at least one } B) = 1 - P(RR)$$

and it is much easier to find this:

$$P(\text{at least one } B) = 1 - \frac{64}{196} = \frac{196}{196} - \frac{64}{196} = \frac{132}{196}$$

A different example: let us say you will play two games of tennis against someone. The probability you win the first game is $\frac{3}{5}$. If you win it, the probability you will win the second game is $\frac{2}{3}$, but if you lose the first game, the probability you will win the second is only $\frac{1}{4}$. Let us draw a probability tree for this situation with the information given. Let us denote a win by W and a loss by L:



Unlike the previous example, we have some gaps to fill, as we were not given the probabilities of losing in any situation. But we can find them remembering that winning and losing are *complements*, that is, they must add up to 1. So, for the first game, as you have

$$P(W) = \frac{3}{5}$$

we can find the probability of losing by taking away $P(W)$ from 1:

$$P(L) = 1 - \frac{3}{5} = \frac{2}{5}$$

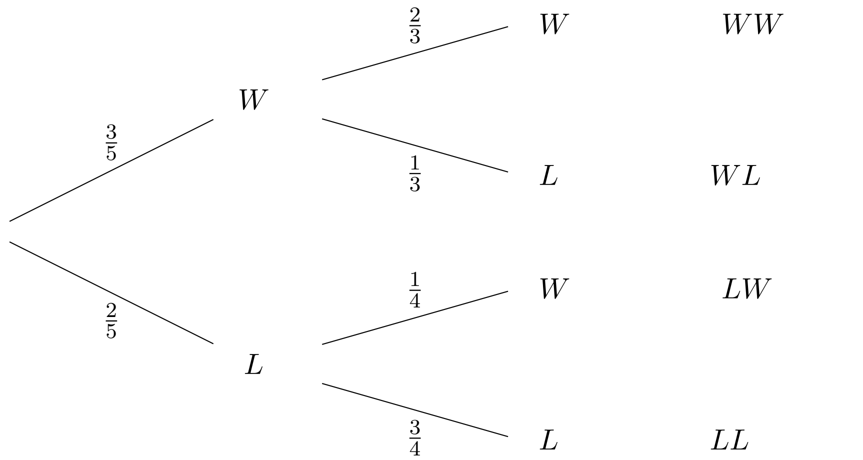
If you win, you have probability $\dfrac{2}{3}$ of winning again, so the probability of losing at the game is:

$$P(L) = 1 - \frac{2}{3} = \frac{1}{3}$$

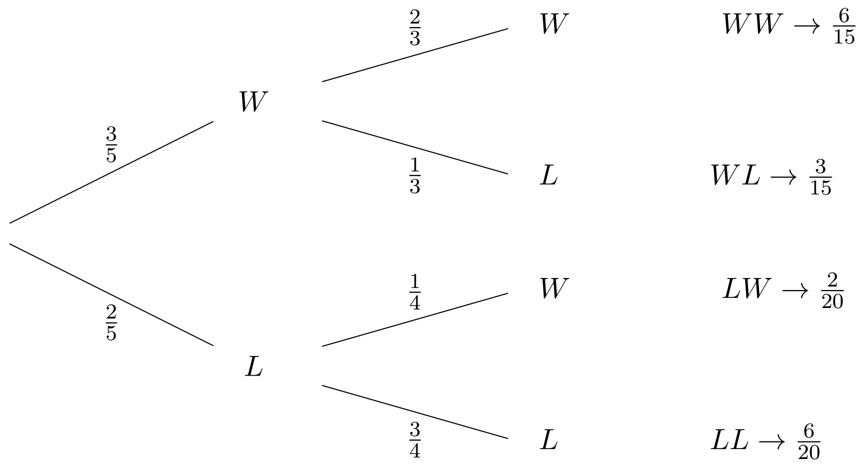
(I am using W and L for all games as it is clear from the context what we are talking about. We could be more pedantic about notation, but why?). Finally, if you lose the first game, the probability of winning the second is $\dfrac{1}{4}$. Thus, the probability of you losing it is

$$P(L) = 1 - \frac{1}{4} = \frac{3}{4}$$

Let us add this information to the tree:



And let us find the probability of each leaf by multiplying the probabilities in each path (remember we are just using the and rule):



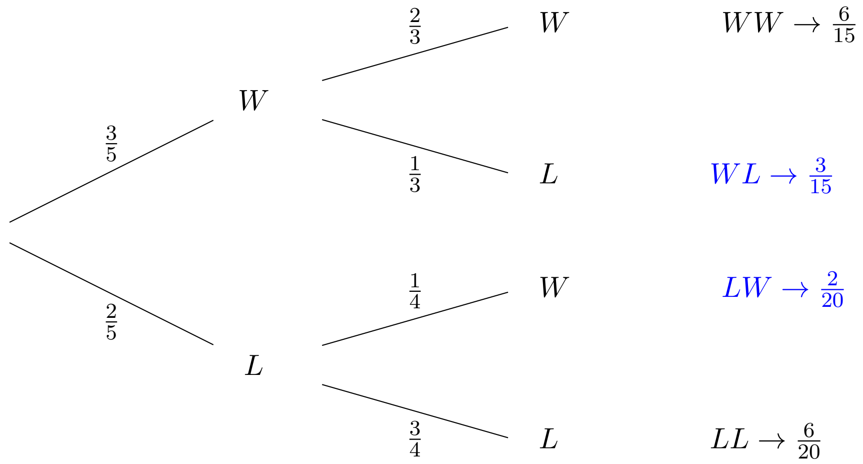
To check, let us add all the probabilities and see if they total to 1:

$$\frac{6}{15} + \frac{3}{15} + \frac{2}{20} + \frac{6}{20} = \frac{24 + 12 + 6 + 18}{60} = \frac{60}{60} = 1$$

Using the tree we can the probability of losing both games, for instance. The outcome LL has probability

$$P(LL) = \frac{6}{20}$$

Now, let us use the tree to find the probability of winning exactly one game. We are interested in the outcomes WL or LW:

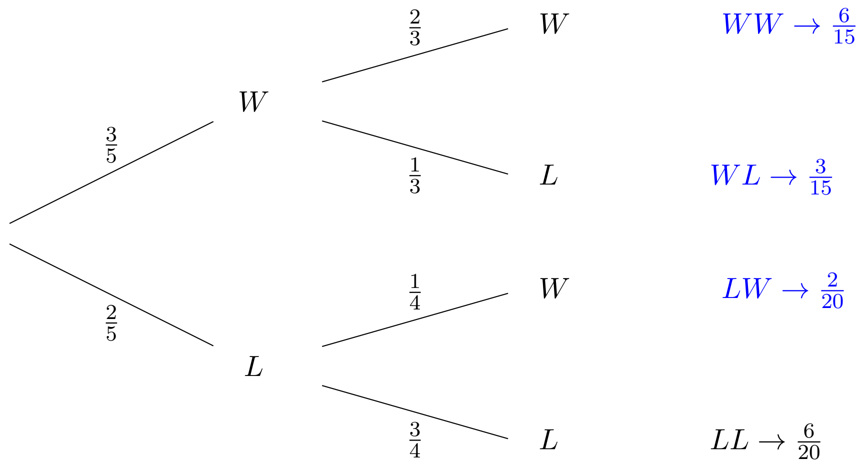


We can add those results to obtain our probability:

$$P(1W) = P(WL \text{ or } LW) = P(WL) + P(LW)$$

$$P(1W) = \frac{3}{15} + \frac{2}{20} = \frac{12}{60} + \frac{6}{60} = \frac{18}{60}$$

What about winning at least one game? Remember that winning at least one game is composed of the outcomes WW, WL and LW:



We can add those probabilities:

$$P(\text{at least 1 } W) = P(WW) + P(WL) + P(LW)$$

$$P(\text{at least 1 } W) = \frac{6}{15} + \frac{3}{15} + \frac{2}{20} = \frac{14}{20}$$

Remember that we could also find this using:

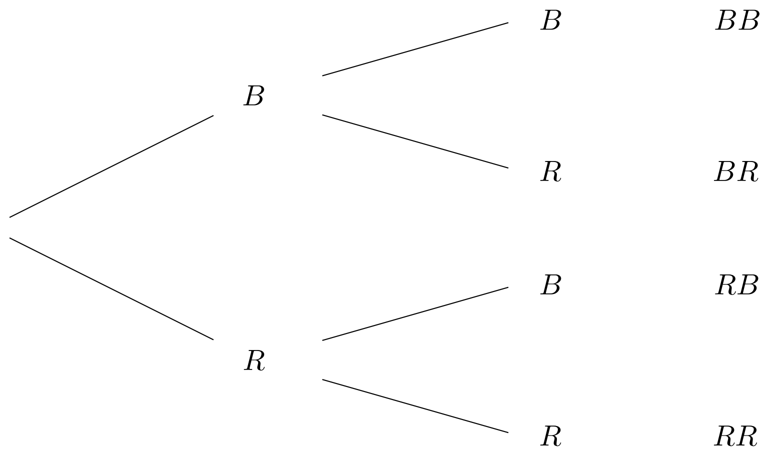
$$P(\text{at least 1 } W) = 1 - P(LL)$$

as losing both games is complementary to winning at least one. Hence,

$$P(\text{at least 1 } W) = 1 - \frac{6}{20} = \frac{14}{20}$$

There is a final type of probability tree problems that may appear which, to be honest, is not of the same kind as these we covered so far. If we were to analyze them formally, they should be in Section 49.10, but I will put them here as in the IGCSE the language for them is not the same.

Say that we have a bag with 5 red balls (R) and 8 blue balls (B). We remove one ball, *but we do not put it back in the bag.* This time, we do not have reposition, so the probabilities will change after we remove one ball. The general format of the tree will be the same:



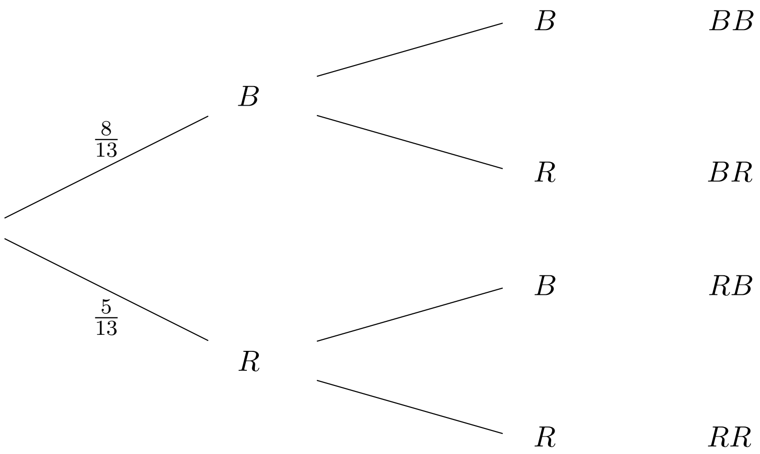
And when we add the probabilities for the first branch, nothing changes: for the first ball, we can either remove a R, with probability

$$P(R) = \frac{\text{number of } R}{\text{total balls}} = \frac{5}{13}$$

or a B, with probability

$$P(B) = \frac{\text{number of } B}{\text{total balls}} = \frac{8}{13}$$

and we can write these probabilities:



For the second branches, though, we have to be careful. Depending on what colour we took on the first draw, the probabilities will change.

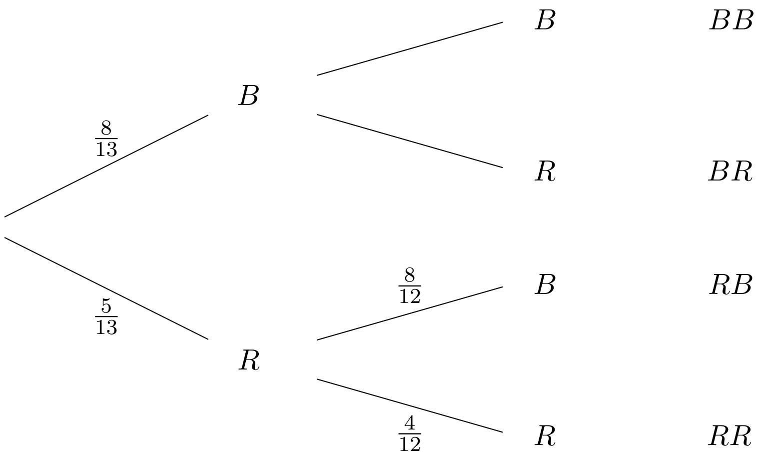
Let us first assume we drew a red ball. Our bag now changes: we have 4 red balls and we still have 8 blue balls. Therefore,

$$P(R) = \frac{\text{number of red balls}}{\text{total balls}} = \frac{4}{4+8} = \frac{4}{12}$$

and for B:

$$P(B) = \frac{\text{number of blue balls}}{\text{total balls}} = \frac{8}{12}$$

Notice that these values still add up to 1, as they must given that they are all possibilities. Let us add these probabilities to the tree:



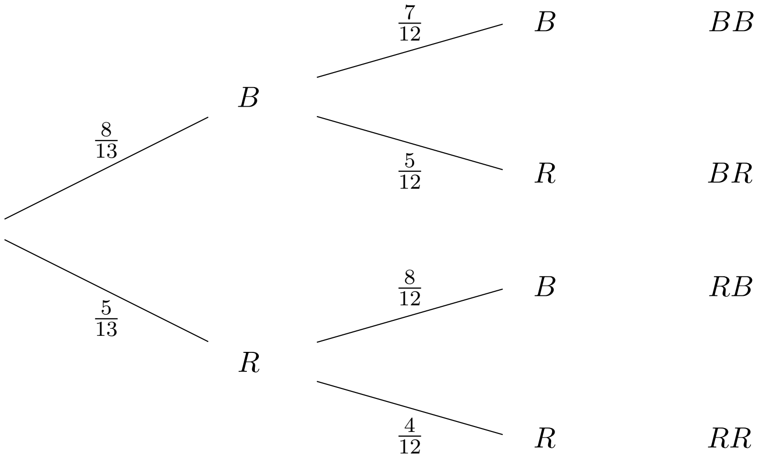
If, on the other hand, we had drawn a blue ball, the bag then has 7 blue balls and still 5 reds. Thus, we have

$$P(B) = \frac{\text{number of blue balls}}{\text{total number of balls}} = \frac{7}{7+5} = \frac{7}{12}$$

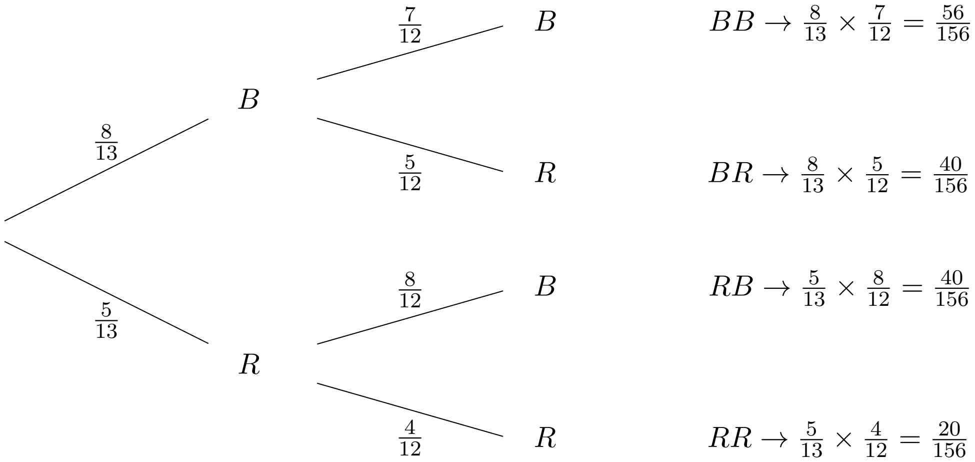
and for R:

$$P(R) = \frac{\text{number of red balls}}{\text{total number of balls}} = \frac{5}{12}$$

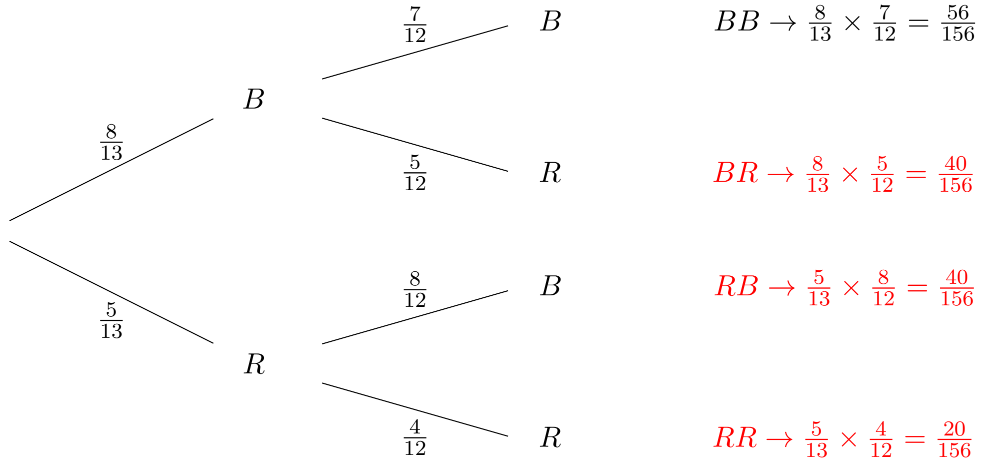
which again add up to 1. We can finish our tree now:



and we can use the "and" rule to find the probability of each path, as we did before, by multiplying each probability in the paths:



Let us use the tree to find the probability of drawing at least one red ball. The leaves we are interested are any that have at least one R:



and we can now use the "or rule" and just add those probabilities:

$$P(\text{at least one } R) = P(BR \text{ or } RB \text{ or } RR) = P(BR) + P(RB) + P(RR)$$

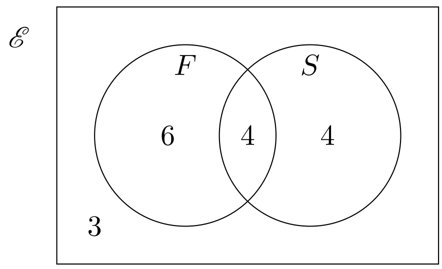
$$P(\text{at least one } R) = \frac{40}{156} + \frac{40}{156} + \frac{20}{156} = \frac{100}{156}$$

which we could also have solved by using the complement of this event, that is, 1 minus the probability of both balls being blue:

$$P(\text{at least one } R) = 1 - P(BB) = 1 - \frac{56}{150} = \frac{100}{156}$$

49.9. Probabilities in Venn diagrams

Another interesting type of problem that may appear involves Venn diagrams (do revise Chapter 10 if you need a refresh). For instance, given this diagram:



in which the set F represents people that speak French and the set S people that speak Spanish.

We can now ask questions such as: what is the probability someone speaks French? We still use our basic idea of "what we want divided by what we have", we just need to get the information from the diagram. Let us count the total number of people by adding all numbers in the Venn diagram (or, using set notation, finding $n(\mathscr{E})$):

$$n(\mathscr{E}) = 3 + 6 + 4 + 4 = 17$$

And now we can find the number of people that speak French by adding both "regions" of the F set (or again, using set notation, finding $n(F)$):

$$n(F) = 6 + 4 = 10$$

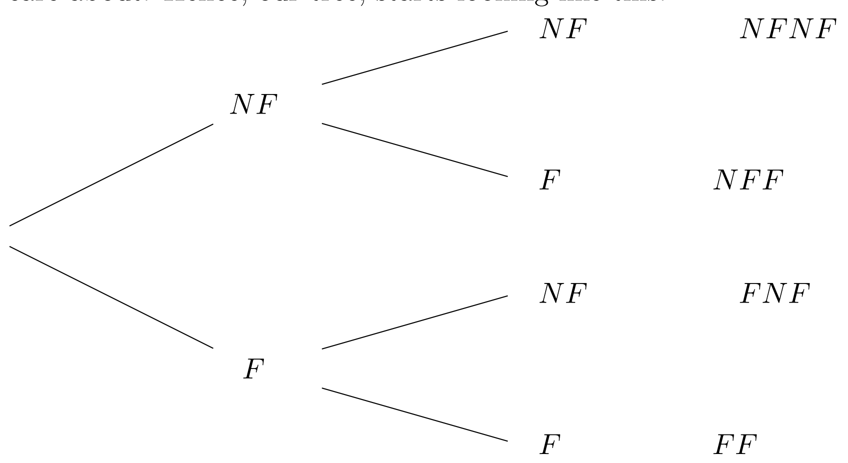
Finally, we can use our basic notion of probability ("what we want over what we have"):

$$P(F) = \frac{\text{number of people that speak French}}{\text{total people}} = \frac{n(F)}{n(\mathscr{E})} = \frac{10}{17}$$

We can also ask something as: what is the probability that someone speaks both French and Spanish? Those are represented in the intersection of the F and S sets ($F \cap S$), so we can find it easily:

$$P(\text{French and Spanish}) = \frac{\text{number of people that speak both}}{\text{total people}} = \frac{4}{17}$$

The most complex problems that can appear involve building a tree diagram. For instance, if we choose two people from the ones in the diagram, what is the probability exactly one of them speaks French? First thing we have to be careful is that we *cannot select the same person twice*, as we want to choose two people. Second is that we do not branch the tree into French and Spanish, but into French and "not French", as that is what we care about. Hence, our tree, starts looking like this:



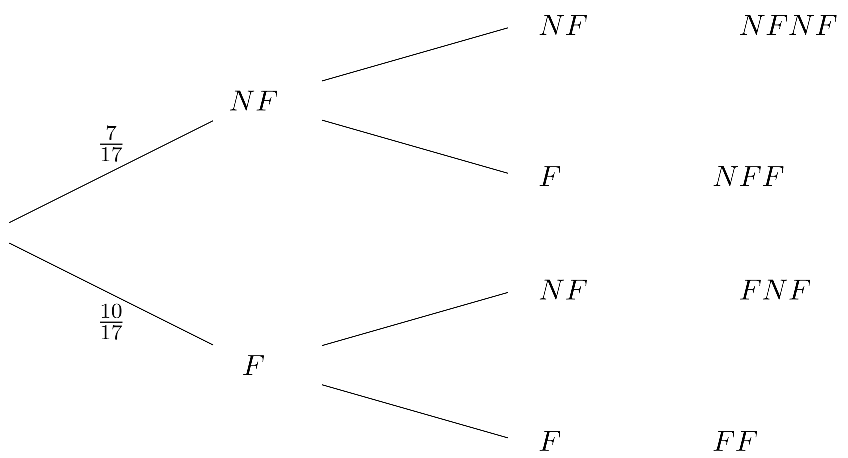
For the first person, that probability of her speaking French is

$$P(F) = \frac{10}{17}$$

and the probability of her not speaking French is

$$P(NF) = \frac{7}{17}$$

(notice that they must add up to 1). Adding those values to our tree:



The second person, though, depends on the first. If the first person speaks French, now we have only 9 people that speak French left, out of a total of 16, so now the probability of choosing someone that speaks French is

$$P(F) = \frac{9}{16}$$

and the probability of choosing someone that does not is

$$P(NF) = \frac{7}{16}$$

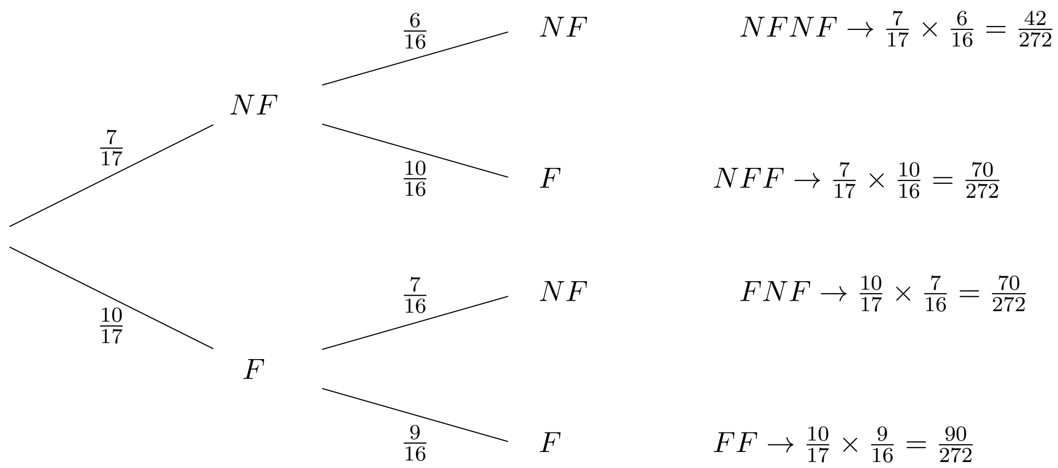
If the first person does not speak French, however, we still have 10 people that speak French out of 16, so we have

$$P(F) = \frac{10}{16}$$

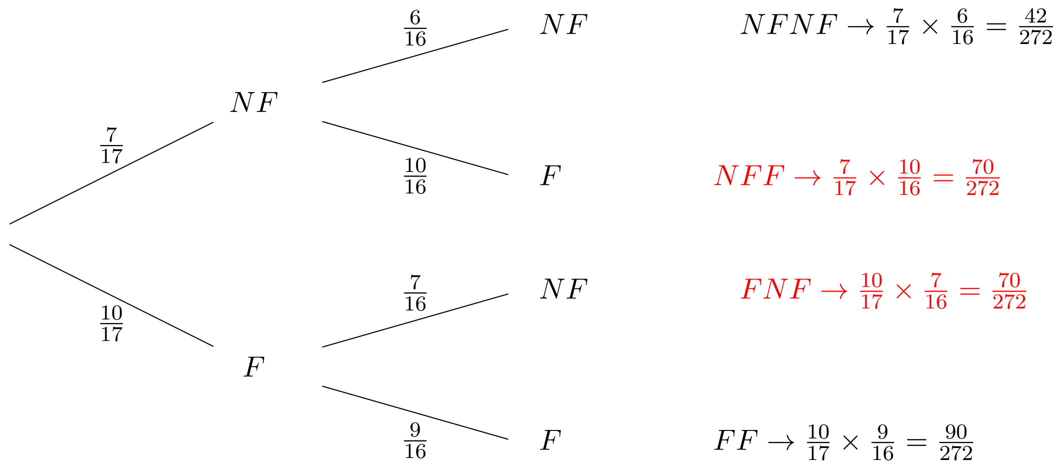
and

$$P(NF) = \frac{6}{16}$$

Let us add those to the tree and find the probability of each leaf by multiplying the probabilities in each path:



Finally, we are interested in the outcomes which have exactly one F:



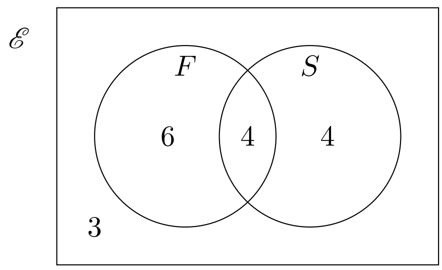
and we can add those two probabilities:

$$P(\text{exactly one person speaking } F) = \frac{70}{272} + \frac{70}{272} = \frac{140}{272}$$

49.10. Conditional probability: "given that" problems

Finally, the last kind of problem that may appear in the IGCSE is very interesting, but very easy when you understand the idea behind it. Let us see an example.

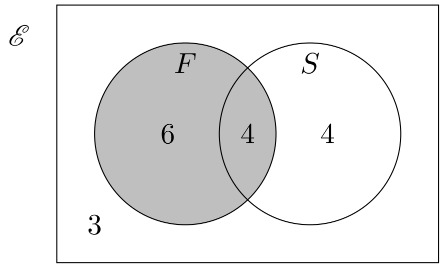
Using the same diagram for Spanish and French speakers from the last section:



The question we want to ask is: you select someone at random. *Given that* they speak French, what is the probability they also speak Spanish?

Whenever you see these terms *given that*, what is happening is simple: you have extra information. So, when you think about "what we want over what we have", the "what we have" part is different, as you will not be looking at the total possibilities, but only the ones that satisfy the condition you are given.

In our case, we know that the person we chose speaks French. Hence, it cannot be anyone from outside the F set! We need to, then, focus our attention to only the people within F:



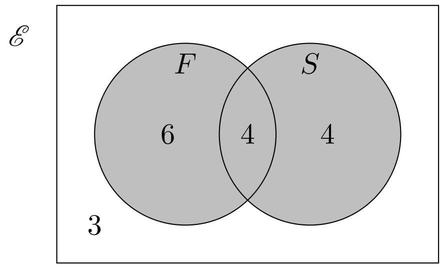
Now, thinking only on the shaded area, we have that 10 people speak French, and out of those, 4 also speak Spanish (the intersection part). Hence,

$$P(\text{Speaks Spanish given French}) = \frac{\text{number of people that speak Spanish and French}}{\text{number of people that speak French}}$$

$$P(\text{Speaks Spanish given French}) = \frac{4}{10}$$

And that's it. Whenever you read *given that* in the instructions, you just need to *restrict* the universe from which possibilities come.

Another example question: we select someone at random. Given that they speak at least one language, what is the probability they speak both Spanish and French? In this case, the "what we have" part is anything within the sets S and F:



Within the shaded region, we have 14 people. Out of those, 4 speak both languages. Hence,

$$P(\text{speaks both languages given that speaks at least one}) = \frac{\text{number of people that speak both}}{\text{number of people that speak at least one}}$$

$$P(\text{speaks both languages given that speaks at least one}) = \frac{4}{14}$$

49.11. Exam hints

Probability questions are extremely common, so it is a very important topic to revise and be fluent on.

Remember that probabilities are always numbers between 0 and 1, which can help you identify mistakes in your solutions.

Be very careful with exercises that have or do not have not reposition (you choose something and either can or cannot choose the same thing again). Usually, if you are selecting more than one person, it is implied you cannot select the same person twice.

Finally, I highly recommend to always draw the trees in questions where you need to select more than thing sequentially. It will help you.

Summary

- *Probabilities* are measures of the likelihood of something random happening;
- The *values* probabilities can have are all numbers between 0 and 1. If you have an *event* X (something random X), the probability of X, denoted $P(X)$:

$$0 \leq P(X) \leq 1$$

- You can represent probabilities on a number line, called a *probability scale*. It is a regular number line from 0 to 1;
- When performing a *random experiment*, such as rolling a dice, you can count how many times each outcome happens. The *relative frequency* of each outcome is the value of

$$\frac{\text{number of times } X \text{ happens}}{\text{total number of times experiment was done}}$$

- The *expected frequency* of an outcome, when repeating a random experiment, is given by

$$\text{number of times experiment is done} \times \text{probability of outcome}$$

- In general, the probability of something can be found by dividing the number of ways the something happens by the total number of ways anything can happen. In other words, "what we want divided by what we have";
- Events are *independent* if the probability of one them happening is not affected by the other happening or not. If one event does change the probability of the other happening, they are called *dependent*;
- Events are *mutually exclusive* if only one can happen;
- The *and rule* says that, if two events A and B are independent, then the probability of A and B happening together is the product of their probabilities:

$$P(A \text{ and } B) = P(A)P(B)$$

- The *or rule* says that, if two events are mutually exclusive, the probability of one or the other happening is the sum of their probabilities:

$$P(A \text{ or } B) = P(A) + P(B)$$

- You can represent outcomes using a *table*, a *tree* or a *Venn diagram*;
- When finding probabilities in a tree, the probability of a path is given by the *product* of the probabilities in that path;
- When solving problems with *given that*, remember you are restricting the possibilities you can focus on.

Formality after taste

Something about measures

I mentioned at the beginning of the chapter that probabilities are *measures*. But what does that actually mean? In a very simplified way, a measure, in mathematics, is a special function.

Say that we have a set S, such as all the possible outcomes of a random experiment. For instance, if we roll one six sided dice,

$$S = \{1, 2, 3, 4, 5, 6\}$$

We can now define a function P, which takes each element from S and attributes a number to it. A general measure function could attribute any non-negative number to each element of S, as long as it satisfied these properties:

- $P(\varnothing) = 0$: that is, the measure of the empty set is 0. In our case, with probabilities, this would mean that the probability of nothing happening is 0;
- $P(X) \geq 0$: that is, all numbers the function gives are non-negative. Nothing changes with probabilities.
- If X and Y are disjoint, that is, $X \cap Y = \varnothing$, then $P(X \cup Y) = P(X) + P(Y)$. For probabilities, it means that if your events are mutually exclusive, the probability of one or the other happening is the sum of their probabilities (the or rule!). Also, this is true for unions of as many sets as you want: if you have sets $X_1, X_2, \ldots, X_n, \ldots$ (yes, infinite sets are fine!), all pairwise disjoint (which means the intersection of any two is empty, or it mathematicianese $X_i \cap X_j = \varnothing$ for $i \neq j$), then

  $$P\left(X_1 \cup X_2 \cup X_3 \cup \ldots \cup X_n \cup \ldots\right) = P(X_1) + P(X_2) + P(X_3) + \ldots + P(X_n) + \ldots$$

  Of course, for this to make sense with probabilities (remember this is true for general measures), an infinite number of events will have to a smaller infinite of events with probability 0.

For probabilities, though, we have an extra restriction: the sum of all possible probabilities need to be 1. That is,

$$P(S) = 1$$

which means that the probability of all outcomes add to 1.

This is very simplified, but just to give you a taste.

Conditional probability

In the "given that" problems section, we treated it very "naturally". There is a way to formalize it, though (of course).

We denote the probability of A given B by

$$P(A|B)$$

which means that, if we know that B has happened, what is the probability of A happening. In the case of the example of French and Spanish, where we asked what is the probability of speaking Spanish given that the person speaks French, we could write

$$P(S|F)$$

Let us think on a general way to find these probabilities. We want to find a formula to find

$$P(A|B)$$

To do that, let us think on

$$P(A \cap B)$$

first. This means the probability of A and B happening together. We can also think of A and B happening together as first A happening, and then B happening *given that* A has happened! In probabilities:

$$P(A \cap B) = P(A)P(B|A)$$

We can now make $P(B|A)$ the subject:

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

which is a general way to find conditional probabilities. However, what it is basically saying is "what we want" (that A and B happen together) divided by "what we have" (we have A, as we are *given* it!).

50. Types of data and basic representation

50.1. Why learn basic data visualization

Why look at numbers and tables when you can look at pictures?

Jokes about our laziness aside, everywhere we look there are visual representations of data. In this chapter we will learn the most used ones, which you will use whatever you choose to do with your life.

50.2. Types of data

Data can be split into two categories: qualitative and quantitative. Quantitative data is anything about *numbers*, that we can measure somehow: sizes, income, grades (in percentages), for instance. Qualitative data is about traits: eye colour, hair colour, colour (creative examples here).

Most of what we will learn in this part of the book only applies to quantitative data, but some applies to both. It is important, then, to distinguish them in order to know what we can or cannot do with each.

50.3. Comparing absolute quantities

The first type of charts will see are about comparing the absolute values of our measures, that is, looking at what appears more frequently.

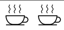
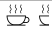
50.3.1. Pictograms

A pictogram is basically a tally chart, but the quantities are represented by pictures.

Let us say we have the following data about how many cups of coffee some teachers drink in one day:

Teacher	Cups of coffee
Mr Virgili	18
Mr Barton	8
Dr Frost	12
Mr Khan	6
Miss Hindley	4

We can represent the data from the table in a pictogram:

Teacher	Cups of coffee
Mr Virgili	
Mr Barton	
Dr Frost	
Mr Khan	
Miss Hindley	

 = 4 cups of coffee

and honestly I see no point in doing this. You can clearly see that Mr Virgili drinks more coffee than anyone else, and that Miss Hindley drank the least. Of course, you could see this from the table with the numbers...

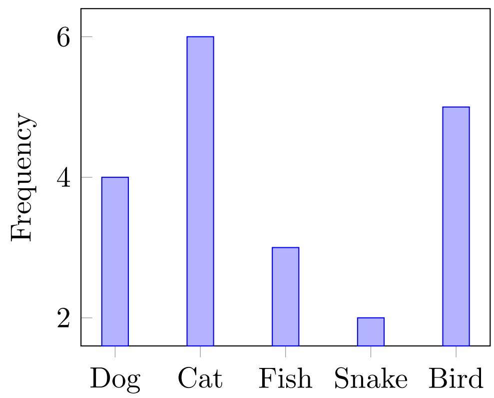
50.3.2. Bar charts

Bar charts are more common, particularly when you have very big or small numbers and want to compare the quantities. Let us use this data set:

Pet	Frequency
Dog	4
Cat	6
Fish	3
Snake	2
Bird	5

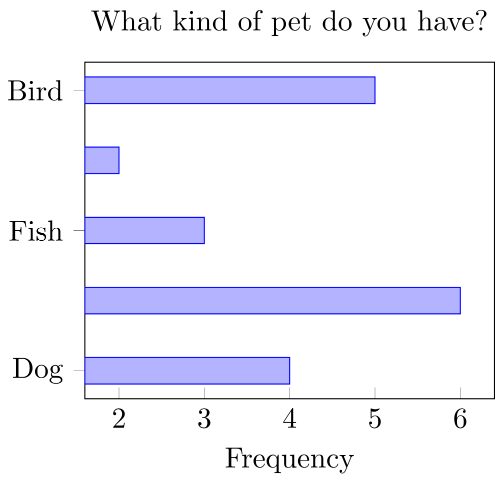
We can make a vertical bar chart, in which the height of each bar represents the frequency of, in our case, a pet:

What kind of pet do you have?



from which it is very clear there are more people with cats, and snakes are least common type of pet.

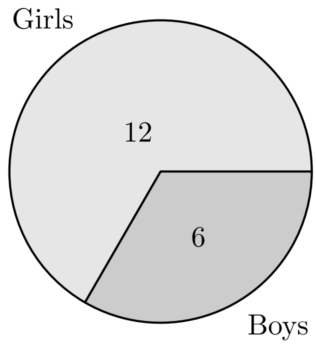
We can also draw the bars horizontally, with the same idea:



50.4. Comparing relative quantities: pie charts

Unlike bar charts and pictograms, sometimes we want to compare the numbers in relation the whole we are encountering.

For instance, say that a class has 12 girls and 6 boys. We can represent these in a pie chart, which looks like this:



The basic idea in a pie chart is that the area of each sector is proportional to the relative number to the total. In our case, we have 18 students in total, so the girls sector will have

$$\frac{12}{18} \text{ out of the total area of the circle}$$

As we learned in Chapter (ADD REF), the area of a sector is proportional to the central angle, so the only thing we need to do is multiply our fraction of the total by 360°, and we obtain the angle of the sector for girls:

$$\text{angle for girls} = \frac{12}{18} \times 360^\circ = 240^\circ$$

and as we only have another sector for boys, its angle must be what is left of 360°:

$$\text{angle for boys} = 360 - 240 = 120^\circ$$

In general, then, if you have a total quantity T and a specific number S of something, the angle for S is

$$\text{angle for } S = \frac{S}{T} \times 360^\circ$$

Let us build a pie chart for our pet data set:

Pet	Frequency
Dog	4
Cat	6
Fish	3
Snake	2
Bird	5

First, let us find the total number of animals, T:

$$T = 4 + 6 + 3 + 2 + 5 = 20$$

Now, let us find the sector for each animal. For dogs, of which we have 4:

$$\text{angle for dogs} = \frac{\overbrace{4}^{\text{quantity of dogs}}}{\underbrace{20}_{\text{total amount of pets}}} \times 360 = 72^\circ$$

for cats we do a similar calculation, but the numerator is 6, as we have 6 cats:

$$\text{angle for cats} = \frac{6}{20} \times 360 = 108^\circ$$

And we repeat the same idea for the remaining animals:

$$\text{angle for fishes} = \frac{3}{20} \times 360 = 54^\circ$$

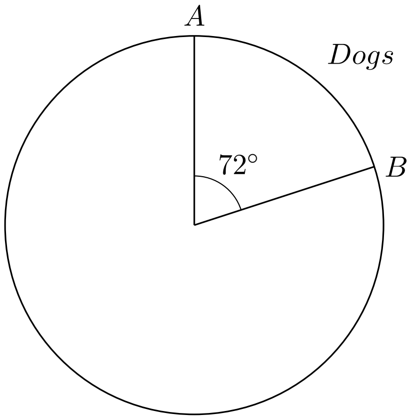
$$\text{angle for snakes} = \frac{2}{20} \times 360 = 36^\circ$$

$$\text{angle for birds} = \frac{5}{20} \times 360 = 90^\circ$$

An important thing to remember is that our sectors need to add up to 360°, as we will use the full circle:

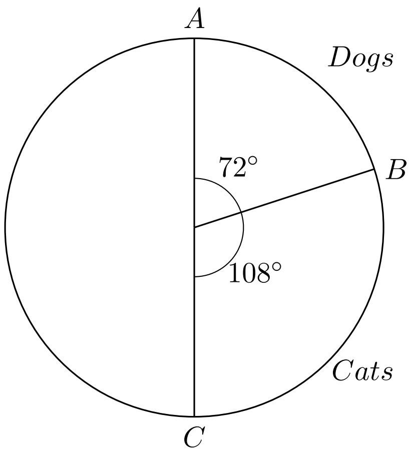
$$72 + 108 + 54 + 36 + 90 = 360^\circ$$

Now, we will divide a circle into 5 sectors, one sector for each animal with the correct angle. Let us start with dogs:



To build this angle, I started with a radius going "north", which I labeled A. From it you use a protractor[1] and measure $72°$, obtaining a new point B on the circle. Then you draw a radius going to B and you have the dogs sector.

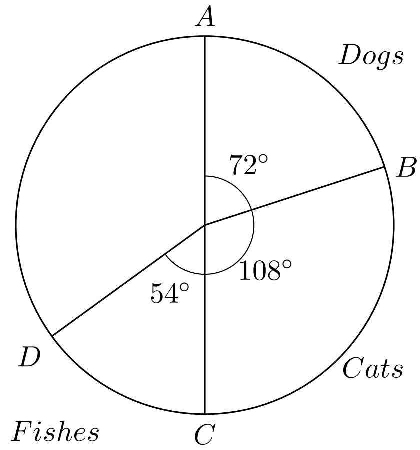
We now repeat the idea, but we measure the angle using the radius going to B. Cats have an angle of $108°$, so we measure $108°$ starting at B and find the point C. We then draw a radius to C and obtain:



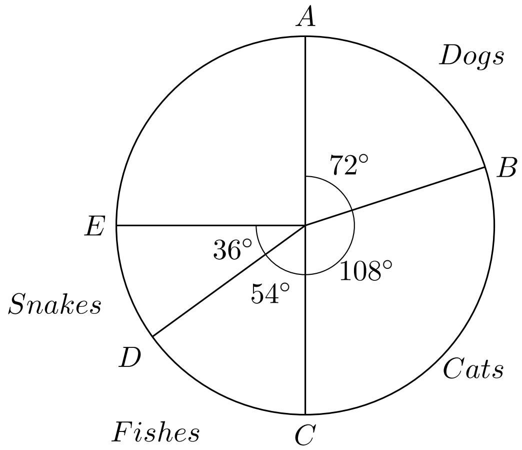
(The fact that dogs and cats add up to $180°$ is a coincidence).

We now do the same for fishes: we measure $54°$ starting at point C, finding a new point D. We connect the centre to D and have our fish sector:

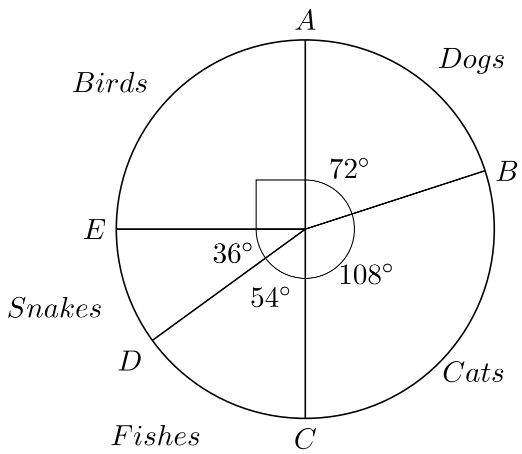
[1] Khan Academy has an excellent video on how to use a protractor if you forgot.



Just one more left, as the last sector always comes for free. The angle for snakes is 36°, so we measure 36° from D, obtain point E and draw a radius to it:



The 90° angle for birds is already drawn, so let us just label it:



And we are done.

Sometimes they may give you a sector angle and a total amount, and ask you to find the quantity. For instance, say they tell you there are 50 people in a room, and the sector in a pie chart for the number of blue eyed people is 36°. How many blue eyed people are there in the room? To find this, we just need to remember that

$$\text{angle for blue eyed people} = \frac{\text{number of blue eyed people}}{\text{total number of people}} \times 360$$

After we substitute the values given, we have a nice equation to solve, as usual. Let us call x the number of blue eyed people:

$$36 = \frac{x}{50} \times 360$$

and let us solve it:

$$\begin{aligned}
36 &= \frac{x}{50} \times 360 \\
36 \times 50 &= 360x && \text{Multiplying both sides by 36} \\
1800 &= 360x \\
x &= \frac{1800}{360} && \text{Dividing both sides by 360} \\
x &= 5
\end{aligned}$$

Therefore, we have 5 blue eyed people in the room.

50.5. Exam hints

Be very careful when reading graphs with their scale, as you do not want to lose marks for misreading something.

In the core papers, it is very common to have pie charts questions, so do revise that carefully if you are taking core.

Summary

- Data can be *quantitative*, when it is numerical, and *qualitative*, when it is about traits or categories;
- We can compare absolute values of data using:
  - *pictograms*, in which a picture represents a quantity of something;
  - *bar charts*, in which the height (or length, if horizontal) of the bars represent the quantity of something;

- *Pie charts* represent the *relative* amount of each element in the data by having *sectors of a circle* with angles given by

$$\text{angle of X} = \frac{\text{number of X}}{\text{total number of things}} \times 360$$

51. Averages

51.1. Why learn averages

According to a google search I've just done, the current estimate is that, by the end of 2020, there will be 44 zettabytes in the digital universe. That is a lot (one zettabyte is 1000^4 gigabytes), and although most of is trash (I am sure you agree your latest TikTok or whatever is the fashion right now is not the most useful digital entity), it is not possible for us to make sense of all of that data.

Hence, we need some techniques to *extract information* from it. Statistics, machine learning and data science are some areas that work on extracting information from data, and in this chapter we start with the very fundamentals of statistics: describing data.

51.2. Measures of central tendency

In this section you'll learn (or revise) the three measures of central tendency we most commonly use. A measure of central tendency is simply a value that we could call *typical*, that describes how the "average" (in the general meaning) looks like.

51.2.1. Mode

The mode is simply the value in the data set that appears more often. Say we have the following numbers:

$$1, 3, 3, 5, 7, 17$$

Here, the number 3 appears two times, and everyone else appears only once. Hence, 3 is the mode. Sometimes you can have more than one modal value:

$$1, 1, 2, 2, 3, 3$$

in which all values $1, 2$ and 3 appear two times. Hence, all of them are the mode.

The most interesting thing about the mode is that it is the only measure of central tendency that you can apply to categorical data (see Chapter ADDREF). For instance, we can say that the mode of

$$\text{blue, blue, red, red, red, green}$$

is "red", as it is the most common colour in the list.

Be careful when finding the mode in a frequency table. For instance, say we have the following table:

Grade	Frequency
4	5
5	12
6	10
7	3

the most frequent grade is 5, of which we have 12. The mode is 5 (the most frequent grade), not 12 (its frequency). This is a classic mistake, so do pay attention to it.

51.2.2. Arithmetic mean or, as we are lazy, just mean

The arithmetic mean, or just mean (yes, there are others, see the formality after taste) can be interpreted as a value that, if every element of the data set were equal, they would be the mean. We find it by *adding every number in the data set and dividing it by the number of elements we added.*

For the numbers

$$1, 3, 5, 7, 10$$

for instance, we can find the mean by adding all of them and dividing the result by 5, as there are 5 numbers in the list:

$$\frac{\overbrace{1+3+5+7+10}^{\text{sum of the numbers}}}{\underbrace{5}_{\text{number of elements}}} = \frac{26}{5} = 5.2$$

In general, then,

$$\text{mean} = \frac{\text{sum of the numbers in the data}}{\text{total number of numbers in the data}}$$

Sometimes we have to find the mean from a frequency table. Using the same table from the previous section as an example

Grade	Frequency
4	5
5	12
6	10
7	3

we can see that we have 5 people with grade 4, 12 people with grade 5 and so on. One way of "looking" at this table is to write at as list of grades:

$$\underbrace{4, 4, 4, 4, 4}_{\text{frequency is 5}}, \underbrace{5, 5, 5, 5, \ldots, 5}_{\text{frequency is 12}}, \underbrace{6, 6, 6, \ldots, 6}_{\text{frequency is 10}}, \underbrace{7, 7, 7}_{\text{frequency is 3}}$$

and then adding all of those values and dividing by $5 + 12 + 10 + 3 = 30$, the total frequency. However, that is very time consuming, and we can remember that adding 4 to itself 5 times is the same as doing 4×5; adding 5 to itself 12 times is the same as doing 5×12 and so on. Thus, I recommend finding the mean "using the table":

Grade	Frequency	Grade×frequency
4	5	$4 \times 5 = 20$
5	12	$5 \times 12 = 60$
6	10	$6 \times 10 = 60$
7	3	$7 \times 3 = 21$
	Total frequency: 30	Total sum: $20 + 60 + 60 + 21 = 161$

and now dividing the total sum by the total frequency:

$$\text{mean} = \frac{161}{30} = 5.37$$

One classic mistake in this type of question is to divide the total sum by 4, as there are 4 rows in the table. If you did that, you would obtain a mean of 40.25, which is not possible: the arithmetic mean is always a number between the minimum and the maximum values in a data set. Therefore, to avoid this, always check to see if the mean you found makes sense.

51.2.3. Median

The median is the guy in the middle, *if the numbers are put into ascending order.* That means that if we write the numbers from smallest to biggest, the median is the number which divides the data set into two parts of equal size.

Let us first look at the case in which the data set has an odd number of elements. Say we have the numbers

$$3, 1, 8, 2, 4$$

and we want to find the median. First, we order them:

$$1, 2, 3, 4, 8$$

Now, we have two ways of finding the median. The first is to "kill from the extremes", which means we "cancel" numbers from both ends of the list. We start with the minimum and the maximum:

$$\not{1}, 2, 3, 4, \not{8}$$

and now we continue from what is left, canceling from the new extremes:

$$\not{1}, \not{2}, 3, \not{4}, \not{8}$$

As we are left with only one number, 3, it is the median of $3, 1, 8, 2$ and 4.

The second method is to first add 1 to the number of elements in the list (in our case, 5), and divide the result by 2:

$$\frac{\text{number of elements in the list} + 1}{2} = \frac{5+1}{2} = 3$$

And, in the ordered list, we find the 3rd number:

$$1, 2, \underbrace{3}_{\text{number in the 3rd position}}, 4, 8$$

We have to be a bit careful when there is an even number of elements in the data set. For example, if we want to find the median of

$$-4, 0, -2, 10, 14, 6$$

we start, again, by writing them in order:

$$-4, -2, 0, 6, 10, 14$$

If we use the "killing from the extremes" method, we will eventually kill everyone, though:

$$\not{-4}, -2, 0, 6, 10, \not{14}$$

$$\not{-4}, \not{-2}, 0, 6, \not{10}, \not{14}$$

$$\not{-4}, \not{-2}, \not{0}, \not{6}, \not{10}, \not{14}$$

So, what we do is, we "kill" until we have two numbers left:

$$\not{-4}, -2, 0, 6, 10, \not{14}$$

$$\not{-4}, \not{-2}, 0, 6, \not{10}, \not{14}$$

and we then find the mean of them:

$$\text{median} = \frac{0+6}{2} = \frac{6}{2} = 3$$

The other method also has this "mean" part. We have 6 numbers in this data set, so we first add 1 to 6 and divide by 2:

$$\frac{\text{number of elements in the list} + 1}{2} = \frac{6+1}{2} = 3.5$$

which means we need to find the number in the "position 3.5", which does not exist. What does exist is the mean of the numbers in positions 3 and 4:

$$-4, -2, \underbrace{0}_{\text{3rd position}}, \underbrace{6}_{\text{4th position}}, 10, 14$$

and we find the mean of those

$$\frac{0+6}{2} = 3$$

In tables, even though you can use the killing method, I highly suggest using the "number of elements plus 1" method. In our friend:

Grade	Frequency
4	5
5	12
6	10
7	3

we have total frequency 30, so the median is the number in position

$$\frac{\text{number of elements in the list} + 1}{2} = \frac{30+1}{2} = 15.5$$

which we can find by finding the *cumulative frequency* (see Section 52.4) of each group in the table:

Grade	Frequency	Cumulative frequency
4	5	5
5	12	$5 + 12 = 17$
6	10	$17 + 10 = 27$
7	3	$27 + 3 = 30$

In grade 5, we already have 17 people (5 with grade 4, 12 with grade 5). Hence, the 15.5 position is one the fives. This means the median is 5.

Why do we need the median, you may ask yourself. The main reason is that the median is less sensitive to "strange" numbers in the data. For instance, say we have this data about incomes:

$$1, 1, 100$$

Its mean is

$$\frac{1+1+100}{3} = 34$$

whereas its median is 1. The mean is changed a lot by the 100, whereas the median is a better measure of a "typical" element.

51.3. Measures of dispersion

By only finding how the data is centered we don't know how it is spread around the central value we found. For instance, it is very different to have the following three people and their yearly incomes as $10, 10, 10$ and $1, 1, 28$: both data sets have the same mean, 10, but the second one is very different from the first.

Hence, we need tools to measure how *disperse* the data is, how "unlikely" the numbers are among themselves.

51.3.1. Range

We define the *range* of a data set as the maximum value minus the minimum. Hence, in the data set $1, 2, 3, 4$, as we have

$$\min = 1$$

$$\max = 4$$

the range is

$$\text{range} = \max - \min = 4 - 1 = 3$$

The higher the range, the more the data is disperse. In our incomes example, the $10, 10, 10$ data set has range 0 (max= 10, min= 10, range= $10 - 10 = 0$), whereas the data set $1, 1, 28$ has range 27 (max= 28, min= 1, range= $28 - 1 = 27$). We can see the data is more varied in the second data set, and the range confirms that.

To find the range in tables, just find the maximum value and the minimum and subtract them. In the grade table we have been using, for instance:

Grade	Frequency
4	5
5	12
6	10
7	3

the maximum grade is 7, whereas the minimum is 4. Therefore, the range is

$$\text{range} = \max - \min = 7 - 4 = 3$$

51.4. Exercises with all averages together

51.4.1. Given the mean and all numbers minus one, find the number

One common type of exercise involving the mean is like this: given the numbers

$$1, 3, 5, 10, x$$

which have mean 8, find x. We just need to remember that the mean is found by adding all numbers and dividing by the number of numbers we added. In this case:

$$\frac{1+3+5+10+x}{5} = 8$$

and now we just a have a nice equation to solve:

$$\begin{aligned} \frac{1+3+5+10+x}{5} &= 8 \\ \frac{19+x}{5} &= 8 && \text{Collecting like terms} \\ 19+x &= 40 && \text{Multiplying both sides by 5} \\ x &= 40-19 && \text{Subtracting 19 on both sides} \\ x &= 21 \end{aligned}$$

51.4.2. Adding a number to a data set

A classic type of problem is something as: a class has 17 people and their mean in maths is 6.3. A new student is now put it into the class, and his grade was 5.8. What is the new mean of the class?

A classic mistake here is to add the old class mean with new student grade and divide the result by 2, which is not how we do it (in other words, the mean of a mean and a new value is not the new mean).

To solve this, we need to remember that the old mean, 6.3, is found by adding the grades of all 17 students and dividing the result by 17:

$$\frac{\text{sum of all 17 grades}}{17} = 6.3$$

which we can multiply both sides by 17 to get the sum of all 17 original grades:

$$\text{sum of all 17 grades} = 6.3 \times 17 = 107.1$$

Now, we know that the new mean of the class is the sum of all 18 grades (the original 17 plus the new student) divided by 18:

$$\frac{\text{sum of all 17 grades} + \text{new grade}}{18}$$

and we can just substitute:

$$\frac{107.1+5.8}{18} = 6.27$$

This new kid, making the mean decrease!

51.4.3. Creating data sets given averages

The type of question I hate the most is when they want you to make up data sets given some information. I don't like it because there is always some guessing (or a lot of equations) involved, and I find it time consuming.

Say they tell us that 5 numbers have mode 3 and median 2 and want us to find one possible data set. As we know the mode is 3, we already have that 3 is the most common number. Also, as the median is 2, we know that if we write the numbers in ascending order, 2 will be the middle (the third) value. We can try something as:

$$\_\_\_, \_\_\_, 2, 3, 3$$

and we just need to fill up the first two gaps. Anything smaller than 2 is fine, as long as we don't have more than one of each, so let us go with 0 and 1:

$$0, 1, 2, 3, 3$$

The hardest ones are the ones which you have to remember that by telling you the mean, the problem is telling you the sum of all numbers in the data set (which you get by multiplying the mean by the number of numbers in the data set).

For instance, a problem like this: the mean of 4 numbers is 3, and their range is 4. Find a possible solution.

The first thing I do is creating variables for each number, and assume they are in increasing order:

$$x_1, x_2, x_3, x_4$$

and we know x_1 is the minimum and x_4 the maximum. As we know the range is 4, we know that

$$x_4 - x_1 = 4$$

and as the mean is 3, we know that

$$\frac{x_1 + x_2 + x_3 + x_4}{4} = 3$$

which means that the sum of the 4 numbers is 12, as we can multiply both sides of the equation by 4:

$$x_1 + x_2 + x_3 + x_4 = 12$$

Now comes the guessing: we need to come up with the values for the x's that satisfy the two restrictions we have. Let us start with the range:

$$x_4 - x_1 = 4$$

so we need to numbers which have difference 4. I think 1 and 5 are a great guess, so we have now:

$$1, x_2, x_3, 5$$

as our numbers. We know that $1 + 5 = 6$, and as the four numbers add up to 12, we know that $x_2 + x_3$ needs to be 6 as well. So we need to guess two numbers which add up to 6, and are between 1 and 5 (we assumed that x_1 and x_4 were the minimum and maximum!). Why be creative when both x_2 and x_3 can be 3? So, our list becomes:

$$1, 3, 3, 5$$

which does satisfy our mean equals to 3 and range equals to 4 restrictions.

51.5. Exam hints

Be careful when finding the mode of a table: the mode is the most frequent number, not its frequency!

When finding the mean of a data set, always check your result and remember: the mean must be between the smallest and biggest values.

Summary

- We know three measures of *central tendency*, which represent a *typical* value in our data set:
  - The *mode*, which is the most common value (or values);
  - The *(arithmetic) mean*, which is found by adding all the values and dividing the result by the number of values added;
  - The *median*, which is the value that divides the data set into two parts of equal size. We can find the median by finding the value in the position $\frac{\text{number of elements} + 1}{2}$, and remember that if we have an *even* number of values, we will need to find the mean of two central numbers;
- We know one measure of *dispersion*, which represents how *varied* the values in our data set are:
  - The *range*, which is found by finding the difference between the maximum and minimum values:

$$\text{range} = \text{maximum} - \text{minimum}$$

Formality after taste

Some different means

In Subsection 51.2.2 we learned about the arithmetic mean, or just mean, as we are lazy. But as its full name is arithmetic mean, that implies there are other means. Let us learn two.

A famous one is the *geometric mean*. Instead of adding and dividing, we multiply and take the n-th root. An example, for the numbers $1, 2, 3$ and 4 (total of 4 numbers):

$$\text{geometric mean} = \sqrt[4]{1 \times 2 \times 3 \times 4} = \sqrt[4]{24} \approx 2.21$$

(the arithmetic mean would be $\dfrac{1+2+3+4}{4} = 2.5$). In general, for a data set

$$x_1, x_2, \ldots, x_n$$

with n numbers, the geometric mean is given by

$$\sqrt[n]{x_1 x_2 \cdots x_n}$$

One interesting thing about the geometric mean is that sometimes it is not defined. For instance, we cannot find the geometric mean of -1 and 2, and it would involve in finding the square root of a negative number, which is not a real number.

You may ask what is the point of having different kinds of means, and it all depends on the application. One that I find amazing is in matching possible couples in dating apps (search on YouTube for "Inside OKCupid: The math on online dating" to see how interesting it is).

Another famous mean is the *harmonic mean*, which is the reciprocal of the arithmetic mean of the reciprocals of the numbers we have. That may sound complicated, but it is very silly. The reciprocal of a number is just "when we flip it" (if you already know indices, the reciprocal of x is x^{-1}); so the reciprocal of $\dfrac{2}{3}$ is $\dfrac{3}{2}$, and the reciprocal of $4 = \dfrac{4}{1} = \dfrac{1}{4}$. Thus, the harmonic mean is just the arithmetic mean of the reciprocals, and we then also do the reciprocal of the result. For instance, for the numbers $1, 2, 3$ and 4, first we find the arithmetic mean of their reciprocals

$$\frac{\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4}}{4} = \frac{\frac{25}{12}}{4} = \frac{25}{48}$$

and now we find the reciprocal of $\dfrac{25}{48}$

$$\text{harmonic mean} = \frac{48}{25} = 1.92$$

In general, for our numbers $x_1, x_2, \ldots, x_n$, their harmonic mean is

$$\frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \ldots + \frac{1}{x^n}}$$

Finally, did you notice that for our examples with $1, 2, 3$ and 4 that the arithmetic mean is bigger than the geometric mean, which is bigger than the harmonic mean:

$$2.5 > 2.21 > 1.92$$

That is actually always true:

$$\text{Arith. mean} \geq \text{Geo. mean} \geq \text{Harm. mean}$$

It is actually quite simple to show that the arithmetic mean is larger than the geometric mean for only two numbers when they are both non-negative (that is, greater than or equal to 0). Say we have two numbers, x_1 and x_2, both equal or larger than 0. We want to show that

$$\frac{x_1 + x_2}{2} \geq \sqrt{x_1 x_2}$$

which is equivalent to showing that

$$x_1 + x_2 \geq 2\sqrt{x_1 x_2}$$

as we multiply both sides by 2. Now, squaring both sides:

$$(x_1 + x_2)^2 \geq 4x_1 x_2$$

To show this last inequality, we start from

$$(x_1 - x_2)^2 \geq 0$$

as the square of a number is always non-negative. We now expand the brackets:

$$(x_1 - x_2)^2 \geq 0$$

$$x_1^2 - 2x_1x_2 + x_2^2 \geq 0$$

$$x_1^2 - 2x_1x_2 + x_2^2 + 4x_1x_2 \geq 4x_1x_2 \qquad \text{Adding } 4x_1x_2\text{on both sides}$$

$$x_1^2 + 2x_1x_2 + x_2^2 \geq 4x_1x_2 \qquad \text{Collecting like terms on the LHS}$$

$$(x_1 + x_2)^2 \geq 4x_1x_2 \qquad \text{Factorising the LSH}$$

which shows what we wanted. Those are only equal $x_1 = x_2 = 0$.

Standard deviation, a common measure of dispersion

We learned about range as a way to analyze how varied the data is, but you have to agree that by only looking at the maximum and minimum values in the data set, we are ignoring everything else! You will learn another measure in Chapter (add ref cumulative freq), but still is a little limited.

A very standard (pun intended) way of measuring spread is called standard deviation. Let us see how we define it and why it is like that.

Let us start with a numeric example first. If our data set is made of

$$1, 2, 3, 4, 5$$

we can find its mean, which we will denote here by $\bar{x}$:

$$\bar{x} = \frac{1+2+3+4+5}{5} = \frac{15}{5} = 3$$

We now can find how *far* each of the values in the data set is from the mean, which we call the *deviation from the mean* of each number:

$$1-3=-2$$

$$2-3=-1$$

$$3-3=0$$

$$4-3=1$$

$$5-3=2$$

A very logical idea would be to find the mean of the deviations as a measure of spread: if this mean is big, there would be a large spread; if it were small, the numbers would be, on average, close to the mean. However, we have a problem: the sum of the deviations of the numbers from the mean is always 0! In our example:

$$-2 + -1 + 0 + 1 + 2 = 0$$

Let us think first intuitively why that is the case: the mean, $\bar{x}$, when multiplied by the number of elements in the data set, has exactly the same value of the of elements in it. In our case, for instance,

$$3 \times 5 = 15 = 1+2+3+4+5$$

When we are adding the deviations of each element, we will subtract the mean from each of them, so we subtract the mean once for each element in the data set. That can be seen as multiplying the mean by the number of elements we are considering, which gives the same sum of all elements in the data set. As we subtract it from the sum of the elements, the sum of the deviations will always be 0.

We can also prove that formally: say we have a data set with n elements

$$x_1, x_2, x_3, \ldots, x_n$$

which has mean

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \ldots + x_n}{n}$$

which we can interpret as the $\bar{x}$ being one n-th of the sum of the values in the data set. Let us, for easiness of writing, write S_n for the sum of the elements:

$$S_n = x_1 + x_2 + x_3 + \ldots + x_n$$

So, we have that

$$\bar{x} = \frac{S_n}{n}$$

The deviation from the mean for x_1, for instance, is $x_1 - \bar{x}$. Adding the deviations for all numbers in the data set gives

$$(x_1 - \bar{x}) + (x_2 - \bar{x}) + (x_3 - \bar{x}) + \ldots + (x_n - \bar{x})$$

which we can rearrange to

$$(x_1 + x_2 + x_3 + \ldots + x_n) - \left(\underbrace{\bar{x} + \bar{x} + \bar{x} + \ldots + \bar{x}}_{n \text{ times}}\right)$$

We have n elements, and we subtract the mean $\bar{x}$ from each of them, so we have n means being subtracted in the second brackets. Let us substitute $\bar{x} = \frac{S_n}{n}$ and write the first bracket as S_n :

$$S_n - \left(\underbrace{\frac{S_n}{n} + \frac{S_n}{n} + \frac{S_n}{n} + \ldots + \frac{S_n}{n}}_{n \text{ times}}\right)$$

In the second bracket we are adding n times $\frac{S_n}{n}$, which is the same as multiplying $\frac{S_n}{n}$ by n:

$$S_n - n \times \frac{S_n}{n}$$

$$S_n - \cancel{n} \times \frac{S_n}{\cancel{n}}$$

$$S_n - S_n = 0$$

So now we are sure that the sum of the deviations is always 0, and it is always good to be sure. But, as the sum of deviations is always 0, the mean of the deviations will always also be 0, and that is useless. We could then add the *squares* of the deviations: as the square of a (real) number is always positive, the sum of the squares of the deviations will definitely not be 0. So, we could find it

$$(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2$$

and then dividing it by n, finding the *mean of the squares of the deviations from the mean.* This result is called the *variance*:

$$\text{Variance} = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2}{n}$$

Which we could use as our measure of spread. However, if our data was about heights, for instance, the units could be in cm. The variance, on the other hand, would be in cm^2, and that is slightly strange. Thus, we can square root the variance, and that is, finally, the *standard deviation*:

$$\text{Std. deviation} = \sqrt{\text{Variance}}$$

$$\text{Std. deviation} = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2}{n}}$$

Instead of squaring the deviations we could have found their absolute values, which gives a different measure of dispersion. What a nice world full of possibilities.

52. Grouped frequency tables, cumulative frequency and histograms

52.1. Why learn about grouped frequency tables and their graph friends

In some situations, it makes sense to group some values of our observations together: the individual differences between the values are not so relevant, but between groups of values they are.

Think about this data set of 20 percentages in an exam:

12	24	33	33	38	39	50	52	58	60
63	71	75	79	81	86	93	94	95	100

To analyze the performance of this class, I would probably group the data, perhaps something as:

Grade	Frequency
Smaller than 40%	6
Between 40% and 55%	2
Between 56% and 65%	3
Between 66% and 85%	4
Larger than 85%	5

Looking at the data like this, I can easily see that 5 students are doing quite well (grades larger than 85%), but another 5 are failing very badly (less than 40%). Looking at the grades in groups like this I can identify patterns and plan better lessons later.

Grouping data, then, can help us understand it better. And as we have so much data, grouping it is sometimes mandatory.

52.2. What is a grouped frequency table and the modal class

A grouped frequency table is exactly what we just saw: a table in which individual values are grouped together in *classes*. So, in the grade example I have given:

Grade	Frequency
Smaller than 40%	6
Between 40% and 55%	2
Between 56% and 65%	3
Between 66% and 85%	4
Larger than 85%	5

we have 5 classes. The first is the one that has the grades smaller than 40%, the second is the one that has the grades between 40% and 55% and so on.

Even though writing the groups as I did is fine, it is more common to use intervals. To do that, we first create a variable to represent our grades: g. Now, instead of writing "smaller than 40%", we will write $g < 40$. We do need to be careful with the classes which have "between", we do not want to accidentally include one value in two different classes.

Hence, the class "Between 40% and 55%" will be written as $40 \leq g < 55$. We are using the smaller than or equal sign, $\leq$, as we need to include 40, given that we did not include 40 in our "smaller than 40" class. For 55, on the other hand, we use the smaller than sign, $<$, as it 55 is already included into the next class, $55 \leq g < 65$. We continue the same reasoning, and we now write the table like this:

Grade (g)	Frequency
$g < 40$	6
$40 \leq g < 55$	2
$55 \leq g < 65$	3
$65 \leq g < 85$	4
$g \geq 85$	5

Notice that we can insert any grade in the table, and no grade can go into two different classes. That is very important when building your own grouped frequency tables!

As we cannot have grades smaller than 0 and larger than 100%, though, we can write all classes as intervals to make the table more consistent:

Grade (g)	Frequency
$0 \leq g < 40$	6
$40 \leq g < 55$	2
$55 \leq g < 65$	3
$65 \leq g < 85$	4
$85 \leq g \leq 100$	5

Thus, any value starting at 40 (including it) up until 55 (not including it) would go into the class $40 \leq g < 55$, as we have

$$40 \underbrace{\leq}_{\text{includes 40}} g \underbrace{<}_{\text{does not include 55}} 55$$

Finally, we are able to define the *modal class*. The modal class is the "mode" of a grouped frequency table: it is the most common class in the table. In our example, the most common class is $g < 40$, with frequency of 6. As with the mode in regular tables, be careful: the modal class is $g < 40$, not its frequency!

52.3. Estimating the mean of a grouped frequency table

One thing we usually want from data is to find its mean. We cannot find an exact mean from a grouped frequency table, though: we do not know exactly how the values are distributed within a class. In our grades example:

Grade (g)	Frequency
$0 \leq g < 40$	6
$40 \leq g < 55$	2
$55 \leq g < 65$	3
$65 \leq g < 85$	4
$85 \leq g \leq 100$	5

we do not know how the 6 students which have grades smaller than 40 did exactly (or in any other class), so the best we can do is *estimate* the mean from table.

To do that, we will be lazy and pretend everyone in each class has the same value: the middle of each class interval. Of course, this can either over or underestimate the contribution from each class, but it is the simplest solution.

Hence, in our grade table, we will pretend the 6 students with grades from 0 to 40 got 20, which we call the *mid value* or *midpoint* of the interval. We obtain it finding the mean of the two extremes of each interval. For $0 \leq g \leq 40$, then, the mid value is:

$$\frac{0 + 40}{2} = 20$$

and for the class $40 \leq g < 55$ the mid value is

$$\frac{40 + 55}{2} = 47.5$$

We do this for every class in the table, which I suggest you do by adding a new column to it:

Grade (g)	Frequency	mid values
$0 \leq g < 40$	6	$\frac{0+40}{2} = 20$
$40 \leq g < 55$	2	$\frac{40+55}{2} = 47.5$
$55 \leq g < 65$	3	$\frac{55+65}{2} = 60$
$65 \leq g < 85$	4	$\frac{65+85}{2} = 70$
$85 \leq g \leq 100$	5	$\frac{85+100}{2} = 92.5$

Now, it is as if we had this table to find the mean:

Frequency	mid values
6	20
2	47.5
3	60
4	70
5	92.5

which we learned how to do in the previous chapter: we multiply each frequency by its corresponding values, in this case the mid values, and divide by the total frequency:

$$\text{estimate of the mean} = \frac{6 \times 20 + 2 \times 47.5 + 3 \times 60 + 4 \times 70 + 5 \times 92.5}{20} = 56.875$$

which some people like to do in table form:

Frequency	mid values	frequency×mid values
6	20	$6 \times 20 = 120$
2	47.5	$2 \times 47.5 = 95$
3	60	$3 \times 60 = 180$
4	70	$4 \times 70 = 280$
5	92.5	$5 \times 92.5 = 462.5$
Total frequency: 20		Sum of above: 1137.5

and then just divide

$$\text{estimate of the mean} = \frac{1137.5}{20} = 56.875$$

Thus, to find an estimate of the mean of a grouped frequency table:

1. Find the mid values of each class by adding the two extremes of each interval and dividing the result by 2;
2. Multiply the mid value by the frequency of each class;
3. Divide the sum of frequency×mid value for all classes by the total frequency of the table

Be very careful on two things. The first is that you *need* to show the mid values in your working out, so either write your extended table in the working out section or write the equation with the mid values. The second is to be careful to divide the sum of frequency×mid by the *total frequency* and not by the number of classes (if your mean is larger than the maximum value in the table, you probably did that!).

52.4. Cumulative frequency

Say we have the following data on rainfall on 40 different days:

Rain (r) in mm	Frequency
$0 \leq r \leq 2.5$	15
$2.5 < r \leq 8$	12
$8 < r \leq 50$	10
$50 < r \leq 75$	3

A natural question to ask is "how many days had less than 25mm of rain?", which our table by itself does not tell us. We could answer how many days had less than 50mm, by adding the all the days with less than 50mm of rain, but there is nothing we can do about 25mm.

To answer questions like this (and others) we will use the *cumulative frequency* of our table.

52.4.1. Finding the cumulative frequency

To find the cumulative frequency of a grouped frequency table, we will build another table. For each class in the original table, we will take the right limit only and have this first column:

Rain (r) in mm
$r \leq 2.5$
$r \leq 8$
$r \leq 50$
$r \leq 75$

and we can now fill the in the rest of the table using the original data. We know that we have 15 days with 2.5mm or less of rain, as the first class tells us. Thus, we can fill the first line of the *cumulative frequency table:*

Rain (r) in mm	Cumulative frequency
$r \leq 2.5$	$\underbrace{15}_{\text{Freq. of } 0 \leq r \leq 2.5}$
$r \leq 8$	
$r \leq 50$	
$r \leq 75$	

The second row needs to be filled with how many days had 8mm or less of rain. We can find this information from our original table as well: the first two classes, $0 \leq r \leq 2.5$ and $2.5 < r \leq 8$ *both* satisfy this restriction. So, to find the cumulative frequency of $r \leq 8$, we *add the frequency of the second class with the cumulative frequency so far*:

Rain (r) in mm	Cumulative frequency (cf)
$r \leq 2.5$	15
$r \leq 8$	$\underbrace{15}_{\text{old cf}} + \underbrace{12}_{\text{freq. of second class}} = 27$
$r \leq 50$	
$r \leq 75$	

The third row needs information about how many days had less than or equal to 50mm of rain. We already know how many days had 8mm or less from the second row of our table, so we just need to add the frequency of the $8 < r \leq 50$ to our old cumulative frequency:

Rain (r) in mm	Cumulative frequency (cf)
$r \leq 2.5$	15
$r \leq 8$	27
$r \leq 50$	$\underbrace{27}_{\text{old cf}} + \underbrace{10}_{\text{freq. of third class}} = 37$
$r \leq 75$	

Finally, the number of days that had 75mm or less of rain is found by adding 37, the cumulative frequency so far, with the frequency from the $50 < r \leq 75$ class:

Rain (r) in mm	Cumulative frequency (cf)
$r \leq 2.5$	15
$r \leq 8$	27
$r \leq 50$	37
$r \leq 75$	$\underbrace{37}_{\text{old cf}} + \underbrace{3}_{\text{freq. of fourth class}} = 40$

Notice that the last row of the cumulative frequency will always have the total frequency, so it is a good way to check if you made a mistake.

We then have this table:

Rain (r) in mm	Cumulative frequency (cf)
$r \leq 2.5$	15
$r \leq 8$	27
$r \leq 50$	37
$r \leq 75$	40

which we will now use to plot a graph.

52.4.2. Plotting the cumulative frequency curve

From the cumulative frequency table, we can find the points of the cumulative frequency curve:

- the x value of each point is the limit of the interval;
- the y value is the cumulative frequency

In our table:

Rain (r) in mm	Cumulative frequency (cf)	Point
$r \leq 2.5$	15	$(2.5, 15)$
$r \leq 8$	27	$(8, 27)$
$r \leq 50$	37	$(50, 37)$
$r \leq 75$	40	$(75, 40)$

To the table, we add the point $(0, 0)$ as well, which makes sense, as there are 0 days with less than 0mm of rain. Therefore, we have the following points to plot:

$$(0, 0)$$

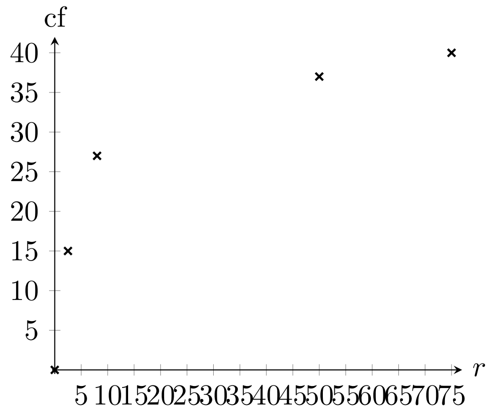
$$(2.5, 15)$$

$$(8, 27)$$

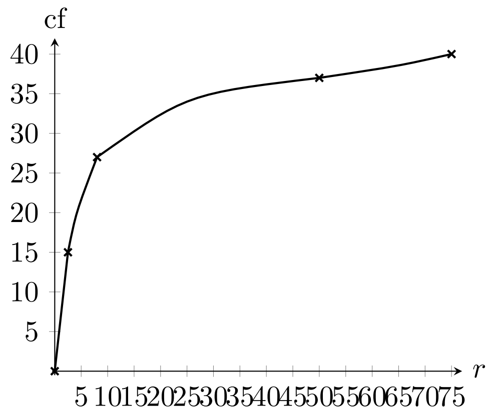
$$(50, 37)$$

$$(75, 40)$$

We now plot those values on the coordinate plane. The x axis is rain (so it is an r axis) and y axis is cumulative frequency:



Finally, we connect all our points *smoothly*. We cannot use a ruler, and we must go through all the points. It takes practice, but you get used to it:



Another example. Say we have the following grouped frequency table about grades of 30 students:

Grade (g)	Frequency
$0 \leq g \leq 30$	3
$30 < g \leq 45$	6
$45 < g \leq 60$	10
$60 < g \leq 80$	9
$80 < g \leq 100$	2

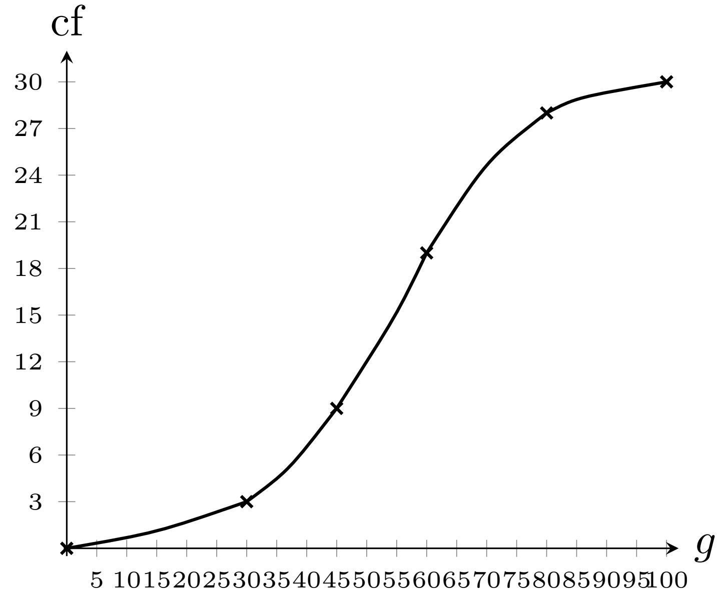
First let us find the cumulative frequency:

Grade (g)	Frequency	Cumulative freq.
$0 \leq g \leq 30$	3	3
$30 < g \leq 45$	6	$3 + 6 = 9$
$45 < g \leq 60$	10	$9 + 10 = 19$
$60 < g \leq 80$	9	$19 + 9 = 28$
$80 < g \leq 100$	2	$28 + 2 = 30$

and let us get our points. Remember that the x coordinate is always the left limit of the interval for each class and the y coordinate the cumulative frequency:

Grade (g)	Frequency	Cumulative freq.	Points
$0 \leq g \leq 30$	3	3	$(30, 3)$
$30 < g \leq 45$	6	$3 + 6 = 9$	$(45, 9)$
$45 < g \leq 60$	10	$9 + 10 = 19$	$(60, 19)$
$60 < g \leq 80$	9	$19 + 9 = 28$	$(80, 28)$
$80 < g \leq 100$	2	$28 + 2 = 30$	$(100, 30)$

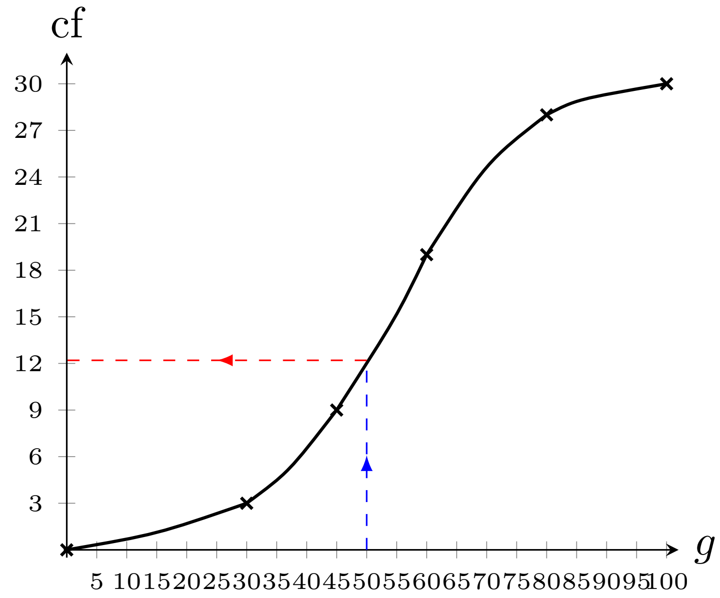
and remember that the point $(0, 0)$ is always part of the graph. Let us plot and connect the points:



52.4.3. Using the curve to find quantities

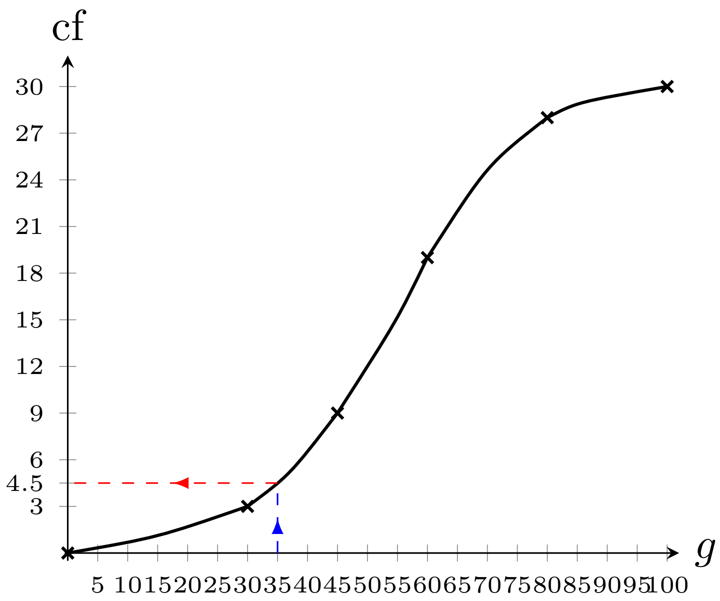
Now that we have our cumulative frequency diagram, we can finally answer "how many students had a grade smaller than 50?" or similar questions.

In a cumulative frequency diagram, the y value of a point on the curve represents *how many whatever are smaller than or equal to the x coordinate*. So, to find out how many students had a grade smaller than 50, we go to $x = 50$ and go vertically up until we meet our cumulative frequency curve and we find the y coordinate of the point where this line meets the cumulative frequency curve by going to the y axis:



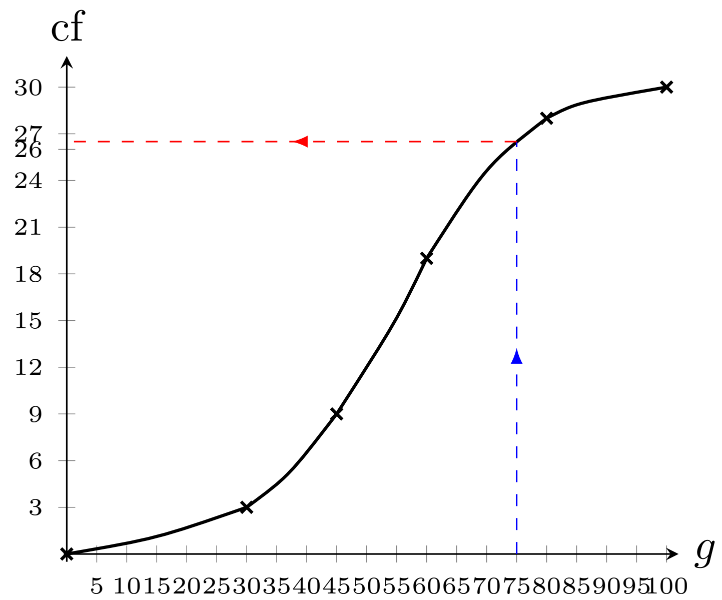
We can see that the red line goes somewhere close to 12 on the y axis (the exact value is 12.2), so we can say that there are 12 students with grades smaller than or equal to 50.

Let us find the number of students with grade smaller than or equal to 35 using the same idea:



so we can say that 4.5 students had a grade smaller than or equal to 35. Of course, we cannot have half a student, but remember we are estimating.

The cumulative frequency diagram also allow us to find how many students had a grade larger than 75. We work indirectly, though: first we first how many students had a grade smaller than or equal to 75:



which is around 26.5. So, as 26.5 students had a grade smaller than or equal to 75, and we know there are 30 students in total, we can find how many had a grade bigger than 75 by subtracting 26.5 from the total:

$$30 - 26.5 = 3.5$$

Thus, around 3.5 students had a grade larger than 75.

52.4.4. Using the curve to find percentiles

One amazing convenience of the cumulative frequency diagram is that the data is now ordered! So it is very easy to find the median and her friends, the other *percentiles*.

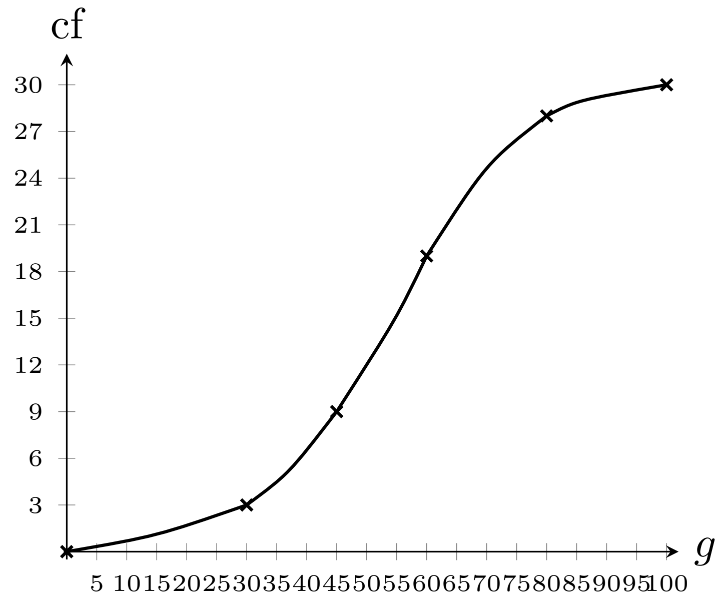
52.4.4.1. Median (50% percentile)

As we say in Chapter 51, the median splits the data into two groups of equal size. The cumulative frequency diagram is very useful to find the median: we can simply find the value in the x axis which divides the cumulative frequency into half, which is precisely the definition of the median!

So, to find the median of our grouped frequency table:

Grade (g)	Frequency
$0 \leq g \leq 30$	3
$30 < g \leq 45$	6
$45 < g \leq 60$	10
$60 < g \leq 80$	9
$80 < g \leq 100$	2

we were able to build this diagram:



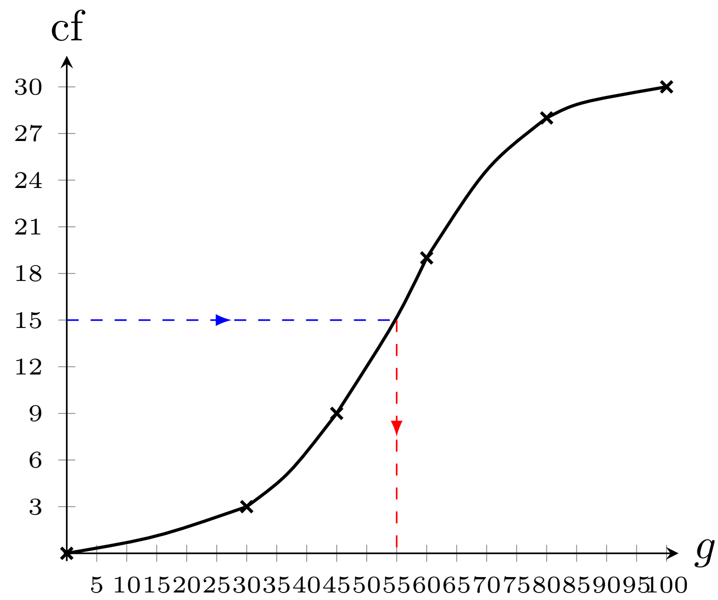
and now, to find the median, we have to split our data in two. As we have 30 total students, the median is the value which splits the data into

$$\frac{30}{2} = 15$$

or, more generally, the value which corresponds to 50% of our total frequency:

$$50\% \text{ of } 30 = 0.5 \times 30 = 15$$

So we go from the y axis, which represents the frequency value of 15 until we reach the curve, and then go down:



so the median is 55[1]

52.4.4.2. Upper quartile (75% percentile)

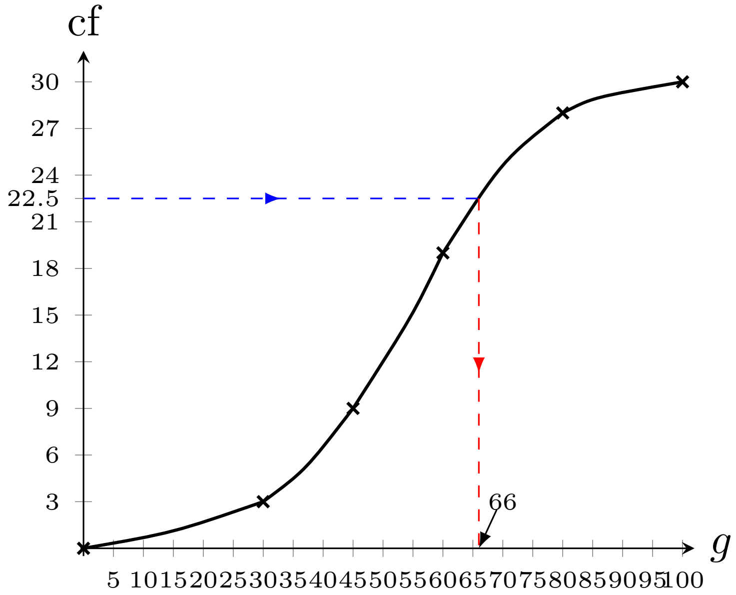
The upper quartile (UQ) is another measure which is easy to find in a cumulative frequency diagram. While the median splits the data into $50\% - 50\%$, the upper quartile splits the data into $75\% - 25\%$: 75% of the data is smaller than or equal to the UQ, and only 25% is larger than it.

To find the UQ, we find 75% of our total frequency:

$$75\% \text{ of } 30 = 0.75 \times 30 = 22.5$$

and do the same as we did for the median: going from $y = 22.5$ to the right until we reach the curve, and then go down

[1]You may be wondering why, in Chapter 51, we had to add 1 to the total frequency before dividing it by 2 (or finding 50% of the value), and here we do not. It is because in this chapter the data is *continuous*, whereas in Chapter 51 the data was *discrete*. Continuous implies that our grades, which range from 0 to 100, can also be any value between those two limits. Discrete data, on the other hand, is not continuous, so it can only assume values in certain steps.



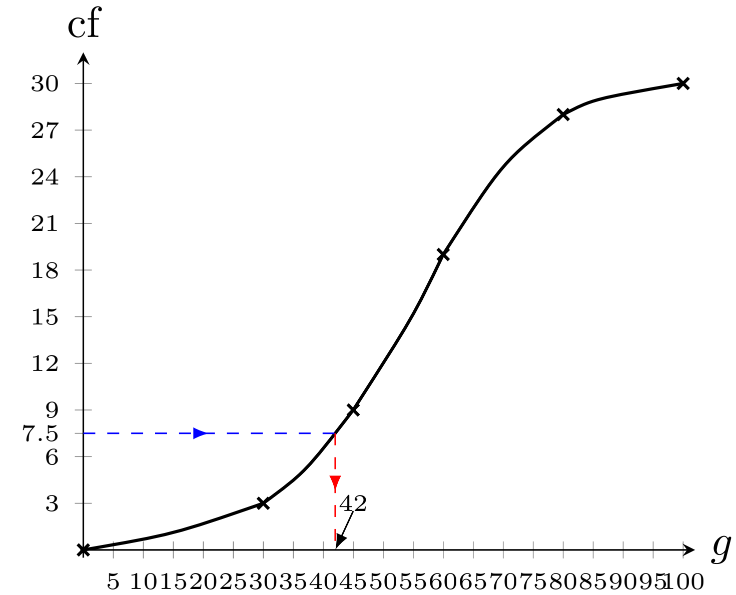
which we can see meets the x axis at 66. Hence, the UQ for this data is equal to 66.

52.4.4.3. Lower quartile (25% percentile)

The lower quartile (LQ) is again similar, but splits the data into $25\% - 75\%$. Thus, to find it, we find 25% of the total frequency:

$$25\% \text{ of } 30 = 7.5$$

and again go from $y = 7.5$ until we meet the curve and go down until the x axis:



Giving that the LQ is equal to 42.

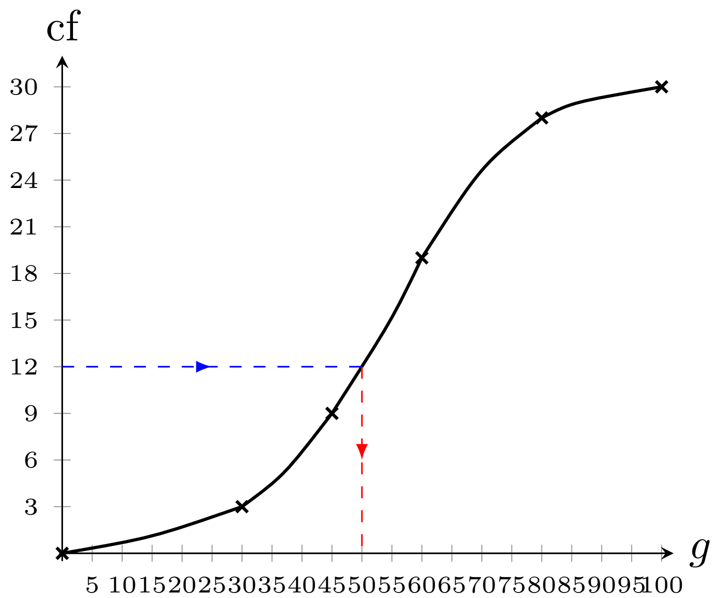
52.4.4.4. General percentiles

As you probably noticed, the LQ, median and UQ are very similar: we find a percentage of the total frequency and then do the "from y to the right until curve, then down until x" technique. They are all *percentiles*, and we can find a general percentile by finding the percentage in it.

Say we wanted to find the 40th percentile of our data, which is the value that divides the data into 40% – 60%. We first find 40% of the total frequency:

$$40\% \text{ of } 30 = 12$$

and then do the classic:



which gives us 50. Thus, the 40th percentile of our grade data is equal to 50.

Summarizing: to find the pth percentile using a cumulative frequency diagram:

1. Find $p\%$ of the total frequency;
2. On the y axis, go from the value you found in Step 1 until you reach the cumulative frequency curve;
3. Go down from the curve until you reach the x axis. That is your percentile.

52.4.5. Interquartile range, another measure of spread

The interquartile range (IQR) is another way to evaluate how the data is spread, much like the range in Chapter 51.

The IQR is defined as

$$\text{IQR} = \text{UQ} - \text{LQ}$$

In our grade example, we found that

$$\text{UQ} = 66$$

$$\text{LQ} = 42$$

so we have

$$\text{IQR} = \text{UQ} - \text{LQ} = 66 - 42 = 14$$

52.4.6. Using the median and the IQR to compare data sets

Let us have a full "use case" of cumulative frequency diagrams.

Pretend we have the following tables of results from one of my exams:

Grade	Girl's freq.	Boy's freq.
0 – 20	2	3
21 – 40	4	8
41 – 50	6	6
51 – 60	8	10
61 – 80	7	8
81 – 100	3	5

The tables may look weird because of the interval notation used, but

$$0 - 20$$

simply means grades from 0 to 20, or $0 \leq g \leq 20$. It is ugly, but sometimes it appears, so better see it at least once!

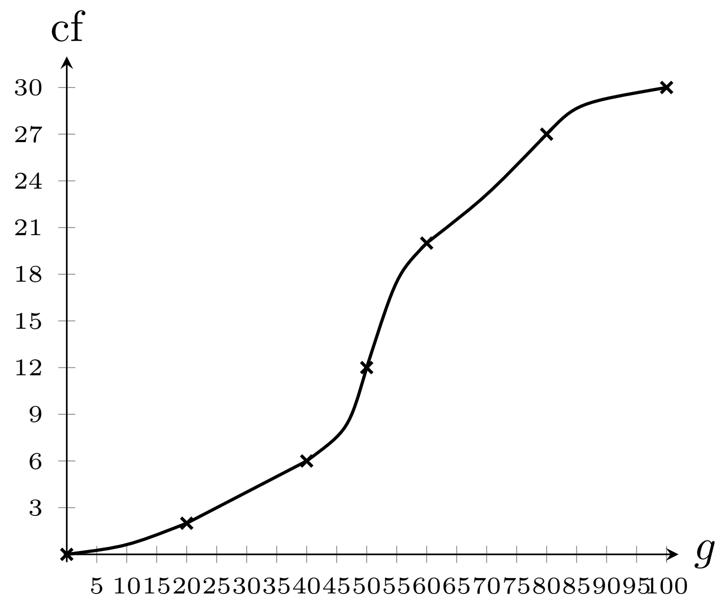
I would like to see which group did better, girls or boys. One simple way to answer that question is to find the median and IQR of both data sets. The median will tell us the average grade of each group, while the IQR how spread they were.

You may ask yourself what is the point of knowing the spread of data, but if the spread is small, we know that the median is a good representative of the data set. If the spread is very large, then the median may be a bad one.

Let us start by finding the cumulative frequency and the points to plot for the girls:

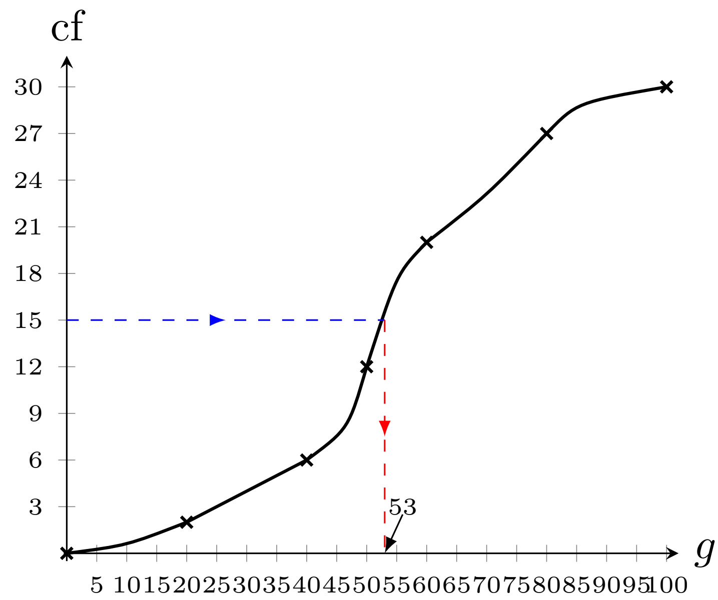
Grade	Girl's freq.	Cumulative frequency	Points
$0-20$	2	2	$(20, 2)$
$21-40$	4	$2+4=6$	$(40, 6)$
$41-50$	6	$6+6=12$	$(50, 12)$
$51-60$	8	$12+8=20$	$(60, 20)$
$61-80$	7	$20+7=27$	$(80, 27)$
$81-100$	3	$27+3=30$	$(100, 30)$

and now we can plot the cumulative frequency graph:



and let us find its median and IQR.
To find the median, we find the 50th percentile:

$$50\% \text{ of } 30 = 15$$

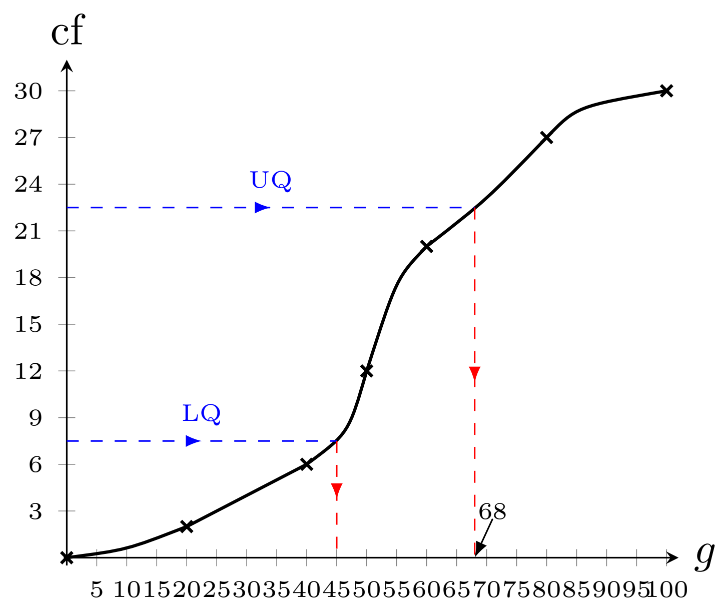


which gives the median to be 53. For the LQ, we need to find the 25th percentile, and for the UQ the 75th percentile:

$$25\% \text{ of } 30 = 7.5$$

$$75\% \text{ of } 30 = 22.5$$

which we can use the graph again to find:



which tells us that the

$$\text{LQ} = 45$$

$$\text{UQ} = 68$$

and allows us to find the IQR:

$$\text{IQR} = \text{UQ} - \text{LQ} = 68 - 45 = 23$$

So we have all the information we need for the girls:

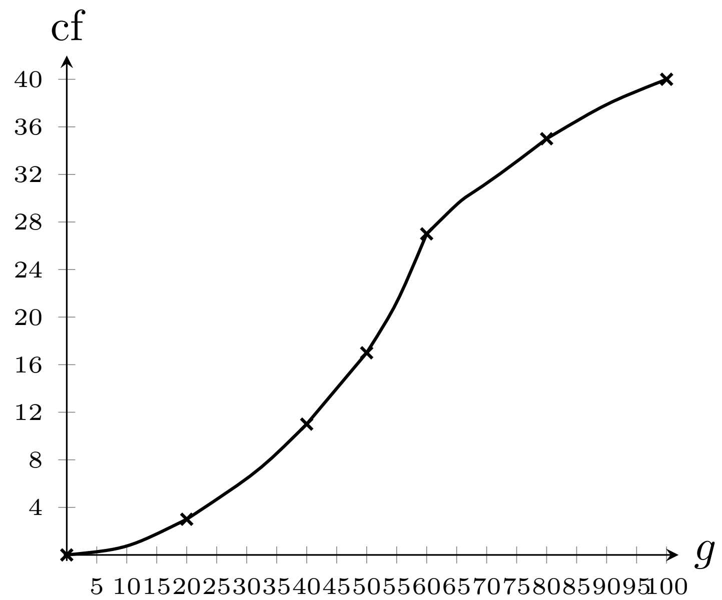
$$\text{median} = 53$$

$$\text{IQR} = 23$$

We now repeat the process for boys. First we find the cumulative frequency and the points to plot:

Grade	Boy's freq.	Cumulative frequency	Points
$0 - 20$	3	3	$(20, 3)$
$21 - 40$	8	$3 + 8 = 11$	$(40, 11)$
$41 - 50$	6	$11 + 6 = 17$	$(50, 17)$
$51 - 60$	10	$17 + 10 = 27$	$(60, 27)$
$61 - 80$	8	$27 + 8 = 35$	$(80, 35)$
$81 - 100$	5	$35 + 5 = 40$	$(100, 40)$

which we can graph:

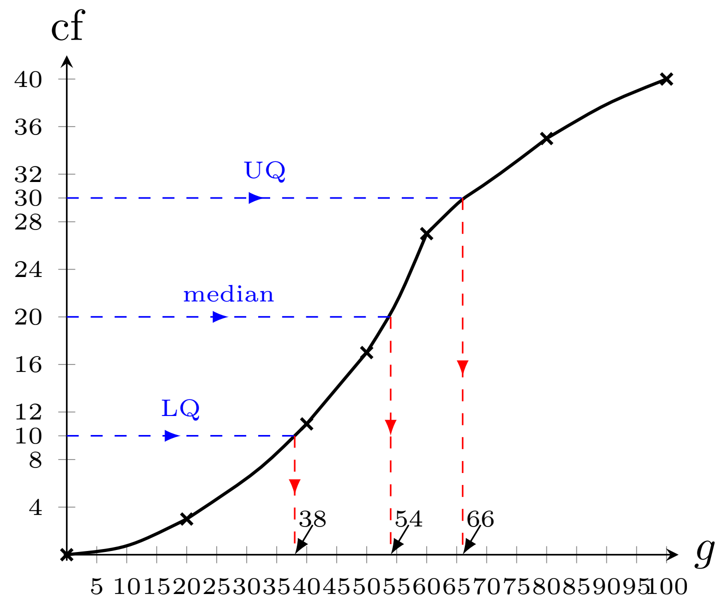


which we can use to find the median, LQ and UQ. Let us first find the y values we need to start:

$$\text{median} = 50\% \text{ of } 40 = 20$$

$$\text{LQ} = 25\% \text{ of } 40 = 10$$

$$\text{UQ} = 75\% \text{ of } 40 = 30$$



which gives for the boys:

$$\text{median} = 54$$

$$\text{LQ} = 38$$

$$\text{UQ} = 66$$

$$\text{IQR} = \text{UQ} - \text{LQ} = 66 - 38 = 28$$

We can now compare the medians and IQR for boys and girls:

Group	Median	IQR
Boys	54	28
Girls	53	23

As the median of boys and girls is very similar, we can say that, on average, they did equally well, with the boys doing slightly better. However, as the IQR for the boys is larger than for the girls, we can say that the boys were less consistent: their grades were more spread.

If the data gave results like this:

Group	Median	IQR
Boys	40	3
Girls	71	25

the analysis would be that, as the median for the girls was much larger than for the boys, they did better. However, they were much more inconsistent, as their IQR was much larger than the boys, which had very little spread. In summary, boys did bad and were consistent at doing bad, whereas girls did well, but were inconsistent, as they had a large spread.

52.5. Histograms

Let us use again this grouped frequency table:

Grade (g)	Frequency
$0 \leq g \leq 30$	3
$30 < g \leq 45$	6
$45 < g \leq 60$	10
$60 < g \leq 80$	9
$80 < g \leq 100$	2

If we wanted to build a bar graph of this data, the graph would not represent much, as the widths of each bar would be different. Hence, we make a compromise: we plot, instead of frequency of each bar, its *frequency density*, defined as:

$$\text{frequency density} = \frac{\text{frequency}}{\text{class width}}$$

Class width is the size of the interval, found by subtracting the left limit from the right limit:

Grade (g)	Frequency	Class width
$0 \leq g \leq 30$	3	$30 - 0 = 30$
$30 < g \leq 45$	6	$45 - 30 = 15$
$45 < g \leq 60$	10	$60 - 45 = 15$
$60 < g \leq 80$	9	$80 - 60 = 20$
$80 < g \leq 100$	2	$100 - 80 = 20$

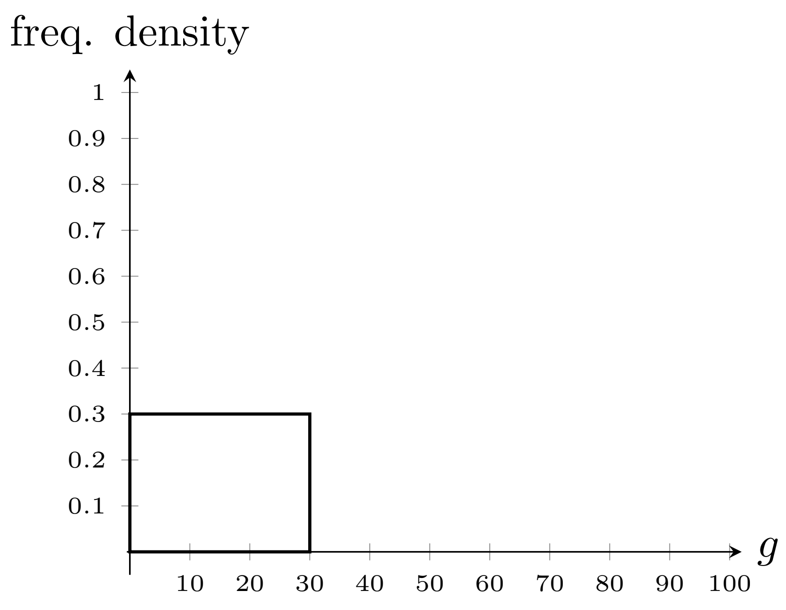
and we can now divide the frequency by the width of each class:

Grade (g)	Frequency	Class width	Frequency density
$0 \leq g \leq 30$	3	30	$\frac{3}{30} = 0.3$
$30 < g \leq 45$	6	15	$\frac{6}{15} = 0.4$
$45 < g \leq 60$	10	15	$\frac{10}{15} = 0.667$
$60 < g \leq 80$	9	20	$\frac{9}{20} = 0.45$
$80 < g \leq 100$	2	20	$\frac{2}{20} = 0.1$

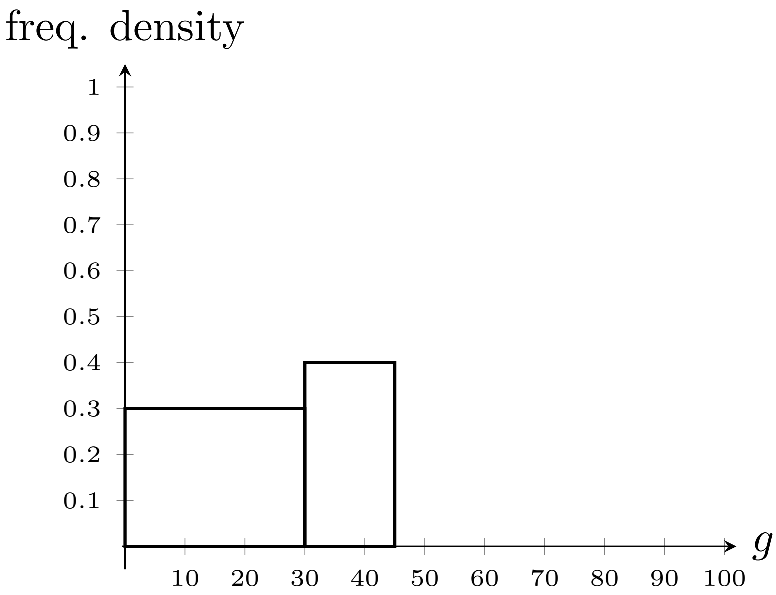
Which we can use to build the histogram, a special bar graph which:

- The height of each bar is its frequency density. The y axis is labeled frequency density;
- The bar width is the class interval. The x axis is whatever we are measuring.

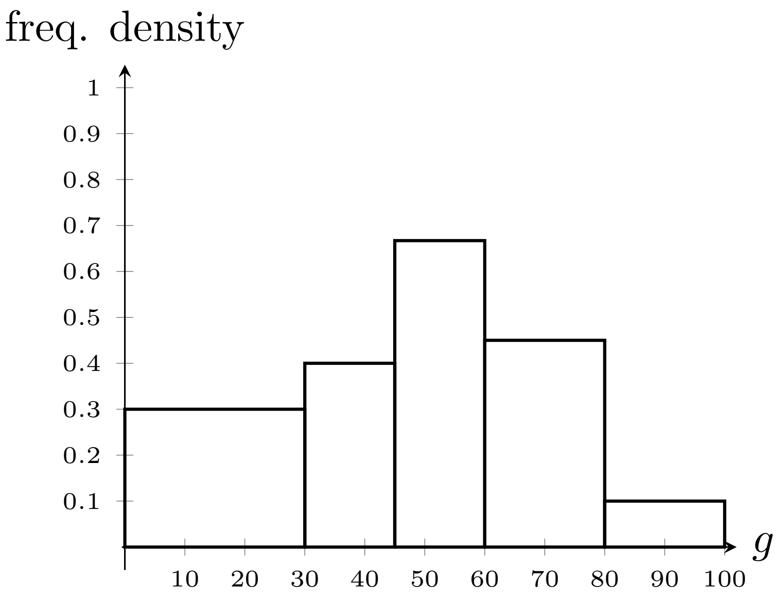
Let us build the histogram for the data above. We start with the first class: it will be a bar from 0 to 30 with height 0.3:



and we continue with the second class, going from 30 to 45 and with height 0.4:



and we continue for each class:



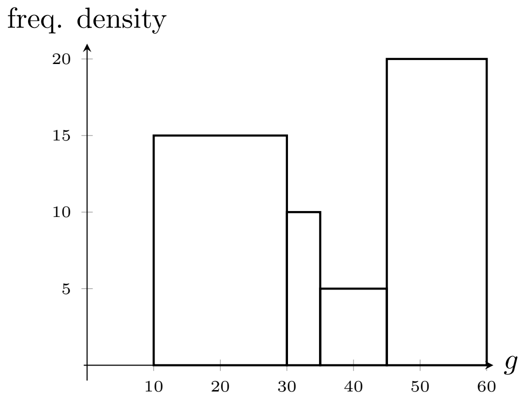
We can use the histogram to see how the data is distributed. In this case, we have a clear "peak" in the 45 to 60 class. There are more worse performing students than top performers, as the first bar has higher frequency density than the last. Also, the distribution is reasonably symmetric, which is nice.

Sometimes, they will give you the histogram and tell you to find the frequency. We do that by making the frequency the subject in the frequency density formula:

$$\text{frequency density} = \frac{\text{frequency}}{\text{class width}}$$

$$\text{frequency} = \text{freq. density} \times \text{class width} \quad \text{Multiply both sides by the class width}$$

which is exactly the same as finding *the area* of the bars: the freq. density is the height of each rectangle, and the class width the size of its base. Thus, if you were given this histogram:



we know that we would have 4 classes. The first would be from 10 to 30, and as it has frequency density 15, its total frequency would be:

$$\text{frequency} = \text{area of the bar} = \underbrace{20}_{\text{base=cw}} \times \underbrace{15}_{\text{height=fd}} = 300$$

and so on for the other bars.

52.6. Exam hints

When finding the estimate of the mean in a grouped frequency table, remember that:

1. You *need* to show the mid values in your working out;
2. Divide the sum of mid values×frequency by the total frequency of the table, which is usually given in the instructions of the question.

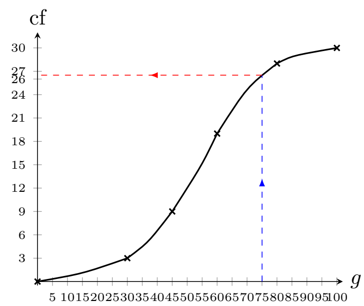
Cumulative frequency diagrams must be drawn in pencil (as all graphs in the IGCSE). Remember to mark your points with little "x" and to have a smooth curve going through all of them.

When answering questions on cumulative frequency graphs, be very careful when they ask you to find “how many X are *more* than” a value: remember that you need to find how many are smaller than that quantity and subtract the result from the total frequency. This kind of question is also usually worth 2 marks, whereas finding “how many X are smaller than” only 1 mark.

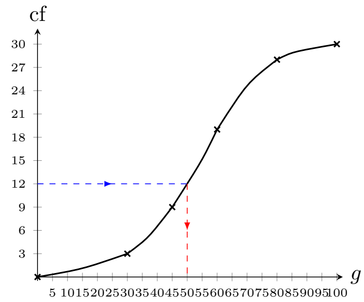
For histograms, you do need to memorize the formula for frequency density.

Summary

- A *grouped frequency table* is a table in which individual observations are *grouped* into *classes*;
- The *modal class* of a grouped frequency table is the *most common class*, that is, the one that has the highest frequency;
- To find an *estimate of the mean* of a grouped frequency table:
  1. Find the mid values of each class by adding the two extremes of each interval and dividing the result by 2;
  2. Multiply the mid value by the frequency of each class;
  3. Divide the sum of frequency$\times$mid value for all classes by the total frequency of the table
- We can find the *cumulative frequency* of a grouped frequency table by adding the frequency of a class to the total frequency so far;
- By plotting the points (right end of the class interval, cumulative frequency) for each class in the plane, and always the point $(0, 0)$, we obtain a *cumulative frequency diagram*;
- We can use a cumulative frequency diagram to find "how many of X" are smaller than a given value. To do that, go from the x axis at the value given until you meet the cumulative frequency curve, and find the y coordinate of where they meet:

  and be careful when asked to find "how many of X are larger than" a value: you have to subtract the quantity you find using the graph from the total.
- We can find *percentiles* using the cumulative frequency diagram. To find the nth percentile, we find $n\%$ of the total frequency and, from that value on the y axis, go right until we meet the cumulative frequency curve. We then find the x coordinate of the point where they meet:



- The *median* is the 50th percentile. Thus, to find it, you first find 50% of the total frequency, then go from that value on the y axis until you meet the cumulative frequency curve. The x value of where you met the curve is the median;

- The *lower quartile* is the 25th percentile. Thus, to find it, you first find 25% of the total frequency, then go from that value on the y axis until you meet the cumulative frequency curve. The x value of where you met the curve is the LQ;

- The *upper quartile* is the 75th percentile. Thus, to find it, you first find 75% of the total frequency, then go from that value on the y axis until you meet the cumulative frequency curve. The x value of where you met the curve is the UQ;

- The *inter-quartile range* is a measure of spread found by

$$\text{IQR} = \text{UQ} - \text{LQ}$$

- A *histogram* is a special bar graph in which the height of each bar is its *frequency density*, found by

$$\text{frequency density} = \frac{\text{frequency}}{\text{class width}}$$

53. Scatter diagrams, stem-and-leaf and box-and-whisker plots

53.1. Why learn even more statistical graphs

In this chapter we will learn about graphical approaches to comparing data sets, which is something quite common.

Stem-and-leaf charts are good organization tools for data while you are collecting it. They also allows us to find some averages very easily.

53.2. Stem-and-leaf diagrams

53.2.1. Interpreting stem-and-leaf diagrams

Stem-and-leaf charts are tables with a special rule: numbers are split into *stems*, which is a part they have in common, and *leaves*, which is everything else. An example is the best way to understand them:

3	3 3 4 5
4	0 0 1
5	2
6	
7	2 5 9 9

Key: 3|2 represents 32

The numbers on the left are the stems and every number on the right are leaves. The most important part is the key, as it tells you how to read the table: you need to take a stem and join it with a leaf:

$$\underbrace{3}_{\text{stem}} | \underbrace{2}_{\text{leaf}} = 32$$

Thus, every leaf gets the stem on the left of its row. The first row, in this case, has stem 3 and leaves $3, 3, 4, 5$, which means we have the numbers

$$33, 33, 34, 35$$

in our data. The second row, which has stem 4, tells us we have the numbers

$$40, 40, 41$$

in our data.

Notice that they key tells us everything we need to know. For instance, take this other stem-and-leaf chart:

3	3 3 4 5
4	0 0 1
5	2
6	
7	2 5 9 9

Key: 3|2 represents 3.2

I just changed the key, but the data is completely different. The first row now represents the values

$$3.3, 3.3, 3.4, 3.5$$

and the second row

$$4.0, 4.0, 4.1$$

The best thing about stem-and-leaf charts is that the data has to be ordered. Thus, it is very easy to find the minimum value: it will always be the first number in the first row. The same thing for the maximum: it will always be the last value in the last row. In our last stem-and-leaf, then, the minimum value is 3.3 and the maximum is 7.9.

We can also easily find the median. First we find out how many numbers are there in the chart by counting how many leaves we have. In our case, there are 12 leaves. To find the median, as we saw in Chapter 51, we add 1 to the number of elements and divide it by 2:

$$\text{median position} = \frac{12+1}{2} = 6.5$$

of course, we do not have position 6.5, so we need to find the mean of the numbers in positions 6 and 7, counting from the top:

3	3 3 4 5
4	0 $\boxed{0}$ $\boxed{1}$
5	2
6	
7	2 5 9 9

Key: 3|2 represents 3.2

but be very careful: remember that the 0 is not a 0, we have to combine it with its stem, 4, and the same for the 1. So the median is

$$\frac{4.0+4.1}{2} = 4.05$$

Going back to our first stem-and-leaf diagram:

3	3 3 4 5
4	0 0 1
5	2
6	
7	2 5 9 9

Key: 3|2 represents 32

we can find its lower quartile by finding the number in position

$$\frac{n+1}{4}$$

where n is the number of elements. Notice that the median is the element in position

$$2 \times \left(\frac{n+1}{4}\right)$$

and the upper quartile the element in position

$$3 \times \left(\frac{n+1}{4}\right)$$

and all of those are easy to find in the stem-and-leaf. Hence, the lower quartile is the element in position

$$\text{LQ position} = \frac{12+1}{4} = 3.25$$

and as we don't have that position, we again find the mean of the numbers in positions 3 and 4:

$$\text{LQ} = \frac{34+35}{2} = 34.5$$

53.2.2. Building stem-and-leaf diagrams

Writing stem-and-leaf diagrams from the data is quite straightforward. Say we have the following data:

$$10, 13, 19, 20, 20, 21, 29, 33,$$

$$35, 36, 40, 61, 61, 61, 69, 77$$

The minimum value is 10 and the maximum is 77, so a good idea for the stem is for it to represent the tens in the numbers, which will go from 1 to 7:

1	
2	
3	
4	
5	
6	
7	

Key: 1|7 represents 17

and now we go row by row, which will represent the numbers with each tens digit. For the first row, we will write the values $10, 13$ and 19, which have 1 in their tens digit, but we just write the leaves (what is missing after we remove the stem):

1	0 3 9
2	
3	
4	
5	
6	
7	

Key: 1|7 represents 17

For the stem 2, we have the numbers $20, 20, 21$ and 29, so we add the leaves $0, 0, 1$ and 9 to the second row:

1	0 3 9
2	0 0 1 9
3	
4	
5	
6	
7	

Key: 1|7 represents 17

and we continue doing that for every stem:

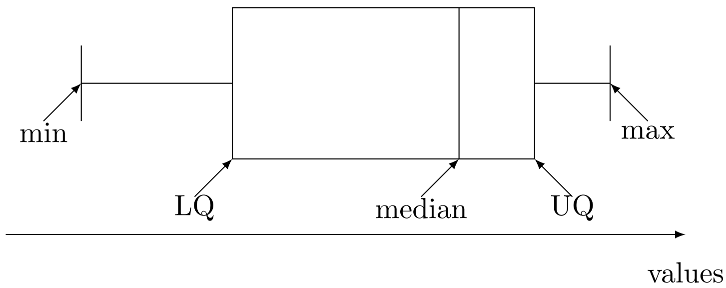
1	0 3 9
2	0 0 1 9
3	3 5 6
4	0
5	
6	1 1 1 9
7	7

Key: 1|7 represents 17

53.3. Box-and-whisker plots

A box-and-whisker plot is an excellent graphical display of 5 pieces of important information about our data: the minimum value, the lower quartile, the median, the upper quartile and the maximum.

We show all of those as such:



As you can see, the box is delimited by the lower and upper quartiles, with the median being a "cut" in the box. The whiskers of the box are made by joining the left of the box to the minimum value and the right of the box to the maximum value.

For instance, say we have some data with the following values:

$$\text{minimum} = 2$$

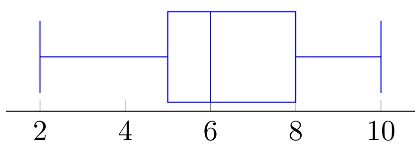
$$\text{LQ} = 5$$

$$\text{median} = 6$$

$$\text{UQ} = 8$$

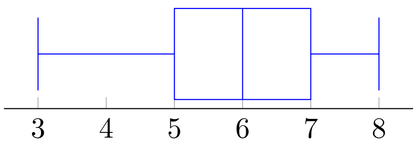
$$\text{maximum} = 10$$

we can build a box-and-whisker plot:



The point of building these diagrams is to easily see how the data is distributed: if the box is quite small, the inter quartile range is small, hence the data has a small spread. If the box is large, the spread in the data is big.

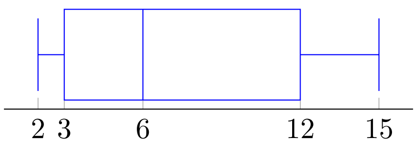
Let us see another example:



from which we can find the values of the minimum, at the left whisker, 3; the LQ is the left border of the box, so 5; the median the line inside the box, so 6; the UQ the right part of the box, so 7; and, finally, the maximum, at the right whisker, which is 8. Hence, we can find the IQR:

$$\text{IQR} = \text{UQ} - \text{LQ} = 7 - 5 = 2$$

Compare the above with this box-and-whisker plot:



In this one, the values given to us are:

$$\text{minimum} = 2$$

$$\text{LQ} = 3$$

$$\text{median} = 6$$

$$\text{UQ} = 12$$

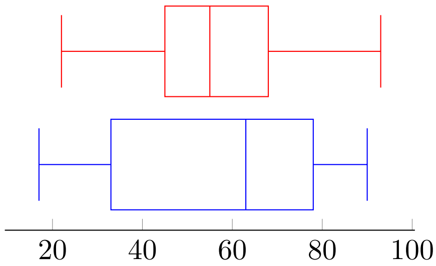
$$\text{maximum} = 15$$

which has IQR

$$\text{IQR} = \text{UQ} - \text{LQ} = 12 - 3 = 9$$

which is a lot bigger. However, we can evaluate the IQR just by looking at the width of the box, which is quite convenient.

Where box-and-whisker plots really shine is to compare data sets, however. For instance,



which represent the grades two groups obtained in a test (out of 100%). We can see that the red group had the best grade, as its right whisker is further to the right than the the right blue whisker. We can also see that the red group is more consistent: their grades have a smaller IQR, as the red box has a smaller width than the blue box. On the other hand, we can see that the blue group has a better median, as the line inside the blue box is to the right of the line inside the red box.

53.4. Scatter plots, correlation and line of best fit

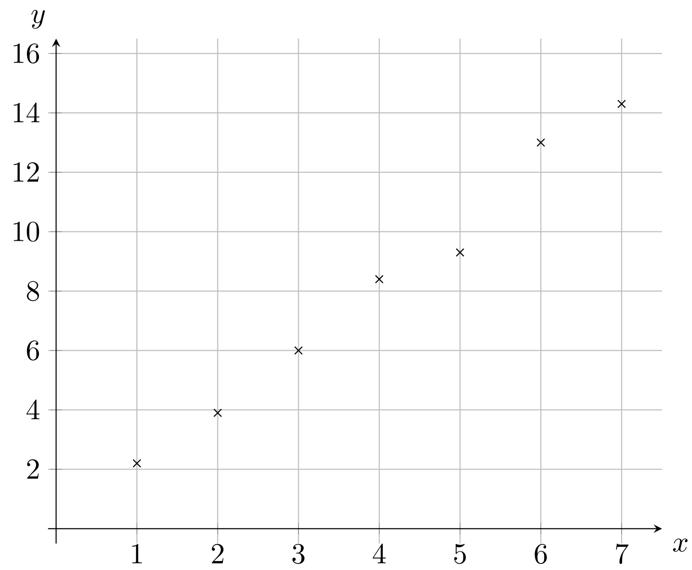
Scatter plots are graphs in which we have data of two different measurements related to the same individual or thing. We group those measurements as an ordered pair and plot them to see if there is any relationship between them: perhaps they grow together, perhaps when one of them increases the other decreases, or even they have no relation at all. If they do show some relationship, we say that the two variables are *correlated* and we can do further analysis on their relationship.

53.4.1. Scatter plots

Let us first do a scatter plots of some artificial data. Say we have the following measures for two variables, x and y:

x	1	2	3	4	5	6	7
y	2.2	3.9	6	8.4	9.3	13	14.3

and let us plot them into the xy-plane:



As you can see, a scatter plot just plots each pair of measures. Let us now see the types of correlation we can have.

53.4.2. Types of correlation

53.4.2.1. Positive correlation

Let us look at a classic example, which is quite interesting. We will produce a scatter plot of the following data[1]:

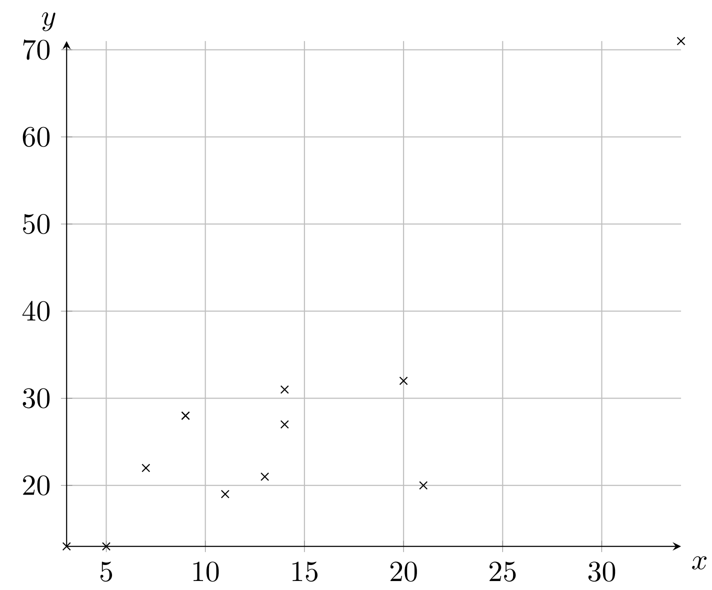
Country	Number of IKEA stores	Number of Nobel prizes
France	34	71
Sweden	20	32
Russia	14	31
Switzerland	9	28
Japan	9	28
Canada	14	27
Austria	7	22
Netherlands	13	21
Italy	21	20
Poland	11	19
Denmark	5	13
Hungary	3	13

[1] Data from:

1. `https://en.wikipedia.org/wiki/List_of_countries_with_IKEA_stores`
2. `https://en.wikipedia.org/wiki/List_of_Nobel_laureates_by_country`

Both accessed on 23/12/2020.

Let us plot pairs (number of IKEA stores, number of Nobel prizes) and see what we can find:



It looks, from the plot, that the more IKEA stores in a country, the more Noble prizes it has! This is a case of *positive correlation*: when one variable increases, the other one does as well. In variable words,

$$\uparrow x \rightarrow \uparrow y$$

In our case:

$$\uparrow \text{IKEA stores} \rightarrow \uparrow \text{Nobel prizes}$$

It is important here to learn something: **correlation does not imply causation**. Just because two values happen to have some statistical coincidence, it does not mean they actually cause one another. I am sure you agree that increasing the number of IKEA stores in a country has no bearing whatsoever in the quality of their research. In cases like this, we would need to do further analysis to discover what is really causing the correlation.

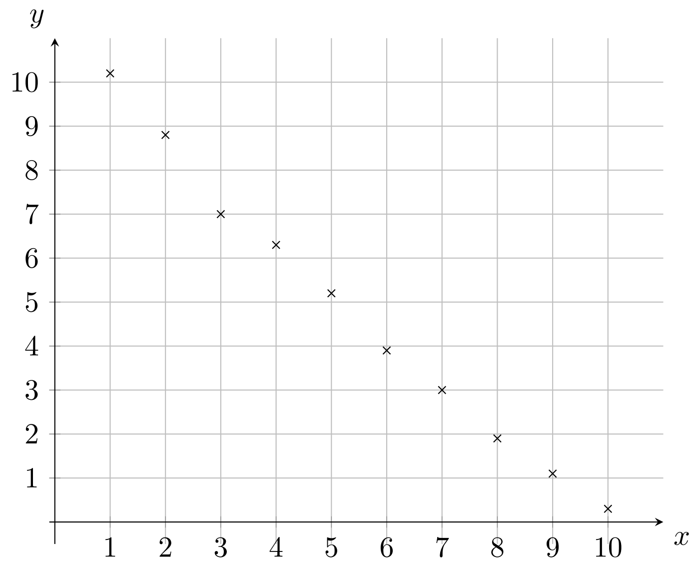
Another important thing: correlation is a statistical property of the data, which means it is a "general behaviour". It is totally possible for the x value to increase and the y value to decrease, as long as the general behaviour remains: as x grows, y grows. For instance, in this example, Denmark has 5 IKEA stores, whereas Hungary has 3. However, both Hungary and Denmark have the same number of Nobel prizes.

So, remember:

- correlation does not imply causation; and
- correlation is a "general behaviour", it does not have to happen for all values of the data.

53.4.2.2. Negative correlation

Negative correlation is when the relationship of the two variables is inverse: when one of them increases, the other one decreases. A fake example:



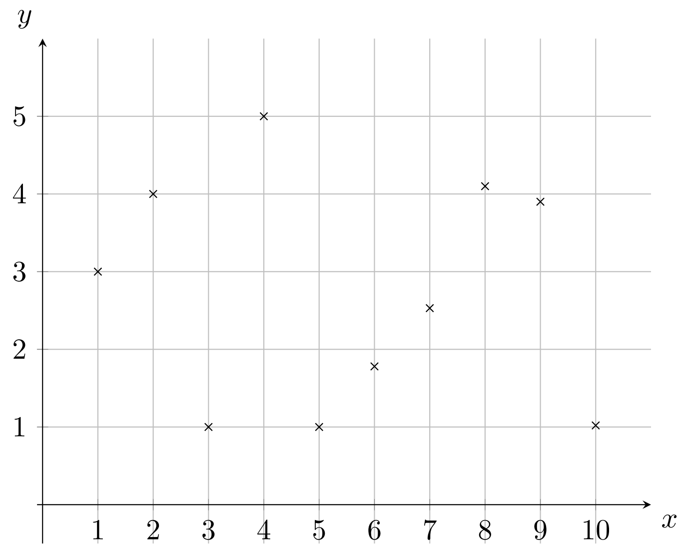
As you can see, when x increases (we move to the right), the values of y decrease. Hence, in variable:

$$\uparrow x \rightarrow \downarrow y$$

Again, this is a statistical observation, and does not have to happen all the time. It is a general behaviour of the data.

53.4.2.3. No correlation

Of course, you can have two variables that have no correlation whatsoever:



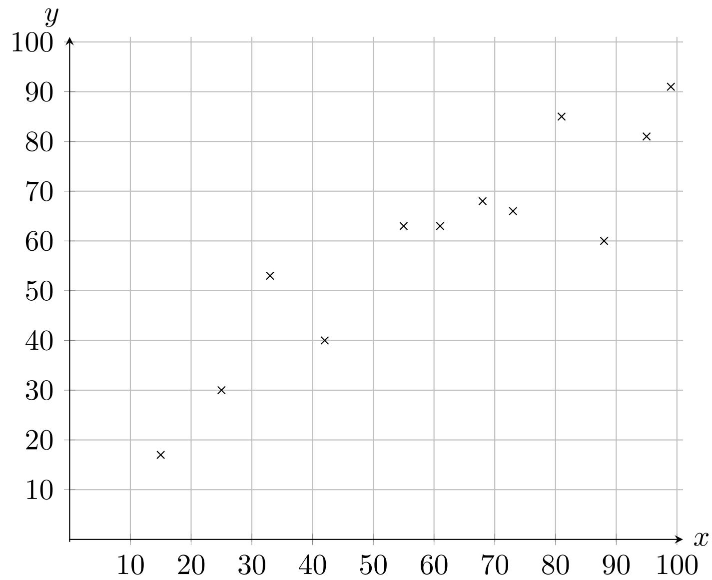
there is no relationship whatsoever between x and y now: sometimes x increases and y increases, sometimes y decreases. When this happens there is no correlation.

53.4.3. Linear correlation and the line of best fit

Say we have this data about the grades in arts and maths of some students:

maths	73	99	55	42	82	61	68	95	33	50	88	15	25
arts	66	91	63	40	85	63	68	81	53	35	60	17	30

Let us plot those pairs of grades, maths on the x axis and arts on the y axis:

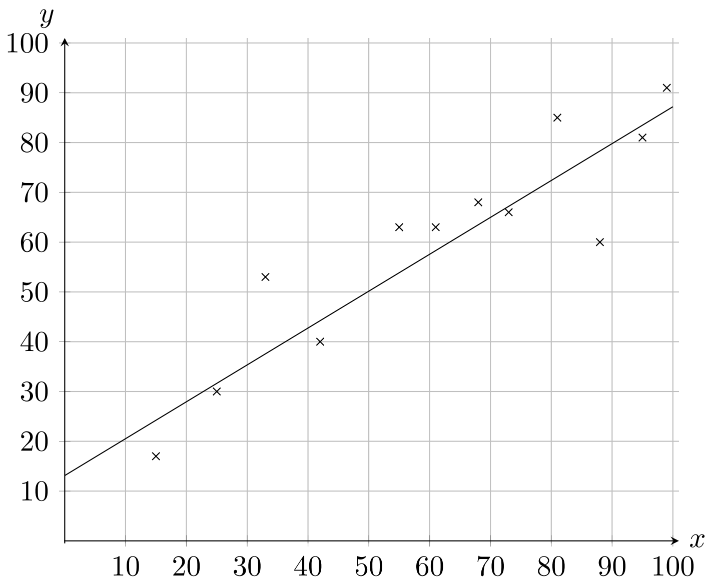


as we learned above, this scatter plot shows there is *positive correlation* between the maths and arts grades: as the maths grades increase, so do the arts grades. However, there is an even stronger observation about this data: there is *linear correlation* between maths and arts grades.

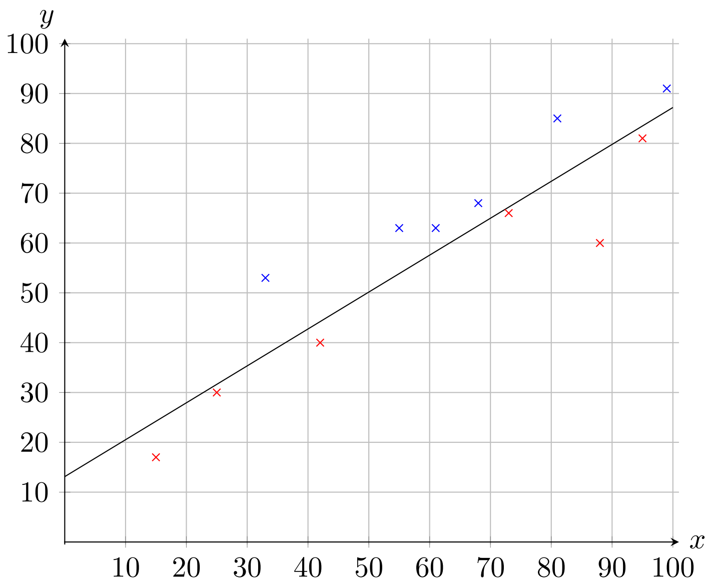
Linear correlation implies that we can relate the x and y variables by a linear graph, which we call the *line of best fit*. There are methods to find the line of best fit (see the Formality after taste of this chapter), but in the exam you just need to eyeball it. On the IGCSE, your lines of best fit must:

- Have approximately the same number of points above and below;
- If we know the mean of the x and y data sets, the line *must go* through the point $(\text{mean of } x, \text{mean of } y)$

It also has to be a straight line, but I suppose that is obvious. Let us add the line of best fit in the scatter plot above:



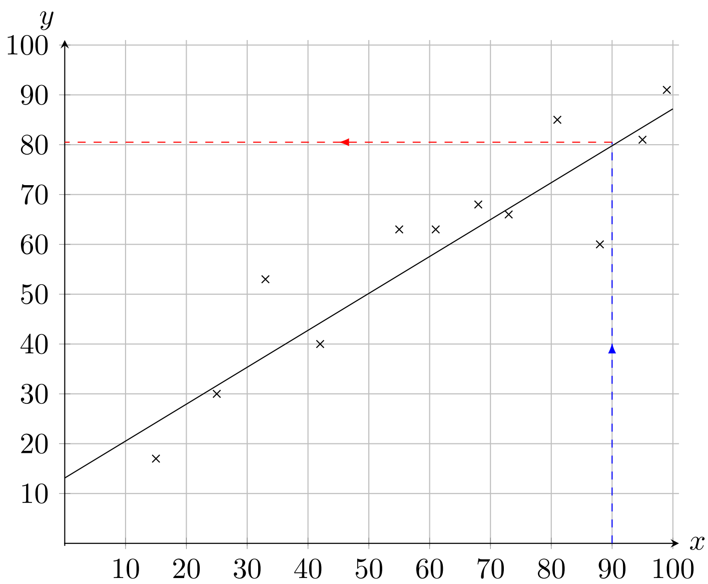
we can count the points above the line (let us mark them blue) and the ones below (in red):



and you can see that there are 6 points below and 6 points above the line. Of course, I cheated and used the computer to find the line of best fit, so it is perfect. In the exam, it is fine as long as those numbers are similar.

Before we continue, a very common misconception is to think that lines of best fit must go through the origin, and you can see in our example that they do not.

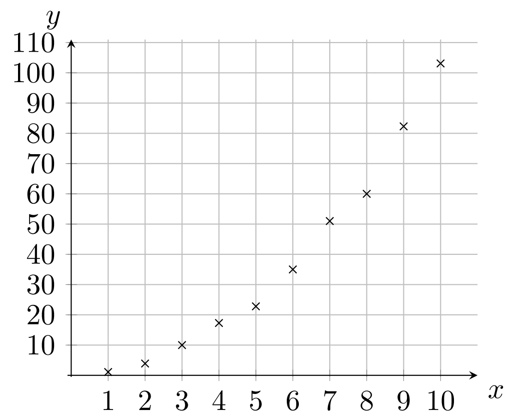
Now, let us use our line to do some predictions: if someone's results in maths were 90, how much would they have in arts? We can use our line to find that: on the x axis, which represents maths, we first find the value 90. Then, we draw a vertical line until we meet the line of best fit (blue traced line), and then we go to the left until we meet the y axis (red traced line). The value on the y axis is our estimate of the arts grade:



Thus, if someone had 90 for maths, we would predict they would have 80 in arts. You can also just get the y coordinate of the point where the blue line meets the line of best fit. Whatever floats your boat.

53.4.4. Other types of correlation

Even though in the IGCSE you will only be assessed on linear correlation (or no correlation at all), there are many instances of two data sets being related to another non linearly. For instance:



Here, we have a correlation: as x grows, y also grows. However, y grows faster than x: we have quadratic correlation (I basically squared x coordinates and changed the results a bit).

53.5. Exam hints

These topics are excellent, as it is very repetitive. You just have to be careful with:

1. Do not forget the key in the stem-and-leaf diagrams: without it the table makes no sense;

2. The line of best fit does not have to go through the origin; and

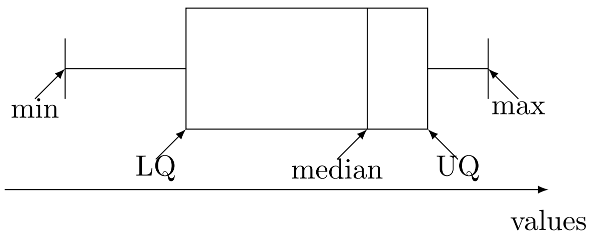
3. Graphs needs to be done in pencil!

Summary

- A *stem-and-leaf* diagram is a table which:
  - stores the data *in order*;
  - all values in it are made of joining the *stem*, which is the first number on the left of each row, with a *leaf*, which is a single number after the | in the table;

   - they *key* tells you how to join the stem with the leaf.

- A *box-and-whisker* plot is a graph that tells you the *minimum, lower quartile, median, upper quartile and maximum* of the data:



- A *scatter plot* is a graph in which we plot two different measures as an *ordered pair* (x, y);

- A scatter plot can show *correlation* between x and y: a statistical relationship between the values. Correlation can be:
   - *positive*: when x and *y move together.* If one increases the other one increases, if one decreases the other one also decreases;
   - *negative*: when x and *y move in opposite directions.* If one increases the other one decreases, if one decreases the other one increases;
   - *no correlation*: when there is *no relationship* between x and y;

- If there is positive or negative correlation between x and y, we may draw a *line of best fit* in the graph. This line must have approximately the same number of points above and below it.

Formality after taste

(This may be the "hardest" formality after taste in the book, so have fun!)

The chain rule

In Chapter ADDREF, we learned how to find derivatives of simple functions. Here, I will introduce you to the chain rule, a way to derive compositions of functions.

If we have a composition

$$fg(x)$$

which we can write as

$$f\left(g(x)\right)$$

and want to find its derivative, the chain rule states that

$$\left(f\left(g(x)\right)\right)' = g'(x)f'\left(g(x)\right)$$

This may look complicated, but it is quite simple. Let us give names to the functions in the composition:

$$\overbrace{f}^{\text{outer function}}\left(\underbrace{g(x)}_{\text{inner function}}\right)$$

where we called f the outer function, as it is outside the big brackets pair, and g the inner function as it is inside the big brackets pair. The chain rule, then, gets labeled as

$$\underbrace{g'(x)}_{\text{derivative of inner}} \quad \overbrace{f'\left(\underbrace{g(x)}_{\text{inner function}}\right)}^{\text{derivative of outer}}$$

Thus, the chain rule states that the derivative of a composition of an outer with an inner function is given by the derivative of the inner function multiplied by the the composition of the derivative of the outer function with the inner function. Lovely, no?

Let us see some examples. Say we would like to derive

$$f(x) = \left(3x^2 + 5x\right)^4$$

The easiest way of doing it is identifying the inner and outer functions:

$$f(x) = \left(\underbrace{3x^2 + 5x}_{\text{inner}}\right)^{\overbrace{4}^{\text{outer}}}$$

Here, the composition is made of two functions: $3x^2 + 5$, the inner function, with the function x^4,the outer function. We can check that writing f as

$$f(x) = \text{outer}(\text{inner}(x))$$

with

$$\text{outer}(x) = x^4 \text{ and } \text{inner}(x) = 3x^2 + 5$$

Let us now use the formula from above, which tells us we need to derive both the inner and outer functions:

$$\text{inner}'(x) = \left(3x^2 + 5\right)' = 6x$$

$$\text{outer}'(x) = \left(x^4\right)' = 4x^3$$

and now we combine them:

$$f'(x) = \underbrace{6x}_{\text{derivative of inner}} \times \overbrace{4\left(\underbrace{3x^2 + 5}_{\text{inner}}\right)^3}^{\text{derivative of outer}}$$

which we can simplify to

$$f'(x) = 24x\left(3x^2+5\right)^3$$

Another example: say we want to derive

$$h(x) = \sin\left(\cos(x)\right)$$

Here $\sin(x)$ is the outer function, and $\cos(x)$ is the inner function. Deriving them:

$$\text{inner}'(x) = (\cos(x))' = -\sin(x)$$

$$\text{outter}'(x) = (\sin(x))' = \cos(x)$$

and we can now combine them:

$$h'(x) = \underbrace{-\sin(x)}_{\text{derivative of inner}} \times \overbrace{\cos\left(\underbrace{\cos(x)}_{\text{inner}}\right)}^{\text{derivative of outer}}$$

which we cannot really simplify, but it does look cute.

Deriving a formula for the line of best fit

In the differentiation chapter, we learned how to find turning points of functions that have a single variable.

There is a similar idea to derive one formula to find the coefficients m and c of the line of best fit given some points.

Let use say we have n observations

$$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$$

and we want to find the line $y = mx + c$ which minimizes the sum of the square of the distances between the correct value for y of each point and the value that the line gives. This means that, for each point (x_k, y_k) in our list, we first find the y value that $y = mx + c$ gives:

$$y = mx_k + c$$

Now, we subtract the correct value of y_k from the value for y above:

$$y_k - (mx_k + c)$$

and square this:

$$(y_k - (mx_k + c))^2$$

We then add these squares of distances for all values of k from 1 to n:

$$(y_1 - (mx_1 + c))^2 + (y_2 - (mx_2 + c))^2 + \ldots + (y_k - (mx_k + c))^2 + \ldots + (y_n - (mx_n + c))^2$$

Now, let us define a function f that has two variables, m and c, which is the sum of the squares above:

$$f(m,c) = (y_1 - (mx_1 + c))^2 + (y_2 - (mx_2 + c))^2 + \ldots + (y_k - (mx_k + c))^2 + \ldots + (y_n - (mx_n + c))^2$$

Our goal, now, is to minimize f, that is, we want to find the values of m and c for which f is the least value possible (it does have a minimum value because f is a positive paraboloid, a "3d happy parabola").

To find the minimum for a function of two variables, we also find its derivative and set it equal to 0. However, as we now have a function of two variables, we have two derivatives, one for m and one for c. The notation changes a bit:

$$\frac{\partial f}{\partial m} = 0 \text{ and } \frac{\partial f}{\partial c} = 0$$

Here, $\frac{\partial f}{\partial m}$ is read as the "partial derivative of f with respect to m", and we find it by imagining every other variable (in our case c) is constant. And when we derive constants (or variables we set as constant), they become 0. Then, it is very similar to the differentiation we did in Chapter ADDREF, but we do it twice: once for the variable m pretending everything else is constant, and once for the variable c, pretending everything else is constant.

Let us see how to proceed here using the first term

$$(y_1 - (mx_1 + c))^2$$

as an example. We have a composition of functions:

$$\text{inner} = y_1 - (mx_1 + c)$$

$$\text{outer} = x^2$$

so we need to use the chain rule. Let us derive this term with respect to m:

$$\frac{\partial\left((y_1 - (mx_1 + c))^2\right)}{\partial m} = \underbrace{-x_1}_{\text{inner}'} \times 2 \left(\underbrace{y_1 - (mx_1 + c)}_{\text{inner}}\right) = -2x_1 (y_1 - (mx_1 + c))$$

All the terms of f, when derived with respect to m, yield a similar expression, only changing the x_k and y_k. Hence, we have:

$$\frac{\partial f}{\partial m} = -2x_1 (y_1 - (mx_1 + c)) + -2x_2 (y_2 - (mx_2 + c)) + \ldots + -2x_n (y_n - (mx_n + c))$$

which we make equal to 0 and obtain our first equation:

$$-2x_1 (y_1 - (mx_1 + c)) + -2x_2 (y_2 - (mx_2 + c)) + \ldots + -2x_n (y_n - (mx_n + c)) = 0$$

When we derive f with respect to x we obtain a similar, but simpler expression. Using the first term again as an example:

$$\frac{\partial\left(y_1-(mx_1+c)\right)^2}{\partial c}=\underbrace{-1}_{\text{inner}'}\times 2\left(y_1-(mx_1+c)\right)=-2\left(y_1-(mx_1+c)\right)$$

and again we only have to exchange the subscripts for the other terms. We then have:

$$\frac{\partial f}{\partial c}=-2\left(y_1-(mx_1+c)\right)+-2\left(y_2-(mx_2+c)\right)+\ldots+-2\left(y_n-(mx_n+c)\right)$$

and make it equal to 0, obtaining our second equation:

$$-2\left(y_1-(mx_1+c)\right)+-2\left(y_2-(mx_2+c)\right)+\ldots+-2\left(y_n-(mx_n+c)\right)=0$$

We now have two simultaneous equations:

$$\begin{cases}-2x_1\left(y_1-(mx_1+c)\right)+-2x_2\left(y_2-(mx_2+c)\right)+\ldots+-2x_n\left(y_n-(mx_n+c)\right)=0 & (1)\\ -2\left(y_1-(mx_1+c)\right)+-2\left(y_2-(mx_2+c)\right)+\ldots+-2\left(y_n-(mx_n+c)\right)=0 & (2)\end{cases}$$

which we want to solve for m and c. Let us first simplify these equations: everything can be divide by -2 and let us fix the

$$\begin{cases}x_1\left(y_1-(mx_1+c)\right)+x_2\left(y_2-(mx_2+c)\right)+\ldots+x_n\left(y_n-(mx_n+c)\right)=0 & (1)\\ \left(y_1-(mx_1+c)\right)+\left(y_2-(mx_2+c)\right)+\ldots+\left(y_n-(mx_n+c)\right)=0 & (2)\end{cases}$$

Let us play with equation (2) first. We can write it like this:

$$\left(y_1+y_2+\ldots+y_n\right)-m\left(x_1+x_2+\ldots+x_n\right)-\left(\underbrace{c+c+\ldots+c}_{n\text{ times}}\right)=0$$

$$\left(y_1+y_2+\ldots+y_n\right)-m\left(x_1+x_2+\ldots+x_n\right)-nc=0$$

Now, let us divide everything by n:

$$\frac{\left(y_1+y_2+\ldots+y_n\right)}{n}-m\frac{\left(x_1+x_2+\ldots+x_n\right)}{n}-c=0$$

Notice that adding all values of y and dividing it by n is the same as finding the mean of y. The same happens for adding all the values for x and dividing it by n, we get the mean of x. Let us introduce some notation:

$$\bar{x}=\frac{\left(x_1+x_2+\ldots+x_n\right)}{n}\text{ and }\bar{y}=\frac{\left(y_1+y_2+\ldots+y_n\right)}{n}$$

and we obtain

$$\bar{y}-m\times\bar{x}-c=0$$

which we can rearrange to make c the subject:

$$c = \bar{y} - \bar{x}m$$

This last equation allows us to find c when have found the value of m. Before we continue, though, let us play a bit with Equation (1):

$$x_1(y_1 - (mx_1 + c)) + x_2(y_2 - (mx_2 + c)) + \ldots + x_n(y_n - (mx_n + c)) = 0$$

$$x_1y_1 - x_1^2m - x_1c + x_2y_2 - x_2^2m - x_2c + \ldots + x_ny_n - x_n^2m - xc = 0$$

$$(x_1y_1 + x_2y_2 + \ldots + x_ny_n) - c(x_1 + x_2 + \ldots + x_n) - m\left(x_1^2 + x_2^2 + \ldots + x_n^2\right) = 0$$

$$(x_1y_1 + x_2y_2 + \ldots + x_ny_n) - \underbrace{(\bar{y} - \bar{x}m)}_{c}(x_1 + x_2 + \ldots + x_n) - m\left(x_1^2 + x_2^2 + \ldots + x_n^2\right) = 0$$

$$(x_1y_1 + x_2y_2 + \ldots + x_ny_n) - \bar{y}(x_1 + x_2 + \ldots + x_n) + \bar{x}m(x_1 + x_2 + \ldots + x_n) - m\left(x_1^2 + x_2^2 + \ldots + x_n^2\right) = 0$$

$$\bar{x}m(x_1 + x_2 + \ldots + x_n) - m\left(x_1^2 + x_2^2 + \ldots + x_n^2\right) = \bar{y}(x_1 + x_2 + \ldots + x_n) - (x_1y_1 + x_2y_2 + \ldots + x_ny_n)$$

$$m\left(\bar{x}(x_1 + x_2 + \ldots + x_n) - \left(x_1^2 + x_2^2 + \ldots + x_n^2\right)\right) = \bar{y}(x_1 + x_2 + \ldots + x_n) - (x_1y_1 + x_2y_2 + \ldots + x_ny_n)$$

$$m = \frac{\bar{y}(x_1 + x_2 + \ldots + x_n) - (x_1y_1 + x_2y_2 + \ldots + x_ny_n)}{\bar{x}(x_1 + x_2 + \ldots + x_n) - (x_1^2 + x_2^2 + \ldots + x_n^2)}$$

What a lovely result, no?

I did all of that to show that the point $(\bar{x}, \bar{y})$ is on the line of best fit

$$y = \frac{\bar{y}(x_1 + x_2 + \ldots + x_n) - (x_1y_1 + x_2y_2 + \ldots + x_ny_n)}{\bar{x}(x_1 + x_2 + \ldots + x_n) - (x_1^2 + x_2^2 + \ldots + x_n^2)}x + \bar{y} - \bar{x}\left(\frac{\bar{y}(x_1 + x_2 + \ldots + x_n) - (x_1y_1 + x_2y_2 + \ldots + x_ny_n)}{\bar{x}(x_1 + x_2 + \ldots + x_n) - (x_1^2 + x_2^2 + \ldots + x_n^2)}\right)$$

$$\bar{y} = \frac{\bar{y}(x_1 + x_2 + \ldots + x_n) - (x_1y_1 + x_2y_2 + \ldots + x_ny_n)}{\bar{x}(x_1 + x_2 + \ldots + x_n) - (x_1^2 + x_2^2 + \ldots + x_n^2)}\bar{x} + \bar{y} - \bar{x}\left(\frac{\bar{y}(x_1 + x_2 + \ldots + x_n) - (x_1y_1 + x_2y_2 + \ldots + x_ny_n)}{\bar{x}(x_1 + x_2 + \ldots + x_n) - (x_1^2 + x_2^2 + \ldots + x_n^2)}\right)$$

and you can see that the

$$\bar{x}\left(\frac{\bar{y}(x_1 + x_2 + \ldots + x_n) - (x_1y_1 + x_2y_2 + \ldots + x_ny_n)}{\bar{x}(x_1 + x_2 + \ldots + x_n) - \left(x_1^2 + x_2^2 + \ldots + x_n^2\right)}\right)$$

cancel each other out and we obtain

$$\bar{y} = \bar{y}$$

Now we can sleep well.

Part VII.

Appendices

Cavalieri's principle

What do we need Cavalieri's principle for?

In order for us to have at least some demonstration of the formulas for volumes, we need either to know calculus or use Cavalieri's principle.

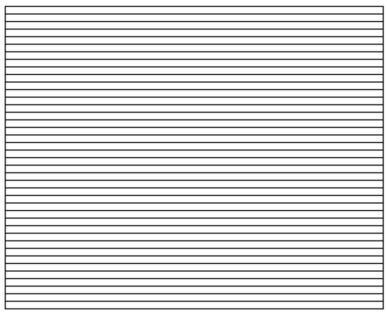
I say "at least" because proving Cavalieri's principle is also not appropriate at this level, so we will use it as an *axiom*, a given truth.

Before we continue, this discussion is based on the book "A Matemática do Ensino Médio", Volume 2, by Lima, Carvalho, Wagner and Morgado. These authors are Brazilian legends in the teaching of maths, and I do recommend the book if you become a maths teacher (which you should :).

The idea

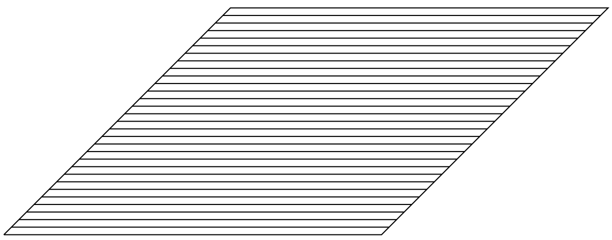
Before enunciating Cavalieri's principle, let us see a specific example to motivate it.

Say that you just opened a pack of printing paper, and you put all the sheets on top of a table:



In this stack, each sheet has a given volume and surface area. Also, the volume of the whole stack is the sum of the volumes of each sheet.

Now, we tilt our stack of paper a bit:



We still have exactly the same number of paper sheets in the stack, and each sheet has the same volume (and area) as in the original stack: thus, the tilted stack has exactly the same volume as the original stack.

This an example of Cavalieri's principle: we start with two solids, both standing on the same table, and we slice both of them with a parallel plane to the table. If all those slices have equal areas in both solids, the solids have the same volume.

Cavalieri's principle

Given two solids and a plane P, if all planes parallel to P determine sections of equal areas in the two solids, then the solids have equal volumes.

53.6. An intuition of why this works

If you think that each of the "slices" we make when cutting the shapes is very very thin, the volume of each slice with same area is approximately equal; the more equal, the thinner the slice. As planes are an abstract notion, they are the thinnest you can get.

Bibliography and further reading

Calculus, Tom Apostol

This book is one that greatly influenced me. After initial courses in Calculus following the usual order (limits, derivatives, integrals), it was very magical to see the integral defined first!

One of the most suggested reading for Calculus enthusiasts.

Naive Set Theory, Paul Halmos

Elementary Geometry from an Advanced Standpoint, Edwin E. Moise