Understanding science: what we cannot know

This item contains selected online content. It is for use alongside, not as a replacement for the module website, which is the primary study format and contains activities and resources that cannot be replicated in the printed versions.

About this free course

This version of the content may include video, images and interactive content that may not be optimised for your device.

You can experience this free course as it was originally designed on OpenLearn, the home of free learning from The Open University:

https://www.open.edu/openlearn/science-maths-technology/understanding-science-what-we-cannot-know/content-section-overview.

There you'll also be able to track your progress via your activity record, which you can use to demonstrate your learning.

Contents

Introduction and guidance

Introduction and guidance

Welcome to this free badged course, *Understanding science: what we cannot know*. It's been developed in collaboration with Marcus du Sautoy, The Simonyi Professor for the Public Understanding of Science, inspired by his book *What We Cannot Know: Explorations at the Edge of Knowledge*.

This course is dedicated to the memory of Professor Uwe Grimm whose enthusiasm and expertise led its production.

It comprises eight weeks with approximately three hours of study time each. You can work through the course at your own pace, so if you have more time after completing one week, there is no problem with pushing on to complete another week. This course covers a broad range of scientific topics: probability, particle physics, quantum physics, space, time, neuroscience, and infinity.

There are opportunities to check your learning throughout the course. At the end of each week, there is a quiz to help you check your understanding. And, if you want to receive a formal statement of participation, at the end of Weeks 4 and 8 there is a quiz which you need to pass. You can read more on how to study the course and about badges in the next sections.

After completing this course, you should be able to:

- understand current knowledge in a diverse range of scientific fields
- describe some events in the history of scientific discovery
- outline where scientific investigation might lead us next and what might be discovered
- describe some potential limitations of human understanding.

Moving around the course

In the 'Summary' at the end of each week, you can find a link to the next one. If at any time you want to return to the start of the course, click on 'Course content'. From here you can navigate to any part of the course. Alternatively, use the week links at the top of every page of the course.

It's also good practice, if you access a link from within a course page (including links to the quizzes), to open it in a new window or tab. That way you can easily return to where you've come from without having to use the back button on your browser.

The Open University would really appreciate a few minutes of your time to tell us about yourself and your expectations for the course before you begin, in our optional start-of-course survey. Participation will be completely confidential and we will not pass on your details to others.

What is a badged course?

While studying *Understanding science: what we cannot know* you have the option to work towards gaining a digital badge.

Badged courses are a key part of The Open University's *mission to promote the educational well-being of the community*. The courses also provide another way of helping you to progress from informal to formal learning.

Completing a course will require about 24 hours of study time. However, you can study the course at any time and at a pace to suit you.

Badged courses are available on The Open University's OpenLearn website and do not cost anything to study. They differ from Open University courses because you do not receive support from a tutor, but you do get useful feedback from the interactive quizzes.

What is a badge?

Digital badges are a new way of demonstrating online that you have gained a skill. Colleges and universities are working with employers and other organisations to develop open badges that help learners gain recognition for their skills, and support employers to identify the right candidate for a job.

Badges demonstrate your work and achievement on the course. You can share your achievement with friends, family and employers, and on social media. Badges are a great motivation, helping you to reach the end of the course. Gaining a badge often boosts confidence in the skills and abilities that underpin successful study. So, completing this course could encourage you to think about taking other courses.

How to get a badge

Getting a badge is straightforward! Here's what you have to do:

- read each week of the course
- score 50% or more in the two badge quizzes in Week 4 and Week 8

For all the quizzes, you can have three attempts at most of the questions (for true or false type questions you usually only get one attempt). If you get the answer right first time you will get more marks than for a correct answer the second or third time. Therefore, please be aware that for the two badge quizzes it is possible to get all the questions right but not score 50% and be eligible for the badge on that attempt. If one of your answers is incorrect you will often receive helpful feedback and suggestions about how to work out the correct answer.

For the badge quizzes, if you're not successful in getting 50% the first time, after 24 hours you can attempt the whole quiz, and come back as many times as you like.

We hope that as many people as possible will gain an Open University badge – so you should see getting a badge as an opportunity to reflect on what you have learned rather than as a test.

If you need more guidance on getting a badge and what you can do with it, take a look at the OpenLearn FAQs. When you gain your badge you will receive an email to notify you and you will be able to view and manage all your badges in My OpenLearn within 24 hours of completing the criteria to gain a badge.

Get started with Week 1.

Week 1: Chance and chaos

Introduction

Mathematics has long been associated with certainty. For example, the equation 2 + 2 = 4 has been a synonym for evident truth since at least the middle of the 16th century, whereas the equation 2 + 2 = 5 has become a synonym for the opposite (most notably in George Orwell's dystopian novel *Nineteen Eighty-Four*). But, as you will see, there are parts of mathematics where certainty is a commodity in short supply.

The two mathematical ideas you'll explore this week are chance and chaos – both of which will be demonstrated through the rolling of dice.

Here's Marcus to introduce this week's topic.

Video content is not available in this format.

Video 1 Introduction

By the end of this week, you should be able to:

- understand the basic concept of probability
- explain the use of probability in the context of rolling dice
- understand the notion of mathematical chaos, and what is meant by sensitivity to initial conditions
- explain the problem of the stability of the solar system, and how it can be modelled mathematically

- appreciate the difficulties inherent in predicting chaotic systems, like weather forecasting.

The Open University would really appreciate a few minutes of your time to tell us about yourself and your expectations for the course before you begin, in our optional start-of-course survey. Participation will be completely confidential and we will not pass on your details to others.

1 Chance

Games of chance – especially those involving the rolling of dice or similar objects – have a long history, dating back centuries across many different cultures. However, it is only in more recent times that games of chance have been analysed mathematically, with some very interesting and powerful mathematics developing as a result, in the area now known as probability theory. The correspondence between two 17th century French mathematicians, Blaise Pascal and Pierre de Fermat, about a problem connected with a dice game is now considered a seminal moment in this development. The problem in question, known as the 'problem of points', concerns the fair division of stakes should a dice game be interrupted before a player has won.

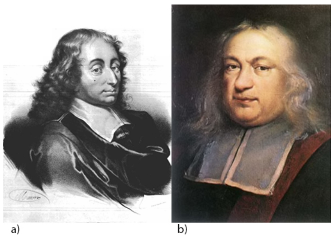


Figure 1 (a) Blaise Pascal (1623–1662), (b) Pierre de Fermat (1607–1665)

You will look at this problem shortly. First, some concepts in probability theory will be discussed. In what follows, it shall be assumed that a die is a perfect cube with each one of its six sides having a different number of spots on it, ranging from 1 to 6.

Figure 2 Some regular six-sided dice

1.1 Dice rolls and probability

Here's Marcus to discuss the idea of probability when rolling dice. Watch the video before calculating some probabilities yourself in Activity 1.

Video content is not available in this format.
Video 2 Probability

Keeping the example numbers of 5 and 12 in mind, let's refer to any total number as *n*. If we denote the event frequency – the number of ways of getting *n* – by F(*n*), then in the first case F(5) = 4, and in the second F(12) = 1. Since in each case the number of possible outcomes is 36 but the event frequency is different, the probability of getting a total of 5 cannot equal the probability of getting a total of 12.

Activity 1 Rolling two dice

Allow about 10 minutes

Complete the following table, where the notation (1, 2) denotes a throw of 1 followed by a throw of 2.

Table 1 Results when rolling two dice

Desired result (*n*)	Ways to obtain the result	Event frequency F(*n*)	Probability of obtaining desired result P(*n*)
1	*Provide your answer...*	*Provide your answer...*	*Provide your answer...*
2	*Provide your answer...*	*Provide your answer...*	*Provide your answer...*
3	*Provide your answer...*	*Provide your answer...*	*Provide your answer...*

4	Provide your answer...	Provide your answer...	Provide your answer...
5	(2, 3), (3, 2), (1, 4), (4, 1)	4	4/36 = 1/9
6	Provide your answer...	Provide your answer...	Provide your answer...
7	Provide your answer...	Provide your answer...	Provide your answer...
8	Provide your answer...	Provide your answer...	Provide your answer...
9	Provide your answer...	Provide your answer...	Provide your answer...
10	Provide your answer...	Provide your answer...	Provide your answer...
11	Provide your answer...	Provide your answer...	Provide your answer...
12	(6, 6)	1	1/36

Answer

Here's a completed version of the table. As a check, the sum of the 12 probabilities in the final column should add up to 1.

Table 1 Results when rolling two dice

Desired result (n)	Ways to obtain the result	Event frequency F(n)	Probability of obtaining desired result P(n)
1	None	0	0
2	(1, 1)	1	1/36
3	(1, 2), (2, 1)	2	2/36 = 1/18
4	(1, 3), (3, 1), (2, 2)	3	3/36 = 1/12
5	(2, 3), (3, 2), (4, 1), (1, 4)	4	4/36 = 1/9
6	(1, 5), (5, 1), (2, 4), (4, 2), (3, 3)	5	5/36
7	(1,6), (6, 1), (2, 5), (5, 2), (3, 4), (4, 3)	6	6/36 = 1/6

8	(2, 6), (6, 2), (3, 5), (5, 3), (4, 4)	5	5/36
9	(3, 6), (6, 3), (4, 5), (5, 4)	4	4/36 = 1/9
10	(4, 6), (6, 4), (5, 5)	3	3/36 = 1/12
11	(5, 6), (6, 5)	2	2/36 = 1/18
12	(6, 6)	1	1/36

As you can see from the completed table, if you roll two dice, the probability of achieving a total of 7 is greater than the probability of achieving any other total. How much more likely are you to achieve a total of 7 than a total of 10?

Answer

Since the probability of achieving a total of 7 is P(7) = 6/36 = 1/6, and the probability of achieving a total of 10 is P(10) = 3/36 = 1/12, and 1/6 equals twice 1/12, you are twice as likely to achieve a total of 7 than you are to achieve a total of 10.

Next, let's see what happens when three dice are rolled.

1.2 Adding more dice to the rolls

Since there are thirty-six possible outcomes for two dice, for each outcome of these two dice there will be six possible outcomes for the third die. This makes a total of 6 x 6 x 6 = 216 outcomes.

There will be three different types of outcome. Those in which:

- all three dice are the same
- two dice are the same and one is different
- all three dice are different.

These three types can be labelled as [a, a, a], [a, a, b] and [a, b, c] respectively, where a, b, c each represent a different number on the dice.

Activity 2 Three types of outcome

Allow about 10 minutes

In the first case [a, a, a], all three dice produce the same number, so there is only one way of rolling this combination. What about the other two cases? Can you work out how many ways you can roll the dice in each of these cases?

Discussion

In the second case [a, a, b], the dice can be rolled in three ways: (a, a, b), (a, b, a) and (b, a, a).

In the third case [a, b, c], the dice can be rolled in six ways: (a, b, c), (a, c, b), (b, a, c), (b, c, a), (c, a, b), (c, b, a).

Table 2 shows the probabilities of obtaining certain results when rolling three dice. For example, there are ten ways of achieving a total of six: three ways of rolling [1, 1, 4], six ways of rolling [1, 2, 3], and one way of rolling [2, 2, 2]. Thus P(6) = 10/216 = 5/108.

Can you calculate P(9) and P(13)?

Table 2 Results when rolling three dice

Desired result (*n*)	Ways to obtain the result using three dice	Event frequency F(*n*)	Probability of obtaining desired result P(*n*)
1	None	0	0
2	None	0	0
3	(1, 1, 1)	1	1/216
4	(1, 1, 2)	3	3/216 = 1/72
5	(1, 1, 3), (1, 1, 2)	3 + 3 = 6	6/216 = 1/36
6	(1, 1, 4), (1, 2, 3), (2, 2, 2)	3 + 6 + 1 = 10	10/216 = 5/108
7	(1, 1, 5), (1, 2, 4), (1, 3, 3), (2, 2, 3)	3 + 6 + 3 + 3 = 15	15/216 = 5/72
8	(1, 1, 6), (1, 2, 5), (1, 3, 4), (2, 2, 4), (2, 3, 3)	3 + 6 + 6 + 3 + 3 = 21	21/216 = 7/72
9	Provide your answer...	Provide your answer...	Provide your answer...
10	(1, 3, 6), (1, 4, 5), (2, 3, 5), (2, 2, 6), (3, 3, 4), (4, 4, 2)	6 + 6 + 6 + 3 + 3 + 3 = 27	27/216 = 1/8
11	(1, 4, 6), (2, 3, 6), (2, 4, 5), (1, 5, 5), (3, 4, 4), (5, 3, 3)	6 + 6 + 6 + 3 + 3 + 3 = 27	27/216 = 1/8
12	(1, 5, 6), (2, 4, 6), (2, 5, 5), (3, 3, 6), (3, 4, 5), (4, 4, 4)	6 + 6 + 3 + 3 + 6 + 1 = 25	25/216
13	Provide your answer...	Provide your answer...	Provide your answer...
14	(2, 6, 6), (3, 5, 6), (4, 4, 6), (4, 5, 5)	3 + 6 + 3 + 3 = 15	15/216 = 5/72
15	(3, 6, 6), (4, 5, 6), (5, 5, 5)	3 + 6 + 1 = 10	10/216 = 5/108
16	(5, 5, 6), (6, 6, 4)	3 + 3 = 6	6/216 = 1/36

17	(5, 6, 6)	3	3/216 = 1/72
18	(6, 6, 6)	1	1/216

Answer

Table 3 P(9) and P(13) results

Desired result (*n*)	Ways to obtain the result using three dice	Event frequency F(*n*)	Probability of obtaining desired result P(*n*)
9	(1, 2, 6), (1, 3, 5), (1, 4, 4), (2, 2, 5), (2, 3, 4), (3, 3, 3)	6 + 6 + 3 + 3 + 6 + 1 = 25	25/216
13	(1, 6, 6), (2, 5, 6), (3, 4, 6), (3, 5, 5), (4, 4, 5)	3 + 6 + 6 + 3 + 3 = 21	21/216 = 7/22

Probability calculations on the roll of three dice date back centuries, with the earliest known appearance being a Latin poem titled ‘De Vetula’, written during the 13th century. It is thought that the reason many copies of this poem survived is because it provided medieval gamblers with a certain way to make money, even though they almost certainly did not understand it!

1.3 Problem of points

You are now going to look at the ‘problem of points’, which was mentioned earlier. Specifically, the problem involves a game of chance with two players. Both players, each of whom have an equal chance of winning, have put up the same stake, and the first player who collects a certain number of points wins the game and collects the entire stake. Suppose, however, that the game is interrupted and cannot continue before either player has won. How then should the winnings be divided? Clearly the player who has already accumulated the greater number of points should receive a greater share of the winnings, but it is not so clear what that share should be. How can this be resolved?

Let’s look at a specific example. Suppose that the two players, Alex and Bill, are rolling a single die. Alex scores a point if the die shows a 4, 5 or 6, and Bill scores a point otherwise. In other words, on each roll of the die they each have an equal chance of scoring a point. The total stakes are £60, which will go to the person who first scores a total of 3 points. If the game is interrupted when Alex has scored 2 points and Bill 1 point, how should they divide the £60?

There are two ways of looking at this. Either you can focus on what has happened in the past or you can focus on what may happen in the future. Taking the first approach here, Alex would walk away with all the winnings and Bill would get nothing, because the only knowledge you have is that Alex is closer to a winning score of 3 points than Bill is.

Question 1

There is a problem with this solution. Can you see what it is?

Provide your answer...

Answer

No value has been attached to the fact that Bill was still in the game when it was interrupted. He still had a chance of scoring 3 points before Alex, and this chance is certainly worth something.

What happens if you focus on the future instead? What possible outcomes can there be? Either Alex wins the next round and wins the game, or Bill wins the next round, and they go to a final round which either one of them can win. There are therefore three possible outcomes, two of which end with Alex as the winner. You might think that this approach means that Alex should get 2/3 of the stake (£40) but it turns out that this too is wrong!

It was Blaise Pascal who worked out the correct solution, which goes as follows. Alex has a 50% chance of winning in one further round, in which case she would get the entire stake of £60. But if Bill wins that round, then, since they are each equally likely to win the final round, they could at this stage (before they play the final round) divide the stake equally between them so that they would get £30 each. In either case, Alex is guaranteed £30. So, the other £30 should be split equally between them, giving Alex a total of £45 and Bill a total of £15. This can be pictured in the following way.

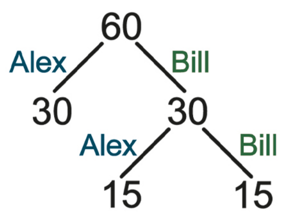


Figure 3 Solving the problem of points

1.4 Pascal's triangle

Remarkably, Pascal further discovered that there is a simple way of working out the division of the stake in more complex situations, e.g. if one player needs 3 points for a win and the other needs 2. The key turned out to be concealed within what today is known as 'Pascal's triangle'.

1
1 1
1 2 1
1 3 3 1
1 4 6 4 1
1 5 10 10 5 1
1 6 15 20 15 6 1
1 7 21 35 35 21 7 1

Figure 4 Pascal's triangle

The triangle is constructed starting with 1, and then numbers are placed beneath it in a triangular fashion, with each number in the triangle being the sum of the two numbers immediately above it. For example: 2 is the sum of 1 + 1; 6 is the sum of 3 + 3; 15 is the sum of 10 + 5; and so it goes on indefinitely.

In the case of an interrupted game where one player needs 2 points to win and the other needs 3, you add the 2 and the 3 together to make 5. Then you look in the 5th row of the triangle, add together the first three numbers (1 + 4 + 6 = 11) and the last two (4 + 1 = 5), and divide the stake according to this proportion. So the first player will get 11/16 of the stake and the second player will get 5/16 of the stake.

In general, if player A needs m points and player B needs n points, you consult the $(m + n)$ th row, add together the first n numbers in the row then add together the remaining m numbers in the row and divide the stake proportionally: player A gets $n/(m + n)$ of the stake and the player B gets $m/(m + n)$ of the stake.

Question 2

Use Pascal's triangle to double check the solution to the game between Alex and Bill.

Provide your answer...

Answer

Alex needs 1 point; Bill needs 2 points. Look in the third row. Add the first two numbers together, 1 + 2 = 3. There is only one last number, 1. Thus Alex gets 3/4 of the stake and Bill gets 1/4 of the stake. Since the stake is £60, they get £45 and £15 respectively as before.

Pascal's triangle was named for Blaise Pascal because of his discoveries about the triangle which he compiled together in a book, *Traité du triangle arithmétique* (trans. 'Treatise on the Arithmetical Triangle'), published in 1665, three years after his death. But Pascal was not the first to study the triangle. It had been studied much earlier by mathematicians elsewhere – notably by Islamic mathematicians in the 10th century and Chinese mathematicians in the early 11th century – in connection with solving certain types of equation. In fact, Pascal's triangle is a particularly useful mathematical tool and turns up in several areas of mathematics.

A common use of the triangle is to find what are known as the 'combinatorial numbers'. If we have n different objects and want to choose k of them, without repetition, and the order of choice doesn't matter, then the combinatorial number, which is written $\binom{n}{k}$, is the number of ways in which that choice can be made. In Pascal's triangle, using the convention that the top row of the triangle is row 0, and the first number in each row is column 0, then the number in row *n* and the number in column *k* is the number $\binom{n}{k}$.

For example, if you're shown 5 different plants but can choose only 3 of them, how many different combinations of plants are there for you to choose from? Or in mathematical terms, what is the value of $\binom{5}{3}$?

You could try writing out all the different combinations, where the plants are labelled *a*, *b*, *c*, *d*, *e*, then the combinations are *abc*, *abd*, *abe*, *acd*, etc., but it is much quicker to use Pascal's triangle. You just look at the entry in the sixth row and the fourth column in the triangle (remember to start counting from zero), and you will see that the answer is 10.

1.5 Close to certainty

It's not possible to predict the outcome of a single roll of a pair of dice. But what if you keep on rolling the dice? In this case, probability can be used to determine how many of each outcome can be expected in the long run.

For example, imagine rolling a pair of dice 10 000 times. Using the probability information from Table 1, you would expect a total of 6 (or a total of 8) to come up approximately 5/36 x 10 000 = 1389 times, or 13.9% of the time, and you would expect a double six to come up approximately one-fifth of that, or 2.78% of the time.

Similarly, while the outcome of a single toss of a coin cannot be predicted, in the long run you would expect it to come up heads and tails in almost equal measure. In either case, if your expectations fail to be fulfilled by a significant margin, you would be justified in questioning the fairness of the dice or the coin.

Activity 3 Rolling many dice

Allow about 10 minutes

To observe what happens to the frequencies when two or more dice are rolled together a large number of times, try using an online dice roller (this is much quicker and easier than using real dice and noting results!)

The site linked below lets you choose the number of dice to roll and provides a bar chart which shows the frequency of the dice roll results. Use this site to roll two and then three dice at least 2000 times each, and watch the bar chart evolve. What do you see?

Dice roll simulator (make sure to open the link in a new tab/window)

Discussion

As the number of rolls increases, the shape of the bar chart becomes more symmetrical about the most common rolls (the rolls of greatest frequency). For two dice, this is 7, and for three dice, this is 10 and 11 (as demonstrated in Tables 1 and 2 earlier). As you increase the number of dice rolled – try six – you will see that the shape of the bar chart becomes increasingly bell-shaped. It approaches what is known as a bell curve (more formally known as 'normal distribution'), an important notion in statistics and one which commonly arises in the study of data, especially in nature.

2 Chaos

When rolling a die, you don't know in advance how the die is going to land. You simply roll the die and hope for the best. As you've seen though, when you roll two or more dice, some outcomes are more likely than others.

But what if we could predict the way the die is going to land? It would cease to be a game of chance at all. For this to be possible, you would need to know everything about the die and the environment in which it is going to be rolled, as well as having the scientific theory and mathematics to predict its motion.

As it turns out, we do at least have the mathematics. This 'differential calculus' – which is concerned with the rate at which quantities such as position and speed change – has been successfully applied to many problems involving motion for several centuries. Such problems vary from calculating the rate of spread of infectious diseases to predicting the positions of the planets, and even designing video games. It was invented independently by both Isaac Newton and Gottfried Leibniz at the end of the 17th century.

Figure 5 (a) Isaac Newton (1643–1727), (b) Gottfried Leibniz (1646–1716)

Of course, each problem has its own particular complexities: the motion of a ball rolling down a smooth slope is quite different to the motion of a ball bouncing on rough ground. And it is one thing to know something in theory, and quite another to apply it in practice. Even getting to grips with the theory can take you in unexpected directions. One of the wonderful things about mathematics is that you can start out asking a question in one area of mathematics, and in the process of looking for a solution you end up somewhere else!

In this case, it turns out that trying to predict the motion of the dice will lead us from the 17th century mathematics of Newton and Leibniz to the 20th century mathematics of chaos theory.

2.1 Predicting the dice roll

If you're going to predict how a rolled die will land, you need data. There are numerous factors here related to the die itself, and to the environment in which it is rolled.

Question 3

Can you think of some factors which will need to be considered?

Provide your answer...

Answer

Factors involving the die:

- material properties of the die
- size
- launch velocity
- angle of launch
- speed and direction of spin
- vertical distance to landing surface

Factors involving the environment:

- material properties of the landing surface(e.g., its friction and bounciness)
- air resistance.

Here's Marcus to discuss a few of the key considerations.

Video content is not available in this format.
Video 3 Predicting a dice roll

Even with the data and the mathematical machinery in hand, there is still something missing – the scientific or physical theory that ties it together.

This comes in the form of Newton's three famous laws of motion which relate an object's motion to the forces acting upon it. Newton explained these laws in his *Principia Mathematica*, first published in 1687. (This is one of the most important scientific books ever to be published, and it nearly didn't make it to the press. Newton was always extremely reluctant to publish his work, and it was published only through the intervention of Edmund Halley, who personally financed its publication after the Royal Society had spent its book budget on a *History of Fishes*!)

Applying Newton's laws to physical systems (like the rolling of a die or the motion of a planet) can be written as mathematical equations. The differential calculus mentioned earlier is the key to solving these equations, and giving us the position of the object in motion at any desired time. However, as you shall see, solving these equations can turn out to be an extremely difficult problem. Sometimes it can even be impossible, or at least impossible to the necessary degree of accuracy.

Newton’s laws can nevertheless be used to make some remarkable deductions. Newton himself applied his mathematics to the solar system, calculating the relative masses of the large planets, the motion of the moon, and much more. He was even able to deduce the shape of the earth as a ‘squashed’ sphere (somewhat like a grapefruit) rather than perfectly spherical, as had been previously believed. Experiments carried out around the globe after Newton’s death proved his deduction to be correct.

2.2 Laplace’s demon

As they stood, Newton’s laws were suitable only for objects that could be treated as ‘point particles’ (meaning that their mass is concentrated in one spot, known as the centre of mass) and wider applicability was sought. This was achieved by the next generation of mathematicians, most notably by the great Swiss mathematician Leonhard Euler, who provided more generalised equations which could be applied in more complex settings where objects are not necessarily rigid.

Such was the success of this next generation in applying Newton’s laws of motion, that it seemed mathematics could be harnessed to describe the motion of everything in the universe: past, present and future. This view, that the universe is completely knowable, is known as determinism. It was summed up by the French mathematician Pierre-Simon Laplace in 1814:

> We may regard the present state of the universe as the effect of its past and the cause of its future. An intellect which at a certain moment would know all the forces that set nature in motion, and all positions of all items of which nature is composed, if this intellect were also vast enough to submit these data to analysis, it would embrace in a single formula the movement of the greatest bodies of the universe and those of the tiniest atom; for such an intellect nothing would be uncertain and the future just like the past would be present before its eyes.
>
> (Laplace, 1814, p.4)

Laplace’s all-comprehending intelligence is now famously known as ‘Laplace’s demon’. It is called a ‘demon’ because it is supposed to be a secular entity and not a divine intelligence. Although it is not known who first used and popularised the term, the notion became well known in the 19th century, feeding into the belief that ultimately no motion could defy prediction.

However, as you are about to see, by the end of the 19th century this belief was to be completely shattered. This will have serious repercussions for the quest to predict the outcome of rolling a die. But first you are going to look at a famous problem that particularly interested Laplace.

2.3 Is the solar system stable?

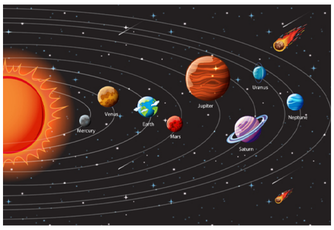


Figure 6 The solar system

As well as working on probability, Laplace worked extensively on celestial mechanics – the branch of astronomy which deals with the motion of bodies in space – and was the author of a monumental five-volume work, *Mécanique Céleste*, published between 1799 and 1825, which was the intellectual successor to Newton's *Principia*. The first two volumes of it were expanded and translated into English as the *Mechanism of the Heavens* (1831) by the Scottish mathematician, Mary Somerville. This translation was highly acclaimed and established Somerville's reputation as a mathematician at a time when very few women had the opportunity to study mathematics.

Figure 7 Mary Somerville (1780–1872)

The problem that particularly interested Laplace relates to the question of whether the solar system is stable. Will the planets continue to follow similar paths to the ones they've travelled along in the past? Could something catastrophic occur, such as a collision or an escape?

The stability of the solar system can be considered a mathematical problem because it can be modelled by what is called the *n*-body problem: given *n* objects or 'bodies' interacting gravitationally with known initial positions and velocities, can you predict their individual motions? Simply put, can you determine their positions and velocities at any time of your choosing? (The way the problem is modelled, the only forces acting on the bodies are the forces of gravity. All other forces, such as those generated by solar winds, are ignored.)

The *n*-body problem, which originated with Newton's law of gravity, is a notoriously hard problem and was attempted by many distinguished mathematicians. Little wonder then, that it was set as a prize competition problem in 1885. The occasion for the competition was the 60th birthday of King Oscar II of Sweden and Norway, who had himself studied mathematics at university. The King's birthday was due to take place in 1889, when the winner would be announced.

The competition was won by one of the most talented mathematicians of the day, the French mathematician Henri Poincaré – even though he hadn't actually solved the problem! Which begs the question: what did Poincaré do that was so good that he won the competition anyway?

Figure 8 Henri Poincaré (1854–1912)

He began by doing what mathematicians often do when they are struggling to solve a problem: he attacked a simpler version of the problem in the hope that if he found a solution, he would be able to generalise it. The version he initially attacked was the 'three-body problem'. But even this version of the problem turned out to be too difficult – there are more variables than there are equations to describe them – and so Poincaré turned to an even simpler version: the 'restricted three-body problem'. (The two-body problem, by the way, had been solved by Newton.)

2.4 Poincaré's error

In the restricted three-body problem, two of the bodies move as a two-body problem and so their motion is known. The third body (often called the planetoid) is considered massless relative to the other two larger bodies – you can think of it as a speck of dust. The motion of the planetoid does not affect the motion of the two larger bodies, but its motion is affected by them. The problem is then to ascertain the orbit of the small body.

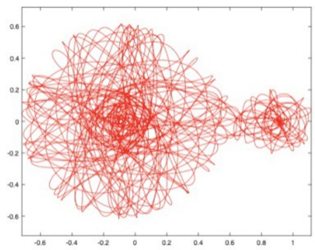


Figure 9 The trajectory of a third body interacting with a large mass (Earth, left) and a small mass (Moon, right)

This might seem a very artificial problem, but in fact it provides a good model for the Sun-Earth-Moon system (in which the Moon is the planetoid). But Poincaré was unable to solve even this version of the problem completely. Nevertheless, he won the prize because he developed a lot of new and important mathematical techniques in his efforts to do so. But his path to glory was not smooth. In fact, it turned out that the first paper he submitted to the competition – the paper for which he won the prize – had a serious error. This was discovered while the paper was being prepared for publication. Fortunately, Poincaré was able to correct the error before the prize ceremony. Unfortunately, he did have to pay for the reprinting of his memoir, which cost him more than he won in prize money!

But Poincaré's error had further-reaching consequences than he could possibly have imagined when he corrected the paper. Originally, Poincaré had assumed (without justification) that the orbit of the small body was stable. What does this mean? Suppose that, given the small body's initial position and velocity, its orbit can be predicted at any time in the future (or in the past). Poincaré's assumption from here was that given very slightly different initial conditions – tiny changes to the small body's position and velocity – its orbit will remain close to before.

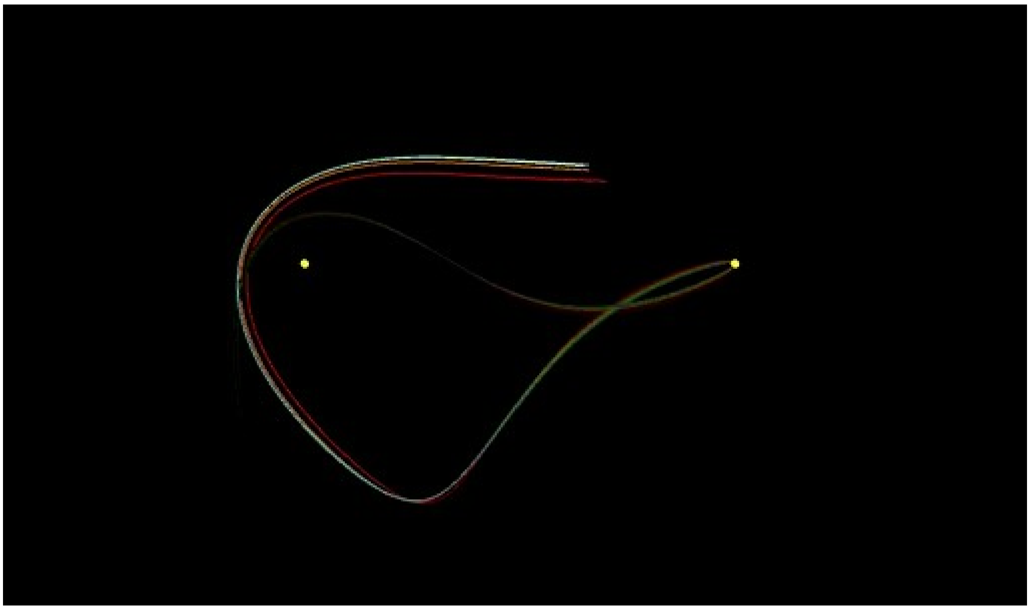
But, as Poincaré himself discovered, this assumption was a mistake. The behaviour of the small body, although governed by deterministic laws, was generally unpredictable, and as Poincaré himself said:

> It may happen that small differences in the initial conditions produce very great ones in the final phenomena. A small error in the former will produce an enormous error in the latter. Prediction becomes impossible.
>
> (Poincaré, 1908, p.68)

In the following video, which simulates the restricted three-body problem, the two large bodies are moving in a rotating reference frame so that they both appear stationary. The small body is in fact five small bodies starting extremely close together. In Poincaré's words, these five bodies exhibit 'very small differences in the [five sets of] initial conditions'. After three minutes the five bodies diverge, demonstrating the system's sensitive dependence on initial conditions.

Video content is not available in this format.

Video 4 The restricted three-body problem (note: there is no spoken audio in this video)



2.5 Chaos theory

In his revised and published paper, Poincaré corrected his error, explaining why the small body's behaviour is in general unpredictable. His insight laid the foundations for the mathematical concept today known as chaos theory, which became established in the second half of the twentieth century.

For most systems, the equations resulting from Newton's laws cannot be solved exactly to obtain the position and velocity of all parts at any given time. In these cases, a computer program must be used to 'solve' the equations. But these computer programs generate tiny rounding errors at each step of the calculation. This combines with Poincaré's observation that such systems are extremely sensitive with respect to initial conditions. Ultimately, even if we knew the initial conditions exactly, our ability to predict the future behaviour of these systems – which we now call 'chaotic systems' – is strictly limited.

Chaos is ubiquitous, both in nature and in human behaviour: from the dripping of a tap to the currents of the oceans, from the beating of the human heart to the function of the brain, from work motivation to trading in financial markets. This is because the mathematical equations which describe physical systems, and the difficulties in solving them, have analogues in other disciplines. A particularly well-studied example is the atmosphere, which you'll take a look at next.

2.6 Weather forecasting

It's no secret that weather forecasting can be less reliable than we would like – particularly when looking beyond a week ahead. Although this field has seen huge improvements in recent decades, through advances in computing and the use of satellites, there are still limits to what can be achieved. Why is this? The answer lies in the data. Specifically, the quantity of data.

In order to make a forecast, we need to know the weather conditions now. Many thousands of observations are recorded around the clock in weather stations worldwide. These data are fed into high-speed computers as the initial conditions for a complex set of equations which model the weather and provide predictions of the weather as time evolves. The problem here though, is that a tiny change in any data – and there's a huge amount of data involved – can lead to a very large change in the output after a relatively short space of time. This means that the model can work well in the short term, but runs into serious difficulties as more time passes. Furthermore, like with the equations resulting from Newton's laws, the computer programs make tiny errors with each step of the calculation. As a result, even if perfectly accurate data for today's weather were available, it would not be possible to predict its long-term changes with accuracy.

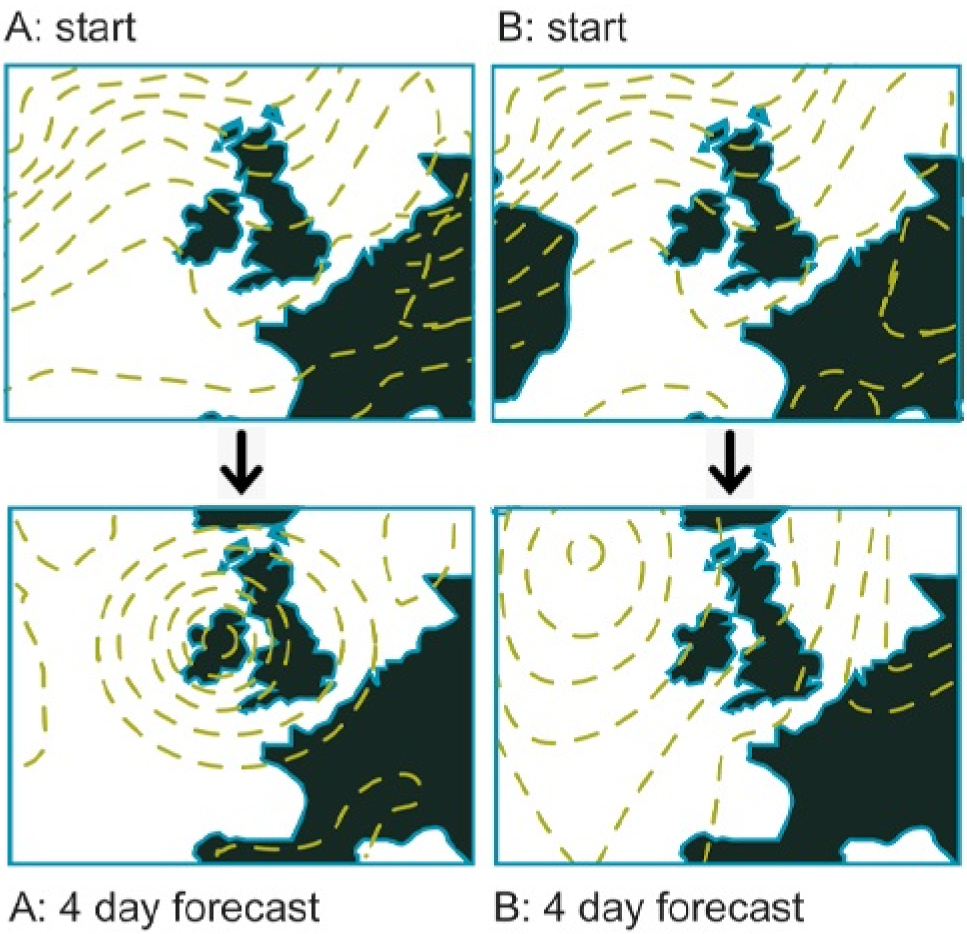


Figure 10 Divergent results in weather forecasting

You've probably heard of 'the butterfly effect'. This is a popular metaphor for sensitive dependence on initial conditions, which actually originated in the context of weather forecasting: it was the title of a talk given by the American mathematician and meteorologist Edward Lorenz. The idea is that a butterfly flapping its wings over the Amazon might set off a tornado in Texas several weeks later. It is not the single act of the butterfly flapping its wings that causes the tornado, but that it might set in chain a sequence of events which eventually ends up with a tornado. The point is that in a complex system, it is almost impossible to know what results a small event might ultimately cause.

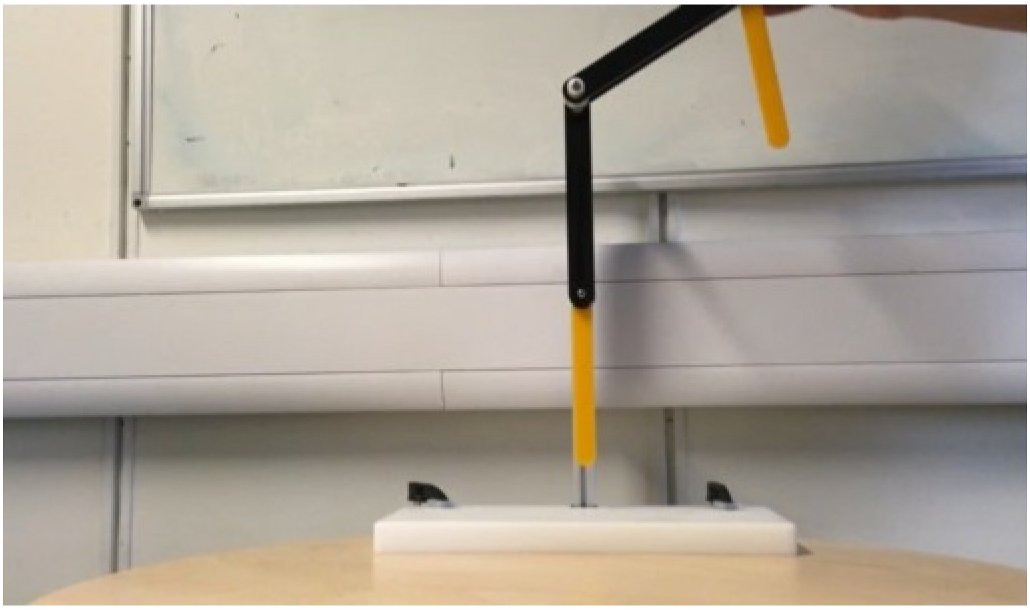
2.7 Double pendula

Another example of chaos is the motion of a double pendulum. This is a pendulum which is suspended from a fixed point, with a second pendulum attached to the end of the first pendulum. In contrast to the regular oscillations of a single pendulum (such as a pendulum on a longcase clock), the oscillations of a double pendulum can rapidly become unpredictable and chaotic, and the motion can exhibit extreme sensitivity to initial conditions.

You can see examples of this behaviour in these two video clips. In the first one, in which a single double pendulum is released from above its resting position, the behaviour of the pendulum soon becomes unpredictable.

Video content is not available in this format.

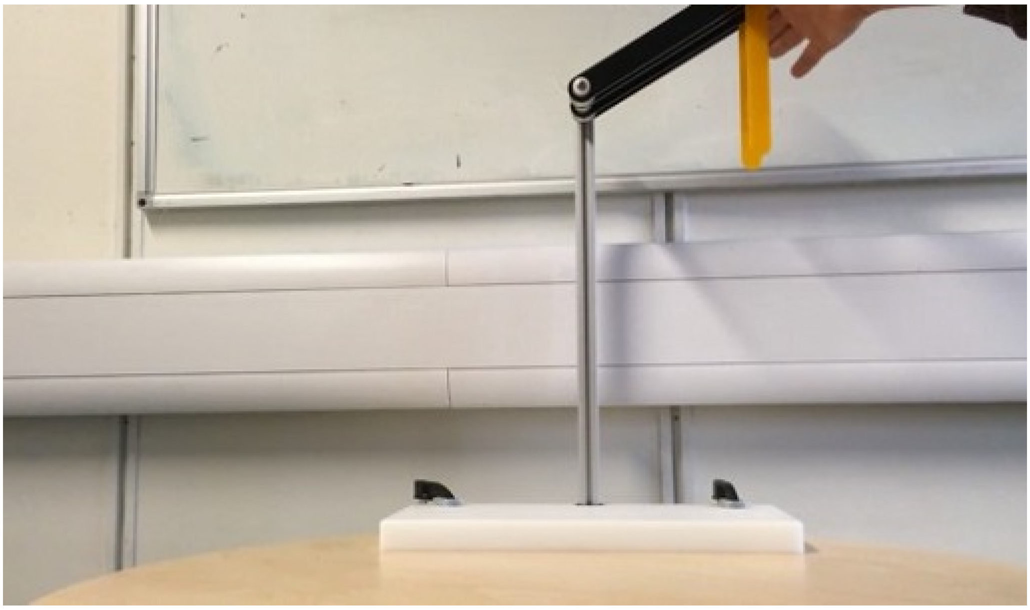
Video 5 Single double pendulum (note: there is no spoken audio in this video)



In the second one, in which two identical double pendula are released from virtually the same starting position, the two pendula begin to behave differently almost immediately.

Video content is not available in this format.

Video 6 Double pendula (note: there is no spoken audio in this video)



2.8 One more roll of the dice

Returning to the start of this week's topic: the 'perfect' die – is its roll chaotic? This question was explored in 2012 by a team of mathematicians in Poland, who used a high-speed camera to capture the trajectory of a die roll at a rate of 1500 frames a second. They found that, in addition to knowing the initial position and velocity of the die, the most important factors are the friction and bounciness of the landing surface. Air resistance on the other hand can be disregarded.

Energy is dissipated each time the die bounces. The number of bounces before the die comes to rest depends on the die's interaction with the surface. On a high-friction rigid surface, the amount of energy dissipated on impact with the surface is low, and the die will tend to bounce around more. On a smooth, low-friction, or soft, surface, the amount of energy dissipated on impact with the surface is higher, resulting in fewer bounces. When a die is rolled on a surface like this, if the initial conditions can be established with sufficient accuracy, then it's theoretically possible for the outcome of the throw to be predicted. But with more bounces, the outcome becomes more chaotic, and harder to predict.

3 This week’s quiz

Well done for reaching the end of Week 1. Check what you’ve learned by taking the end-of-week quiz.

Week 1 practice quiz

Open the quiz in a new window or tab then come back here when you’ve finished.

4 Summary of Week 1

What are the implications of this week's study of chance and chaos? Can we predict dice rolls and beat the casino? Unfortunately, probably not! While it's perhaps achievable in theory, the initial position of the die would need to be known with a degree of precision impossible to realise in practice.

Finally, returning to the question of the stability of the solar system. Does Poincaré's discovery mean that stability is under threat? Luckily for us this isn't the case. Or at least, not for several million years. Although the solar system is chaotic in a mathematical sense, recent numerical integration of the relevant equations over a period of several billion years have shown that, although it is impossible to predict that stability will persist in perpetuity, the solar system is stable on a human timescale.

You can now move on to Week 2.

Week 2: Particle physics

Introduction

What are the fundamental ‘building blocks’ from which our universe is made? This week aims to explore whether that’s a question that can even be answered, beginning with the history of scientific discoveries relating to matter. You’ll learn about the smallest constituents of matter we’ve found, and the forces that act between them. But before you get there, you’ll take a close look at something more familiar: a triangle.

Here’s Marcus to introduce this week’s topic.

Video content is not available in this format.

Video 1 Introduction

By the end of this week, you should be able to:

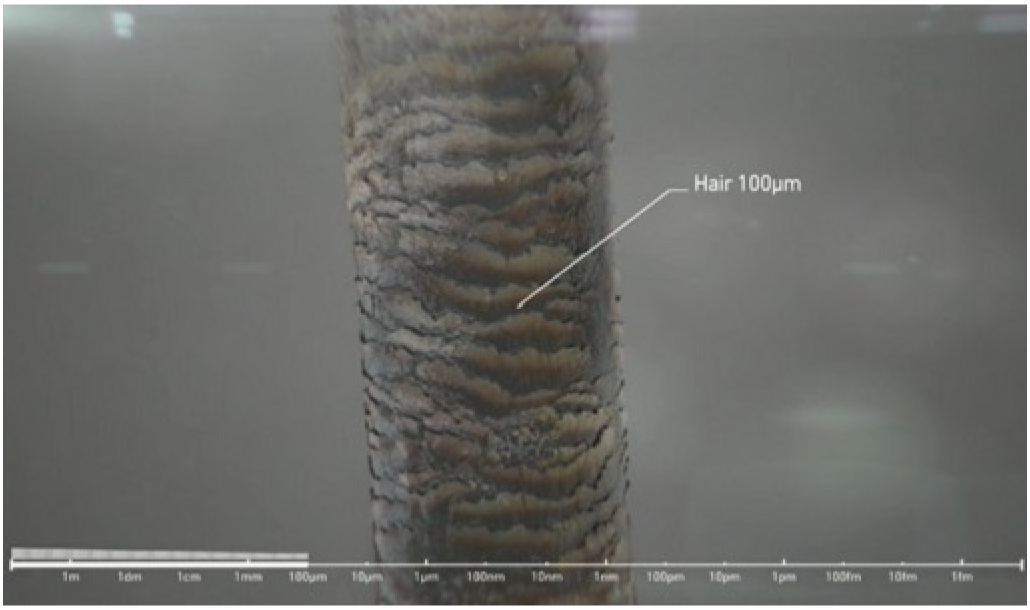
- outline the elementary constituents of matter
- understand the notion of symmetry
- combine simple rotation and reflection symmetries
- appreciate some basic properties of the quark model.

1 What is matter?

What is matter made of? This question has been puzzling humans throughout history, and it's been a significant driver of scientific discovery. Around 2500 years ago, Greek philosophers speculated that matter consists of atoms (small indivisible objects – from the Greek word 'atomos' meaning uncuttable). Direct evidence for the existence of atoms was only obtained much later, with innovations in physics in the early 20th century finally providing us with a detailed understanding of the nature of atoms.

Video content is not available in this format.

Video 2 Voyage into the world of atoms (note: there is no spoken audio in this video)



However, this is not the end of the story. It was quickly realised that atoms themselves are comprised of two parts: a small atomic nucleus, which carries a positive electric charge and most of the mass of the atom, surrounded by a cloud of negatively charged electrons. The nucleus itself is then comprised of two types of particles: positively charged protons, and neutrons which carry no electric charge.

These discoveries completely changed our picture of what constitutes matter. They provide us with explanations for the regularities observed in the periodic table of elements, as well as for properties such as radioactive decay and chemical bonding.

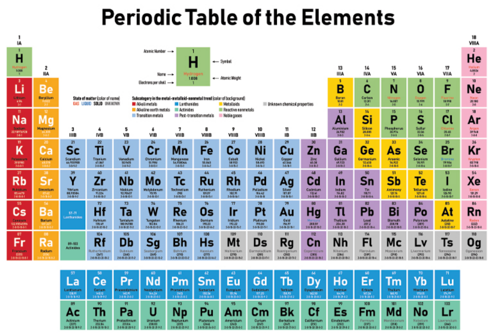


Figure 1 Periodic table of elements. Elements are lined up by atomic number (the number of protons they have)

This might have been the end of the quest for the smallest constituents of matter. However, physicists kept discovering ‘new’ particles – which were not constituents of atoms – in cosmic rays and high-energy experiments. With each new particle, significant questions arose about how these discoveries fit together, and how many more were yet to be unearthed.

Physicists found themselves in a situation similar to one a century before, with the periodic table of elements. Those patterns were eventually explained by the structure of atoms. Here too, researchers observed striking patterns in the properties of newly discovered particles, which pointed to the presence of an underlying pattern. By assuming that these patterns would continue into unexplored areas, they predicted the existence of particles with certain properties. Their predictions turned out to be correct!

The pattern suggested that hadrons – meaning protons, neutrons and some similar particles – are in fact composite particles themselves, with a substructure consisting of quarks. These quarks are particles with intriguing properties that are very different from anything seen before. For instance, it is not possible to observe any single quark on its own. So how can we know that quarks exist, even though nobody has ever ‘seen’ one? The quark substructure of hadrons is inferred indirectly from analysing data from particle accelerator experiments.

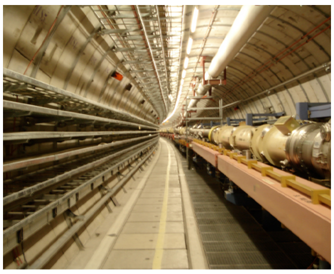


Figure 2 Particle accelerator

Given the story so far, you may now be wondering: have we found the elementary constituents of matter here? Or will quarks be discovered to consist of even smaller particles? Can we ever possibly know whether we’ve reached the end of this quest?

Before delving deeper into the physics of elementary particles, let’s consider symmetry, which is a key plank in our current understanding of the universe’s elementary particles and their interactions.

2 Symmetry

Symmetry is encountered often in nature, whether that's in the patterns of a beautiful flower, the structure of a crystal or the arrangement of cells in a honeycomb.

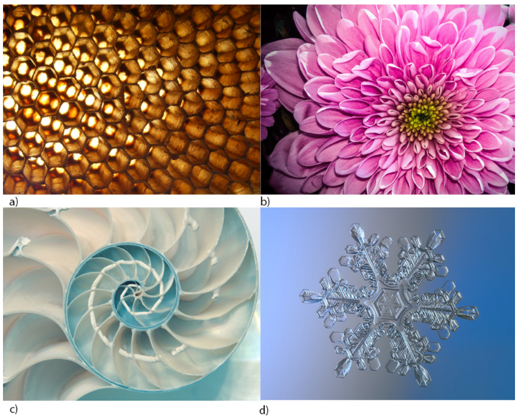


Figure 3 Patterns in nature: (a) apis florea nest, (b) chrysanthemum flower, (c) nautilus shell, (d) snowflake

In fact, you might argue that science is essentially about detecting, describing and explaining the patterns and symmetries found in nature, beginning with sequences like the rising and setting of the sun and moon and the rotation of seasons, and developing into what modern science can tell us about the world today.

2.1 Rotations of a triangle

Let us discuss the notion of symmetry through the example of a triangle. This one is very regular, in that all its sides have equal length and they all meet at equal angles of 60 degrees.

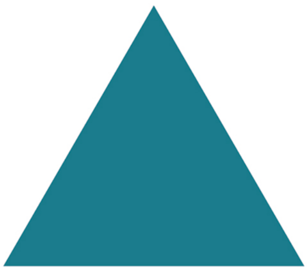


Figure 4 Equilateral triangle

This is called an equilateral triangle. Because it is so regular, this triangle possesses symmetries.

Question 1

Can you see any symmetries upon immediate observation?

Provide your answer...

Answer

You may observe lines of reflection from a corner to the midpoint of the opposite edge. You may also note that if you rotate the triangle, it doesn't take a full rotation to come back to the same position.

You will now explore the symmetries of this triangle more closely. Start by preparing a triangle to play around with.

Activity 1 Creating your triangle

Allow about 15 minutes

Take a piece of paper and cut out an equilateral triangle. Label the lower left corner C1, the lower right corner C2 and the top corner C3. Repeat these labels on the back side. Label the front side of the paper with the letter A and the back with the letter B.

You can find an example here. (If you print this out, make sure it's double-sided and properly lined up for cutting out.)

These downloads are unavailable in this format, please refer to the online version.

Now, take another piece of paper. Draw an equilateral triangle the same size as before, and label the corners P1, P2 and P3, again starting from the lower left corner and proceeding anticlockwise. This time though, write the labels outside the triangle. No need to cut this one out.

You can find an example of this second triangle here.

These downloads are unavailable in this format, please refer to the online version.

Place the cut-out triangle on top of the drawn triangle, with the top side (A) facing up and the corners matching the ones underneath, so that C1 is at position P1, C2 at position P2 and C3 at position P3. This will be the starting position for some experimentation.

Now, pick up the triangle and turn it around anticlockwise, until it is occupying the same space again, with all edges and corners meeting an edge and corner on the paper underneath. How far do you have to rotate the triangle for this to happen?

Answer

You have to rotate it 1/3 of the way round, which is 120 degrees (a full rotation being 360 degrees). Once you have turned it by this amount, the three corners have moved around, and the triangle looks the same as before. This is what geometry calls a symmetry: a transformation that, when performed on an object, leaves the object unchanged.

The corners were labelled to help you describe precisely what happens when the triangle is rotated. Compare the labels in the corners of your triangle with those on the underlying paper. If you labelled everything as described, you should find that corner C1 is now at position P2, corner C2 has moved to position P3, and corner C3 has moved to position P1. This can be written symbolically as $1 \to 2$, $2 \to 3$ and $3 \to 1$, or also as a cycle. This gives us a nice shorthand notation for the effect of the rotation.

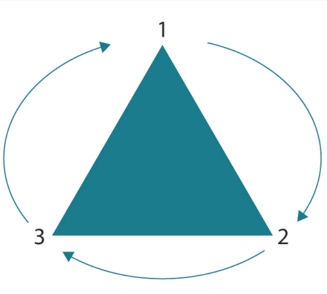


Figure 5 Cycle of the triangle

Based on the results so far, what will happen when you rotate your triangle anticlockwise by a further 120 degrees?

Answer

Following corner C1 as it moved to P2 in the first rotation, it now moves to P3. This result can be written as $1 \to 3$. Note that this corresponds to two steps on the cycle diagram above. Check that this also happens for the other two corners, which means that $2 \to 1$ and $3 \to 2$.

Rotating the triangle a further 120 degrees – a total of 360 degrees, considering all three rotations – puts all corners back in their original positions.

This triangle can be said to have a ‘threefold rotational symmetry’ – there are three different orientations of the triangle where the object looks the same, corresponding to rotations of 0, 120 and 240 degrees.

What happens then, if you rotate by 120 degrees in the clockwise direction? How could you write that in terms of the cycle diagram?

Answer

This rotation is $1 \to 3$, $2 \to 1$, $3 \to 2$.

There are two ways to see this. You either go backwards on the cycle diagram, reversing all the arrows (corresponding to rotating in the opposite direction), or you consider the fact that rotating clockwise by 120 degrees is the same as rotating anticlockwise by 240 degrees, corresponding to two steps on the cycle diagram.

Mathematicians call the structure formed by these rotations a ‘group’, which is a collection of transformations with the following properties:

- You can combine any two transformations to obtain another transformation from the same group.
- For any transformation, there is a transformation in the group that reverses it.

Groups provide the fundamental mathematical structure for the description of symmetries. The strength of mathematical group theory lies in its application to any system with transformations that demonstrate these properties.

Question 2

If the equilateral triangle is rotated anticlockwise by 120 degrees, which transformation reverses this?

Provide your answer...

Answer

A clockwise rotation of 120 degrees. (Of course, the same result can be obtained with a further anticlockwise rotation of 240 degrees.)

So, have the symmetries of this triangle been fully explored? Not quite. You'll look next at reflections.

2.2 Reflections of a triangle

The equilateral triangle can be reflected and yet remain unchanged – can you see where the reflection line must be drawn for this? The triangle must be cut into two equal halves, with the line running from one corner to the midpoint of the opposite edge. In fact, you'll notice three such lines can be drawn.

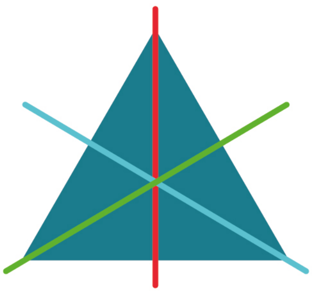


Figure 6 Lines of reflection

Let's focus on the reflection line that cuts through the corner in position P3. What effect will this have? Use your triangles from Activity 1 to test it out.

Activity 2 Reflecting your triangle

Allow about 10 minutes

Put the triangle back into the original position from Activity 1. Now turn it over, keeping C3 at position P3. How have the labels moved?

Answer

In this case, the top corner stayed where it was, so you have 3 → 3, while the other two corners swapped places, so 1 → 2, 2 → 1.

Is this the same as a rotation?

Answer

No. Think back to Activity 1's rotations. If any corner stayed put, they all stayed put. If any moved, they all moved. This is a new symmetry of the equilateral triangle.

What happens if you do the same reflection twice? Since one corner stays put and the other two swap positions, performing the same transformation twice brings us back to the original situation.

What about combining two different reflections? Try this out and see what happens.

Put the triangle back into the original position again. Now turn it over, keeping the top corner in the same position. Then turn it over again, this time keeping the left-hand corner in the same position. How have the labels moved?

Answer

As before, the first transformation swapped C1 and C2, keeping C3 fixed in place at the top.

When you turn it over again, keeping the corner in the lower left fixed, you swap the corners in the lower-right and top positions. You now have corner C2 in the lower-left, corner C3 in the lower-right and corner C1 in the top position. This means you've moved 1 → 2 → 3, 2 → 1 → 1 and 3 → 3 → 2. The overall result of the two reflections is 1 → 3, 2 → 1 and 3 → 2.

What single transformation is this equivalent to?

Answer

Looking back at the cycle diagram, this is equivalent to an anticlockwise rotation by 240 degrees, or a clockwise rotation by 120 degrees.

So, what you've seen here is that combining two reflections produces a rotation! As it turns out, this is true for all combinations of two reflections. You might like to continue testing this for yourself.

This all leads to another interesting question: what happens if you combine a rotation and a reflection?

2.3 Combining rotations and reflections

This week's activities have shown that combining two rotations results in a rotation, as does combining two reflections. But what about both at once? Let's try an example.

Activity 3 Rotating and reflecting

Allow about 5 minutes

Starting again from the original position, rotate the triangle anticlockwise by 120 degrees, and then reflect in the line through the top corner. What happens?

Answer

The rotation moves the edges 1 → 2, 2 → 3, 3 → 1, and the reflection then swaps the two lower corners, keeping the top fixed, which means that 1 → 2 → 1, 2 → 3 → 3 and 3 → 1 → 2.

The net effect is that corner C1 is back in its initial position, while C2 and C3 have swapped places.

What single transformation is this equivalent to?

Answer

This is a reflection with respect to the symmetry line that goes through the corner in position P1.

In this case, combining a rotation with a reflection resulted in a reflection. You can check that this is always true, irrespective of your choice of rotation and reflection, and the order in which you perform them. Together, these transformations form a group.

Looking at all the possible rotations and reflections, you will find that there are only six different transformations in total. You can confirm this mathematically: you're moving the triangle's three corners to three available positions, so there are three choices of position for the first corner, two for the second, and just one for the last one, so in total there are 3 x 2 x 1 = 6 possible moves. Any of these six possibilities represents a possible rotation or reflection of the triangle. This includes the transformation that keeps everything in place.

2.4 When order matters…

If you rotate a triangle twice – first by one angle, and then by another – it does not matter which rotation is done first. Either way, the net rotation is the same, being the sum of the two angles. The two rotations can be said to 'commute' with each other, meaning that the order in which you perform them does not matter.

This probably isn't an unfamiliar concept to you – it applies when adding or multiplying numbers. The order here does not matter either. Whether you take 2 x 3 or 3 x 2, the answer is 6 in both cases.

This property is true for any pair of rotations in the plane about the same centre, whether they are symmetries of our triangle, symmetries of some other regular polygon (such as a regular pentagon) or any pair of rotations with arbitrary angles. The only difference will be the number of rotations you get, which depends on the number of corners on the polygon. Groups that have the property that any two transformations commute with each other are called ‘Abelian’ groups, named after the Norwegian mathematician Niels Henrik Abel.

So, going back to our triangle, what about order when reflections are involved? Let’s check what happens when the transformations from Activity 3 are performed in the opposite order.

Activity 4 Testing the order

Allow about 5 minutes

Starting from the original position, reflect the triangle in the symmetry line through the top corner, and then rotate the triangle anticlockwise by 120 degrees. What happens?

Answer

Reflecting first, the corners are moved according to 1 → 2, 2 → 1, 3 → 3. Rotating by 120 degrees now gives us 1 → 2 → 3, 2 → 1 → 2 and 3 → 3 → 1. The net effect of the combined transformation is that corner C2 is back in its initial position, while C1 and C3 have swapped places.

Recall how in Activity 3, the combined transformation fixed C1 at position P1, while C2 and C3 swapped places.

It turns out the order in which the transformations are performed matters – in other words, they do not commute! Both orders result in a net reflection, but they are reflections in different symmetry lines.

If you like, see what happens when you experiment with rotations of three-dimensional objects. You will find that for rotations around different axes, it matters in which order you rotate.

Watch Video 3 to see Marcus testing this out.

Video content is not available in this format.

Video 3 Order of operations

Now, let's take these learnings back to particle physics.

3 Symmetries in particle physics

In physics, symmetry groups are fundamental ingredients of models that describe anything from subatomic particles to the universe. One central result that connects symmetry and physics is Noether's theorem, discovered by German mathematician Emmy Noether. This states that any symmetry of a physical system implies that there is a quantity that is conserved, meaning that it does not change. Examples in mechanics are, for instance, conservation of momentum for a system that is symmetric under translations in space, and the conservation of angular momentum for a system with rotational symmetry.

Another example of the key role of symmetry in physics is Albert Einstein's theory of relativity. Here, Einstein starts from postulating a symmetry, namely that the physics does not change under certain changes of the coordinate system used to describe it. From this assumption of symmetry alone, he derives his famous theory, including all the mind-boggling consequences about the structure and interaction of space and time. More will be said about relativity later in this course.

3.1 Elementary particles

One of the most striking examples of symmetry in physics occurs in particle physics, which looks at the smallest constituents of matter and their interactions.

The first identified subatomic particles were the electron, discovered by Joseph John Thomson in 1897, and the proton, discovered by Ernest Rutherford in the late 1910s. In the following years, the detection of new particles in experiments continued rapidly. Scientists soon started to look at patterns that would allow them to organise this so-called 'particle zoo', particularly hadrons (as mentioned in Section 1, these are composite particles like protons, neutrons, etc.).

Scientists noticed that these particles, when arranged according to their electric charge and another observed property called 'strangeness', give rise to patterns that they could link to a particular type of symmetry (which are related to higher-dimensional rotations).

Building on this symmetry, in 1961 Murray Gell-Mann and Yuval Ne'eman independently arrived at a model for hadrons, which has become known as the 'eightfold way' (derived from the title of Gell-Mann's publication, an allusion to the Noble Eightfold Path of Buddhism). However, there was one issue – the symmetry predicted a particle that had not yet been observed. From the model, Gell-Mann could predict the specific properties of the particle. It was eventually found in 1964, corroborating the model's predictions, and earning Murray Gell-Mann the 1969 Nobel Prize in Physics.

3.2 Quarks

Gell-Mann's model stipulated that hadrons were in fact composite particles, consisting of what are now known as quarks (this name was apparently inspired by the line 'Three quarks for Muster Mark' in James Joyce's 1939 novel *Finnegans Wake*). There are three different types – often referred to as 'flavours' – named u ('up'), d ('down') and s ('strange').

Each quark also has its own antiparticle, denoted in writing by a bar, giving us three antiquarks: $\overline{\mathrm{u}}$, $\overline{\mathrm{d}}$ and $\overline{\mathrm{s}}$. These quarks have unusual properties. While all hadrons have whole numbers of electric charge, quark electric charges are +2/3 for the up quark and -1/3 for the down and strange quarks. The antiquarks have opposite charges, so that's -2/3 for $\overline{\mathrm{u}}$, and +1/3 for $\overline{\mathrm{d}}$ and $\overline{\mathrm{s}}$. Here are some examples to demonstrate these numbers.

A proton has electric charge of +1, and it consists of three quarks: two up quarks and one down quark. So, it can be represented as *uud*. Note that you can easily check the maths here: 2/3 + 2/3 - 1/3 = 1.

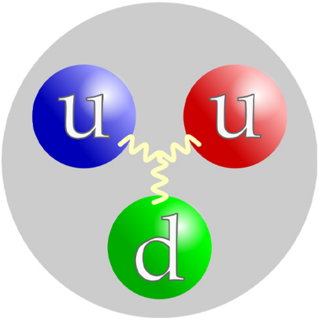


Figure 7 Proton quark structure

The neutron has no charge, and it too consists of three quarks: one up quark and two down quarks. So, it can be represented as *udd*. Adding up the charges, you find the expected result here too: 2/3 - 1/3 - 1/3 = 0.

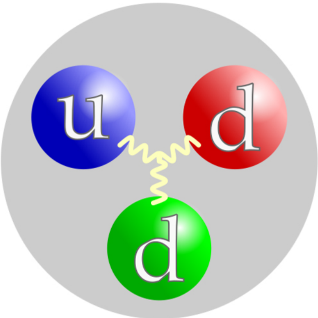


Figure 8 Neutron quark structure

The up and down quarks are the two quarks with the smallest mass and make up all the natural nuclear matter on earth.

Protons and neutrons are examples of 'baryons', a term which refers to particles consisting of three quarks. Their antiparticles, which contain three antiquarks, are known as 'antibaryons'. In addition to these, there are also hadrons of a different composition, which are called 'mesons'. They are made from a quark and an antiquark.

While symmetry played a significant role in detecting the pattern and deriving the model, the symmetry is in fact not perfect, because the three quarks have different masses. So, changing the quark content (a 'transformation' akin to the rotation of the triangle earlier) changes the mass of a particle (so the properties are not exactly the same as before).

But that's not all! More particles were soon found which revealed additional quarks. The current quark model describes a total of six different flavours. There are the three that have been mentioned, plus: b ('bottom' or 'beauty'), c ('charm') and t ('top') quarks.

The six quarks are organised into two groups: the u, c and t quarks have electric charge +2/3, while the d, s and b quarks have charge -1/3. In addition, they all have their own antiquarks, so there are already 12 fundamental constituents here. Remember too, that quarks only account for the hadrons, not for any other types of particles such as leptons (the electron being the most familiar example). Even in terms of these basic constituents, nature appears to be rather complex!

The optional interactive below allows you to explore the world of hadrons further, by combining quarks into baryons, antibaryons and mesons. It details their electric charge, as well as the quantities known as strangeness and charm.

Interactive content is not available in this format.

You may be wondering why quarks always appear in baryons, antibaryons and mesons, and you don't observe other combinations, such as particles made up of two quarks. This explanation lies in the force that acts between these particles, as you will learn about next.

3.3 Forces and interactions

Elementary particles interact with each other in various ways. Right now, we know of four fundamental forces in nature. Only two of these can be experienced directly in everyday life. These two are gravitation (which is described by Newton's law of gravitation, and more precisely by Einstein's general theory of relativity) and electromagnetism (which is described by a set of equations known as Maxwell's equations).

Gravitation is the attraction between any two objects that have mass. This is the force that shapes our universe and holds our solar system together. Electromagnetism covers the electric and magnetic forces between charged or magnetic objects. Both gravitation and electromagnetism also act on particles at the atomic and subatomic scale. For instance, the nucleus and the electrons in an atom are attracting each other due to the electric force because the nucleus has positive charge and the electrons are negatively charged, and opposite charges attract each other. It is this force that holds atoms together, because the gravitational force is tiny.

However, at the subatomic scale, two more forces become relevant. These are known as the weak nuclear force and the strong nuclear force. The strong force keeps the atomic nucleus together – it is strong enough to overcome the repulsion of its positively charged constituents – but it does not affect electrons. The weak force acts on all the elementary particles that constitute an atom, but as the name suggests, it is much weaker than the strong force. The weak force is responsible for a particular type of radioactive decay that is observed in some unstable elements (such as for Carbon-14, used in dating ancient materials). It is the weak force that allows the heavier quarks to decay into the lighter quarks.

3.4 The strong force

The current theory of electromagnetic, weak nuclear and strong nuclear forces (but not gravity) in the world of subatomic particles is called the 'Standard Model'. It is based on three symmetry groups, which are generalisations of rotation groups. One of these is Abelian (electromagnetic interactions), while the other two are non-Abelian groups.

Perhaps the most intriguing of these is the strong nuclear force, which acts on quarks. The interaction is much more complicated than gravitation or electromagnetism. For the strong force, there are six different 'charges' which are commonly labelled as primary colours: the quarks are r (red), r (green) and b (blue), and the antiquarks are denoted by the corresponding ‘anticolours’ $\overline{\mathrm{r}}$, $\overline{\mathrm{g}}$ and $\overline{\mathrm{b}}$. Note that the use of ‘colours’ here is nothing more than an analogy to help us describe the properties of this force. It was also chosen because light of these three basic colours combines to produce white light. In the proton, three quarks of different colours combine to form a ‘colourless’ particle. Indeed, the theory predicts that all particles that can be observed are ‘colourless’ – which means that we can never encounter a single quark on its own.

As well as mixing all three colours, there’s another way to produce colourless particles: by matching a colour with its corresponding anticolour.

This combination into a ‘colourless’ system that cannot be broken up may seem strange, but here’s another analogy for this behaviour. If you have a magnet, it will have a north pole and a south pole. Imagine cutting it down the middle – you might expect that you’d end up with two pieces, one being a north pole and the other a south pole. However, this is not the case – there is no magnetic ‘monopole’ (single pole). Instead, both pieces are magnetic, each with a north and a south pole. The colour charge of quarks is similar, in that the proton as a whole is colourless. Even when smashed in a particle accelerator, the products observed will again be colourless. Nevertheless, the substructure can be detected indirectly, by its effects on other particles – much like observing the effect the magnet’s two poles have on other magnets.

4 This week's quiz

Well done for reaching the end of Week 2. Check what you've learned by taking the end-of-week quiz.

Week 2 practice quiz

Open the quiz in a new window or tab then come back here when you've finished.

5 Summary of Week 2

This week leaves us with a lot of lingering questions. Are quarks really the ‘elementary’ particles, in the sense that they do not possess a substructure? Are the symmetries used in the description of the forces indeed fundamental symmetries of nature? Are the four fundamental forces all there is, or are there more?

We don’t know the answers to these questions, and perhaps we never will. All we can say is that experimental results agree with the predictions from the models we’ve developed based on our knowledge. But even if discrepancies between theory and experiment are discovered at some point in the future, and these point to another layer of complexity, we could never be certain that we’ve reached the ultimate goal of identifying the fundamental building blocks of matter.

Next week stays on the ‘small scale’, as you move from particle physics to quantum physics.

You can now move on to Week 3.

Week 3: Quantum physics

Introduction

This week, you will learn about some of the more mysterious physical discoveries of the 20th century, as you enter the world of quantum physics. You will learn how quantum theory helps explain experimental observations that are not predicted by classical physics, such as how light behaves simultaneously as both a particle and a wave. You will see how quantum theory is fundamentally probabilistic in nature and how this explains radioactive decay. And you'll find out how a famous thought experiment involving cats and deadly poison can help us understand quantum superposition.

Here's Marcus to introduce this week's topic.

Video content is not available in this format.

Video 1 Introduction

By the end of this week, you should be able to:

- understand the basic notion of wave motion
- understand the notion of wave-particle duality
- appreciate the consequences of Heisenberg's uncertainty principle
- demonstrate a general level knowledge of quantum theory.

1 The gaps in classical physics

In the late 19th century, it started to seem as if the fundamental laws of physical science had all been established. These laws constitute what's now referred to as 'classical physics'. However, there were a few early warning signs that classical physics may not yet cover everything.

James Clerk Maxwell, whose contributions to the theory of electrodynamics and the kinetic gas theory are among the pillars of classical physics, was one of the first scientists to observe a discrepancy that classical physics failed to explain. The discrepancy – which concerned the specific heat of gases (like oxygen or nitrogen) – deviated from the predictions of kinetic gas theory. Maxwell famously said in a lecture:

> I have now put before you what I consider to be the greatest difficulty yet encountered by the molecular theory.
>
> (Maxwell, 1875)

Another important experimental observation that defied classical physics was the photoelectric effect, which was studied by Heinrich Hertz in 1887. The photoelectric effect is the emission of electrons when light hits a material. Experiments showed that low-frequency (low-energy) visible light would not lead to the emission of electrons, no matter how intense the irradiation, while ultraviolet (high-energy) light would. Classical physics could not explain this behaviour.

In 1905, Albert Einstein proposed an explanation of the photoelectric effect. He employed a concept that was first put forward by Max Planck, which assumed that light consisted of tiny bundles of energy (quanta). While his work at the time was not immediately recognised by the community, it is now considered as a key step in the development of a new kind of theory – 'quantum mechanics' or 'quantum theory' – that describes nature at the atomic and subatomic scale. Experiments carried out in 1914 by Robert Millikan provided support for Einstein's model, and in 1921 Einstein was awarded the Nobel Prize in Physics 'for his services to theoretical physics, and especially for his discovery of the law of the photoelectric effect'. Within the decade, quantum theory would be fully established and become the standard theory of atomic physics.

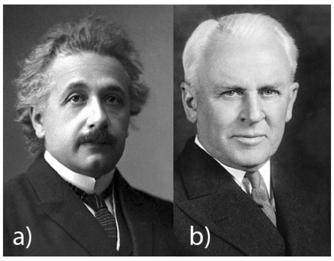


Figure 1 (a) Albert Einstein (1879–1955), (b) Robert Millikan (1868–1953)

Quantum theory explains our observations in the world of atoms and subatomic particles, but aspects of the theory's interpretation have led to challenging discussions among scientists, which continue to this day. In a 'classical world' (which means the world of macroscopic objects that are much larger than atoms), things can become predictable. Think back to Week 1: if you know enough about the initial configuration of a system – like the positions and velocities of particles and their interactions – you can predict the future using the classical equations of motion. The date of the next solar eclipse and where it will be visible on Earth can be calculated with absolute precision. In a quantum world, certainty is replaced by probability, and we can only predict the probabilities of various outcomes of experiments. Many scientists are struggling with this idea and its implications, particularly with the question of when the subatomic world of probability crosses over into the macroscopic world of certainty. This scepticism is captured in Einstein's reaction in a 1926 letter to Max Born, who had championed the probabilistic interpretation of quantum theory:

The theory says a lot, but does not really bring us any closer to the secret of the 'old one'. I, at any rate, am convinced that He does not throw dice.

(Born, 1971)

The famous physicist Richard Feynman wryly stated in a 1964 lecture:

I think I can safely say that nobody understands quantum mechanics.

(Cornell University, 1964)

Nevertheless, quantum theory has been extremely successful, and the probabilistic interpretation has generally been accepted. However, it puts clear and fundamental limits on what we can possibly know about any quantum system.

2 Waves or particles?

One of the fundamental features of quantum physics is the ‘wave-particle duality’, which refers to the fact that, depending on the experiment or type of observation, a system can exhibit either particle- or wave-like behaviour.

The following video introduces this section’s concepts (and takes it a step further, in arguing that the question of whether an electron is a particle or a wave is actually quite a meaningless one).

Video content is not available in this format.

Video 2 The wave/particle paradox

2.1 Water waves

Let’s start with a familiar example: waves in water. When you throw a stone into a pond, a circular wave pattern emanates from the point of impact. When a duck swims across a lake, it leaves a triangular wave in its wake. Waves like these consist of a change in the water level, with a regular pattern of crests and troughs. This basic idea is represented in Figure 2. How would you measure a wave like this?

Figure 2 Diagram of a wave

There are four quantities you might use to measure the wave’s motion.

The first is the distance between crests, which is called the ‘wavelength’. Wavelengths of water waves can vary hugely: from a few centimetres for the waves created by your pebble, to tens of metres for waves driven by strong winds, to hundreds of kilometres for tsunami waves triggered by earthquakes.

The second quantity describes how quickly crests and troughs alternate if you look at the wave at a fixed position. The number of cycles of crests and troughs in one second is called the ‘frequency’ of the wave. For our duck-made water waves, this is usually small (less than 1 per second). Cycles can take several seconds, for normal water waves, up to several minutes for tsunami waves.

The third quantity is the ‘wave speed’, which tells you how quickly the wave is moving. Again, this varies enormously: wind-generated ocean waves move at tens of kilometres per hour, whereas tsunami waves can move at hundreds of kilometres per hour.

As you might have noticed, these three quantities are related. The wave speed is the product of the wavelength and the frequency: speed = wavelength × frequency.

Question 1

What is the wave speed of a tsunami wave with a 200km wavelength and frequency of 4 per hour (which means one cycle takes 15 minutes)?

Provide your answer...

Answer

The speed is the product, so 4 x 200 = 800 kilometres per hour. This wave is moving about as fast as a commercial jet plane.

The fourth and final quantity describes how much the water level changes. This is called the ‘amplitude’ of the wave, and it’s measured by the difference between the crest level and the average (calm) surface.

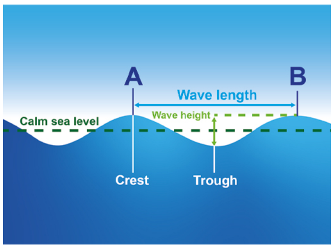


Figure 3 Diagram of a wave (labelled)

Let’s now consider some other types of waves.

2.2 Sound waves

Sound waves are another type of wave motion that you’re probably aware of already. Of course, you cannot actually see these waves. They are still quantified by their wavelength, frequency and amplitude, much like the waves in the last section.

With a sound wave in air, it’s the density of air that’s changing, going through cycles of compression and expansion. Sound waves in air travel at a speed of about 343 metres per second. Light travels almost a million times faster, hence the rule of thumb for measuring your distance from a lightning strike: you see the flash almost instantly, while it takes about three seconds per kilometre for the sound of thunder to reach you.

When a plane travels faster than the speed of sound, its sound waves form a cone behind it, like the water waves produced by the swimming ducks. The compression at the edge of this cone generates sound energy, which sounds to an observer like an explosion or thunderclap. This is known as a sonic boom.

Figure 4 Animation of a sonic boom

Wavelengths of audible sounds vary from centimetres (high-pitched notes, high frequency) to several metres (low-pitched sound, low frequency). The audible frequency range for humans is typically from about 20 to 20,000 oscillations per second.

For sound waves, the amplitude measures the change in density. This is perceived as the loudness of the sound. The larger the amplitude of the sound wave, the louder the sound is perceived.

2.3 Electromagnetic waves

Electromagnetic waves come in many forms, and they’re hugely important for modern life and technology. Radio waves transmit information across a range of frequencies, be it radio and TV programmes or mobile phone calls. Microwaves are used to heat food and to navigate planes and ships by radar. Visible and near visible (infrared, ultraviolet) light keep us warm and allow us to orientate ourselves in the world. X-rays and gamma rays are used in medical imaging and radiological therapy.

The travelling electromagnetic wave consists of oscillating electric and magnetic waves. The difference between radio waves, microwaves or visible light is just their wavelength or frequency. The size of an antenna gives you an idea of the wavelength for radio waves, while the wavelength of visible light is around a thousandth of a millimetre. Regardless of their wavelength or frequency, all electromagnetic waves travel at the speed of light (which is defined to be exactly 299,792,458 metres per second, so about 300,000 kilometres per second, in vacuum). Figure 5 shows the electromagnetic spectrum, so you can see some of this in context.

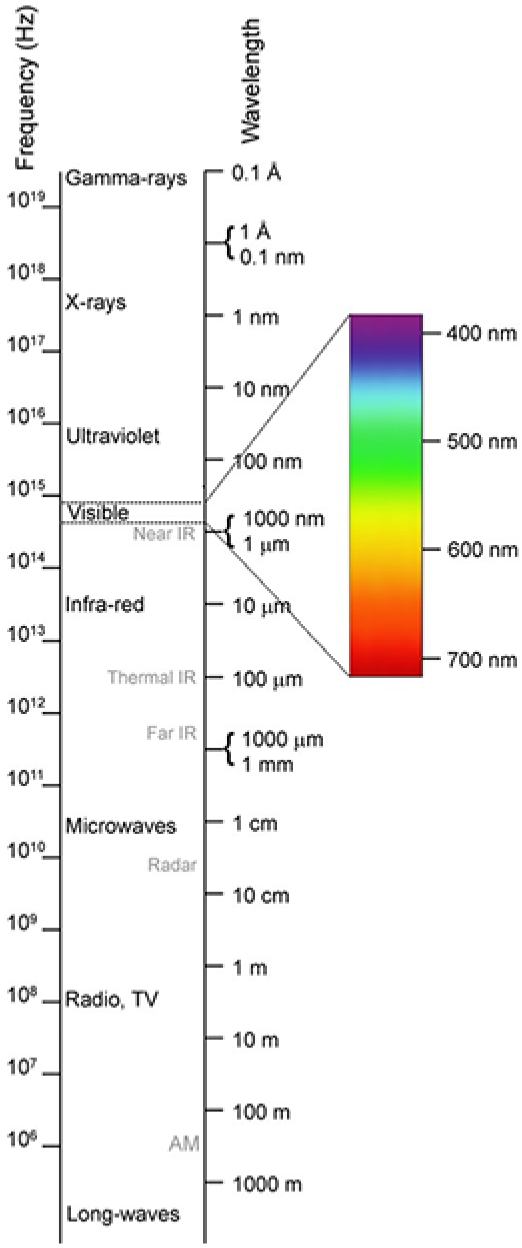


Figure 5 The electromagnetic spectrum

French physicist Louis de Broglie hypothesised in 1924 that all matter demonstrates wave properties; such behaviour was soon demonstrated in experiments, and de Broglie won the Nobel Prize in Physics in 1929. The 'de Broglie wavelength' is given by Planck's constant divided by the momentum (mass times speed) of the particle.

But what is 'Planck's constant'? It's a value derived from the energy of a photon, relating its energy to its frequency. That is, if you take the energy of any photon (measured in Joules) and divide it by the photon's frequency (measured in meters per second), you get a constant. This is called Planck's constant, and is usually denoted by the letter *h*. It's one of the fundamental constants of nature which, like the speed of light, has been defined with an exact value. This value is $h = 6.62607015 \times 10^{-34}$ J-s (Joule-seconds), which is a very small number indeed (10^{-34} is a kind of shorthand; to write it out fully, there would be 33 zeroes after the decimal point before the first non-zero digit).

The wavelength of these 'matter waves' is absolutely tiny as a result, and cannot be observed outside the world of molecules, atoms and subatomic particles. This brings us to the realm of quantum theory. But before entering this intriguing world, let's look at a crucial property of wave motion, which is important both in the classical and in the quantum world.

2.4 Wave motion

How can we capture both the heights of the wave crests and troughs as well as the undulation of the wave, without using complex mathematics? It can be done! Let's say we have a wave of a certain amplitude. Now think of the wave as being represented by an arrow of a length that equals the amplitude and which rotates in a circle as the wave progresses, such that it points up at a crest and points down in a trough.

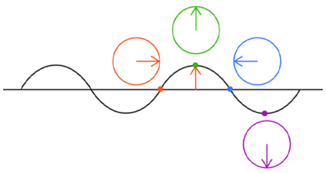


Figure 6 Wave represented by an arrow

You may be wondering how this picture is useful. It helps us to understand what happens when two waves meet. If you throw two stones into a pond, you can see circular waves emanating from both impact sites. Eventually these waves will overlap. In fact, they'll essentially pass through each other, with the heights of the two waves combining to form the resulting wave pattern.

Let us consider a situation with two waves in terms of our arrow picture. For simplicity, let's also assume that the amplitudes are the same, so there are two arrows of equal length. As the waves progress, our arrows rotate around. How do the waves add up? If at any place and time both arrows point up, the waves add together and produce a crest of double the height. If both point down, this produces a trough of double the depth. However, if both arrows point in opposite directions, the waves cancel out. The way the two waves interact is determined by the angle between the two arrows. This is often referred to as the 'phase' or 'phase difference'.

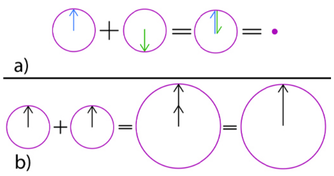


Figure 7 (a) Waves cancelling out (b) waves adding together

So, in the right circumstances – with the correct phase difference – waves can cancel out. There are modern electronic devices that make use of this fact: noise-cancelling headphones. They are in fact making noise that cancels out the noise around you!

These kinds of interactions between waves are described with specific terms: the most common being 'interference' and 'superposition'. The former is commonly used for sound waves or electromagnetic waves. Superposition is a term that is more commonly applied to quantum waves. This type of interaction is characteristic of waves.

Observing an interference pattern is seen as clear evidence of wave motion. A typical setup for observing such a pattern is the double-slit experiment.

The double-slit experiment is an experiment where light is projected onto a screen. An opaque barrier with two thin slits cut into it is then placed between the light source and the screen. When this experiment is performed correctly, an interference pattern of light and dark stripes appears on the screen.

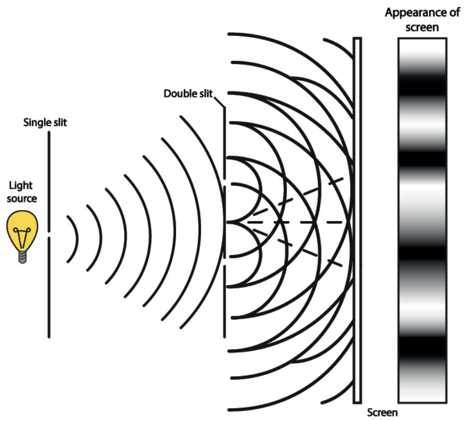


Figure 8 The double-slit experiment

If we think in terms of light, the interference pattern consists of bright and dark stripes (or if we just consider a two-dimensional slice, bright and dark spots). The reason that these stripes form is due to the phase difference. This result comes from the difference in the path lengths of light that travelled through one slit compared with light that travelled through the other slit. Whenever the phase difference causes the waves to cancel, it produces a dark stripe. When they point in the same direction, it produces a bright stripe.

Note that the phase difference is all that matters here. The brightness of the light itself does not impact the direction of our arrow. Note that in order to observe this interference effect, the double-slit setup must have a size that is comparable with the wavelength of the light – so it has to be very small for visible light. Also, you will need a nice beam of light, which consists of a single wave rather than a mixture of many waves. Such light is provided by a laser as a light source.

Having covered these concepts, it's time to ask how things work in the realm of quantum physics.

3 Quantum theory

So far, this week has shown how the general principles of quantum theory can help explain experimental observations. In this section, you'll see how quantum waves provide a powerful mathematical description for the quantum state of a system, and how this allows us to predict the behaviour of quantum systems.

3.1 Quantum waves

This is where the quantum world gets weird. Contrary to the previous examples of waves, quantum waves do not appear to represent a physical quantity, but instead describe something like a probability. A quantum wave describing an electron in a hydrogen atom, say, will tell you how likely it is to find the electron in a particular region. Note that this already mixes up both the particle and the wave point of view.

Even weirder, the most popular interpretation (called the Copenhagen interpretation, largely devised between 1925 to 1927 by Niels Bohr and Werner Heisenberg) asserts that physical systems generally do not have definite properties prior to being measured. Quantum theory can predict the probability of a given measurement's results, but it is the actual act of measurement which affects the system, causing the measured quantity to assume a definite value.

What happens if we send a beam of electrons (rather than light) through a double-slit setup, like the one seen in Figure 8 earlier? The electrons can pass through either slit. If they were just particles, you would presumably get two shadows (maybe a bit diffuse at their borders) of the two slits. However, if they behave as waves, they will interfere in the manner of classical waves, with the bright and dark stripes corresponding to regions with high and low probability of finding an electron, respectively. Indeed, that is what is observed. Even if you do the experiment with electrons very slowly, so you can register the point where electrons arrive one-by-one, you still find that this eventually builds up the expected interference pattern.

However, if you try to set up the experiment in a way that allows you to find out which slit the electron passed through, you will find that the interaction required to detect the electron's path destroys the interference pattern. The same thing happens if you close one slit while keeping the other open, and it doesn't matter how often you switch this round. The electron, while having to pass through one of the slits (when seen as a particle), somehow 'knows' about the presence of both slits.

It may not come as a surprise that the probabilistic nature of the information provided by quantum theory has troubled many scientists, including some of the most famous minds of the time, like Albert Einstein. Discussions about the interpretation of quantum theory are continuing to this day. Regardless, quantum theory is an incredibly successful step forward in describing the microscopic world. There are many applications for it – the rest of this week will introduce a few examples of scientific processes, theories and thought experiments.

3.2 Radioactive decay

One of the natural processes which best demonstrates the probabilistic nature of the quantum world is radioactive decay. This process occurs in unstable atomic nuclei. Depending on the type of decay, different particles are ejected from the nucleus, which can be detected and counted. However, the time it takes for a specific nucleus to decay cannot be predicted, regardless of how long the atom has existed. Quantum theory will only provide bigger-picture statistical predictions, like how long it takes on average for an atom to decay. While this is broadly useful information when you're looking at many atoms at once, it can't be used to make any predictions about individual atoms. Experimental observations confirm this. The counts from radioactive decay appear completely random, and the statistics are in agreement with quantum theory.

You may be aware of the notion of 'half-life', which refers to the time required for half of the atoms in your sample to decay (assuming that you have a large number of atoms to start with). Figure 9 shows a simulation of this radioactive decay, starting with either 4 atoms per box (left) or 400 (right). The number at the top records how many half-lives have elapsed.

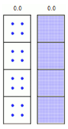


Figure 9 A half-life simulation

Be careful not to get confused here: this isn't the same as the mean lifetime of a single atom. For example, radium-226 has a half-life of 1,602 years. If you put a block of radium-226 into your drawer, the amount of radium-226 atoms in it will have halved after 1,602 years (and it will keep halving from there, so that after 3,204 years you're left with a quarter of your original radium-226). However, if you take single radium-226 atoms, record how long they each take to decay, and take the average, you'll find that the mean lifetime is 2,311 years. This difference has nothing to do with quantum effects – it's just a consequence of different ways of taking an average.

Radioactive decay is used in many areas – one of which is radiocarbon dating. While most of the carbon atoms on Earth are the stable ^{12}C atoms (carbon-12, whose nuclei have 6 protons and 6 neutrons each) and ^{13}C atoms (whose nuclei have one additional neutron), there is a very small amount of unstable carbon atoms (about one in a trillion) whose nuclei have two additional neutrons. Measuring the amount of these ^{14}C atoms that are left in a sample from a dead plant or animal provides information that can be used to calculate how long ago the plant or animal died. This is done by picking up and counting the signals from decaying atoms in the sample, which allows you to estimate the number of atoms of this type. Because the half-life of ^{14}C is about 5,730 years, this method allows the reliable dating of materials up to around 50,000 years old – there are too few ^{14}C atoms left in samples older than this, and the statistics are not good enough to draw firm conclusions.

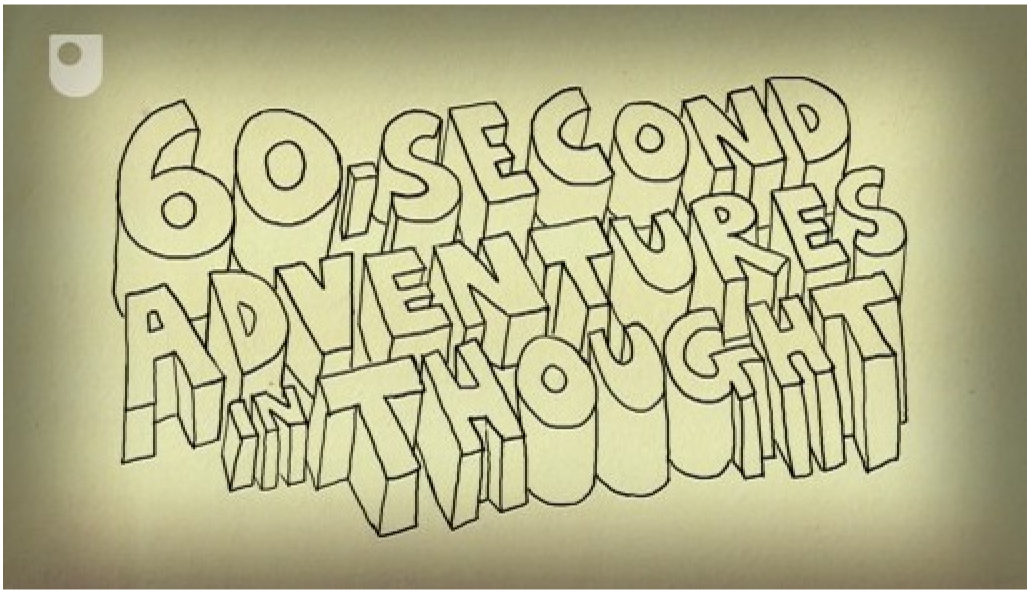
3.3 Schrödinger's cat

The most famous example of the puzzling implications of quantum theory is known as 'Schrödinger's cat'. This is a thought experiment in which an imaginary cat is sealed in a box, along with a device that will release a deadly poison once a radioactive atom has decayed. As quantum mechanics only tells us about the probability that an atom has decayed after a certain amount of time, it's not possible to know for sure whether the cat is still alive without checking inside the box. All we know is that the longer we wait, the worse the cat's chances become.

In the framework of quantum theory, this is described as a 'superposition' of two states: one in which the cat is alive, and one in which it is dead. So, in a sense, until the outcome is confirmed, Schrödinger's cat is both alive and dead at the same time.

Video content is not available in this format.

Video 3 60-Second Adventures in Thought: Schrödinger's Cat



This popular paradox relies on a link between the quantum world of the radioactive atom and the device that releases the poison. They are linked in a way that transports the quantum probability related to the atomic scale to the macroscopic scale, by triggering the device.

The wider issue this thought experiment raises concerns the measurement process, by which we infer what happens at the atomic scale. Necessarily, this has to be done by linking the quantum system to a 'macroscopic' device, something that can be counted (e.g. the clicks of a Geiger counter) or otherwise recorded. Once this is done, what was a quantum system with a probabilistic description has been assigned a definite outcome – the atom has either decayed (i.e. we heard the click of the Geiger counter) or it hasn't.

Understanding the measurement process and what happens is one of the key issues in quantum theory. One way out of Schrödinger's cat paradox is to take the macroscopic device that detects the decay to be the observer that measures the quantum system. In this case, there is no more uncertainty about the outcome, whether or not you've looked in the box.

3.4 Heisenberg's uncertainty principle

One of the fundamental consequences of the wave nature of matter is known as Heisenberg's uncertainty principle. There are limits to what can be definitively known about certain pairs of physical properties of a particle. The most commonly used example is the position and the momentum of a particle: the more precisely the position of some particle is determined, the less precisely its momentum can be known, and vice versa. You can know either one of them reasonably precisely while having very little information on the other, or you can know both approximately. This is linked to the final complication discussed about the double-slit experiment earlier – if you try to determine the position of the electron by detecting which slit it passes through, you affect its momentum and destroy the interference pattern.

Heisenberg's uncertainty principle, then, puts a fundamental limit on what we can know. And there are other strange consequences. One of these is known as the 'zero-point energy' or 'zero-point motion'. If you put an electron into a very small box, then its position is determined up to the size of the box. Its momentum has an uncertainty that directly relates to the size of the box: the smaller the box, the larger the uncertainty in momentum, and the quicker the electron will move around in the box.

A similar consequence is that in the ground state (lowest energy state) of a hydrogen atom (the simplest atom consisting of a single proton as its nucleus and an electron), the electron is not sitting on the atomic nucleus, but is moving in a region close to the nucleus. You can estimate the lowest energy from the uncertainty principle alone. More precisely, quantum theory will determine the energy as well as the probability to find an electron at a given distance from the nucleus at any one time. Unlike the planets in our solar system orbiting the sun, the electron does not stay on a circular or elliptical orbit, but can, with a certain probability, be found over a range of distances.

3.5 Quantum fluctuations

Another mind-boggling consequence of Heisenberg's uncertainty principle is that there's no such thing as empty space, or a complete lack of motion. Everything is subject to 'quantum fluctuations'. In fact, even if a space was initially completely empty, particles would spontaneously form and vanish in it.

Surprisingly, the effects of quantum fluctuations can be measured, and they have been experimentally observed. The most prominent example of this is the 'Casimir effect', named after the Dutch physicist Henrik Casimir. Two uncharged conductive plates are placed a few nanometres apart in a vacuum. They experience a force, due to the effect that the plates have on the quantum fluctuations between them.

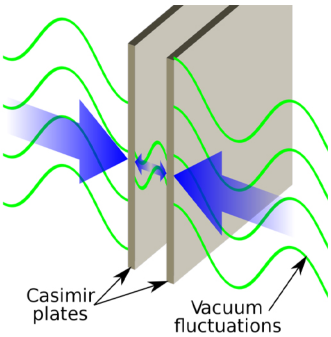


Figure 10 Casimir effect

This raises some fundamental questions about our understanding of the world. In classical theory, you start with an empty space, and consider the motion of particles in this space, all subject to certain forces. In quantum theory, this basic assumption of an empty space is already flawed: it contains fluctuations, and the particles you're studying will interact with these fluctuations. This makes the system much more complicated to analyse.

3.6 Planck time and length

In 1899, Max Planck considered combinations of three fundamental constants – the speed of light, the gravitational constant and Planck's constant – as a possible basis for a 'natural' system to measure time and length. This led to the notions of a 'Planck time' and a 'Planck length' (the distance travelled by light in one Planck time). Both of these units are tiny, the Planck time being approximately 5.4×10^{-44} seconds, and the Planck length about 1.6×10^{-35} metres. To put these into perspective, a proton is about 10^{20} Planck lengths in diameter (written out in full that's a 1 with 20 zeros). Because these units are so small, the idea to base our measuring units on them has not taken off. Nevertheless, in combining properties of quantum theory (Planck's constant) and gravitation (the gravitational constant), these quantities have become a central aspect of speculation about what happens beyond the range of current physics. You'll revisit this later in Week 5.

As mentioned earlier, quantum effects on motion of particles are only apparent in the world of molecules, atoms or subatomic particles. While there currently is a theory of gravitation that is fully consistent with quantum mechanics, it is expected that around the scale of the Planck length, quantum gravitation effects take over. To measure anything the size of a Planck length, the momentum needs to be very large due to Heisenberg's uncertainty principle. The energy required in such a small space would potentially create a tiny black hole the size of a Planck length. Any attempt to investigate shorter distances by performing even higher-energy collisions would result in the production of black holes, which means that length scales smaller than the Planck length would be completely inaccessible.

This has led to the notion of the Planck length as a minimum length of space, beyond which we cannot know anything. It's important to remember though, that such arguments are based on combining constants from quantum theory (which works at the subatomic scale) with the theory of gravitation (which works at macroscopic scales). It's entirely possible that this picture just isn't complete yet. A consistent theory of quantum gravitation could involve other constants and new physics that completely change the behaviour seen at such small scales.

4 This week’s quiz

Well done for reaching the end of Week 3. Check what you’ve learned by taking the end-of-week quiz.

Week 3 practice quiz

Open the quiz in a new window or tab then come back here when you’ve finished.

5 Summary of Week 3

Quantum theory has greatly improved our understanding of the world of atoms and subatomic particles. Nevertheless, it poses conceptual challenges which question our intuitive understanding of the structure of space and matter. It turns out that our minds, trained by the observable world around us, aren't well suited to understanding how an electron can behave like a particle and a wave simultaneously. We must therefore carefully observe and develop explanatory theories for why things in the subatomic world behave very differently to our naive intuition.

Our current understanding of quantum theory severely limits what we can know about systems at the subatomic scale. Heisenberg's uncertainty principle shows that we cannot pin down the location of anything too precisely without increasing the uncertainty of its other measured quantities. In particular, there is a limit as to how precisely we can know both the position and momentum of a particle. Measuring one property of a quantum system automatically affects others – and this can extend even over large distances. Einstein called this phenomenon the 'spooky action at a distance'. The effect of measurements on a quantum system may be exploited in secure quantum communication, where any eavesdropping would be detectable.

Going further into smaller and smaller distances, the Planck length may well prove to be the ultimate limit beyond which we cannot gain any information.

Now you'll be moving from the very small to the very large, as next week looks at our knowledge of the universe.

You can now move on to Week 4.

Week 4: Space

Introduction

Last week explored how the quantum nature of the world puts a limit on what we can know, especially on very small scales. This week you will turn your attention to the largest scales of all and ask some big questions about the universe as a whole. Does the universe go on forever? If not, how big is it? Can we ever know these things?

Here's Marcus to introduce this week's topic.

Video content is not available in this format.

Video 1 Introduction

By the end of this week, you should be able to:

- appreciate that looking out into space equates to looking back in time
- describe the effective boundary of the observable universe
- appreciate that observations of galaxies show how the whole universe is expanding at an accelerating rate
- understand the implications of Einstein's theory of general relativity on the expansion of the universe.

1 Our place in space

We know that the Sun is a star, one of several hundred billion stars that make up our galaxy, the Milky Way. The Milky Way is part of a cluster of about 40 galaxies known as the Local Group, which is itself associated with the much bigger Virgo Cluster containing more than 2000 galaxies. This and other clusters are bound up superclusters. It is these superclusters that make up the landscape of the universe on the largest scales. Figure 1 shows an artist's impression of the spiral structure of our galaxy, the Milky Way, based on recent observations. The Sun is just below the 'o' in the Local Arm. The image is about 100 000 light years across.

Figure 1 The Milky Way

Figure 2 depicts the superclusters that make up our neighbourhood in the universe. The Local Group lies just to the left of the Virgo cluster in the centre of the image, which is about 2 billion light years across.

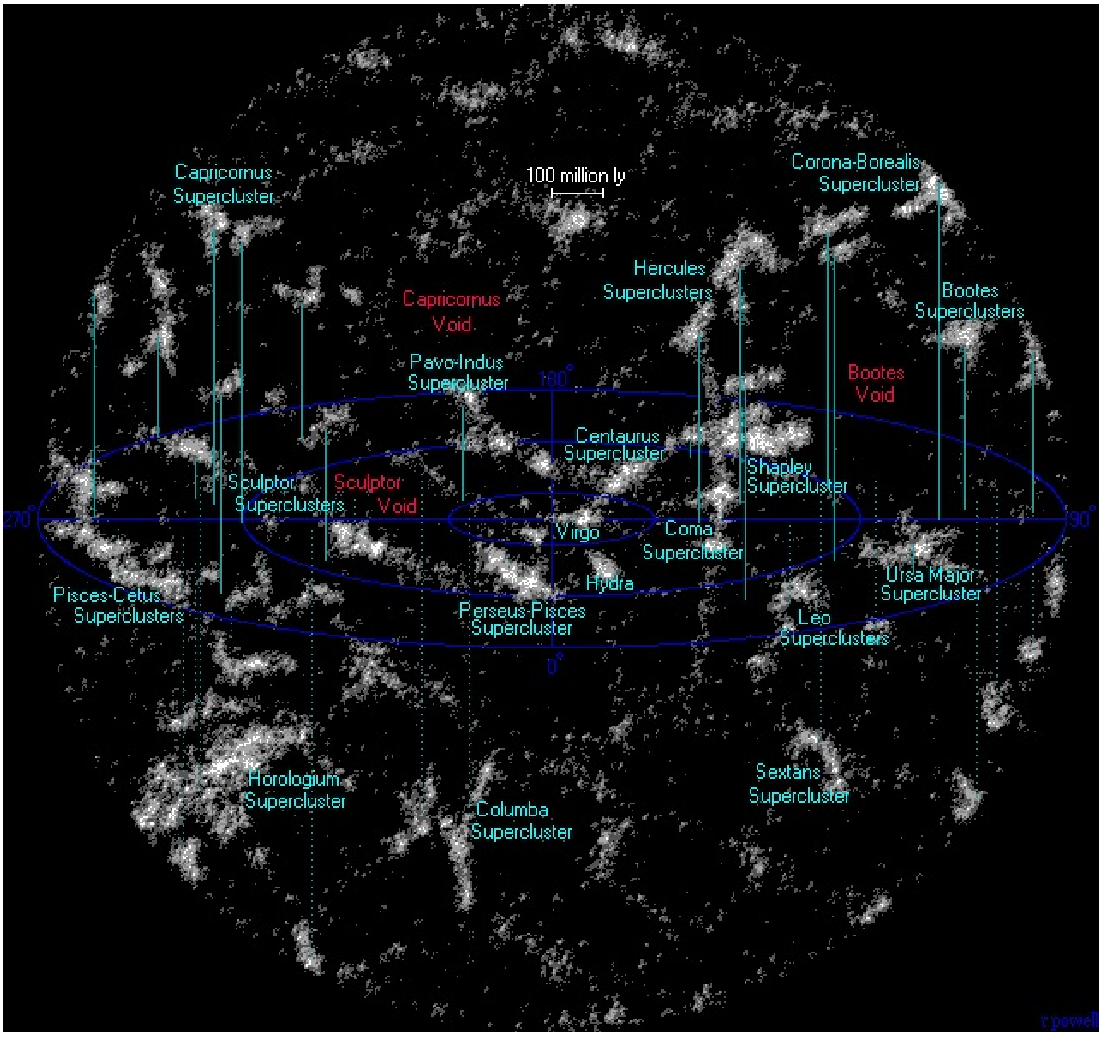
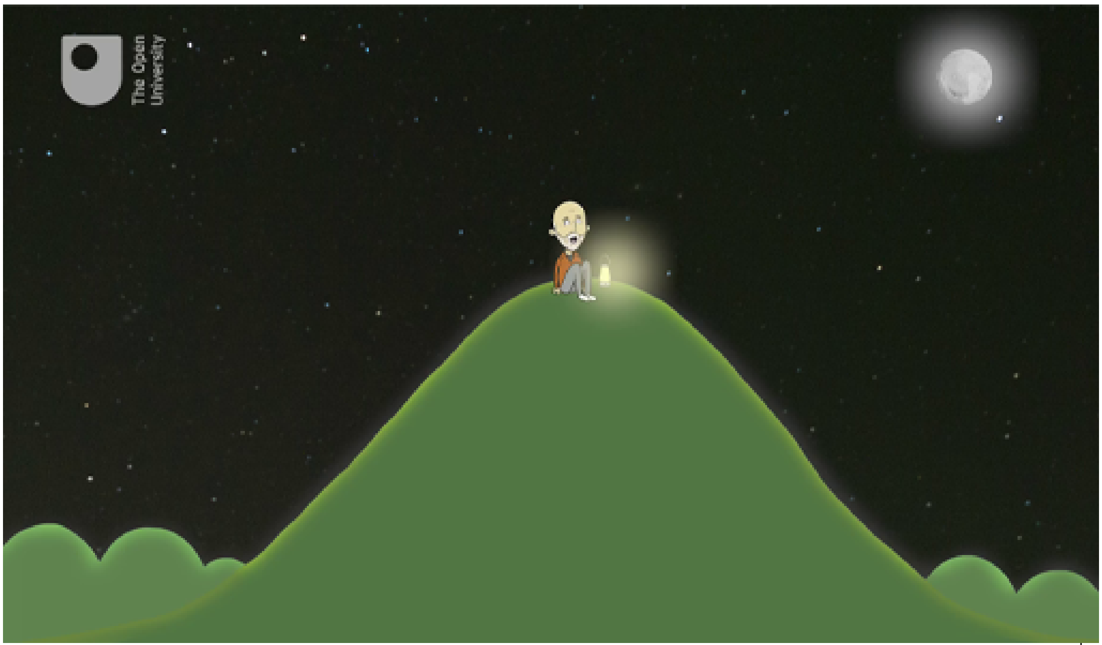


Figure 2 Superclusters

Our clear picture of the universe is relatively recent. Barely a century ago the universe was thought to be a much smaller place. The common belief was that the Milky Way was the entire universe and sat in an otherwise empty void. The misty patches now known to be other galaxies were thought to be part of the Milky Way itself. It was only in the early 1920s that astronomers began to realise the enormous scale of the universe, and our knowledge of its extent has grown ever since. Video 2 explores our history of looking into space, and introduces a few of this week's ideas.

Video content is not available in this format.

Video 2 Timeline of our understanding



The distances involved in this topic are staggering when you think about them. The furthest superclusters in Figure 2 lie about a billion light years away. That's a huge distance – but how much further can we go?

2 How far can we see?

Gazing into the sky might make you feel like space goes on forever, but in one sense, we know it cannot. When you look out into space, you're looking back in time. How can this be? It's because light takes time to get to us, and everything you see carries information from the past, not the present.

Box 1 Distances in space

Light and other types of electromagnetic radiation (e.g., radio waves) travel through space at precisely 299 792 458 metres per second (m s^{-1}), commonly approximated as 300 000 000 m s^{-1} or, in scientific notation, 3×10^8 m s^{-1}. In one year, light will travel 9.46×10^{15} m – this distance is known as a light-year (ly). In other places you may see astronomical distances measured in parsecs, where one parsec is about 3.26 ly, but this course will mainly use light-years.

The light coming to us from the Moon started out about 1.3 seconds ago. When the Apollo astronauts were exploring the Moon between 1969 and 1972, the controllers on Earth had to wait almost three seconds for a reply. That was the minimum time it took for radio waves – travelling at the speed of light – to reach the astronauts on the Moon, and for their reply to come back to Earth.

Likewise, we see the Sun not as it is now, but as it was over eight minutes ago. We see the nearest star, Proxima Centauri, as it was 4.2 years ago. We see the centre of our own galaxy, the Milky Way, as it was about 27 000 years ago, when our Stone Age ancestors were painting animals on the walls of caves in what is now France.

The nearest large spiral galaxy similar to our own is known as Messier 31 (M31). It is the largest member of our Local Group of galaxies. You may have seen pictures of it (Figure 3), but did you know that on a dark night you can see it yourself without a telescope?

Figure 3 M31

M31 is in the constellation of Andromeda. Although this is in the northern sky, it's visible at some time from most parts of the world. It is more difficult to see from southern latitudes though.

The galaxy is best seen in dark skies when the Moon is absent. The very best views will be in autumn from the northern hemisphere but it can be glimpsed at other times of year, depending on your latitude. In the spring it is too close to the Sun in the sky to be seen at all.

In this activity you'll use the Stellarium Web online sky chart to see if M31 is visible from your location, and find out the best time to look for it.

Activity 1 Glimpsing Andromeda

Allow about 20 minutes

Follow these steps:

1. Go to the Stellarium Web website. If you are happy for the site to access your location, allow it to do so. Otherwise click the location button at the bottom of the screen and find your position on the map. Then click 'Use this location'.
2. You will now see a view of the sky from the place you have chosen (it may be in daylight, of course). From the box at the bottom right of the screen you can set the date and time to display the sky at any time you wish. Today's date and soon after nightfall would be a good choice to start with.
3. To find the Andromeda galaxy, type 'M31' in the search box at the top of the screen. From the drop-down list select 'M 31 (Andromeda Nebula)'.
4. The chart will then adjust to put M31 near the centre of the screen together with some information, including when it will rise and set. Use these times to set the chart to a convenient time when M31 will be above the horizon during the night.
5. On a date and time when M31 will be visible in a dark sky, go outside and use the nearby constellation patterns to identify where to look. (If it helps, try turning on the constellation outlines with the button at the bottom of the screen.) The galaxy is just visible to the unaided eye as a misty patch, but it's much clearer through binoculars. If you glimpse it, congratulations! You are seeing light that started on its journey 2.5 million years ago.

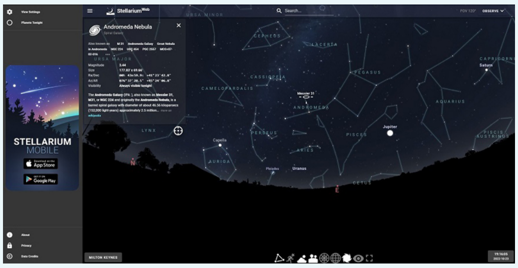


Figure 4 Screenshot from Stellarium Web

M31 is the most distant object visible to the unaided eye. So, if someone asks how far you can see in the dark, you can now give the answer: 2.5 million light-years.

3 The observable universe

You might be thinking: with ever bigger and more advanced telescopes, surely our view of the universe will keep increasing with no limit? But that's not quite how it works.

We are reasonably confident that the universe is around 13.8 billion years old. It began in a state of extremely high density and has been expanding ever since. This initial expansion is known as the 'Big Bang'. How far light travels over time influences the distance we're able to see. Given the universe's finite age, then it follows logically that this distance also has a limit. This means that however big the universe might be, the part that we can actually see – known as the 'observable universe' – is strictly limited. The universe as a whole may well be infinite, but the observable universe is not.

If the observable universe is finite, then how big is it? You might logically presume that if light from the most distant parts has been travelling for 13.8 billion years, then the 'edge' of the observable universe must be 13.8 billion light-years from us. But it's not quite that simple either. This would be correct if the universe were static – if all the galaxies were stationary. But in fact, the universe is expanding, and the galaxies have been moving away from each other ever since the Big Bang.

That makes our estimates somewhat difficult. The expansion means that distant galaxies were much closer to us when their light was emitted. They're much further away from us now. To work out where they are today, some assumptions need to be made about the speed of the universe's expansion throughout its life. Recent estimates indicate that if the universe is 13.8 billion years old, the very edge of the observable universe will now be about 46 billion light-years away. You can think of us sitting in the middle of a sphere 46 billion light-years in radius that contains everything that we can ever know about.

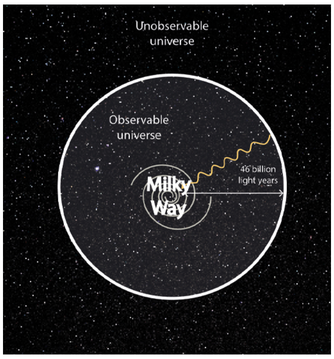


Figure 5 The observable universe

Establishing the limits of the observable universe raises another question. How far can we see within the observable universe?

4 The most distant objects

The further into space you look, the older the objects you see. The Hubble Space Telescope, an observatory that has been orbiting around the Earth since 1990, has been accumulating images of small areas of sky known as ‘deep fields’. One of these images, known as the ‘Extreme Deep Field’ (XDF), depicts an area in the constellation of Fornax, just south of Orion. (In fact, it sits within an area called the ‘Ultra Deep Field’, which is itself within a mere ‘Deep Field’!)

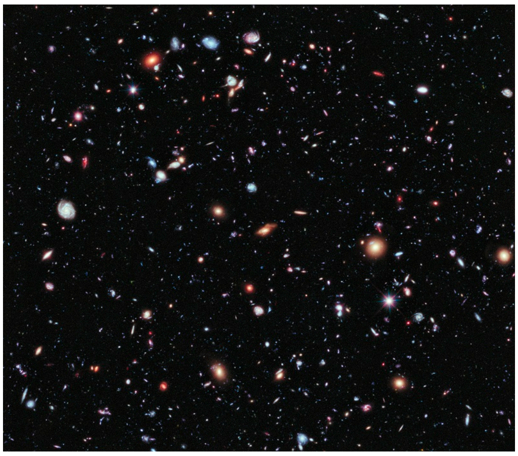


Figure 6 The Extreme Deep Field

The XDF is a composite of over 2000 images, made over a ten-year period. The image covers a very small patch of the sky as seen from Earth (just 0.6% of the size of the Moon) which nevertheless contains an estimated 5500 galaxies. These distant galaxies are very young but they were formed hundreds of millions of years after the Big Bang. The oldest of these formed when the universe was around 450 million years old.

These galaxies are some of the most distant objects ever seen – but they’re still not at the ‘edge’ of the observable universe. What does that edge look like?

5 Light from the edge

Keep in mind that light from near the edge was emitted when the universe was very young, and it would not look the same as it does now. The further we look back in time, the closer to the Big Bang we are seeing. This raises the intriguing question of whether it's possible to see all the way back to the Big Bang itself. That, surely, would mark the edge of the observable universe!

Unfortunately, the answer is no – we can't see all the way back to the Big Bang. But it's possible to get very close! In fact, we can see back to about 380 000 years after the Big Bang. That's still a very long time in human terms, but only 0.0028% of the age of the universe, taking us back to its relative 'infancy'. At that point, the sky becomes covered in a glowing fog which prevents us from seeing any further. The fog is very faint in visible light but easier to see at wavelengths around one millimetre (in the microwave part of the spectrum). For that reason, it is known as the 'cosmic microwave background radiation' (CMBR). It is very smooth and covers the whole sky.

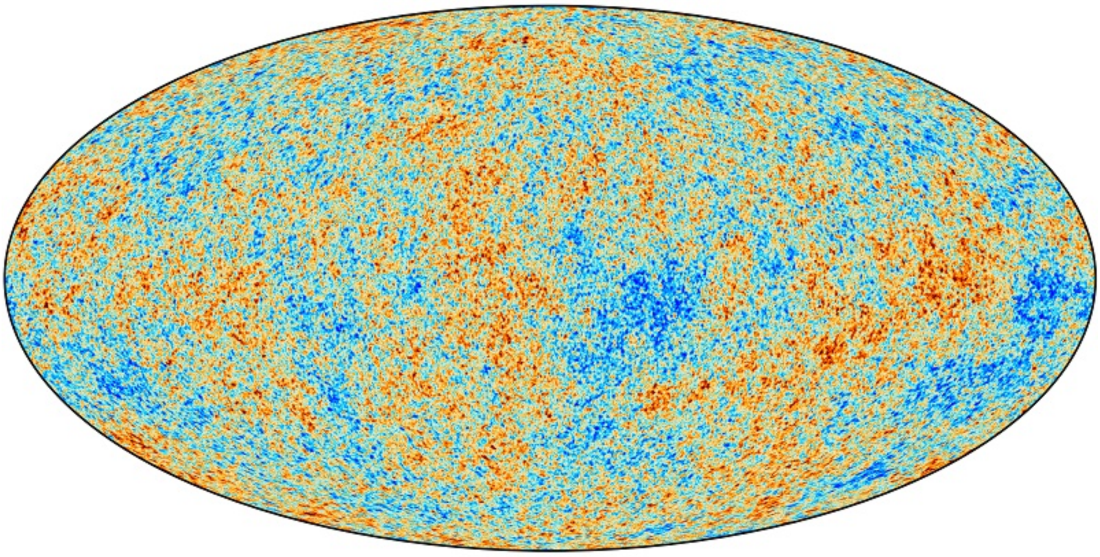


Figure 7 A map of the CMBR

Figure 7 shows an image of the microwave 'fog' covering the whole sky as observed by the Planck space observatory. It is exceptionally smooth – the red and blue areas represent minute differences in brightness of around one part in 10 000. We are looking at the universe when it was just 380 000 years old. The CMBR marks the edge of the observable universe.

What happened when the universe was 380 000 years old? Before that time there were no stars and not even atoms. The universe was full of ionised gas – a plasma – mainly nuclei and electrons, as it was still too hot for atoms to form. The free electrons in the plasma prevented light travelling very far and so space was opaque. As the universe expanded and cooled, the electrons and nuclei combined to form atoms of hydrogen and helium until – at around 380 000 years – the fog cleared. The universe has been transparent ever since.

For all practical purposes, then, the foggy glow of the CMBR marks the edge of our observable universe – currently some 45 billion light years away, and so slightly closer than the theoretical limit. Light now arriving from the CMBR has been travelling towards us for 13.8 billion years, meaning it is the oldest light in the universe.

6 Beyond the edge

What lies beyond the edge of the observable universe? Again, it's important to be careful and take the time into account. The ionised gas that we now see as the CMBR will have since condensed into atoms, stars and galaxies.

Cosmologists have long been guided by what is called the 'cosmological principle'. This states that on the very largest scales the universe is generally 'homogeneous' (all places are alike) and 'isotropic' (all directions are alike). The observations of the very smooth CMBR have only strengthened this view. From this principle, it can be inferred that those distant regions beyond the observable universe will be much the same as our own neighbourhood. Of course, this principle could well be incorrect, and distant parts of the universe may be rather different from our own. If that were the case, much of our cosmology would have to be revised.

There will be stars, galaxies, clusters and superclusters of galaxies out there which we cannot see, because they lie beyond the edge of the observable universe – light from them is on its way to us, but hasn't yet arrived. Likewise, if there are astronomers in those galaxies they can never learn about us, because we lie beyond the edge of their observable universe. You can think about this in terms of individual circles, as in Figure 8. In (a) the observable universes are quite separate. In (b) they overlap, and some of the same galaxies can be seen in both.

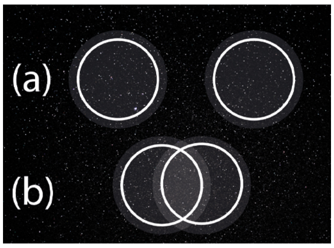


Figure 8 Everyone is at the centre of their own observable universe

One thing that has not yet been discussed here is the extent of the wider universe. Our observable universe – bounded by the cosmic background radiation – is finite. But beyond that 'edge', how big is the universe as a whole? Could this be infinite?

7 The wider universe

It's been established that the universe is expanding. This was first discovered by Edwin Hubble in the 1920s, who noticed that almost all the galaxies in the sky appeared to be moving away from us. And the more distant the galaxy, the faster it was moving away.

This isn't because the Earth is uniquely unpopular in the universe! Indeed, the whole of the universe is expanding uniformly – every cluster of galaxies is moving away from every other cluster. Wherever you may be in the universe, you will see galaxies moving away at a speed proportional to their distance from you.

To help you visualise uniform expansion, have a go at this practical activity.

Activity 2 An elastic analogy

Allow about 10 minutes

Find a piece of elastic, or an elastic band. Use a marker pen or ballpoint to make several dots along its length to represent galaxies in space. Then stretch the elastic out. As you do this, what do you notice?

Discussion

As you stretch the elastic, you should notice how the dots move apart. Each dot sees the neighbouring dots moving away at a speed proportional to their distance. The expansion of the universe works similarly. You can think of the galaxies being carried along by the expansion of space.

Figure 9 demonstrates how this works. In (a) 'galaxies' are marked along a piece of elastic. As the elastic is stretched (b) the galaxies move apart. In (c) all the distance have doubled as the 'universe' continues to expand. Note that each galaxy will see the others moving away at a speed proportional to their distance.

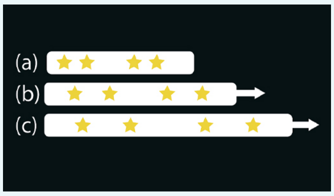


Figure 9 Analogy for the expansion of the universe

Today, this expansion is understood through the theory of general relativity, devised by Albert Einstein in the early 20th century. He demonstrated a direct connection between the average density of matter and energy in the universe and its fate. If the density is lower than a certain 'critical density', this is indicative of an 'open' and infinite universe which will expand forever. If, on the other hand, the density is greater than the critical value, this is indicative of a 'closed' and finite universe, where the expansion will eventually stop and go into reverse. Measuring the average density would tell us whether our universe is open or closed.

This critical density threshold is very small in itself – equivalent to just six atoms of hydrogen in a cubic metre. Until the late 1990s, the average density of all the matter we could see appeared to be somewhat less than that, indicating that the expansion would go on forever. This is still true when taking ‘dark’ matter into account, which is invisible to us but detectable from its gravitational pull. But the measured density was close enough to the critical density to present an intriguing puzzle.

Then, in 1998, two research groups discovered from observations of distant supernovae that the expansion of the universe was slower in the past than it is now. This is quite the opposite of the theoretical predictions – why should the expansion be speeding up? This implies that something is continuing to drive the expansion. That ‘something’ has been given the name ‘dark energy’. There’s not enough room in this course to discuss what that energy might be, but the idea is that it permeates all of space. (Oddly enough, Einstein proposed something very similar about a century ago but later discounted it as implausible.)

It’s now thought that dark energy accounts for about 70% of the density of the universe, with the other 30% being due to matter. Together they add up to a value very close to the critical density. But the accelerating effect of dark energy means that the critical density is no longer a simple criterion for whether the universe will keep expanding. The discovery of dark energy almost certainly points to a future of indefinite and accelerating expansion.

8 The future of the observable universe

It was established earlier this week that the observable universe – which is finite – is defined by the region of space in which light can travel to us since the Big Bang. What will that region look like in the future? This is quite tricky to think about. Several effects must be taken into account.

The hot plasma that is now visible to us as the background radiation – the effective edge of the observable universe, as discussed in Section 5 – will gradually cool, condense and give rise to stars and galaxies that will become visible to us. This will happen over the course of a few hundred million years.

At the same time, the 'edge' is moving outwards at the speed of light. Every year, the light from the background radiation is coming from further away and more galaxies are coming into view in front of it. The observable universe is getting bigger!

That would be true even if the universe were static. The next consideration is the expansion of the universe. As the galaxies continue to move apart (which you can think of as the expansion of space itself), the distances increase, and the observable universe becomes larger and emptier. At the same time, the light from these distant galaxies becomes weakened, moved towards the red end of the spectrum in a phenomenon known as the cosmological 'redshift'. As time goes by, galaxies will become fainter and further apart – even as more of them come into view.

On top of all that, the expansion is accelerating. That changes things in quite a dramatic way. In the distant future, due to the rapidity of the expansion, the galaxies will be receding so quickly, and the redshift will be so significant, that they will effectively fade from our view. They will still be within the bounds of our 'observable' universe, but we won't be able to see them!

Box 2 How to think of the cosmic expansion

Some of this week's concepts may be making your head hurt! This is one part of a relatively short, wide-ranging course, so it's not possible to explain every idea in depth. But here are a few thoughts that may ease your pain:

- The proper description of the expanding universe is a mathematical one derived from the theory of general relativity. The kind of maths required here is very difficult. But if you like to think visually, try this. Don't imagine the galaxies flying outwards through space – think instead of space expanding uniformly, carrying the galaxies with it. This picture helps to answer a couple of puzzling questions.
- Where did the Big Bang take place? You might have a picture of a gigantic explosion with material flying off in all directions into empty space. But this is misleading. If space is expanding, then the answer to 'where' the Big Bang took place is: everywhere. There is no centre to the universe, and no edge. Wherever you are, you will see the galaxies expanding away from you.
- What is the universe expanding into? All of space came into being with the Big Bang, so there is no 'outside' space for the universe to expand into – and indeed, no need for it. Space is part of the universe and takes part in the expansion.

9 Is the universe finite or infinite?

The current consensus, since the discovery that the expansion is accelerating, is that the universe will expand forever. But we're still no closer to knowing for sure whether the universe as a whole is infinite – the proximity to the critical density threshold discussed earlier means that the notion of a closed, finite universe is not entirely ruled out. If the universe were closed, what would it look like? And what does it actually mean for a universe to be 'closed'?

Einstein's theory of general relativity, which gives us the mathematical tools to make calculations about how the universe is expanding, also tells us that space can be curved. In this view, the gravitational pull of one object on another is due to the curvature of space, not a mysterious force acting at a distance. Einstein also showed that the whole universe could be curved due to the matter within it, and the curvature will determine whether the universe is finite or infinite.

9.1 Curved space

Curved space is hard to imagine, but you can use the analogy of a curved surface. A sphere is a familiar example of a curved surface and is said to have 'positive curvature'. If space is positively curved, like the surface of a sphere, then it is finite and we have a 'closed' universe. If space has zero curvature, like a flat surface, then it is infinite and we have a borderline 'open' universe.

Figure 10 Positive curvature

It is also possible for space to have a negative curvature, in which case it will also be infinite and open. An analogy for this is harder to visualise – but a popular potato snack comes close. You just have to imagine the edges extending to infinity!

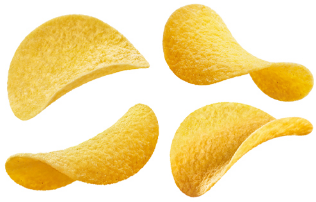


Figure 11 Negative curvature

A finite universe need not have an edge. Consider the surface of the Earth – it is finite, and its surface area can be measured precisely. And yet you can travel any distance across the Earth's surface, in any direction, and you'll never reach an edge. The same may apply if the universe is closed and positively curved. Despite having a finite and precisely measurable volume, you could travel around and never encounter an edge.

But this is not the whole story, of course. If you travel far enough across the surface of the Earth, you'll eventually return to the point from where you started. The same thing could happen in a closed universe. Given enough time, a starship sent out in a straight line to explore the depths of space could eventually find itself back on Earth. Now, we're not likely to embark on such an expedition anytime soon, but one thing does travel fast enough through space to manage it, and that's light.

9.2 What if the universe were finite?

This week began by asking how far we could see out into the universe. Let's imagine for a moment that the universe is indeed closed and finite, just like the surface of the Earth. By looking deep enough, could we – metaphorically speaking – see the backs of our own heads?

If the universe is finite but larger than the observable universe, then this wouldn't be possible, even in principle. Light from the back of our heads could not get back to our eyes within the age of the universe; the CMBR would get in the way.

On the other hand, some researchers have speculated that the universe might actually be smaller than the observable universe. In this case, the Hubble deep fields might reveal the same patterns of galaxies – including our own Milky Way – recurring at increasing distances. This would mean that we were seeing the same regions more than once between us and the background radiation, albeit at different stages in their evolution. But so far, there is no evidence of such patterns. Video 3 brings together and illustrates some of this week's concepts.

Video content is not available in this format.

Video 3 The size of the universe

In the end, we can be confident that the universe is expanding, but perhaps we can never be certain whether it is infinite. It may well be one of those things we cannot know.

10 This week's quiz

Well done for reaching the end of Week 4.

Now it's time to complete the Week 4 badged quiz. It's similar to previous quizzes, but this time instead of answering five questions there will be fifteen, covering material from the first four weeks of the course.

Week 4 compulsory badge quiz

Remember, this quiz counts towards your badge. If you're not successful the first time, you can attempt the quiz again in 24 hours.

Open the quiz in a new window or tab then come back here when you've finished.

11 Summary of Week 4

This week you've learned how our view of the universe is affected by the speed of light – looking out into space means looking back in time – and how the age of the universe limits the size of the observable universe. No knowledge of objects at greater distances is possible, as their light hasn't had time to reach us.

The effective boundary of the observable universe is the cosmic background radiation which originated about 380 000 years after the Big Bang. It covers the whole sky like a fog, preventing us from seeing any further. Observations of galaxies within the observable universe show not only that the whole universe is expanding, but that the expansion is accelerating. Einstein's theory of general relativity tells us that such a universe will expand forever.

General relativity has been fabulously successful, not only in modelling the universe, but in many other areas of astronomy and physics. But it cannot explain everything. Next week will probe the limits of the theory, starting by considering what it says about time, and then by exploring a place where this all gets broken apart.

You are now halfway through the course. The Open University would really appreciate your feedback and suggestions for future improvement in our optional end-of-course survey, which you will also have an opportunity to complete at the end of Session 8. Participation will be completely confidential and we will not pass on your details to others.

You can now move on to Week 5.

Week 5: Time

Introduction

In Week 4 you looked at the limits of space, and whether the universe could be infinite. This week, you will look more closely at what's meant by 'time' and how it's connected to space. Then you'll look more deeply (very deeply, in fact!) into one place where it all changes, and we can already outline what we cannot know.

Here's Marcus to introduce this week's topic.

Video content is not available in this format.

Video 1 Introduction

By the end of this week, you should be able to:

- appreciate that the passage of time is related to how quickly one is moving
- understand the implications of Einstein's theory of general relativity on gravity, space and the flow of time
- outline some of our current knowledge of black holes, their structure, and what happens inside them
- describe how general relativity clashes with quantum mechanics at the centre of a black hole, and what this means.

1 What is time?

Most people probably don’t think too much about what time ‘is’. It just flows. Ray Cummings, the American science fiction writer, commented that time ‘is what keeps everything from happening at once’, which may be as good a definition as any.

Whatever it is, time certainly seems quite different from space. Time has a direction – it moves forward, and there’s no choice but to move forward with it – whereas in space you can choose to move in any direction. Time is measured with a clock; space is measured with a ruler. An event can be assigned and described by an instant in time and a location in space, but the two are quite distinct. This was the scientific view too, until about a century ago. This week, you’ll learn how time and space are no longer regarded as separate, but as inextricably intertwined.

2 How time is measured

Time cannot be discussed for long without an understanding of how it is measured. How long is one second? If there are 60 seconds in a minute, 60 minutes in an hour and 24 hours in a day, then a second could be defined as 1/86400 of a day. But then, how long is a day?

Traditionally that question was answered by astronomers. For many centuries they had the job of measuring the length of the day (strictly the 'mean solar day') by observing the apparent motions of the Sun and stars as the Earth rotated. But by the 20th century, it was becoming clear that that the rotation of the Earth was not completely smooth, and the length of the day was not constant. So what does that mean for the measurement of time?

A way forward appeared in the 1950s with the invention of the first practical atomic clocks. Rather than relying on pendulums or electronic oscillators – as the best clocks then did – atomic clocks count the vibrations of microwave radiation from atoms (in particular the caesium atom).

Figure 1 shows the world's first practical atomic clock, built at the UK National Physical Laboratory under the leadership of Louis Essen, who had earlier made accurate measurements of the speed of light. It is now at the Science Museum in London.

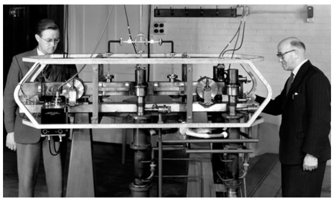


Figure 1 Louis Essen (right) and the first practical atomic clock

With these clocks, physicists soon discovered that 'atomic time' is far more stable than time kept by the rotating Earth, or even the best conventional clocks. What is known as International Atomic Time (or TAI, from the French *Temps atomique international*) officially began on 1 January 1958 and has been running steadily ever since.

In 1967, the second was redefined as the duration of 9 192 631 770 periods of the radiation from the caesium atom. The definition, part of the International System of Units (or SI, from *Système international*), was chosen to match the duration of the astronomical second as measured by atomic clocks in 1958. Since 1972, a time scale derived from TAI, known as Coordinated Universal Time (UTC), has become the basis for all civil and scientific timekeeping. UTC attempts to keep in step with the rotation of the Earth by introducing an occasional 'leap' second. At time of writing in 2022, UTC is 37 seconds behind TAI.

But there is no one atomic 'master clock'. TAI is actually the average of more than 450 atomic clocks maintained by laboratories around the world, with the duration of the second determined by about a dozen ultra-precise machines designated as primary frequency standards. TAI is occasionally 'steered' to ensure that the length of the second remains as close to the formal definition as it possibly can. The best atomic clocks will gain or lose no more than a second in 100 million years.

So, what is time? A pragmatic definition is that time is what is measured by clocks – and the most accurate time is TAI, measured by atomic clocks.

3 Speed of light revisited

With the length of the second firmly tied to atomic physics, rather than the rotation of the Earth, surely time is now well and truly nailed down? Unfortunately, no it isn't! To find out why, you will need to revisit the speed of light.

Suppose you're a police officer wanting to use a radar or laser gun to measure the speed of traffic. It's no good using the gun when you're driving along, as that would only give you the relative speed between your car and the other cars on the road. To measure the speed of a suspect vehicle correctly, you must be standing motionless beside the road.

Watch this short video in which Marcus introduces relative speed, then have a go at the calculations yourself in Activity 1.

Video content is not available in this format.

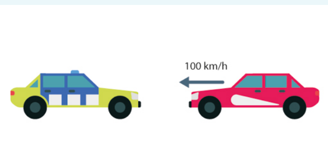
Video 2 Relative speed (part 1)

Activity 1 Relative speeds

Allow about 10 minutes

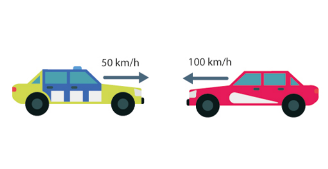
With that in mind, imagine the three scenarios below. What is the relative speed in each case?

(a) A stationary police car at the roadside observes a car travelling at 100 km/h towards it.



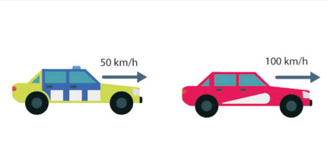
- 50 km/h
- 100 km/h
- 150 km/h
- 200 km/h

(b) A police car travelling at 50 km/h observes a car travelling at 100 km/h towards it.



- 50 km/h
- 100 km/h
- 150 km/h
- 200 km/h

(c) A police car travelling at 50 km/h observes a car travelling at 100 km/h away from it.



- 50 km/h
- 100 km/h
- 150 km/h
- 200 km/h

Discussion

Relative speeds are calculated by adding the two speeds when they are in opposite directions and by subtracting when they are in the same direction. Therefore the relative speeds, as measured by the radar gun, are calculated as follows:

a. 0 + 100 = 100 km/h
b. 50 + 100 = 150 km/h
c. 100 - 50 = 50 km/h

These results line up with the intuition that a head-on collision has a much higher impact than hitting a car driving in the same direction. But only in the first scenario above does the relative speed match the actual speed of Car B.

Returning to the speed of light: where should we stand to measure it? Again, imagine three scenarios:

a. A stationary physicist measures the speed of a light beam travelling toward them.

b. A physicist moving at half the speed of light measures a light beam travelling toward them.
c. A physicist moving at half the speed of light measures a light beam travelling away from them.

Following on from the police car scenarios, you might imagine that the relative speeds here would be 1.0, 1.5 and 0.5 times the speed of light. This would seem logical, but it's not right! In all three cases the physicist would register the speed of the light beam as exactly the same value, measuring it at 3×10^8 m/s. As this is a special value which is used all the time in science, it is generally written as just *c*.

Here's Marcus again to discuss this development.

Video content is not available in this format.

Video 3 Relative speed (part 2)

This will all begin to make sense again – but not before you reconsider the nature of time.

4 Your time is not my time

The consequences of Einstein’s insight are profound. This section will look at just one of them, namely that time flows differently for different observers. The following video shows how an imaginary ‘light clock’ can be used to illustrate this rather strange idea.

Video content is not available in this format.

Video 4 60-Second Adventures in Astronomy: Special Relativity

To examine this a bit further with the use of some maths, have a go at the following activity.

Activity 2 The ‘light clock’ thought experiment

Allow about 10 minutes

Imagine you have two mirrors, set four metres apart. A pulse of light bounces back and forth between the two, as shown in Figure 2.

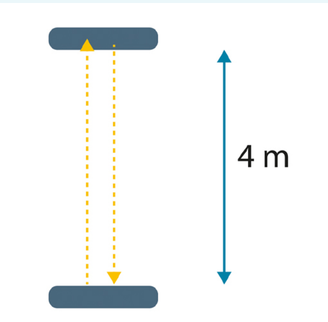


Figure 2 Two mirrors set apart

If the speed of light is 3 × 10^8 m/s, how long will it take the pulse to make the round trip of 8 metres?

Provide your answer...

Answer

$$\text{time} = \frac{\text{distance}}{\text{speed}} = \frac{8\text{m}}{3 \times 10^8 \text{m/s}} = 26.7 \times 10^{-9}\text{s}$$

This round trip of 8 metres is completed in about 27 billionths of a second, or 27 nanoseconds. By keeping count of the number of round trips, as the light 'ticks' between the mirrors, you've got a ticking clock!

Meanwhile, an astronaut on a passing spaceship (here moving at 60% of the speed of light) watches our clock as she zooms past. Her view at three instants (Figure 3) is rather different. Because she sees the mirrors in motion, the pulse of light now has to travel a longer distance, which can be calculated.

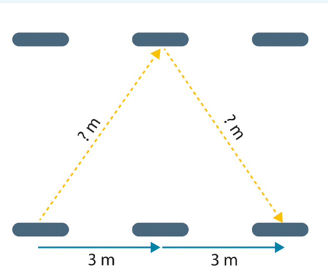


Figure 3 An astronaut passes the clock

In the time for a return trip, the mirrors have moved by 6 metres. How far does the pulse now have to travel? Hint: note that the pulse is travelling along the hypotenuse of two identical right-angled triangles.

Provide your answer...

Answer

Pythagoras' theorem tells us that the square of the hypotenuse is equal to the sum of the squares of the other two sides:

$$a^2 + b^2 = c^2$$

So for the first triangle, the pulse travels the following distance:

$$\sqrt{3\text{m}^2 + 4\text{m}^2} = \sqrt{25} = 5 \text{ metres}$$

The round trip distance, then, is 10 metres.

How long does the pulse take to travel 10 metres?

Discussion

The speed of light for the astronaut is still 3×10^8 m/s, so the time for the round trip is now

$$\text{time} = \frac{\text{distance}}{\text{speed}} = \frac{10\text{m}}{3 \times 10^8 \text{m/s}} = 33.3 \times 10^{-9}\text{s}$$

This is approximately 33 nanoseconds.

In these two scenarios, two people observe the same clock. One sees it tick every 27 nanoseconds, while the other sees it tick every 33 nanoseconds. A moving clock ticks more slowly than a stationary clock.

Although this activity used an imaginary light clock, this difference has nothing to do with the construction of the clock. It stems from the nature of space and time. This effect is known as ‘time dilation’ and it means that, in general, no two observers will agree on their measurements of time. But it gets stranger still. Because it’s the relative motion that matters, you would see an identical clock on the moving spaceship running slow for the very same reason.

That might be surprising (and even hard to believe), but time dilation is firmly within the realm of what we can and do know about time, and won’t be explored in more depth here. (If the topic intrigues you though, you may like to do some further independent learning – see what you can find out about the ‘twin paradox’.)

5 Curved space and gravity

Einstein saw that if motion through space affects measurements of time, then space and time must no longer be considered separate. In modern physics, space and time are now regarded as being two aspects of a four-dimensional entity called ‘spacetime’. Spacetime has three dimensions of space, and one of time.

Keep in mind though, this doesn’t mean that space and time are equivalent – there is a common misconception that time is a ‘fourth dimension’ of space – but rather that they intermingle with one other. Special relativity says that space and time are still different ‘things’ but what appears to you as time may appear to someone else as space, and vice versa, depending on how you are moving. If that wasn’t enough, Einstein went much, much further. In 1915 he published his theory of general relativity, which was an attempt to explain gravity.

In special relativity, spacetime is flat, rather like a sheet. As you saw last week, general relativity tells us that space can be curved, and the curvature is what we perceive as gravity. Mass distorts spacetime, curving the space in its vicinity.

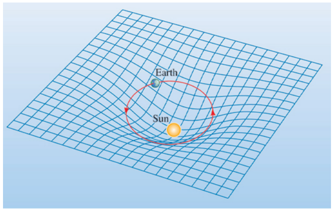


Figure 4 The mass of the Sun curves the space around it

In light of this, Earth orbits the Sun not because the Sun is pulling on the Earth, but because the Sun is curving the space around it so that the Earth follows a closed orbit rather than a straight line. The Earth is responding to the local curvature of space – not a pull from a distant Sun. Here’s Marcus to discuss these developments in our understanding.

Video content is not available in this format.

Video 5 Curved spacetime (part 1)

Evidence for general relativity came quickly. Einstein predicted that light from distant stars passing through the curved space close to the Sun would be bent, making the stars appear to be further from the Sun than they really are. But in terms of verifying this prediction, the Sun is so bright that stars cannot be seen anywhere near it – except during a total eclipse, when the Sun is completely covered by the Moon.

Sir Arthur Eddington, a prominent astrophysicist of the day, led a project to measure this tiny displacement during the total eclipse of 29 May 1919, when stars became briefly visible in the darkened sky close to the Sun. Figure 5 shows a photographic negative taken by Eddington's team. The white disc is the Moon covering the Sun. The two stars indicated by faint horizontal lines are among those seen to be displaced by the curvature of space. Eddington's results matched Einstein's prediction, as has every other test of general relativity in the 100 years since.

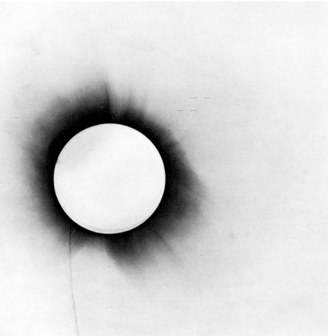


Figure 5 A photographic negative of the eclipse of 1919

6 Time and gravity

The theory of special relativity revealed that time flows at different rates for different observers according to their relative speeds. The general theory goes further: time also flows at different rates where the strength of gravity differs – strong gravity makes clocks run slow. This is known as 'gravitational time dilation' and is quite distinct from time dilation due to relative speeds.

Video content is not available in this format.

Video 6 Curved spacetime (part 2)

The effect was first demonstrated in a laboratory experiment in 1959, and in a high-altitude rocket experiment in the 1970s. The theory predicts that clocks will run faster on top of a mountain than they would in the valley below, as the top is further from the centre of the Earth and therefore gravity is weaker. Of course, the difference here is very small. Near the surface of the Earth a clock will gain just 9.4 nanoseconds a day for each kilometre of height. But that's enough to be taken into account when readings of atomic clocks around the world are combined to form International Atomic Time.

Such tiny differences might not sound too significant, but they have practical consequences. For example, they have an impact on satellite navigation. The satellites of the Global Positioning System (GPS), used by satnavs and smartphones, orbit the Earth at a height of around 20 000 kilometres. At that distance, the gravitational field is weaker than that experienced on the surface of the Earth, and consequently their onboard clocks gain 45 microseconds (millionths of a second) a day compared to clocks at sea level. On the other hand, because of their orbital speed, the clocks lose 7 microseconds a day. Together these effects – from general and special relativity – cause the satellite clocks to gain a net 38 microseconds a day compared to clocks on the ground. Fortunately, when your smartphone computes its position from signals received from GPS, these different rates are taken into account. If relativity were ignored, you would literally not know where you were.

Having touched on the rudiments of general relativity – curved spacetime and gravitational time dilation – it's time to explore a place where it's all going to go wrong.

7 Black holes

You've probably heard of black holes. Put very briefly, a black hole is an object – perhaps a collapsed star – so massive that not even light can escape its gravitational pull. The idea dates back to the late 18th century, but the theory was only properly developed with the advent of general relativity.

The radius of a black hole (known as the Schwarzschild radius) is the distance from its centre at which its escape velocity – the minimum speed necessary for a body to escape the gravitational pull – is equal to the speed of light. For example, if the mass of the Sun were compressed into a black hole, its Schwarzschild radius would be just 3 kilometres. For the Earth, it would be just 9 millimetres. It's named after Karl Schwarzschild, a German astronomer who was one of the first scientists to use general relativity to understand the gravitational field around stars.

The imaginary surface surrounding the black hole at the Schwarzschild radius is known as the event horizon. Objects can fall inwards through the event horizon but nothing can ever come out, because nothing can move through space faster than the speed of light. That means that once a black hole forms, perhaps from the collapse of a massive star at the end of its life, it can draw in other material and grow to be very large indeed. Some black holes are believed to have a mass several billion times that of the Sun.

So, light cannot escape a black hole as the escape velocity is too high. By comparison, the escape velocity of the Earth is 11 km/s. An object thrown into space faster than this will never fall back down again. Now, this might suggest an inaccurate mental image of light trying to get out of a black hole and falling back in. That's not what happens. Think instead of the space around a black hole flowing into the hole – much like water going down a drain, but from all directions. The flow gets faster closer to the hole, reaching the speed of light at the event horizon. Light emitted at the event horizon travels through space at the speed of light, but space is flowing into the hole at the same speed, so the light makes no outward progress. As nothing can move through space faster than light, anything (or anyone) that crosses the event horizon is swept down the hole with no hope of escape. Figure 6 demonstrates this idea visually.

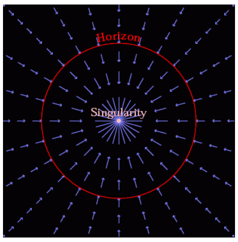


Figure 6 Space flowing into a black hole

The physicist Brian Cox develops a similar analogy in the following video, explaining how a waterfall helps us visualise what happens at the event horizon of a black hole.

Video content is not available in this format.

Video 7 Wonders of the Universe – Black Holes

7.1 A picture of a black hole

You might think that a black hole, being black and absorbing all the light that falls on it, would be invisible. In a sense that's true. But there are ways to 'see' a black hole.

Since the 1950s, astronomers have been intrigued by distant, powerful sources of radio waves known as radio galaxies. It's now known that these are otherwise normal galaxies with a supermassive black hole – billions of times heavier than the Sun – in the centre. Material from the galaxy is accelerated to high speed by the intense gravity and heated as it swirls around the black hole, before ultimately disappearing through the event horizon. Vast amounts of energy are released in the process (in the form of x-rays and radio waves), and jets of material are often thrown out in opposite directions. So, although the black hole itself is not visible, the violent activity in its surroundings gives away its location. Figure 7 shows an example of this – the radio galaxy Cygnus A lies at the centre of this image showing radio emission. Two high-energy jets are ejecting plumes of material from a black hole at its centre.

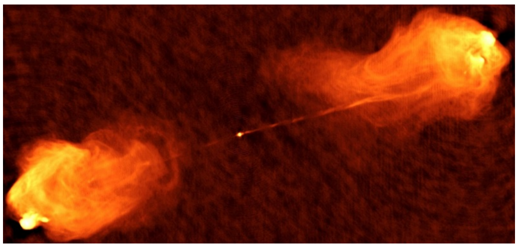


Figure 7 Cygnus A

It's now thought that most, if not all galaxies – including our own Milky Way – have a massive black hole at the core. Fortunately for us, the Milky Way black hole, with a mass about 4.5 million times that of the Sun, appears to be sitting quietly without any sign of such violent outbursts.

In 2019, astronomers published the first picture that revealed the event horizon of a black hole (Figure 8). It was made by eight radio telescopes around the world working together to form the Event Horizon Telescope.

This black hole, in the centre of an active galaxy called M87, has a mass about 6.5 billion times that of the Sun, and the event horizon is 40 billion kilometres in diameter. The central dark area in the picture is the shadow of the black hole itself, extending to about three times the radius of the event horizon. Radiation from the extremely hot gas swirling around the hole is distorted into a ring shape by the strong curvature of space in the vicinity.

Figure 8 The black hole in the centre of the galaxy M87

Before this image, astronomers were reasonably confident that the black hole theory of active galaxies was correct. The image confirmed it, along with the predictions of general relativity.

None of this makes the idea of entering a black hole sound particularly appealing. But what would actually happen if you did?

7.2 Falling into a black hole

The event horizon is the 'public face' of a black hole, a one-way surface through which space flows at the speed of light. But it's a blank face that gives nothing away. What lies inside the event horizon? And what happens at the centre?

One way to find out would be to cross the event horizon and fall into the hole. Once you're inside though, you would never be able to tell anyone at home what you'd discovered. What would happen to a plucky astronaut who took on that challenge? It would depend on whose view you take – the astronaut or a distant observer.

The gravity around a black hole is so strong that time passes noticeably slowly, even to a distant observer, in accordance with gravitational time dilation. The closer to the event horizon, the slower it passes. The observer far away from the strong gravity will see the astronaut approach the event horizon ever more slowly but never get there. At the horizon itself, as judged by the distant observer, time will appear to stop.

The astronaut's experience, on the other hand, is very different (and quite unpleasant). Eventually the tidal forces near the hole will become so strong the astronaut will be shredded by stretching from head to toe and squeezing from side to side (astrophysicists have coined the charming term 'spaghettification' for this process). The point at which that happens will depend on the mass of the hole. Approaching a smaller hole, the astronaut will be shredded well before reaching the event horizon, so they'll never learn what lies within. But for supermassive holes, like those in active galaxies, the tidal forces are relatively gentle near the event horizon, and it would be possible to fall through unscathed. Indeed, our astronaut wouldn't notice anything unusual when crossing the horizon. There's no marker to warn of the point of no return. But once they've made it within the event horizon, there's no way to transmit any discoveries back to the outside world. And about a tenth of a second before arriving at the centre, spaghettification awaits.

Figure 9 The process of ‘spaghettification’

8 The end of time?

Until now, general relativity has been spoken of in glowing terms. It accounts for the motions of celestial bodies better than Newton's theories. It has passed every test that physicists have devised to try to prove it wrong. But inside the event horizon, this theory is running out of space.

At the centre of a black hole, general relativity predicts that the entire mass of the collapsed star will be compressed into a mathematical point – a 'singularity'. Density will become infinite as will the curvature of space. Physicists don't like quantities becoming infinite, so what does this mean? Has something gone wrong with general relativity?

So far, this week has only considered general relativity and ignored quantum mechanics, which you met in Week 3. Each of these theories has its own realm. General relativity is important for very large masses; quantum mechanics is important at very small sizes. But what happens when a very large mass has a very small size? We don't yet have a theory of 'quantum gravity' that would tell us what happens when general relativity and quantum mechanics clash, though there have been a number of contenders. For that reason, we cannot say what really happens at the centre of a black hole. But we can make some estimates on where and when such a theory might be necessary.

Theoretical physicists have estimated a distance, the Planck length – you came across this in Week 3 – at which general relativity must break down. It's very roughly where the Schwarzschild radius of an object becomes equal to its quantum wavelength and is about 1.6×10^{-35} metres. The time it takes light to travel that distance, the Planck time, is about 5×10^{-44} seconds.

The Planck length and the Planck time represent fundamental limits on our knowledge of space and time. It doesn't really make sense to talk about distances shorter than the Planck length, or times shorter than the Planck time. So, we ultimately don't know – and for the time being we cannot know – what happens when an astronaut falling into a black hole encounters the singularity at the centre. But if you are already spaghettified, that probably won't matter much to you.

9 This week's quiz

Well done for reaching the end of Week 5. Check what you've learned by taking the end-of-week quiz.

Week 5 practice quiz

Open the quiz in a new window or tab then come back here when you've finished.

10 Summary of Week 5

This week, you've been thinking about time. Einstein's theory of special relativity holds that everyone will measure the speed of light to have the same value, no matter how fast they are moving. It implies that time flows at different rates for different observers. His theory of general relativity meanwhile explains gravity as a curvature of space, and shows that time flows more slowly in a strong gravitational field.

A black hole is a region of space where gravity is so strong that not even light can escape. At the centre of a black hole, general relativity clashes with quantum mechanics, and space and time lose their identity. There's not yet a theory that helps us understand what will happen in these conditions.

Next week, you will be leaving the realm of physics and exploring the boundaries of what is going on inside our own heads. What is consciousness? And where can the study of neuroscience lead us?

You can now move on to Week 6.

Week 6: Consciousness

Introduction

In the past weeks you have explored the furthest reaches of our universe, considered the beginning of time and whether it is possible to predict the future, and delved deep into the smallest constituents of matter. The thing that makes this entire remarkable journey possible is the human brain. So, our destination now is inside our own heads, to ask how this lump of cells, proteins and water manages such extraordinary feats of thought. This week, your brain will be learning about itself, as you encounter some very intriguing questions.

How are the workings of our brain connected to our personalities, our sense of ourselves and our own inner world of experiences, thoughts and feelings? Could equations be found to describe 'us'? Or is our own consciousness something that just can't be explained by science?

As Francis Crick, Nobel laureate for his work on the structure of DNA, said:

There is no scientific study more vital to man than the study of his own brain. Our entire view of the universe depends on it.

Here's Marcus to introduce this week's topic.

Video content is not available in this format.

Video 1 Introduction

By the end of this week, you should be able to:

- understand how scientists study the brain
- explain how the brain works in simple terms, including the role of neurons
- discuss some technological advances related to simulating brain function
- appreciate the difficulties involved in studying and understanding consciousness
- describe some current research ideas on consciousness.

1 What is consciousness?

The human brain is a truly remarkable machine, but unlike a computer it also produces a sense of self – it is conscious.

This short video clip poses some of the compelling questions that you'll explore this week.

Video content is not available in this format.

Video 2 Introducing consciousness

1.1 What are we actually talking about?

In order to examine a concept in science, a very good place to start is by defining it. But while you might have a good intuitive sense of what it means to be conscious, it turns out that reaching a clear definition is not so straightforward. Before going any further, think about the meaning of the word 'conscious'.

Activity 1 The meaning of the word

Allow about 5 minutes

Can you think of some examples of how we use the word 'conscious' in everyday language? Make some brief notes.

Provide your answer...

Discussion

The word has numerous uses and meanings, such as the following:

> knocked unconscious / regaining consciousness

> being self-conscious
> a conscious decision
> a subconscious fear
> an unconscious bias
> collective consciousness
> raising consciousness.
>
> You may well have thought of some others.
>
> These terms all relate to various aspects of being awake, and having awareness of ourselves and our thoughts. They derive from changing historical meanings of the word ‘conscious’.

The question of consciousness has historically been a philosophical one. It’s only quite recently that it’s emerged as an area that science can even begin to examine. This fundamental change has come about through the development of new ways of studying the brain, which you will look at shortly. One important thing to know up front is that this new field of study has yet to develop fully consistent terms, because it’s still unclear precisely what it is we’re talking about!

Broadly speaking, ‘consciousness’ encompasses having thoughts, perceptions, feelings and awareness. But are these different aspects of a single phenomenon, or are there different types of consciousness? That isn’t clear yet either. A precise and universally accepted definition will probably have to wait a while, as more continues to be learned.

The idea of ‘conscious experience’ is a major focus in current scientific research. Our experiences have a particular internal ‘feel’ or ‘character’ to them. These are subjective and can be quite hard (or even impossible) to describe, and yet they’re distinctive and instantly recognisable. For example, when we see something vivid red in colour, our brain doesn’t just ‘tick the red box’ as a robot might. There is an associated quality that we experience internally – and this might vary quite substantially from one individual to another. So, how could that sensation of perceiving ‘red’ be described to someone who has never experienced colour before? How can you know that your experience of red is the same as anyone else’s? These qualities that make up a subjective experience are often called ‘qualia’; some other examples would be things like the smell of freshly ground coffee, the pain of stubbing a toe, and the thrill of riding a roller coaster. This aspect of consciousness is broadly referred to as ‘phenomenal consciousness’.

Note that ‘self-consciousness’ has two different meanings. In psychology, it commonly refers to the preoccupation with one’s own appearance or actions that’s typically heightened during teenage years. It also describes our concept of ‘I’ as a distinct individual.

The term ‘self-aware’ might be used to describe some similar ideas; we’re aware of our own existence in a way that, for example, our television isn’t. We all have this feeling of being an individual person – with our own thoughts, hopes and desires – and observing the external world from behind our eyes.

That clears all that up then! Well, maybe not, but it’s a start. Establishing some terms and fundamental questions paves the way forward in examining consciousness.

1.2 The ‘hard problem of consciousness’

Why do we even have these subjective conscious experiences? Can they possibly be explained by what happens physically inside our brains? This is often regarded as the ‘hard problem of consciousness’. It’s an intriguing question. As you found in the previous section, it’s hard enough defining consciousness, let alone explaining it!

(To be clear, there are ‘easy problems’ too – relatively speaking! These problems involve cognitive functions. For example, identifying the mechanisms in our brain involved in processing visual information and how we react to that information, rather than how and why we have a subjective experience of seeing something.)

It certainly seems very mysterious that these experiences could all come from purely physical processes. How does whatever it is that makes ‘us’ arise from the action of biological cells in the brain? After all, these brain cells are much like any other cells that make up our bodies and the rest of the living world.

While we don’t yet have a universal notion of consciousness, let’s look at how it might be studied – and how we might tackle the so-called hard problem.

2 Inside the brain

Consciousness must have some connection with the processes in the brain. This is well established, because aspects of consciousness are dependent on the brain, as shown by brain scans, and the effect of things such as anaesthetics or brain damage.

So, let’s begin by taking a tour of the brain.

2.1 Anatomy of the brain

There’s a reason our brains look the way they do. The distinctive ridges and grooves of the outer brain which you can see in Figure 1 greatly increase its total surface area, allowing billions of cells to be contained within the tight confines of the skull. The cerebrum is the largest and most highly-developed part of the brain where conscious thought takes place (Figure 1a). Internally, the cerebrum is made up of two hemispheres (Figure 1b). These are largely separate, but communicate with each other via a large bundle of nerves called the ‘corpus callosum’ (Figure 1c).

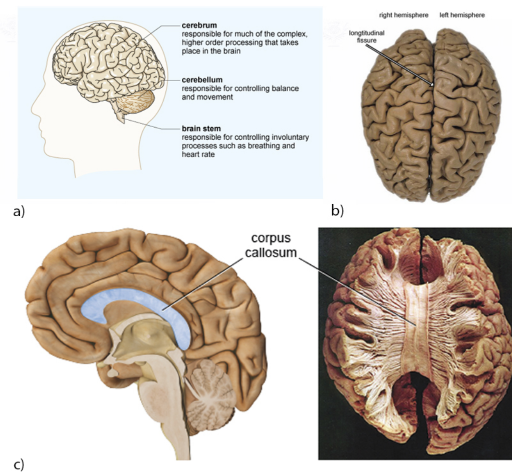


Figure 1 (a) the human brain viewed from the side and (b) in a photograph from above. (c) Left: a cut-away view through the middle of the brain. Right: A post-mortem human brain sample with top layer removed.

Activity 2 External structures of the brain

Allow about 10 minutes

In the following video you will learn a little more about the main structures of the brain and their functions, including the four main lobes of the cerebrum.

Watch the video and then attempt the following questions.

Video content is not available in this format.

Video 3 External structures of the brain

Which lobe is largely responsible for language and reasoning?

- ○ Brain stem
- ○ Frontal lobe
- ○ Occipital lobe
- ○ Parietal lobe

Which part of the brain is responsible for unconscious bodily functions such as controlling breathing and heart rate?

- ○ Brain stem
- ○ Frontal lobe
- ○ Occipital lobe
- ○ Parietal lobe

Is the white matter located in the outer region or inner region of the brain?

- ○ Inner region
- ○ Outer region

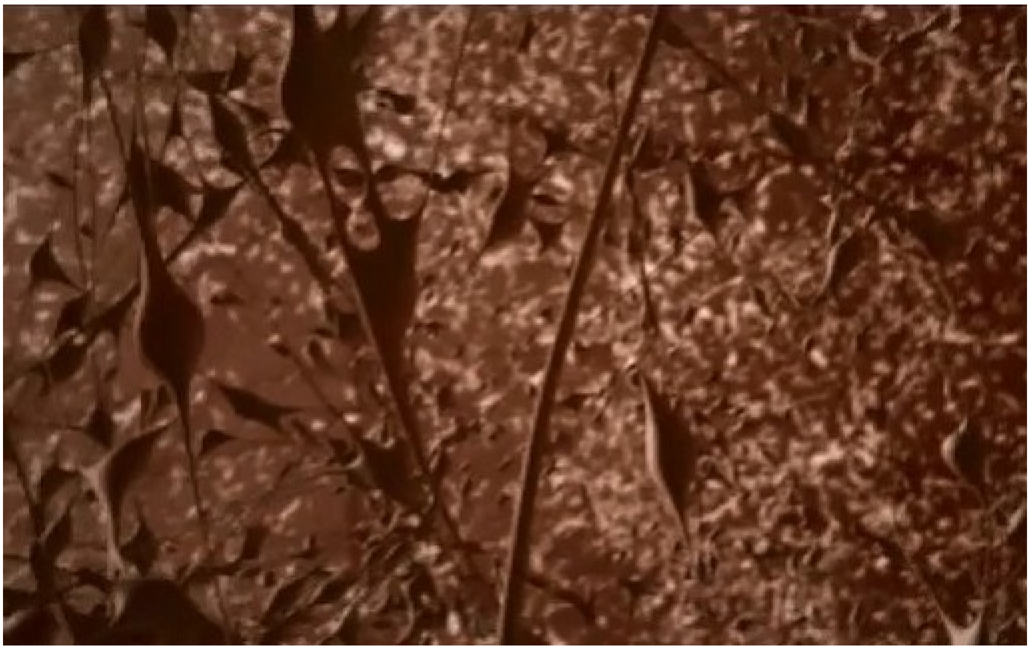
Now you'll delve deeper into how the brain actually works.

2.2 Neurons and neurotransmitters

Meet 'neurons'. These are what would colloquially be referred to as 'brain cells', although the brain is made up of many other types of cells as well. This short video introduces the cells' structure and function.

Video content is not available in this format.

Video 4 Neurons



As you've just seen, neurons are special cells in the brain which transmit electrical pulses. This is how the different parts of the brain 'talk' to each other, and indeed to the rest of the body via other neurons in the 'nervous system'. You can think of this as being like the body's electrical wiring.

Neurons in the brain form a communication network, as illustrated in this short video clip.

Video content is not available in this format.

Video 5 Neuron networks

The diagram in Figure 2 illustrates a typical neuron. The microscopic cell body contains important cellular components, such as the nucleus that houses genetic material.

Typically, the branching dendrites receive incoming messages from neighbouring neurons, which are then integrated together. If the neuron 'fires', the axon carries the signal as an electrical pulse to the axon terminals. The length of an axon varies enormously, from a fraction of a millimetre to a metre or more. The longest axons in humans belong to neurons in the sciatic nerve, which connects the base of the spine to the toes.

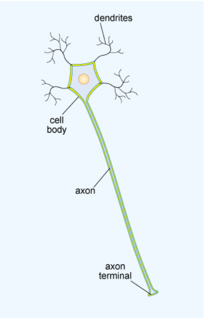


Figure 2 The main structural features of a typical neuron

The connections between neurons are called 'synapses'. In most cases, a signal is transmitted across the small gaps between neurons by chemicals called neurotransmitters, rather than by electrical pulses. Some neurotransmitters act to 'excite' neurons while others have an inhibitory effect, but if the overall chemical signal is strong enough then adjacent neurons will 'fire'. There are many different neurotransmitters, but you may be familiar with a few names already, like adrenaline, endorphin, dopamine and histamine.

The diagram in Figure 3 shows one neuron releasing chemical neurotransmitters, triggered by the electrical pulses travelling along the axon. The surface membrane of the next neuron contains specific receptors for these neurotransmitters, acting somewhat like a lock-and-key mechanism. These chemical messages can be passed to the branching dendrites of more than one neuron, like in the simple network shown in Figure 4.

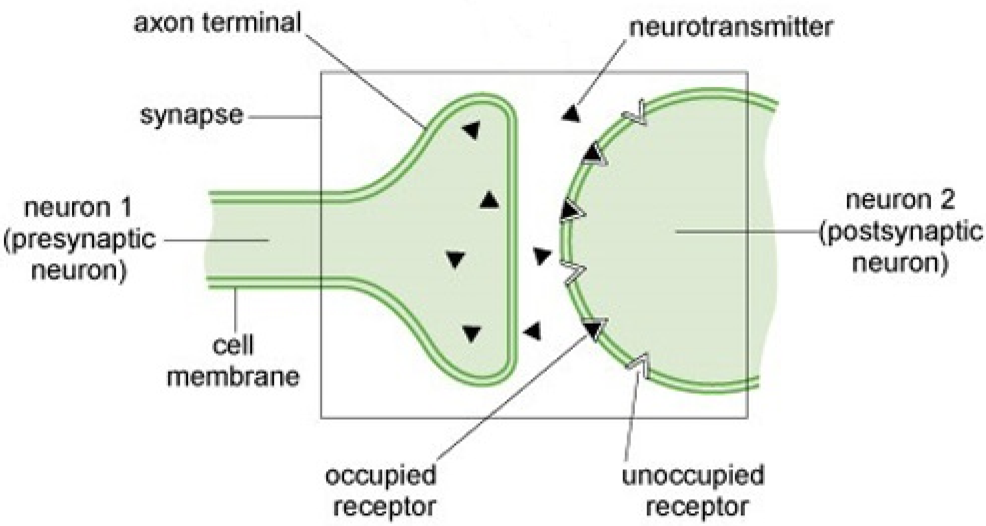


Figure 3 A schematic representation of a synapse

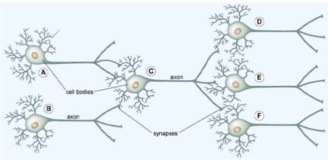


Figure 4 A simple network of neurons, in which information is passed from left to right

Some neurons can ‘pass on the message’ to hundreds or even thousands of other neurons in this way. With around 100 billion neurons in the brain – similar to the number of stars in an entire galaxy – this makes many trillions of connections in total! And it isn’t just the number of connections which is astounding, but the number of messages which are being sent – a neuron can fire off hundreds of messages each second. Video 6 visually demonstrates a few of this section’s concepts.

Video content is not available in this format.

Video 6 How neurons communicate

3 Studying the brain

Neuroscience is the study of the brain and the nervous system. The earliest studies involved people with brains that had been damaged. Correlating the damage with observed impairments allowed scientists to build up a map of the brain, noting which regions are responsible for different tasks.

A pioneer in this field was a French physician, Pierre Paul Broca. In 1865, he published his results from studying twelve patients who all had difficulty in articulating speech, but who were otherwise able to comprehend language. His autopsies showed damage in the same area of the brain, suggesting its link to speech production. This region is now called Broca's area.

Figure 5 Broca's area (shown in red)

While specific brain regions are principally devoted to certain functions (as seen in Figure 6, and in Video 3 earlier), neuroscientists today are discovering that complex functions require many different regions across the entire brain to work together, and that these locations in the brain are more flexible than previously believed.

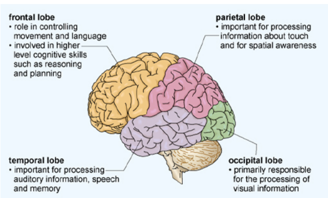


Figure 6 Principal functional regions of the human brain – the four lobes of the cerebrum

So far, the brain has been discussed on the macroscopic scale. But what about the microscopic scale?

3.1 Microscopy and silver nitrate staining

Brain tissue can be treated with chemicals to stain some of the individual neurons, allowing their finely detailed structure to be seen with a microscope. This technique was discovered in the 1870s by Camillo Golgi, and later used and improved by Santiago Ramón y Cajal. It enabled major advancements in neuroscience, particularly in identifying neurons as the key to brain function. The biologists were awarded the Nobel Prize 'in recognition of their work on the structure of the nervous system' in 1906. Figure 7 shows 'Golgi stained' brain tissue viewed under a microscope, with Figure 7a revealing the structure of individual neurons, and 7b showing the patterns of neurons in a section of a mouse brain.

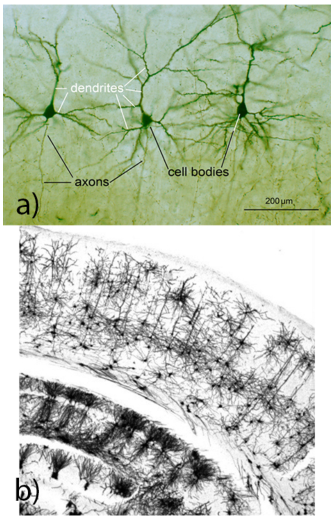


Figure 7 ‘Golgi stained’ brain tissue viewed under a microscope

Figure 8a shows a single ‘Purkinje’ neuron, one of the largest cells in the brain, with an elaborate structure of branched dendrites. Figure 8b shows interconnected neurons in the cortex, appearing in layers.

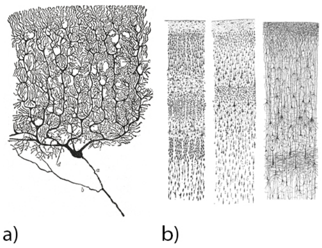


Figure 8 Neurons in ‘Golgi stained’ brain tissue (drawings by Santiago Ramón y Cajal)

The electron microscope was invented in the 1930s and developed over subsequent decades. Rather than light, a beam of subatomic electrons is used to image objects, allowing for much greater magnification, and visualisation on the nanometre scale. This invention was a significant step for neuroscience. At the end of the 1950s, individual synapses were imaged, cementing the neuron theory of brain function.

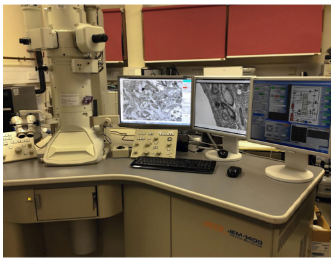


Figure 9 Electron microscope

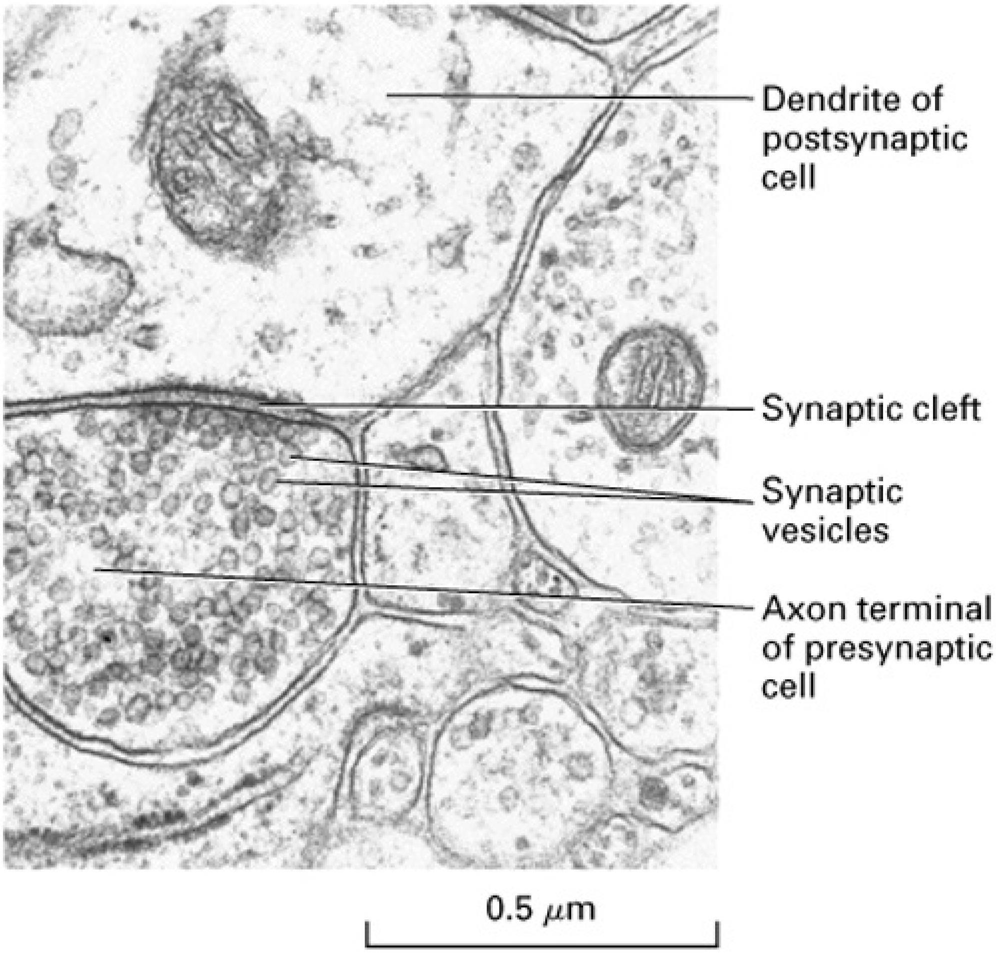


Figure 10 An electron micrograph showing synapses. The synaptic cleft is the gap between the neurons, and the synaptic vesicles in the presynaptic neurons contain neurotransmitters.

3.2 Technological advances

The development of other revolutionary technologies in recent decades have enabled much greater understanding of the brain. This section will explore a few of these advances.

EEG

As discussed in Section 2, neurons transmit electrical signals around the brain. The electrical activity of large numbers of neurons synchronise into large-scale oscillations called brain waves, which can be measured directly by placing electrodes on the scalp. These electrodes detect voltage fluctuations due to the brain waves, which are then recorded as an electroencephalogram (EEG). The result is a picture of how the brain waves change over time.

This technique has shown that different frequencies of brain waves are associated with different mental states, from highly active learning to deep sleep (Figure 11).

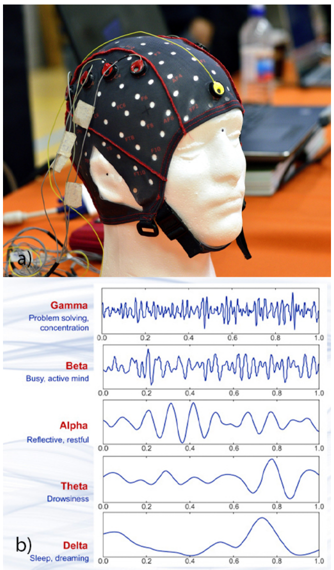


Figure 11 (a) EEG headcap, used to apply electrodes to the surface of the head; (b) different types of brain wave activity measured using EEG

While the EEG technique can detect very rapid changes in brain activity, it can't measure the precise locations of activity.

Imaging

New imaging techniques give us more detailed 'snapshots' of what's going on inside the brain. These don't measure the electrical activity of neurons directly, but measure other things associated with brain activity. Active regions in the brain require more oxygen, which is supplied by an increased blood flow. This can be measured in various ways, but the most common technique uses the fact that the increased oxygenation changes the magnetic properties of the blood. This is detected non-invasively by an fMRI (functional Magnetic Resonance Imaging) scanner. Figure 12 shows such a scanner, and a scan showing regions in the brain that become active in response to a visual stimulus.

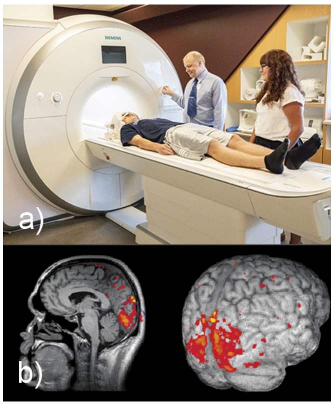


Figure 12 (a) fMRI scanner, (b) fMRI scan

Researchers can observe active regions in the brain associated with carrying out tasks or having different experiences, including sensing pain, using language, storing memories, feeling particular emotions, and so on. These same regions are even activated when a subject is asked to imagine doing something. In Figure 13, brain activity in an area associated with movement is seen when the subject imagines playing tennis, and activity in areas associated with movement and memory is seen when the subject imagines walking through their own home. Similar brain activity is also seen when some apparently non-responsive patients are asked to imagine these tasks.

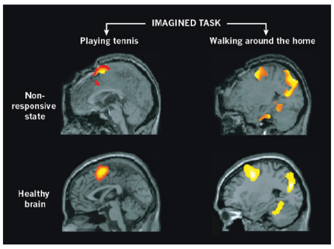


Figure 13 fMRI scans showing brain activity while tasks are imagined

Each individual ‘voxel’ (like a 3-dimensional pixel) of current fMRI images still covers a broad region of the brain, encompassing many thousands of neurons. But this resolution is improving over time, allowing more and more detailed pictures of brain processes.

Artificially stimulating the brain

While the other techniques in this section record natural brain activity, researchers can also use magnetic fields to induce an electric current in specific regions of the brain, artificially activating or suppressing the neurons there. This ‘transcranial magnetic stimulation’ (TMS) can be used therapeutically, for instance in the treatment of some cases of severe depression.

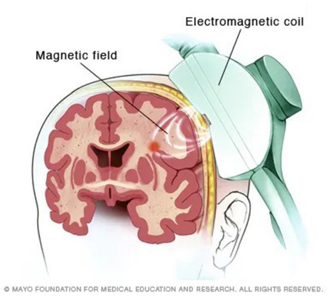


Figure 14 TMS

You can follow the link below to see a demonstration of how TMS induces movement in the body.

Video 7 TMS demonstration (open the link in a new window/tab so you can return here easily)

3.3 Into the future

A triumph of modern biological science has been mapping the human genome – the complete genetic code that is contained in our DNA, with over 3 billion units of information. The Human Genome Project was completed in 2003, after 13 years of unsurpassed international collaboration. Another related quest is the Human Connectome Project, which is ongoing at time of writing. It aims to construct a map of all the neural connections in the brain, termed a ‘connectome’. This could be thought of as a ‘wiring diagram’ for the brain.

Limited mapping of neural circuits is currently possible, using the microscopy techniques on brain tissue seen in Section 3.1. But even with recent refinements in methodology, mapping the whole brain and its 86 billion neurons would be an impossibly vast task. It would take over a year to acquire the data for just 1 cubic millimetre of brain tissue, and the raw data for the whole brain would require computing storage of about 175 exabytes – that's 175 billion gigabytes!

What has been possible, though, is the mapping of the entire connectome of a much simpler creature: a transparent worm just one millimetre long.

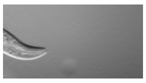


Figure 15 The roundworm *C. elegans*

This worm doesn't have a brain – it's controlled by a nervous system containing just a few hundred neurons. Nevertheless, it took researchers over a decade of work to produce this 'circuit diagram'.

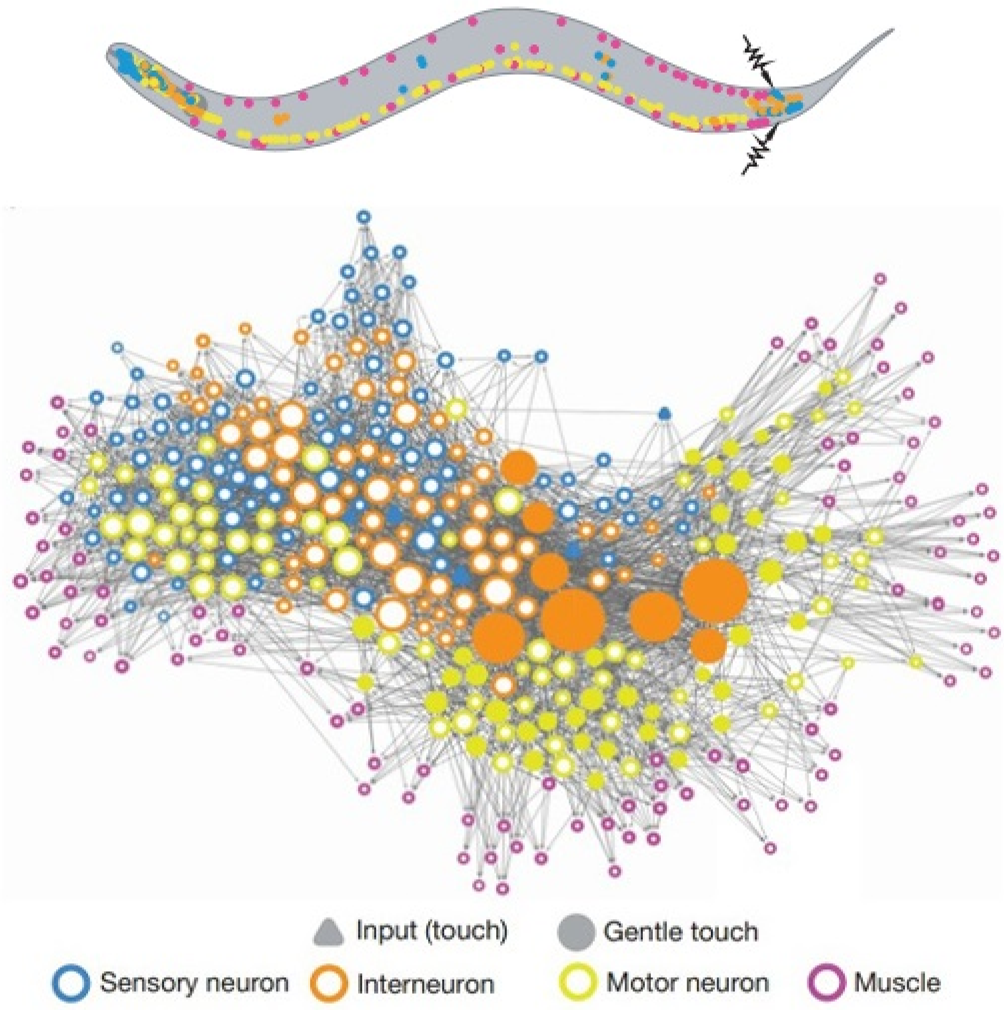


Figure 16 The roundworm connectome

The connective architecture of animal brains can be studied on larger scales with a different approach. This involves injecting 'tracer molecules' that are then transported along the axons of neurons. Typically, these compounds are fluorescent, rendering the pathways visible in a conventional microscope.

Figure 17 shows neuron (in red) made visible by the introduction of a dye using a microelectrode. A second neuron is traced using a fluorescent protein – the green dots show the locations of individual synapses.

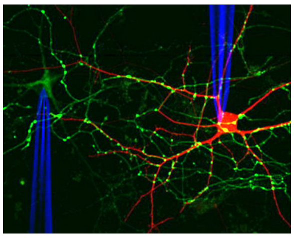


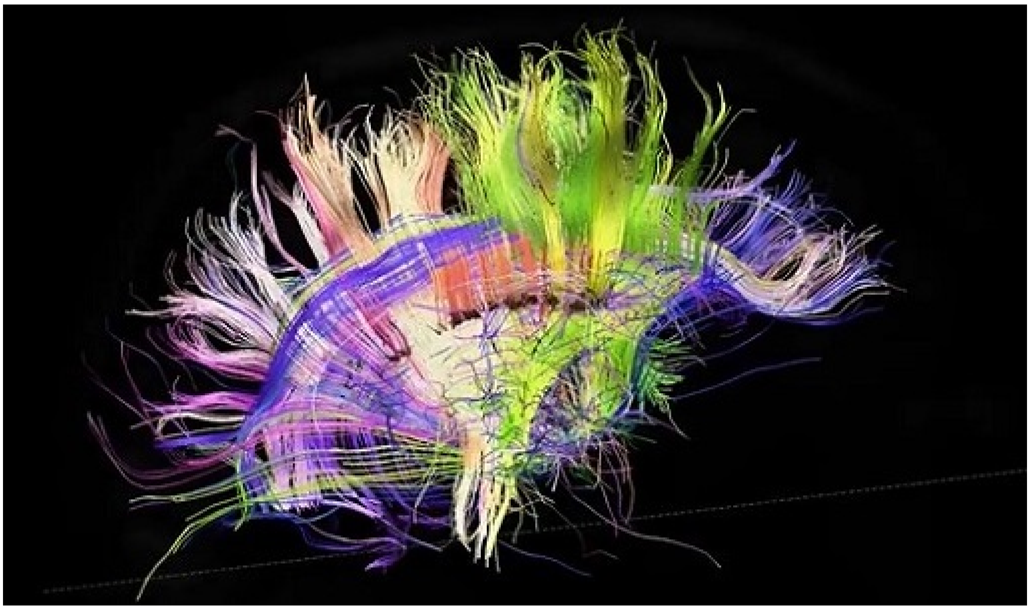
Figure 17 Tracing neurons

A further technique for studying human brains is to use non-invasive MRI (magnetic resonance imaging) to track how water diffuses through the brain, and so trace the main neural pathways (consisting of hundreds of thousands of axons).

It turns out that the connective structure of the brain is remarkably ordered, as shown in this video.

Video content is not available in this format.

Video 8 Brain wiring (note: there is no spoken audio in this video)



3.4 Brain study summary

You've looked at a variety of methods for studying the brain in this section. Complete the following activity to recap what you've learned.

Activity 3 Summarising techniques

Allow about 15 minutes

For each technique covered in this section, fill in some basic details in the table. Which aspect of the brain does it study, and what is produced as the end result? Try to also include some aspect of how the technique works (a detailed explanation is not needed). One row has been completed as an example.

Table 1 Study of the brain

Technique	Aspect	Output	How it works
Autopsy of damaged brains	Functional regions of the brain	Map of macroscopic brain regions and their principal functions	Correlates functional impairments observed in patients with locations of damaged brain tissue
Light microscopy with staining	*Provide your answer...*	*Provide your answer...*	*Provide your answer...*
Electron microscopy	*Provide your answer...*	*Provide your answer...*	*Provide your answer...*
EEG – electroencephalography	*Provide your answer...*	*Provide your answer...*	*Provide your answer...*
fMRI – functional magnetic resonance imaging	*Provide your answer...*	*Provide your answer...*	*Provide your answer...*
TMS – transcranial magnetic stimulation	*Provide your answer...*	*Provide your answer...*	*Provide your answer...*
Tracer molecules	*Provide your answer...*	*Provide your answer...*	*Provide your answer...*

Answer

Don’t worry if your answers are expressed a little differently or have less detail.

Table 1 Study of the brain (completed)

Technique	Aspect	Output	How it works
Autopsy of damaged brains	Functional regions of the brain	Map of macroscopic brain regions and their principal functions	Correlates functional impairments observed in patients with locations of

			damaged brain tissue
Light microscopy with staining	Neurons	Magnified photographs or drawings of individual neurons and patterns of neurons	Brain tissue is treated with chemicals to stain individual neurons
Electron microscopy	Synapses	Highly magnified images	Uses a beam of electrons instead of light
EEG – electroencephalography	Large scale changes in brain activity – brain waves	Graph of brain activity fluctuations over time (electroencephalogram)	Detects voltage changes using electrodes on the scalp
fMRI – functional magnetic resonance imaging	Locations of brain activity	Map of brain activity	Detects changes in magnetic properties of the blood due to increased blood flow and oxygen
TMS – transcranial magnetic stimulation	Functional locations in the brain	Observed changes in patient behaviour	Magnetic fields are used to stimulate or suppress neurons
Tracer molecules	Connectivity between individual neurons	Magnified image of specific neurons and locations of synapses	Injected fluorescent tracer molecules travel along axons which are observed using a microscope

4 The brain at work

Using the variety of techniques that you met in the previous section, scientists have found that neurons are connected to form neural pathways and circuits. Usually, neural circuits in different regions of the brains are interconnected in complex large-scale networks, working together to accomplish tasks such as talking. As a result, contemporary neuroscience is moving further away from the picture of small and isolated regions of the brain being solely responsible for different roles, as discussed at the beginning of Section 3.

Connections between neurons are constantly being either reinforced or abandoned depending on how much they are used. Pathways in the brain are formed in this way, allowing new skills to be mastered and memories to be created. This is called 'synaptic plasticity'. This remarkable ability of the brain to change with experience, continually adapting and learning, is a key area of current neuroscience research.

Here it is in action in a very striking way:

Video content is not available in this format.

Video 9 The rubber hand illusion

It's perhaps possible to appreciate how something relatively simple – like learning to count – could be programmed into the brain. But what about something more difficult, like learning to play a violin concerto from memory? It takes years of practice with the instrument, of course, but just think of everything the brain must store to achieve this task: the minute muscle movements, the sounds, the extraordinarily complex patterns of notes. It's truly wondrous to consider the brain at work.

5 Who is conscious?

Having delved into the structure and study of the brain at some length, let's now return to what is surely the hardest question in neuroscience: consciousness. This section will start by considering human consciousness, before moving on to other examples of consciousness and conscious behaviour. Can these things be conclusively observed in animals, for example? How about non-living things?

5.1 How conscious are we?

Try this quick test of concentration before moving on.

Activity 4

Allow about 5 minutes

Take a look at this video – paying close attention – then reveal the discussion below.
Count the passes (open the link in a new window/tab so you can return here easily)

Discussion

So, this wasn't a test of concentration after all! If you saw the gorilla straight away, you might wonder how anyone could possibly miss it. But in fact, about half of people don't see it.

(By the way, if you already knew about this 'invisible gorilla', you may like to try this one out – open the link in a new window tab so you can return here easily.)

The point here isn't to show how inadequate our brains are, but rather how remarkable they are. Imagine we were equally aware of every detail in our field of vision, along with all the sounds or sensations around us, every bodily movement we make, and all our thoughts on top of that – the experience would be utterly overwhelming. It's not that our eyes couldn't detect the gorilla, but that our brains 'decided' it wasn't important enough to tell us about it, knowing the particular task we were already engaged in. Our brains are equipped to filter out unwanted sensory input like this. They make a lot of unconscious decisions for us!

This can materialise in various ways. Perhaps you've experienced the sensation of driving home and finding you can't then remember the journey. Or maybe you've had a sudden flash of inspiration for a problem you'd been trying to solve hours before. There is a great deal going on inside our brains that we don't directly know about.

Another remarkable example of the brain's independence is a very unusual condition known as 'blindsight'. Particular damage to the brain causes people with this condition to think they are blind. However, tests show that their brain can process some visual information and respond to it. For example, when visually exposed to a threat they will react physically, but without the accompanying feeling of being afraid. They can actually see without conscious awareness. This shows that different regions of the brain are responsible for different aspects of 'seeing'. It seems that only the more recently evolved regions allow for conscious experience.

The question of whether consciousness has an evolutionary advantage is keenly debated. Our brains could still sense our environment and respond to it without us feeling anything. Perhaps it allows us to envisage outcomes for ourselves based on our actions – it helps us to plan ahead. On the other hand, it may just be a by-product of evolutionary brain development.

5.2 Animals and babies

It's commonly understood that animals have less advanced brains than humans, but do they demonstrate any form of consciousness? We might like to believe that animals – whether it's gorillas, cats or dolphins – have rich internal worlds of thought and emotion, but how can we really know? Animals like squid and octopuses have been shown to have pain receptors, but we just don't know if they 'experience' pain in the same way as us, accompanied by emotional distress, or if their physical response is an unconscious reflex directed by their nervous system.

Think back to how we humans have this feeling of being an individual person, seeing the world from behind our eyes. How can we know this about animals? Can we be sure there is 'someone there' inside their heads, and that it's not completely 'dark'?

It's not (yet) possible to ask animals what it's like to be them. Instead, scientists have to work within the limitations and devise ways to discern this indirectly. For example, recognising yourself in the mirror demonstrates that you're aware of your existence as an individual being. Scientists have so far been able to observe this characteristic of self-consciousness in only a few animals, including chimpanzees, elephants, dolphins and magpies. They do this by putting a mark on the animal, then seeing how they react to their reflection. If they try to wipe the mark off, the argument can be made that they realise the animal in the mirror with the strange mark is themselves. You can find footage of such tests online (there's an example linked in the Further Reading section). This method does come in for criticism, however. For one thing, it is biased towards animals whose vision is their primary sense.

Humans don't pass this test until they are around 18 months old. This could well be the source of a common superstition, which says that it's bad luck for a baby to see its own reflection.

In terms of measurable brain activity, the brain patterns in mammals and birds have been measured by EEG. They have been shown to be analogous to those in parts of the human brain. However, this doesn't prove that they're thinking like a human. There is still much investigation to be done here.

So, what about consciousness in non-living things?

5.3 Machines

Artificial intelligence has been a recurring theme in science fiction for a long time – since at least Samuel Butler's 1872 novel *Erewhon* – whether it's depicted as a dystopian threat to human civilisation, or a utopian ideal. But what's the reality in the present day?

Can a computer program actually 'think'? Can it be truly conscious? Or can it only ever simulate conscious behaviour? These are increasingly relevant questions in the 21st century. To begin to unpack them, this activity introduces you to the 'Chinese room' thought experiment, proposed by American philosopher John Searle in 1980.

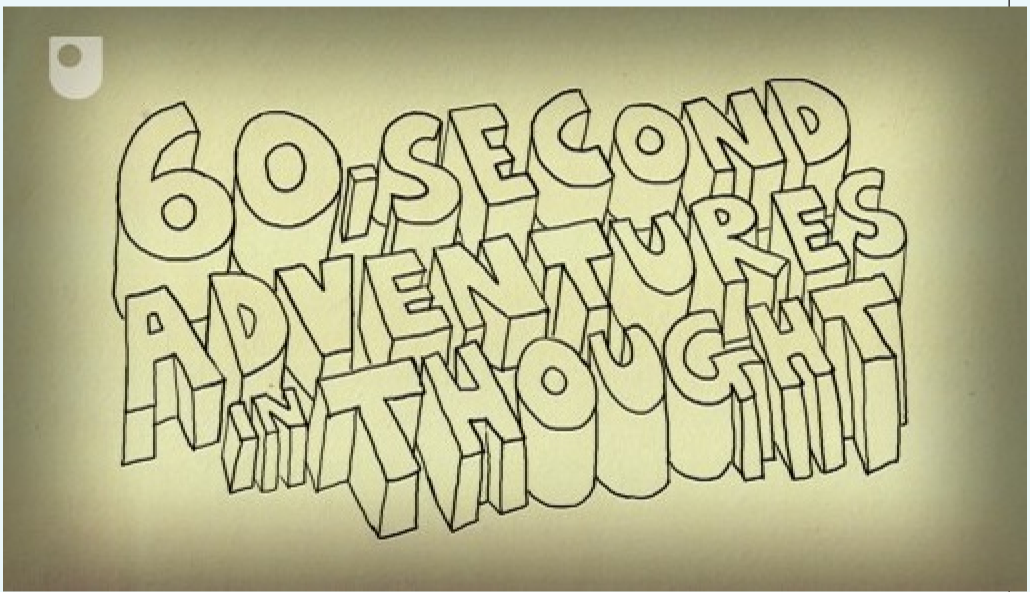
Activity 5 The Chinese room

Allow about 10 minutes

Watch the video, and see what you make of the argument. Then consider the question beneath and make a few notes.

Video content is not available in this format.

Video 10 60-Second Adventures in Thought: The Chinese Room



Do you agree with the conclusion that a computer program could only ever ‘simulate’ intelligent thought and language comprehension? Or do you agree with Alan Turing, that a computer which can pass itself off as human (thereby passing the famous ‘Turing test’) should be said to be intelligent?

Provide your answer...

Discussion

This continues to be keenly debated, so there’s certainly no easy answer here. But it’s true that the philosopher in the thought experiment doesn’t understand the conversation, despite outward appearances. He’s just following step-by-step instructions. Computer programs work in a similar way – meaning a computer could engage in intelligent conversation without actually understanding it.

If you’d like to investigate the arguments further, there are some resources in the Further Reading section you might like to explore.

The focus here is the demonstration of intelligence and understanding, but this idea can be extended to the related concept of consciousness. John Searle has argued that a programmed computer model of consciousness wouldn’t actually be conscious. Others have suggested that we just couldn't be certain about this either way. And this lack of certainty extends past machines, to a related thought experiment of sorts about 'philosophical zombies'. This argument posits that other people around us act like normal human beings, and display the outward characteristics of being conscious. But, without getting inside their heads, we don't really know whether they're feeling anything, or truly experiencing the 'qualia' mentioned earlier. All of their responses could be just following instructions, like the computer program.

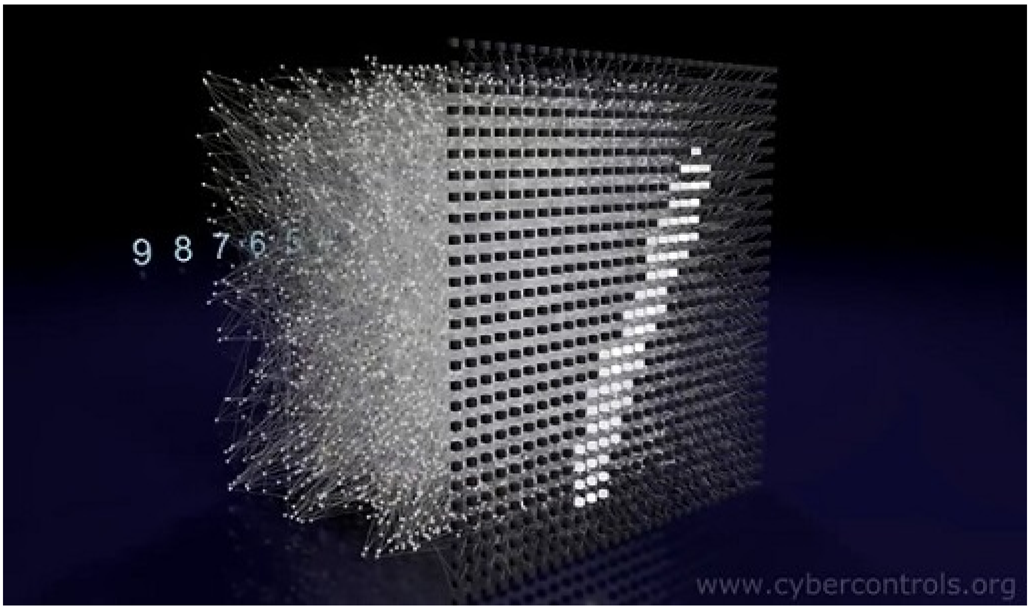
5.4 Neural networks

Artificial neural networks are mathematical systems that loosely emulate how biological brains learn, namely from example and experience. Computers with neural network programming can learn to recognise patterns without being specifically programmed with each decision-making step. For example, they can be trained to recognise facial expressions or individual voices.

The 'neurons' in an artificial network are 'nodes' – these are the points at which a computation takes place. The inputs to each node are either amplified or dampened mathematically, depending on how useful they are to the task (this is analogous to the strengthening or weakening of neural connections in the brain, as described in Section 4). If the sum of these weighted inputs is high enough, the node 'switches on' – like a neuron being fired – and the signal is passed on to the next node. Layers of nodes become the inputs for subsequent layers, and so on, until the final output of the network. The network can then assess how well it did, and adjust the computations to make improvements. Video 11 shows a visual demonstration of this.

Video content is not available in this format.

Video 11 Neural network simulation (note: there is no spoken audio in this video)



Neural networks are one method of 'machine learning', in which computer algorithms improve automatically through experience. This ability is at the heart of developing artificial intelligence. However, returning to the Chinese room argument, this could be thought of as just mimicking learning rather than actually understanding the process.

Another fundamental question is whether artificial intelligence could ever use common sense and intuition to solve problems.

5.5 Artificial neurons and neuromorphic computing

The nodes in artificial neural networks are mathematical, but real-world progress is being made. In 2019, a team led by researchers at the University of Bath announced that they had created artificial neurons on silicon chips whose electrical behaviour accurately mimics biological neurons. The potential application for this technology is remarkable. They could be used in medical implants to treat conditions in which neurons have degenerated or been damaged, such as Alzheimer's.

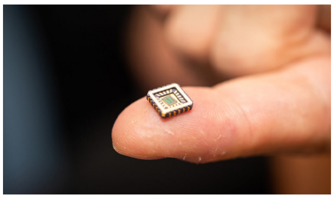


Figure 18 Artificial neurons

The grand ambition here is modelling the neuronal circuitry of the brain, rather than individual neurons. This is the approach of 'neuromorphic computing'. The objective of the European Human Brain Project is to simulate an entire human brain. The SpiNNaker machine at The University of Manchester is currently the world's largest neuromorphic supercomputer, comprising one million computer core processors, with the aim of modelling connections between a billion biological neurons. This is equivalent to 1% of a human brain.

Will a complete neuromorphic model of the human brain be conscious? Would it be a 'person', with its own personality? These are challenging scientific, philosophical and ethical questions without clear answers.

Intelligence and consciousness are, of course, not the same thing. It may be that machines can become ever more intelligent, but never conscious. Perhaps only complex biological entities can be truly sentient, as a product of keeping that biological system alive and functioning.

6 Theories of consciousness

Tentative scientific theories of consciousness are starting to be developed. This week will conclude with a short introduction to these theories, which focus on the brain activity required for different conscious experiences – whether that's activity in specific regions of the brain, or particular global patterns.

6.1 An intrinsic property (Integrated Information Theory)

The Italian neuroscientist Guilio Tononi has considered what's required for an experience to be conscious. One important facet of this theory is that different aspects of the experience are initially processed separately in the brain, then somehow 'integrated' together into a single conscious experience.

Take the experience of picking up an object, for example. The brain will process visual information about the colour, shape and position of the object, as well as other sensory information like touch. This information is shared between different regions of the brain to create the unified conscious experience. Imagine picking up a tomato. In doing so, we aren't aware of each separate attribute – that it's red, looks round and feels smooth – unless we're focusing on them specifically. We just experience it as a tomato.

This is the basis of Tononi's 'Integrated Information Theory'. The more information sharing and integration that takes place, the higher the level of consciousness. This requires the neurons in the brain to be highly connected, as indeed many are.

Experiments using transcranial magnetic stimulation (TMS) suggest that Tononi may be on the right track. When a localised region of an awake person's brain is stimulated, other regions far from the stimulated site respond in complex feedback patterns, which is suggestive of integrative processing and conscious experience. However, if a subject is under induced anaesthesia or in a state of non-dreaming sleep – and therefore without any awareness – then the brainwaves that are generated are more confined and have a much simpler form.

Tononi has derived a mathematical formula that can measure how integrated, and hence how conscious a system is. An intriguing aspect of this theory is that the connected components don't have to be biological neurons for this intrinsic property of consciousness to emerge from the system. They could be artificial neurons, or silicon transistors. The question, then, is whether an artificial physical system with a high level of information integration could actually be conscious. Another implication of this theory is that a computer program can't be conscious, however sophisticated its simulation might be – just as the Chinese Room idea argues.

This theory may give us a way of measuring consciousness, which is certainly highly significant. But what would remain unanswered is how that consciousness is actually created.

6.2 The spotlight of consciousness (Global Workspace Theory)

Another prominent theory of consciousness is called the 'Global Workspace Theory', which was proposed by Bernard Baars and extended by Stanislas Dehaene. The theory posits that information from various input sources comes together into a 'global neuronal workspace', made up of a constantly changing network of neurons. Imagine this as a kind of 'theatre stage' for the mind, with metaphorical actors moving in and out, making speeches or interacting with each other. The ‘spotlight’ of our consciousness shines a bright spot on the stage, which can then be broadcast to the areas in the brain associated with specific tasks.

6.3 Studying consciousness in the laboratory

A complete explanation of consciousness is still far out of our reach. But using our current knowledge, we may be able to discern experimentally between these and other theories.

For example, a key difference between these two theories is the expected focus of brain activity for different conscious experiences. This can be studied by exposing someone to stimuli which are above and below the limits of conscious perception, such as an image of a word. When consciously perceived, there is a burst of activity across different brain regions, but when detected only subconsciously, the brain activity is much more limited.

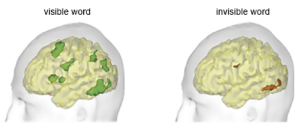


Figure 19 A ‘signature of consciousness’ – patterns of brain activity recorded as a response to conscious (left) and non-conscious (right) visual stimuli (Dehaene and Changeux, 2011)

A complete explanation of consciousness will depend on continuing experimental investigation and technological advances, in parallel with further theoretical developments.

6.4 Why might we never have an explanation for consciousness?

Conscious experience is subjective. It may never be possible to know what it’s like to be another person or animal, or indeed artificial intelligence that is functioning as a conscious being (unless we develop the ability to mind meld like the Vulcans in Star Trek!)

Some thinkers even suggest that our experience of consciousness is just an illusion. Some believe that the hard problem of consciousness is intrinsically unsolvable – this is the philosophical position of ‘mysterianism’.

The 17th century philosopher René Descartes believed that the mind of conscious thought and the physical brain are fundamentally different in nature, and that the interaction between them can’t be explained. This is the so-called ‘mind-body problem’. Contemporary experiments show that there are correlations between physical processes in the brain and the subjective experience of the mind. But could there be a true causal connection, and can we prove it?

Studying such a vastly complex and subjective phenomenon through objective experimental means certainly poses an immense challenge. But scientists working in the field are optimistic that we will be able to bridge the gap between the physical and mental. This may well be achieved in due course. But on the other hand, perhaps when it comes down to it, our brains just aren’t capable of fully understanding themselves!

7 This week's quiz

Well done for reaching the end of Week 6. Check what you've learned by taking the end-of-week quiz.

Week 6 practice quiz

Open the quiz in a new window or tab then come back here when you've finished.

8 Summary of Week 6

This week, you've delved inside your own brain to examine the concept of consciousness. For so long purely the realm of philosophy, scientific study in this fascinating field is now blossoming with technological advances, for example in brain imaging.

Much scientific study currently focuses on what gives rise to individual conscious experiences such as the qualitative experience of seeing the colour red. The ambition is to eventually solve the 'hard problem of consciousness' – what processes in the physical brain actually give rise to these subjective experiences?

Along the way you've learned about how your brain works, by transmitting messages through immense networks of neurons via electrical pulses and chemical neurotransmitters. Ambitious scientific projects aim to map these brain networks, and to build simulated neurons and neural networks. Could these 'simulated brains' ever show aspects of consciousness themselves?

Next week gets really big! You'll be exploring the concept of infinity, and you'll see the power of mathematics in action, as it helps us to approach topics that initially seem too vast to be knowable at all. At the other end of the spectrum, the infinitely small will take you into the beautiful world of fractals.

You can now move on to Week 7.

Week 7: Infinity

Introduction

The concept of infinity can seem difficult to grasp, whether you're looking at the infinitely big or the infinitely small. Intuitively, the infinite can be thought of as something limitless in size. Mathematical progress on understanding the infinite was hindered by clashes with religious and establishment views, as infinity was felt to be God's realm, not man's. However, as you'll see, working with infinity can simplify problems considerably. Sometimes it's easier to work with the infinitely big than it is to work with the very big.

Here's Marcus to introduce this week's topic.

Video content is not available in this format.

Video 1 Introduction

By the end of this week, you should be able to:

- recognise that there are different sizes of infinity
- understand the idea of proof by induction
- describe how sequences (and their limits) help with understanding the infinitely small
- appreciate the paradoxical qualities of fractals.

1 Picturing infinity

Our galaxy, the Milky Way, is roughly 100 000 light years in diameter (which is nearly 1 billion billion miles). Our minds are simply not equipped to understand this sort of scale. Analogies can be helpful in putting this into context, for example: if our solar system was scaled to the size of a full stop, then the galaxy would be about 50 miles in diameter. Nevertheless, it can be easier to picture an infinite space – a space without a boundary – than it is to visualise 100 000 light years.

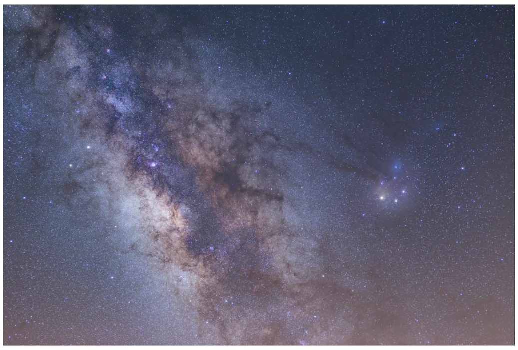


Figure 1 The Milky Way

Infinity is a useful idea, but it does not exist only in our imaginations – there could be real instances of infinity in our universe. For example, the 2020 Nobel Prize for Physics was jointly awarded to Sir Roger Penrose, Reinhard Genzel and Andrea Ghez for their work on black holes. In 1965, Penrose gave a mathematical argument that black holes must arise as a consequence of Einstein's General Theory of Relativity. His theory predicts a gravitational 'singularity' in black hole formation. (As Week 5 introduced, this is the point where density and gravity become infinite.)

Figure 2 (a) Reinhard Genzel, (b) Andrea Ghez, (c) Sir Roger Penrose

The ideas in Week 5 give rise to speculation on the nature of infinity. Working in an infinite space without a boundary gives rise to some strange consequences, as you will see in the next section.

2 Hilbert's Infinite Hotel

Mathematician David Hilbert wanted to make the idea of infinity more accessible to the public. To do this, he devised the paradox of 'Hilbert's Infinite Hotel'. You could think of the motto of this hotel as 'even when it is full, it always has room for you'.

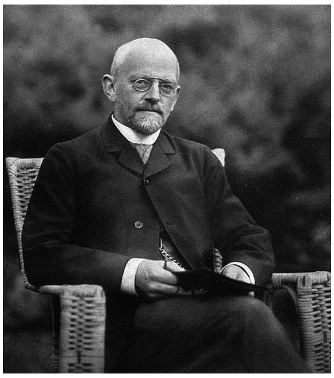
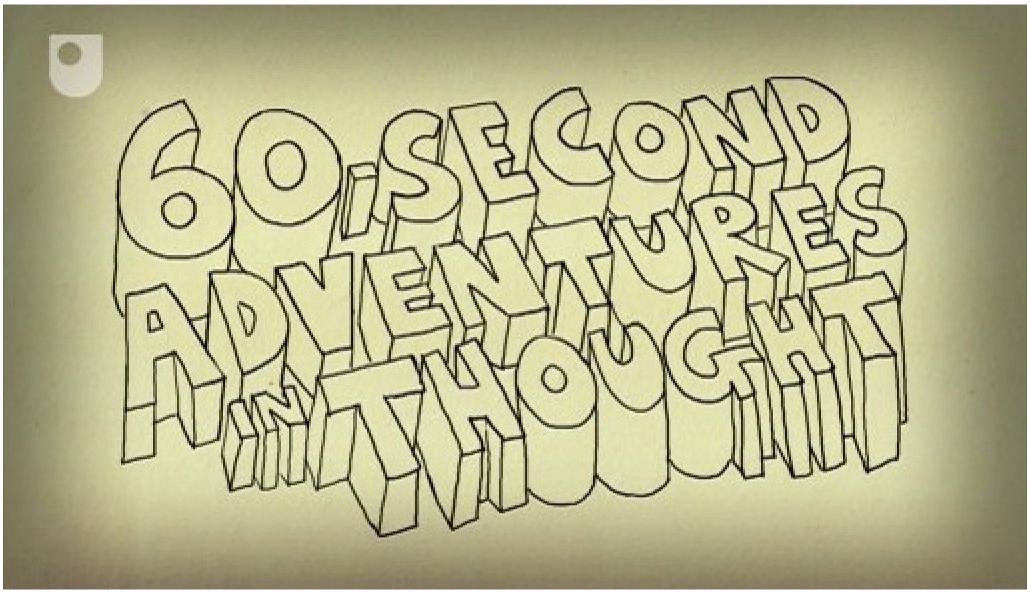


Figure 3 David Hilbert (1862–1943)

Watch Video 2 which introduces the paradox.

Video content is not available in this format.
Video 2 60-Second Adventures in Thought: Hilbert's Infinite Hotel



The paradox shows that infinity is somehow stretchable and squashable. As explained in the video, when the hotel is full, a room can always be found for a new guest. It's simply a case of asking each guest to move up a room number. We know that every number has a successor – and only one. (Likewise, each successor room only has one predecessor – so two guests aren't going to end up in the same room.) In mathematical terms it can be said that any number, n, is 'mapped' (more on this shortly) to the next number, n+1. Now the first room is empty, and the new guest can be accommodated here.

Indeed, infinitely many new guests can fit in the full hotel with no trouble! In this case, each guest is asked to move to the room which is double the room number they are currently in. Each number has one and only one number which is its double, so every guest will have somewhere to go. No two guests will end up in the same room here either, since every even number only has one number which is half its value. With this done, all the odd numbered rooms are now empty. There are infinitely many of these for the infinitely many new guests. In this case, it can be said that *n* is mapped to its double, *2n*. You'll look more closely at what this means in the next section.

There's something else demonstrated by this argument: there are as many even numbers as there are counting numbers (1, 2, 3, …). There are not half as many, as you might expect! This may seem counter-intuitive. Imagine all these numbers laid out in two (infinitely long!) rows. You can 'stretch out' the counting numbers and match them up with the even numbers. Or you can 'squash together' the even numbers and match them up with the counting numbers. You will never reach a point where the numbers stop matching up. Of course, you would never complete this experiment if you tried it! Perhaps this is still a bit difficult to picture. Mathematical notation can help with that, as the next section will show.

Thinking about things like how many even numbers there are leads to a key question: how do you measure the size of something infinite?

3 Measuring infinity

Are all infinite things the same size? The answer, as you'll see, is no! Mathematician Georg Cantor showed how you can think about sizes of infinity, and how you can measure them.

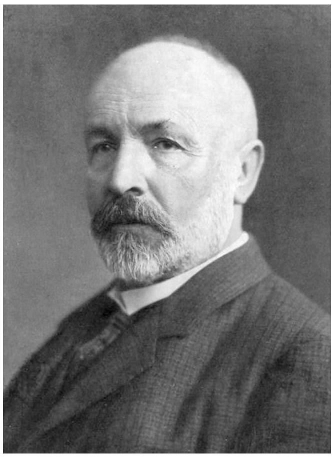
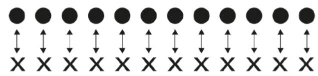


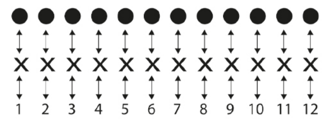
Figure 4 Georg Cantor (1845–1918)

In order to unlock the way to count infinite sets of things, it helps to look closely at how you count finite sets. Consider the picture below. Are there as many dots as crosses?

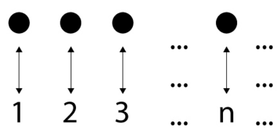
You can check by pairing them up.



Another way to check is counting the set of dots, then counting the set of crosses. This is essentially the same as pairing them up with counting numbers.



In the same way, you can 'count' infinite sets by matching them up with the counting numbers. For example, if you have an infinite set of dots in a line you could pair them up with the counting numbers, like this:



By specifying a ‘map’ in which each number, *n*, is paired with its double, 2*n*, you can formally demonstrate the claim from Section 2 – that there are as many even numbers as counting numbers. This map is illustrated by the following figure.

$$\begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & \dots & n & \dots \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \dots & \updownarrow & \dots \\ 2 & 4 & 6 & 8 & 10 & \dots & 2n & \dots \end{array}$$

Activity 1

Allow about 5 minutes

Can you show there are as many multiples of 4 as there are counting numbers? What should *n* be mapped to, in order to show this?

Answer

n is mapped to 4*n*

$$\begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & \dots & n & \dots \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \dots & \updownarrow & \dots \\ 4 & 8 & 12 & 16 & 20 & \dots & 4n & \dots \end{array}$$

Can you show there are as many odd numbers as there are counting numbers? What should *n* be mapped to, in order to show this?

Answer

n is mapped to 2*n*–1

$$\begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & \dots & n & \dots \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \dots & \updownarrow & \dots \\ 1 & 3 & 5 & 7 & 9 & \dots & 2n\text{-}1 & \dots \end{array}$$

The size of the set of counting numbers can be shown to be the ‘smallest’ infinity, and you have just seen that the set of even numbers and the set of odd numbers are the same size. Other sets which have the same size are:

- all whole numbers both positive and negative
- all the prime numbers
- all numbers that can be written as fractions (this includes whole numbers as they can be written as a fraction over 1).

The last set in that list might’ve been unexpected – you shall see a proof of it shortly.

However, numbers which can’t be written as fractions do exist. The number π (pi) is a well-known example. Once you add in all the numbers like π and $\sqrt{2}$, then you make a set which is bigger than the counting numbers. You can no longer pair up the members of this set with the counting numbers.

4 Proofs

In mathematics, statements are frequently proven true for infinite sets of numbers. For example, every number which is a multiple of 4 is also a multiple of 2. Of course, there are infinitely many multiples of 4, so how can you check that each one is also a multiple of 2? You could work it out, number by number, over and over again. But how high do you need to count to be certain it's true for all numbers? 100? 1000? 10000? Ultimately, to be absolutely sure via this method, you're going to be checking every number from here until eternity.

Instead, you want a proof that works for all numbers. So, first of all: what is a proof? Intuitively, a proof is an argument that leads you from true assumptions, through truth-preserving steps, to a true conclusion. You can make direct proofs about statements concerning infinite sets of numbers, as you would with the example 'every number which is a multiple of 4 is also a multiple of 2'.

Box 1 Example proof

Let n be a number which is a multiple of 4. Then n can be written as $4k$, where k is some whole number. That is,

$n = 4k$

But now you can write n as $2\text{x}2k$ where $2k$ is also a whole number. A number which can be written as 2x something is a multiple of 2. So, n is a multiple of 2.

Another type of proof which is used with infinite sets of numbers is called 'induction'. Proof by induction takes only two steps. Suppose you have a statement you want to prove. You start by proving the statement for the first number. This is called proving the 'base case'. The next stage is to prove what's called the 'induction step'. You look at an arbitrary number, call it k. If you can prove that when the statement is true for k it is also true for the successor of k, then you are finished. You can think about this step by step:

- You have shown the statement is true for the first number.
- You have also shown that when the statement is true for one number, it is automatically true for the following number.
- This means the statement is true for the number following the first number – in other words, it's true for the second number.
- But this also means the statement is true for the number following the second number, so you know it is true for the third number.

See how this logic shows you that the statement is true for all numbers.

To make another practical analogy, imagine dominoes lined up.

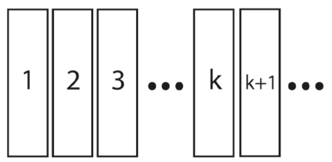


Figure 5 Dominoes in a line

Suppose that:

- The first domino falls.
- If domino *k* falls, then the next domino *k*+1 will fall.

These two facts together tell us that all the dominoes will fall.

Proof by induction is a powerful tool in mathematics, used particularly in Number Theory. Just two steps can be all you need to be acquainted with an infinite string of numbers or facts. This is one of the ways in which infinite sets can actually be simpler than finite ones.

4.1 Primes

A prime number is a number greater than 1 which is only divisible by itself and 1. The first primes are 2, 3, 5, 7, 11, and so on. Earlier it was stated that there are infinitely many primes. This was proved by Euclid around 300 BC. Once again, it initially seems counterintuitive that the primes should carry on forever. As numbers get bigger there are more potential divisors, so you might think there's eventually a point where primes just stop. Here you will look at one proof that there are indeed infinitely many primes.

The first thing you need to know is that every number can be written as the product of 'prime divisors'. Either it's a prime, and therefore is its own prime divisor, or it isn't prime, and it can be broken up and written as the sum of its primes (prime 1 x prime 2 x …). The second important fact is that this prime factorisation is unique to each number. This bold-sounding statement isn't all that difficult to prove, but it's a little beyond the scope of this course. Here are a couple of examples to demonstrate the idea:

15 = 3 x 5

There are no other prime numbers you can multiply together to make 15. You can use the same primes more times to get different prime factorisations:

75 = 3 x 5 x 5

Try this out for yourself.

Question 1

What is the prime factorisation of 22?

Provide your answer...

22 = 2 x 11

What is the prime factorisation of 126?

Provide your answer...

126 = 2 x 3 x 3 x 7

Once you're happy that any number can be written uniquely as the product of primes, you can look at how to prove there are infinitely many primes. You use something call 'proof by contradiction'. You can argue by instead assuming the opposite of the claim, and showing this leads us to an impossible conclusion. In mathematics, if you go through a series of steps which takes you to a contradiction, you know there is something wrong with an earlier step. There are no contradictions allowed in mathematics.

To prove there are infinitely many primes, then, you first assume the opposite – that there are finitely many primes – and you start by listing them. To make this proof easier to picture, label the primes p_1, p_2 , p_3, p_4 and so on, all the way up to the 'last' prime. So p_1 = 2, p_2 = 3, p_3 = 5, p_4 = 7, and as we don't know how big our last prime is, we just use a letter to stand for it, so that's p_r.

Now, think about the number you get when you multiply the whole list of primes together and add 1.

$p_1 \times p_2 \times \ldots \times p_r + 1$

It may take a minute to believe it, but none of the primes in our list can divide this number. Each one leaves a remainder of 1. Now, every number can be written uniquely as a product of primes, so there must be some prime that's not in our list which divides the result above. It's *not* one of the primes in our list, and that list supposedly contained *all* primes. This shows that the assumption that there was a finite list of primes was false. Therefore, there are infinitely many primes.

To confirm that there are as many primes as there are counting numbers, you first observe that all the primes are contained in the counting numbers, so it's not a bigger infinite set than the counting numbers. To give a mapping to the counting numbers (a pairing between all the counting numbers and primes) you simply enumerate the primes: first prime, second prime, third prime and so on. As before, this shows that each set of numbers has the same size.

1	2	3	4	...	n	...
$\updownarrow$	$\updownarrow$	$\updownarrow$	$\updownarrow$	...	$\updownarrow$	...
2	3	5	7	...	nth prime	...

This is a powerful piece of information. The fact that there are infinitely many primes forms the base of a great deal of cryptography, as staggeringly large primes are used to protect data. Finding larger and larger primes is hard and can use complex theories. We don't actually know what all the prime numbers are, and the next prime can't currently be predicted, which is part of the security of using prime numbers.

There is an open problem in mathematics called the Riemann hypothesis. If this were proved, then we could predict how the prime numbers are distributed (but it still won't tell us which the next prime number is!). In fact, proving the Riemann hypothesis is one of the Millennium Prize Problems, a set of important unsolved mathematical problems which each carry a million-dollar prize.

4.2 Fractions

It was stated earlier that there are as many fractions as counting numbers. At face value, this is even harder to swallow than the matching between the even numbers and the counting numbers. Between any two whole numbers there are infinitely many fractions. This is because between any two numbers (fractions or whole numbers) there is another fraction. No matter how small the gap between any two fractions, there is always another one in between.

Think about the gap between 0 and 1/2. There is another fraction halfway between them, 1/4. Then you can find the point in the middle of these fractions. Now you've added in 1/8 and 3/8. You can keep adding points in the middle of the fractions you already have, and you'll never reach an end.

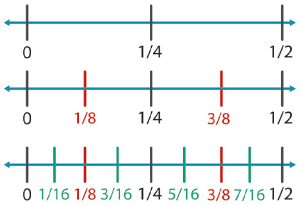


Figure 6 Fractions between fractions

Because the fractions are densely packed in this way, you might think there would be 'more' of them than the counting numbers. So, how can you give a matching between the fractions and the counting numbers?

Here you will just look at the positive numbers and the positive fractions. Cantor argued this in the following way. First, lay out all the fractions in a grid. A fraction is any number divided by another number. The grid is constructed so that the first row has one over every number, the second row has two over every number, and so on. The first column is every number over one, the second column is every number over two, and so on. The first part of the argument is to realise that every fraction will appear somewhere in this grid.

The matching between the fractions and the counting numbers is then given by the diagonal 'walk' shown in Figure 7. Simply match 1 with the first step in the walk, match 2 with the second step, etc. Every fraction will be visited. Occasionally you'll hit a number which is a repeat of what you had before. Skip over these and continue with your matching.

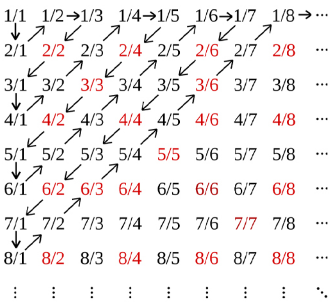
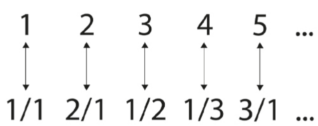


Figure 7 Fraction grid

Following the diagonal walk gives a mapping between the counting numbers and the fractions, confirming that the set of fractions is the same size as the set of counting numbers.



There are bigger infinite sets of numbers. As stated earlier, once you include the numbers which cannot be written as fractions, the 'irrationals', you can no longer make a matching between these numbers and the counting numbers. Sets of this size are called 'uncountable'. Cantor's proof of this again uses a diagonal argument. This time he uses the technique 'proof by contradiction', just as you did with the proof that there are infinitely many primes. This proof starts by assuming that the set of irrationals can be counted – and it reaches a contradiction. This shows that the set of irrationals cannot be mapped to the counting numbers and is therefore uncountable. So, infinite sets can be different sizes!

It was said earlier that you can ‘count’ infinite sets by matching/pairing them with the counting numbers. Similarly, you can compare sizes of infinite sets by finding a matching/pairing between the members of those sets. There are other mathematical techniques for comparing the sizes of infinite sets as well.

It turns out that whatever infinite set you have, you can make a genuinely bigger one out of it. This means there is no ‘biggest’ size of infinity.

5 The infinitely small

So far, you've looked at the infinitely big. You've seen that, while it initially seems counter-intuitive, it can be tamed with methodical proofs. The infinitely small is a different beast though. It's not 'nothing' and it's not 'something'. It can have a position, but it cannot have any size. The infinitely small is akin to the coordinates on a graph or map, which have locations but no width (as they would then be in another point's location).

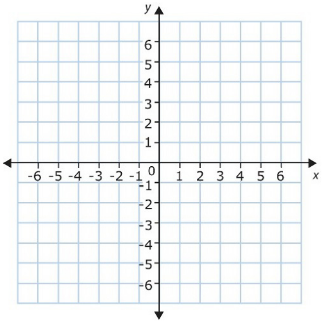


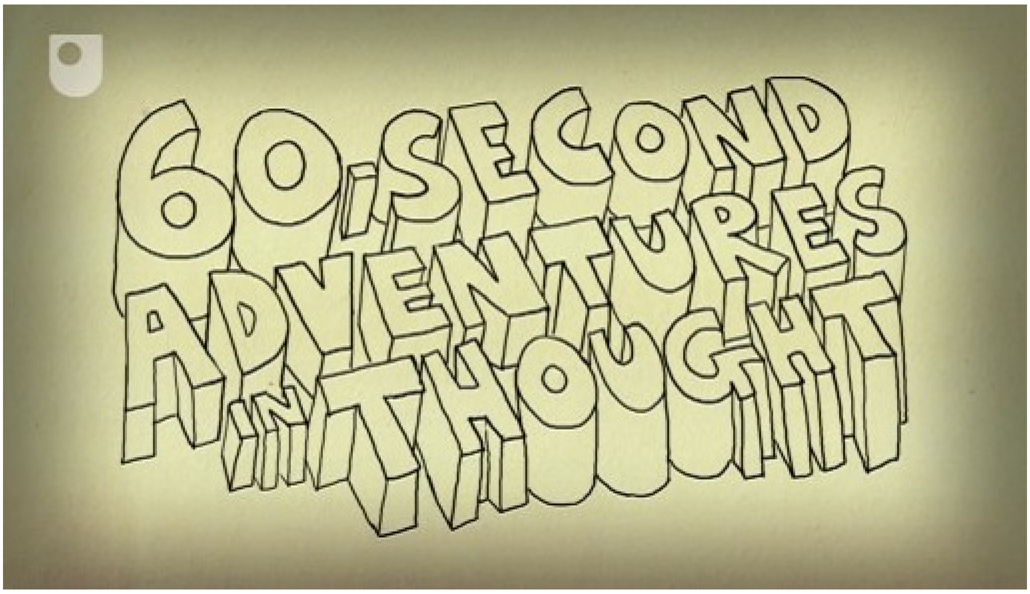
Figure 8 Coordinates on a graph

5.1 Achilles and the Tortoise

Here's one way to begin thinking about the infinitely small. Consider the paradox of 'Achilles and the Tortoise' which was devised by the Greek philosopher Zeno in the 5th century BC. Watch Video 3 for an introduction to this paradox, which took hundreds of years to be logically refuted.

Video content is not available in this format.

Video 3 60-Second Adventures in Thought: Achilles and the Tortoise



Zeno argued in this paradox that Achilles can never catch up to the tortoise in the race. Even though the distances involved are ever smaller, each time Achilles closes the gap, the tortoise has always moved on a tiny bit further. This contradicts what our senses tell us and therefore, Zeno argued, our senses are deceiving us.

It took the work of mathematicians such as Leibniz and Newton to advance our understanding of the infinitely small, and thereby resolve this paradox. What thinkers at the time of Zeno didn't realise is that you can take something finite and divide it (mathematically, if not physically) into infinitely many parts.

5.2 Dividing into infinitely many pieces

Here's an illustration of a finite object being divided into infinitely many pieces: picture a chocolate bar, and an obstinate mathematician who eats half of what's left every day. On day one they eat half of the bar; on day two they eat a quarter of it; on day three they eat an eighth; on day four they eat a sixteenth, and so on. Because you can always divide a fraction in two, this process never ends (though before long, the mathematician isn't eating much chocolate per day).

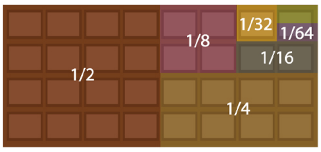


Figure 9 Dividing a chocolate bar

Now, when you come to add up the infinite string of fractions in this scenario, hopefully it's clear that these fractions add up to 1.

$$\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{1}{16}+\frac{1}{32}+\ldots=1$$

1 can be imagined here as the whole bar of chocolate you started with. This illustrates that you can have an infinite addition which gives a finite answer. This seems unexpected at first – so often the case where infinity is concerned! Now, Zeno's paradox may cease to be a paradox at all, if the infinite sum of all those tiny distances the tortoise travels has a finite answer (the point at which Achilles passes the tortoise).

What do you need to make an infinite addition give a finite answer? This is where the work of Newton and Leibniz helps us. They invented 'calculus' – arguably the most powerful piece of mathematics ever gifted to the sciences. You could fill many books describing calculus, and it's a topic for extensive study in itself. This course will give a quick flavour of what calculus is and what it allows us to do.

5.3 Decreasing sequences

In order to examine the infinitely small – whether instants in time or precise positions in space – Newton and Leibniz looked at sequences of numbers. Some sequences get closer and closer to one number. Some do not, and instead they grow forever. Newton and Leibniz looked at these sequences of numbers and asked: what are they going towards? If you know that, you can develop a formal and robust way of talking about what happens in the instant, or in the infinitely small point. You do this by talking about how sequences of numbers that approach this value behave. Look again at the sequence of numbers you used earlier:

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \frac{1}{32} + \dots$$

The fractions in this sequence are getting smaller and smaller. The sequence is going towards zero – but if you stop the sequence at any finite point, it hasn't yet reached zero. Not all sequences that get smaller and smaller go towards zero. So how do you know that this one does? Newton and Leibniz gave us a formal way of saying what a sequence goes towards. This is called the 'limit' of a sequence.

The limit of the sequence above is zero. The way this is proved is by showing that, no matter how small a gap you look for, there's a number in the sequence which is closer to zero. All subsequent numbers in the sequence are closer than that gap to zero.

In order for an infinite list of numbers to give us a finite answer when added together, you first need to know that the sequence goes towards zero. If not, then the sum is going to be unbounded and infinite.

There's a second condition that's necessary for an infinite list of numbers to give us a finite answer when added together. This condition is that the infinite list needs to go towards zero quickly enough. There are infinite lists that go towards zero, but nevertheless cannot be added together to give a finite sum. One example is:

$$\frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \dots$$

Although the numbers go towards zero, they do not do so quickly enough, and you can show that their sum is infinite.

Using calculus, you can look at the infinitely small by looking at the behaviour of infinite sequences that get smaller and smaller. Calculus is used to look at any process that happens over time. It's used in so many different fields: all the physical sciences, computer science, statistics, economics, engineering, medicine, and more. You can look at change over time, change in the instant of time, and you can look at the accumulation of what has happened over a period of time.

Typically, people meet calculus for the first time when asked to look at a rate of change for a phenomenon that has a changing rate of change: speed, rates of infection, value of share prices, etc. In the background there is the mathematics of infinite sequences that get smaller and smaller, allowing us to look at ever smaller intervals of time, and to know what's happening at the infinitely small instant of time. It's quite remarkable that the mathematics of the infinite provides practical solutions for everyday life in this way, as well as solutions to age-old paradoxes.

6 The paradox of fractals

Fractals are beautiful objects that combine ideas of the infinitely big, the infinitely small, dimensions of space and practical real-world applications. Here are a few examples.

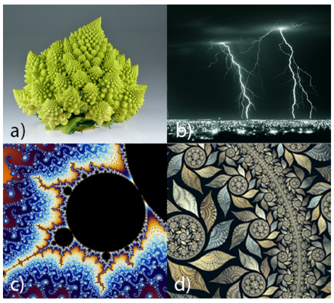


Figure 10 Fractals: (a) Romanesco broccoli, (b) lightning strike, (c) partial view of the Mandelbrot set, (d) an abstract computer generated fractal

Fractals are objects which contain copies of themselves. They appear similar at all levels of magnification. It's sort of like Doctor Who's TARDIS being 'bigger on the inside', but in this case each TARDIS has another TARDIS inside. Or you could think of them like nesting Russian dolls that go on forever.

Figure 11 Russian dolls

One relatively simple fractal is the Koch snowflake, displayed in Figure 12. A Koch snowflake is obtained by taking an equilateral triangle and removing the middle third of each edge and replacing it with a smaller equilateral triangle. You repeat the process with the resulting shape. You remove the middle third of each edge and replace it with an equilateral triangle. This process is repeated and repeated. The Koch snowflake is the mathematical object you have after infinitely many repetitions.

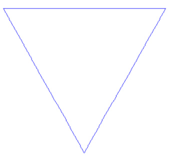


Figure 12 Koch snowflake

It has a finite area. You can show this by enclosing it in a finite circle (or any other shape). But its perimeter is infinite. It's infinitely 'wiggly'. Here is where mathematics gets really mind-bending, as the boundary is said to have a 'fractional dimension', which works differently than the 1-dimensional or 2-dimensional description you might expect for a flat object (it actually has a dimension of 1.26 in this case). Fractals are an interesting subject in their own right – one that quickly becomes mathematically complex and counter-intuitive.

Another relatively simple fractal to describe is the Sierpinski triangle, seen in Figure 13. To make a Sierpinski triangle, you start the first stage with a solid shaded equilateral triangle.

At the second stage, you split the triangle into four equal equilateral triangles and remove the middle triangle. At the third stage, you split the three remaining shaded triangles into four equal equilateral triangles each and remove the middle triangles. The final object is made once this process has run infinitely.

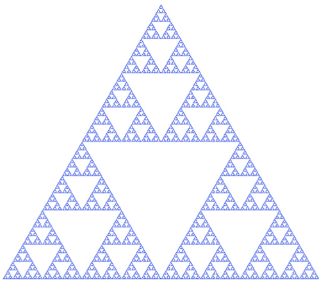


Figure 13 Sierpinski triangle

Activity 2 The making of a Sierpinski triangle

Allow about 5 minutes

Using the information above, see if you can complete this table with the number of triangles at each stage.

Table 1 Triangles by stage

Stage	Number of shaded triangles
1	1
2	3
3	*Provide your answer...*
4	*Provide your answer...*
5	*Provide your answer...*

Answer

Table 1 Triangles by stage

Stage	Number of shaded triangles
1	1
2	3
3	9
4	27
5	81

Figure 14 shows these five stages in action.

Figure 14 The making of a Sierpinski triangle

How many shaded triangles will there be at the nth stage?
(Hint: note that at stage 2, you have 3^1. At stage 3, you have 3^2. Now check the other stages in terms of powers of three.)

Provide your answer...

Answer

The answer is 3^{n-1}.

Apart from giving rise to attractive images, the fractal process was a game changer in the animation industry. Before the advent of this process, animations were somewhat flat and incredibly laborious. Loren Carpenter realised he could use the fractal process to give any desired level of complexity to the images in his animations. The process is used to make surfaces have texture, as you can see in Video 4. This work revolutionised animated film.

Video content is not available in this format.
Video 4 Fractals in animation

7 This week's quiz

Well done for reaching the end of Week 7. Check what you've learned by taking the end-of-week quiz.

Week 7 practice quiz

Open the quiz in a new window or tab then come back here when you've finished.

8 Summary of Week 7

At first glance, infinity seems like an esoteric and impractical subject. But as it turns out, thoughts of infinity gave rise to calculus, which is applicable in every physical science.

Many of the concepts in this course are counter-intuitive and hard to visualise. They show that mathematics will always be a key tool in understanding such mind-bending ideas. Our brains may have evolved to deal with the 'here and now' and the finite, but mathematics can be used to prove notions that we might otherwise struggle to accept or comprehend, like which infinities are the same, and which are bigger than others.

In the final week of the course you will look at the history and future of mathematics, before bringing together what you've learned about the limits of scientific knowledge.

You can now move on to Week 8.

Week 8: What can we know?

Introduction

Eugene Wigner, a theoretical physicist who received the Nobel Prize in Physics four years later, concluded a 1959 lecture on ‘The unreasonable effectiveness of mathematics in the natural sciences’ with the words:

The miracle of the appropriateness of the language of mathematics for the formulation of the laws of physics is a wonderful gift which we neither understand nor deserve. We should be grateful for it and hope that it will remain valid in future research and that it will extend, for better or for worse, to our pleasure, even though perhaps also to our bafflement, to wide branches of learning.

(Wigner, 1959)

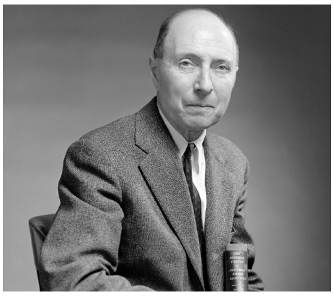


Figure 1 Eugene Wigner (1902–1995)

Having seen the power of mathematics demonstrated throughout this course, you may be surprised to see an eminent particle physicist who obtained the Nobel prize (‘for his contributions to the theory of the atomic nucleus and the elementary particles, particularly through the discovery and application of fundamental symmetry principles’) express his astonishment at the fact that mathematics provides an appropriate ‘language’ to describe the laws of nature. How can that be?

In this final week, you’ll take a broader – and somewhat more philosophical – look at the link between mathematics and the sciences, and the nature of mathematics itself. In particular, this week will explore what’s meant by a ‘theory’, and how we decide which ‘theory’ to accept.

Here’s Marcus to introduce this week’s topic.

Video content is not available in this format.

Video 1 Introduction

By the end of this week, you should be able to:

- appreciate the role of mathematics in science, and its limitations
- reflect on whether mathematics was invented or discovered
- understand what is meant by a ‘theory’ of nature
- discuss philosophical questions relating to a mathematical description of nature
- appreciate the limits to what we can know.

1 Mathematics

The word ‘mathematics’ is derived from the ancient Greek word μάθημα (máthēma), which means knowledge, study and learning. Traditionally, mathematics as a subject has been very closely associated with the sciences, particularly physics. You’ll find mathematics and natural sciences under the same umbrella at many universities (e.g., the Open University’s STEM faculty, which stands for Science, Technology, Engineering and Mathematics).

Many advances in mathematics throughout history were inspired by the study of natural processes. Just think of Newton’s laws of motion and the development of calculus, as discussed earlier in this course. The reverse is also true; an example of this was seen in Week 3, with the existence of particular elementary particles being predicted by mathematical symmetry.

Another classic example came from irregularities in the orbit of the planet Uranus observed in the 1840s. Assuming that Newton’s laws still correctly describe planetary motion at such large distances, the orbit of Uranus could only be explained by the presence of another planet. Its existence and position in the sky were predicted by the laws of planetary motion. The discovery of the planet Neptune in 1846 slotted in perfectly with these mathematical predictions.

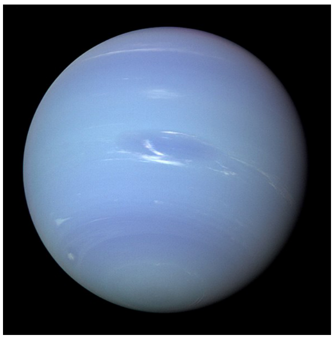


Figure 2 Neptune – a mathematical discovery

Despite such close links, mathematics itself is not considered a ‘natural science’, which the Oxford English Dictionary (OED) defines as:

The branch of knowledge that deals with the natural or physical world; a life science or physical science, such as biology, chemistry, physics, or geology; (in plural) these sciences collectively, in contrast to the social sciences and human sciences.

Nevertheless, the close link is apparent when looking at the definition for ‘mathematics’, which reads:

Originally: (a collective term for) geometry, arithmetic, and certain physical sciences involving geometrical reasoning, such as astronomy and optics; spec. the disciplines of the quadrivium collectively. In later use: the science of space, number, quantity, and arrangement, whose methods involve logical reasoning and usually the use of symbolic notation, and which includes geometry, arithmetic, algebra, and analysis; mathematical operations or calculations. (…)

When the modern subject is studied as an abstract deductive science in its own right, it is often referred to more fully as pure mathematics; when applied to the modelling of physical objects and processes (e.g. in astronomy, various branches of physics, engineering, etc.) and random processes (in probability), and to the handling of data, its full name is applied mathematics. (…)

So, just what is mathematics really? The word ‘science’ is used in the definition here. But if science is something that describes nature, then is mathematics ‘natural’? Or is it an invention of humans?

The question of whether mathematics was ‘invented’ or ‘discovered’ has led to extensive discussion and debate, and there are good arguments on both sides. On the one hand, it’s very hard to conceive how some fundamental mathematical objects, such as the aptly named ‘natural numbers’ (1, 2, 3, 4, 5…) would not be part of any quantitative description of nature. On the other hand, the foundations of mathematics consist of a number of axioms. Going back to the OED, an ‘axiom’ is:

A self-evident proposition, requiring no formal demonstration to prove its truth, but received and assented to as soon as mentioned.

This means that our entire foundation of mathematics is based on a set of statements that we, as humans, see as self-evidently true, and which are used to deduce further statements using logical arguments. So, we might consider mathematics to be a ‘language’ with a vocabulary that’s particularly suited for quantitative statements. If you look back at Wigner’s statement at the beginning of this week, this is exactly what he does when talking about the ‘appropriateness of the language of mathematics for the formulation of the laws of physics’.

Question 1

What do you think? Was mathematics invented or discovered? Can you think of more arguments for either view?

Provide your answer...

Discussion

You might have considered this in a number of different ways.

In an essay on the topic, the distinguished mathematician Timothy Gowers discusses the difference between discovery and observation, beginning with a variety of non-mathematical analogies about the kinds of thing we ‘discover’ – Tutankhamun’s tomb, geographical expeditions, the electron – and those we ‘invent’, such as steam engines, art movements (e.g. cubism), and the rules of cricket. This is in large part a linguistic/semantic discussion, but it’s also psychological, in that some instances simply ‘feel’ more like discovery, and others more like invention. He observes:

The nonmathematical examples suggest that discoveries and observations are usually of objects or facts over which the discoverer has no control, whereas inventions and creations are of objects or procedures with many features that could be chosen by the inventor or creator.

(Gowers, 2012)

Focusing the question on mathematics, Gowers says in reference to the quadratic formula:

> Whoever first derived that formula did not have any choice about what the formula would eventually be … the formula itself was a discovery [but] different people have *come up with* different ways of expressing it.
>
> (Gowers, 2012)
>
> But it turns out to be difficult to condense the history of any specific mathematical inquiry into solely discovery or invention. The Pythagorean theorem is a classic example of this conflict. This theorem had to be discovered (as an aside: it was known to the Mesopotamians long before Pythagoras was supposed to have lived), but many proofs of it have been invented. Indeed, a book on the theorem by Elisha Scott Loomis published in 1927 listed over 350 different proofs.
>
> Within the essay conclusion, Gowers determines:
>
> … there does seem to be a spectrum of possibilities, with some parts of mathematics feeling more like discoveries and others more like inventions. It is not always easy to say which are which, but there does seem to be one feature that correlates strongly with whether we prefer to use a discovery-type word or an invention-type word. That feature is the control that we have over what is produced.
>
> (Gowers, 2012)

2 A brief excursion into philosophy

Relying on facts that appear self-evident can lead us astray. To demonstrate this, take a look at the work of one of the great philosophers, Immanuel Kant. His *Kritik der reinen Vernunft* (‘Critique of Pure Reason’) was first published in 1781.

Figure 3 Immanuel Kant (1724–1804)

There are many different translations of the text – the quotes here are from the 1855 translation by John Miller Dow Meiklejohn (who remarkably translated this momentous work as a teenager).

Being a work of philosophy from two and half centuries ago, it’s written in a style that may be challenging to understand out of context. Don’t worry if you find the quotes a little hard to follow.

This work contains extensive discussion of the relationship between mathematical theory and philosophy. Kant discusses several ‘antinomies’ or contradictions which, in his view, arise necessarily from our attempts to conceive the nature of reality. In his work, these contradictions take the form of a ‘thesis’ and ‘antithesis’ expressing propositions that are mutually exclusive and collectively exhaustive. In other words, both of them cannot be true, and both of them cannot be false, and yet it seems both of them can be proved! Here’s an example of one such antinomy:

Thesis: The world has a beginning in time, and is also limited in regard to space.

Antithesis: The world has no beginning, and no limits in space, but is, in relation both to time and space, infinite.

(Kant, 1787, p. 266)

Kant proceeds to prove both of these statements, following a few basic assumptions to their logical conclusions. As an example, let us look at the ‘time’ part. Kant’s proof for the thesis above is:

Granted, that the world has no beginning in time; up to every given moment of time, an eternity must have elapsed, and therewith passed away an infinite series of successive conditions or states of things in the world. Now the infinity of a series consists in the fact, that it never can be completed by means of a successive synthesis. It follows that an infinite series already elapsed is impossible, and that consequently a beginning of the world is a necessary condition of its existence.

(Kant, 1787, p. 266)

He then proves the antithesis (i.e., that the world has no beginning) as follows:

For let it be granted, that it has a beginning. A beginning is an existence which is preceded by a time in which the thing does not exist. On the above supposition, it follows that there must have been a time in which the world did not exist, that is, a void time. But in a void time the origination of a thing is impossible; because no part of any such time contains a distinctive condition of being, in preference to that of non-being (whether the supposed thing originate of itself, or by means of some other cause). Consequently, many series of things may have a beginning in the world, but the world itself cannot have a beginning, and is, therefore, in relation to past time, infinite.

(Kant, 1787, p. 266)

From a modern perspective, it's fairly clear to see that the issue causing the contradiction lies with the concepts of space and time that Kant employed. Our understanding of space and time has moved on with the development of Einstein's theory of relativity, which links space and time in an intricate manner. The simple view of space and time as static entities in which everything develops is no longer appropriate – space and time interact with each other in a non-trivial way. We can now contemplate the possibility of the universe being a closed space that is finite without having a boundary – much like the surface of the Earth does not have a boundary, despite being evidently finite. However, this is something we have learned from science; we cannot directly experience this link between space and time. Our brains are trained by experience, and in our everyday experience space and time are separate entities. As a result, we cannot visualise the 'true' structure of space-time.

Similarly, we base mathematics on what we perceive as self-evident assumptions. However, what we see as self-evident is arguably influenced by the way our thinking has developed, based on our experience of the world around us. We might naively assume that any alien intelligence would use the same mathematics as we do, and that mathematics is in this sense 'universal', but this is far from assured.

Question 2

Can you imagine what Kant's argument for the 'space' part of this antinomy looks like? Consider this, then reveal the discussion below.

You might want to look closer at the source if this interests you. It's freely available online from Project Gutenberg: The Critique of Pure Reason (open in a new tab or window so you easily return here). These particular arguments are located in 'Section II. Antithetic of Pure Reason'.

Discussion

Here are the relevant bits of text. It's important to note Kant's use of proof by contradiction here, beginning by taking the opposite for granted.

Proof for the thesis:

… let us take the opposite for granted. In this case, the world must be an infinite given total of coexistent things. Now we cannot cogitate the dimensions of a quantity, which is not given within certain limits of an intuition, in any other way than by means of the synthesis of its parts, and the total of such a quantity only by means of a completed synthesis, or the repeated addition of unity to itself. Accordingly, to cogitate the world, which fills all spaces, as a whole, the successive synthesis of the parts of an infinite world must be looked upon as completed, that is to say, an infinite time must be regarded as having elapsed in the enumeration of all co-existing

> things; which is impossible. For this reason an infinite aggregate of actual things cannot be considered as a given whole, consequently, not as a contemporaneously given whole. The world is consequently, as regards extension in space, not infinite, but enclosed in limits.
>
> (Kant, 1787, p. 266)
>
> Proof for the antithesis:
>
> … let us first take the opposite for granted—that the world is finite and limited in space; it follows that it must exist in a void space, which is not limited. We should therefore meet not only with a relation of things in space, but also a relation of things to space. Now, as the world is an absolute whole, out of and beyond which no object of intuition, and consequently no correlate to which can be discovered, this relation of the world to a void space is merely a relation to no object. But such a relation, and consequently the limitation of the world by void space, is nothing. Consequently, the world, as regards space, is not limited, that is, it is infinite in regard to extension.
>
> (Kant, 1787, p. 266)

3 What is a theory?

What do we mean when we talk about a ‘theory’? Think about the examples you’ve seen in this course: Newton’s theory of gravity and Einstein’s theory of gravity (part of his general theory of relativity). In the end, a theory is a model (usually framed in the language of mathematics) of some aspects of the natural world, which is meant to describe how things behave, and explain the outcome of experiments. Such experiments play a key role here – they allow us to test our model, and are often used to inspire new models. Throughout the history of science, our theories have evolved. New experiments test our models in more and more detail, and we find deviations that show that the model is, at best, incomplete.

The two theories of gravity are a good example of this process. Einstein’s theory can explain small deviations observed by experiments, such as providing a precise explanation for the motion of the planet Mercury. It can be seen as a generalisation of Newton’s theory, in the sense that the latter is a good approximation of Einstein’s theory for many everyday purposes.

But how do we choose one theory over another? If one theory successfully explains more observations than another, you would assume that it provides more insight into nature, and is therefore superior. But what if you have two (or more) different theories in your hands, each of which explain some observation equally well? What can you do then? One approach to this dilemma is to come up with an experiment which gives different answers depending on which theory has been applied, and then use this experiment to make a choice. Similarly, when a new theory predicts something which has not yet been observed, and you then go out and observe it exactly as predicted, it lends strong credence to the theory. A good example of this approach, discussed earlier in the course, is how support was gathered for Einstein’s theory of general relativity.

Another approach would be to look at the power of prediction of the theory. Here’s how this works. Suppose you have two different theories which both explain the behaviour of a system equally well. One of the theories uses a few fundamental constants (such as the speed of light in vacuum, or Planck’s constant) while the other contains tens or hundreds of parameters (numbers that enter the equations in the theory) which all have to be chosen correctly for the theory to work, without any information as to where these parameter values come from. You can argue that the first theory possesses more ‘predictive power’ than the second, because it requires less input to describe the behaviour of the system. Trying to reduce the number of parameters is a common approach to find a more ‘fundamental’ theory.

In fact, one reason why scientists are not completely happy with the current Standard Model of particle physics is that it includes quite a few parameters (many of them related to particle masses and the way that particles ‘mix’ when considering different interactions). The theory does not predict values for these parameters; they have to be determined by experiment. Many scientists believe that there should be a ‘reason’ for the values that these parameters take, and that there should be a theory that will predict their values.

Another highly significant aspect is the relationship between theories explaining different aspects of nature. There is a belief that there should be a ‘theory of everything’ that describes all natural processes consistently. Currently, we have quantum theory, which works well to predict physics at the smallest scales, and Einstein’s theory of general relativity, which works well to describe the structure of the universe. However, as far as we know right now, there seems to be no way to consolidate these two theories into a single theory that would encompass both. If an alternative theory were found for either that would allow for such a unified theory – while agreeing with experiments – it would certainly be favoured.

The discussion of theories can reach into more subjective territory. For example, the ‘beauty’ of the mathematical theory has become a matter of some controversy since the late 2010s, as will be discussed in the next section.

4 Lost in maths?

Sabine Hossenfelder's 2018 book *Lost in Math: How Beauty Leads Physics Astray* (described as 'provocative' by the New York Times) argues that striving for mathematical beauty in physics is an aberration that has used up enormous resource but failed to produce any tangible results.

Figure 4 Sabine Hossfelder

An example of this is the pursuit of 'supersymmetry' in particle physics. This approach starts from a larger symmetry than the current model, which has some mathematical advantages, and provides a relationship between the two fundamental types of particles: bosons and fermions. This is appealing, because it may allow us to consider bosons and fermions together rather than separately. It could also explain some of the parameter values in the Standard Model, such as the mass of the famous Higgs boson.

However, there is a problem. Theories based on supersymmetry predict the presence of additional particles, with each particle in the Standard Model possessing a 'partner' particle. None of these partner particles has ever been observed, and it's becoming increasingly difficult to accommodate the experimental limits on the presence of such particles.

It's important to note that we've seen examples of theories being proposed because of their mathematical appeal, which have then proved successful (such as the eightfold way, discussed in Week 2, which predicted the presence of a particle that was later discovered). There is a clear case for mathematical structure driving the development of theories. Ultimately, however, experimental verification must be obtained.

In an article adapted from her book, Hossenfelder writes:

The philosophers are certainly right that we use criteria other than observational adequacy to formulate theories. That science operates by generating and subsequently testing hypotheses is only part of the story. Testing all possible hypotheses is simply infeasible; hence most of the scientific enterprise today—from academic degrees to peer review to guidelines for scientific conduct—is dedicated to identifying good hypotheses to begin with. Community standards differ vastly from one field to the next and each field employs its own quality filters, but we all use some. In our practice, if not in our philosophy, theory assessment to preselect hypotheses has long been part of the scientific method. It doesn't relieve us from experimental test, but it's an operational necessity to even get to experimental test.

(Hossenfelder, 2018b)

This highlights a conundrum: even if you were to dismiss 'mathematical elegance' as a criterion, you still need some way to choose which theories you may consider worthy of attention. Implicitly, scientists apply some ‘quality filter’ (to use Hossenfelder’s term) based on their experience and intuition.

5 The limitations of mathematics

This course has considered various scientific fields, and the possible limitations to our knowledge in each case. The key question throughout has been: what is it that we cannot know? Mathematics is different from the sciences as it is ‘axiomatic’ (based on a set of axioms), and all mathematical statements and results arise as logical conclusions from these axioms. So, does this mean that there’s nothing in mathematics that we can never know?

David Hilbert, whose ‘Infinite Hotel’ was discussed in Week 7, believed that this was the case. In 1900, at the International Congress of Mathematicians in Paris, he made a speech setting out the 23 greatest unsolved problems for mathematics in the twentieth century, in which he said:

This conviction of the solvability of every mathematical problem is a powerful incentive to the worker. We hear within us the perpetual call: There is the problem. Seek its solution. You can find it by pure reason, for in mathematics there is no *ignorabimus*.

(Reid, 1996, p. 81)

Here Hilbert used the Latin word *ignorabimus*, which means ‘we will not know’ (the English verb ‘ignore’ derives from the Latin verb). Hilbert strongly believed that there’s nothing in mathematics that we cannot ever know. He still held this view in 1930, giving a talk in the Prussian city of Königsberg where he said:

The true reason why [no one] has succeeded in finding an unsolvable problem is, in my opinion, that there is no unsolvable problem.

(Reid, 1996, p. 196)

Hilbert didn’t know it, but just one day earlier, at a conference held in the same town, a young Austrian logician named Kurt Gödel had proved him wrong.

6 Gödel's incompleteness theorems

Language contains many paradoxes and contradictions. But when it comes to numbers, we expect a statement to be either true or false. Hilbert believed mathematics could reach all the answers within one overarching axiomatic system. Kurt Gödel didn't.

Gödel was a student of mathematics in Vienna in the 1920s, with an interest in mathematical logic and philosophy. He had been gaining international stature when he informally introduced his incompleteness theorems at another Königsberg event in 1930. His subsequent work in this area went on to be enormously influential.

Figure 5 Kurt Gödel (1906–1978)

Gödel disputed the idea that an axiomatic system could produce mathematical proof for everything – that a formalised mathematical system could ever be 'complete'. Remember that an 'axiomatic' system is one that's described by a set of pre-determined truths held as self-evident, from which everything else is logically derived.

Gödel's work was extremely complex and cannot be explained in detail here, but he essentially produced a system of code numbers for representing mathematical axioms, and statements about those axioms. Whether a statement is true or false was now translated into solving numerical equations.

Gödel's 'incompleteness theorems' highlight the limitations of mathematics, and indeed of any axiomatic system. They show that within the system, there will be true statements which we will never be able to prove are true. And impressively, it's a mathematical proof that demonstrates this mathematical boundary.

In short, his first incompleteness theorem shows that no consistent system of axioms whose theorems can be listed by an effective procedure, such as an algorithm, can prove all true statements about the arithmetic of natural numbers. The second incompleteness theorem shows that such a system cannot demonstrate its own consistency.

With these theorems, Gödel showed that it's impossible to create a set of axioms that explains everything in maths. That, in a manner of speaking, our knowledge will never be complete – it's like a poorly designed jigsaw puzzle, where we might have all the pieces yet they don't quite fit together.

7 This week's quiz

Well done for reaching the end of Week 8.

Now it's time to complete the Week 8 badged quiz. It's similar to previous quizzes, but this time instead of answering five questions there will be fifteen, covering material from the last four weeks of the course.

Week 8 compulsory badge quiz

Remember, this quiz counts towards your badge. If you're not successful the first time, you can attempt the quiz again in 24 hours.

Open the quiz in a new window or tab then come back here when you've finished.

8 Summary of Week 8

This week has explored the limits of mathematics, and the important questions that arise around scientific theory. Our description of nature is based on theories which are mathematical models of nature that have been corroborated by experimental observations.

So, can we ever know whether a theory is actually correct? And can we be certain that it will never need to be refined or replaced? The answer is a resounding ‘no’! All we can say is that a theory explains all experimental observations we have made so far. We’ll never be certain that discrepancies won’t arise as we look deeper into nature, whether this involves shorter/longer times and distances than we could access before. Theories have been developing in line with scientific progress, and will continue to develop, through experimental discoveries that contradict predictions from the current theories, and our continuing efforts to improve and unify existing theories.

We can never be sure that we have arrived at the final stage of this endeavour, but with each step we gain additional insight into the mysteries of nature.

9 Conclusion: what we cannot know

Congratulations on completing this badged course, *Understanding science: what we cannot know*.

During this course, you've explored the edges of our knowledge in a variety of scientific fields. You've reflected on what we currently know, and what we don't. Going beyond that, you've examined our potential knowledge and considered the fundamental underlying question: is there anything we cannot know?

Figure 6 What can we know?

Here's Marcus to close out the course.

Video content is not available in this format.

Video 2 Course conclusion

Where next?

If you've enjoyed this course, you can find many more free resources and courses right here on OpenLearn.

Considering University study? You may be interested in our courses on Science or Mathematics.

Making the decision to study can be a big step. The Open University has over 50 years of experience supporting its students through their chosen learning paths. You can find out more about studying with us by visiting our online prospectus.

Tell us what you think

Now you've come to the end of the course, we would appreciate a few minutes of your time to complete this short end-of-course survey (you may have already completed this survey at the end of Session 4). We'd like to find out a bit about your experience of studying the course and what you plan to do next. We will use this information to provide better online experiences for all our learners and to share our findings with others. Participation will be completely confidential and we will not pass on your details to others.

References

Laplace, P.-S. (1814) *A Philosophical Essay on Probabilities*. Translated from the French by F.W. Truscott and F.L Emory. New York: J. Wiley.

Poincaré, H. (1908) *Science and Method*. Translated from the French by F. Maitland. London: T. Nelson.

Wheeler, J. A. (1998) *Geons, Black Holes and Quantum Foam: A Life in Physics*. New York: W.W. Norton & Co.

Gowers, T. (2012) 'Is Mathematics Discovered or Invented?' in M. Pitici (ed.) *The Best Writing on Mathematics 2012*. Princeton University Press, pp. 28–36.

Hossenfelder, S. (2018a) *Lost in Math: How Beauty Leads Physics Astray*. New York: Basic Books.

Hossenfelder, S. (2018b) 'A Theory with No Strings Attached: Can Beautiful Physics Be Wrong?', *Scientific American*, 11 June. Available at: https://www.scientificamerican.com/article/a-theory-with-no-strings-attached-can-beautiful-physics-be-wrong-excerpt (Accessed: 5 August 2022).

Kant, E. (1787) *The Critique of Pure Reason*. Translated from the German (1855) by J.M. D. Meiklejohn. London: Henry G. Bohn. Available at: https://www.gutenberg.org/ebooks/4280 (Accessed: 5 August 2022).

Reid, C. (1996) *Hilbert*. New York: Springer-Verlag.

Wigner, E.P. (1959) 'The unreasonable effectiveness of mathematics in the natural sciences' (Courant Lectures, New York University, 11 May 1959), *Communications in Pure and Applied Mathematics*, 13, pp. 1–14. Available at: https://doi.org/10.1002/cpa.3160130102 (Accessed: 5 August 2022).

Further reading

Section 5.2: An example of the mirror self-recognition test in elephants.

Section 5.3: You might be interested in another OpenLearn course, Machines, minds and computers. Section 4.3 'What computers can't do?' is particularly relevant.

Section 5.3: Further discussion of the Chinese Room Argument from the Stanford Encyclopedia of Philosophy.

Acknowledgements

This free course was written by June Barrow-Green, Katie Chicot, Uwe Grimm, Tony Jones and Amanda Smith. It is adapted from the book *What We Cannot Know* by Marcus du Sautoy.

Images

Introduction and guidance

Course image: adapted from https://www.pexels.com/@marek-piwnicki-3907296

Week 1

Figure 1:

(a) Blaise Pascal (1623–1662): https://commons.wikimedia.org/wiki/File:Pascal-old.png

(b) Pierre de Fermat (1607–1665): https://commons.wikimedia.org/wiki/File:Pierre_de_-Fermat.jpg

Figure 2: VitalyEdush/Getty Images

Figure 5:

(a) Isaac Newton (1643–1727): https://commons.wikimedia.org/wiki/File:Portrait_of_-Sir_Isaac_Newton,_1689.jpg

(b) Gottfried Leibniz (1646–1716): https://commons.wikimedia.org/wiki/File:Christoph_-Bernhard_Francke_-_Bildnis_des_Philosophen_Leibniz_(ca._1695).jpg

Figure 6: brgfx/Freepik

Figure 7: Mary Somerville (1780–1872): https://commons.wikimedia.org/wiki/File:Thomas_Phillips_-_Mary_Fairfax,_Mrs_William_Somerville,_1780_-_1872._Writer_-on_science_-_Google_Art_Project.jpg

Figure 8: Henri Poincaré (1854–1912) Photograph: Smithsonian Institution from United States: https://commons.wikimedia.org/wiki/File:Portrait_of_Henri_Poincar%C3%A9_(1854-1912),_Mathematician_(2551042945).jpg

Figure 9: The trajectory of a third body interacting with a large mass (Earth, left) and a small mass (Moon, right) from July 19, 2019 by DAVID D. NOLTE: Figure 1 from Getting Armstrong, Aldrin and Collins Home from the Moon: Apollo 11 and the Three-Body Problem: https://galileo-unbound.blog/2019/07/19/getting-armstrong-aldrin-and-collins-home-from-the-moon-apollo-11-and-the-three-body-problem

Week 2

Figure 1: Periodic table of elements. Humdan/Shutterstock Images

Figure 2: Particle accelerator (c) Jason Schwartz: https://commons.wikimedia.org/wiki/File:HERA_Tunnel.JPG https://creativecommons.org/licenses/by-sa/3.0/deed.en

Figure 3: Patterns in nature:

(a) apis florea nest, Sean Hoyland: https://commons.wikimedia.org/wiki/File:Apis_florea_nest_closeup2.jpg;

(b) chrysanthemum flower, monika1607/pixabay;

(c) nautilus spiral, Rpsycho/Getty Images;

(d) snowflake, Alexey Kljatov in Flickr https://creativecommons.org/licenses/by/2.0

Figures 7 and 8: Proton/neutron quark structure: Jacek rybak/Category:CC-BY-SA-4.0 - Wikimedia Commons

Week 3

Figure 1:

(a) Albert Einstein (1879–1955): https://en.wikipedia.org/wiki/File:Albert_Einstein_(Nobel).png;

(b) Robert Millikan (1868–1953),: https://commons.wikimedia.org/wiki/File:Robert_Andrews_Millikan_1920s.jpg

Figures 2 and 3: Diagram of a wave (not labelled and labelled): NOAA U.S. National Oceanic and Atmospheric Administration

Figure 4: Animation of a sonic boom: Jacopo Bertolotti/Creative Commons CC0 1.0 https://commons.wikimedia.org/wiki/File:Sonicboom_animation.gif

Figure 5: The electromagnetic spectrum: Penubag/ Creative Commons — Attribution-ShareAlike 3.0 Unported — CC BY-SA 3.0.5 https://commons.wikimedia.org/wiki/File:Electromagnetic-Spectrum.png

Figure 9: A half-life simulation: Sbyrnes321: https://commons.wikimedia.org/wiki/File:Halflife-sim.gif

Figure 10: Casimir effect: Emok: https://commons.wikimedia.org/wiki/File:Casimir_-plates.svg

Week 4

Figure 1: The Milky Way: Xing-Wu Zheng & Mark Reid BeSSeL/NJU/CFA Copyright © 2019 by Xing-Wu Zheng and Mark Reid

Figure 2: Superclusters: Richard Powell: https://commons.wikimedia.org/wiki/File:Superclusters_atlasoftheuniverse.gif Creative Commons — Attribution-ShareAlike 2.5 Generic — CC BY-SA 2.5

Figure 3: M31 Andromeda: Luca Argalia: https://www.flickr.com/photos/skiwalker79/5309249905/ //creativecommons.org/licenses/by-nc-sa/2.0/

Figure 4: Screenshot of the Stellarium Web interface: Fabien Chéreau (creator of Stellarium) / Stellarium developer group https://stellarium-web.org/

Figure 6: The Extreme Deep Field: NASA, ESA, G. Illingworth, D. Magee, and P. Oesch (University of California, Santa Cruz), R. Bouwens (Leiden University), and the HUDF09 Team

Figure 7: A map of the CMBR: European Space Agency (ESA) and the Planck Collaboration

Figure 10: Positive curvature: Naypong/Getty Images

Figure 11: Negative curvature: Marat Musabirov/Getty Images Plus

Week 5

Figure 1: National Physical Laboratory (NPL): https://www.npl.co.uk/famous-faces/louis-essen

Figure 5: A photographic negative of the eclipse of 1919: https://commons.wikimedia.org/wiki/File:1919_eclipse_positive.jpg

Figure 6: Space flowing into a black hole: Courtesy Professor Andrew Hamilton https://jila.colorado.edu/~ajsh/insidebh/waterfall.html

Figure 7: Cygnus A: National Radio Astronomy Observatory (NRAO) / Associated Universities, Inc. (AU): https://astronomynow.com/2016/10/25/hotspots-in-cygnus-a-an-active-galactic-nucleus

Figure 8: The black hole in the centre of the galaxy M87; Event Horizon Telescope Collaboration: https://eventhorizontelescope.org https://creativecommons.org/licenses/by/4.0/

Figure 9: The process of 'spaghettification' NASA / Laura A. Whitlock, Kara C. Granger, Jane D. Mahon: https://en.wikipedia.org/wiki/File:Spaghettification_(from_NASA%27s_Imagine_the_Universe!)

Week 6

Figure 1:

(a) the human brain viewed from the side © The Open University;

(b) in a photograph from above © Unidentified;

(c) Left: a cut-away view through the middle of the brain © The Open University; Right: A post-mortem human brain sample with top layer removed: Dr. Terence Williams, University of Iowa.

Figure 5: Broca's area (shown in red): Polygon data were generated by Database Center for Life Science (DBCLS) https://creativecommons.org/licenses/by-sa/2.1/jp/deed.en

Figure 7: 'Golgi stained' brain tissue viewed under a microscope:

(a) The OU;

(b) Anna Kamitakahara and Richard Simerly, PhD, from The Saban Research Institute of Children's Hospital Los Angeles

Figure 8: Neurons in 'Golgi stained' brain tissue (drawings by Santiago Ramón y Cajal):

(a) Santiago Ramón y Cajal / public domain: https://commons.wikimedia.org/wiki/File:Purkinje_cell_by_Cajal.png;

(b) Santiago Ramon y Cajal / public domain: https://en.wikipedia.org/wiki/File:Cajal_cortex_drawings.png

Figure 10: An electron micrograph showing synapses. In Lodish et al., 4th edition

Figure 11:

(a): EEG headcap, used to apply electrodes to the surface of the head; Min Jing/Shutterstock;

(b): different types of brain wave activity measured using EEG: ScienceDirect Technological Basics of EEG Recording and Operation of Apparatus: Priyanka A. Abhang, Suresh C. Mehrotra, in Introduction to EEG- and Speech-Based Emotion Recognition, 2016 Figure 2.1. Brain wave samples with dominant frequencies belonging to beta, alpha, theta, and delta bands and gamma waves.

Figure 12:

(a) fMRI scanner, courtesy: Neuroimaging Core, Center for Integrative Neuroscience, University of Nevada, Reno;

(b) fMRI scan Dr Krish Singh

Figure 13: fMRI scans showing brain activity while tasks are imagined: Courtesy: Adrian M. Owen, OBE, PhD

Figure 14: TMS: Mayo Foundation for Medical Education and Research

Figure 15: The roundworm *C. elegans*: https://commons.wikimedia.org/wiki/File:CrawlingCelegans.gif https://creativecommons.org/licenses/by-sa/3.0/deed.en

Figure 16: The roundworm connectome: from: Controlling the *C. elegans* neural network (1c) from authors of Nature article: Gang Yan, Petra E. Vértes, Emma K. Towlson, Yee Lian Chew, Denise S. Walker, William R. Schafer & Albert-László Barabási

Figure 17: Tracing neurons: courtesy: Dr. Juan Burrone and Prof. Venkatesh Murthy

Figure 18: Artificial neurons: https://www.bath.ac.uk/announcements/world-first-as-artificial-neurons-developed-to-cure-chronic-diseases / University of Bath

Figure 19: A ‘signature of consciousness’ – patterns of brain activity recorded as a response to conscious (left) and non-conscious (right) visual stimuli (Dehaene and Changeux, 2011) Dehaene, S. and Changeux, J.P. (2011) ‘Experimental and theoretical approaches to conscious processing’, *Neuron*, 70(2), pp. 200–227.

Week 7

Figure 1: The Milky Way: AstroAnthony https://creativecommons.org/licenses/by/4.0/deed.en https://commons.wikimedia.org/wiki/File:Milky_way_nebula.jpg

Figure 2:

(a) Reinhard Genzel: https://commons.wikimedia.org/wiki/File:Reinhard_Genzel.jpg https://creativecommons.org/licenses/by-sa/3.0/deed.en;

(b) Andrea Ghez: Sipa US/Alamy Stock Photo;

(c) Sir Roger Penrose: Biswarup Ganguly https://commons.wikimedia.org/wiki/File:Roger_Penrose_9552.JPG https://creativecommons.org/licenses/by/3.0/deed.en

Figure 3: David Hilbert (1862–1943) Image: University of Göttingen

Figure 6: Fractions between fractions: adapted from https://www.basic-mathematics.com/density-property.html Basic-mathematics.com

Figure 7: Fraction grid: Cronholm144 https://commons.wikimedia.org/wiki/File:Diagonal_argument.svg https://creativecommons.org/licenses/by-sa/3.0/deed.en

Figure 9: Dividing a chocolate bar: Hyacinth https://commons.wikimedia.org/wiki/File:Eye_of_Horus_square.png https://creativecommons.org/licenses/by-sa/3.0/deed.en

Figure 10: Fractals:

(a) Romanesco broccoli: Ivar Leidus https://en.wikipedia.org/wiki/File:Romanesco_broccoli_(Brassica_oleracea).jpg https://creativecommons.org/licenses/by-sa/4.0/deed.en;

(b) lightning strike: José Eugenio Gómez Rodríguez https://www.flickr.com/photos/jegomezr/2926143475 https://creativecommons.org/licenses/by-nc-sa/2.0/;

(c) partial view of the Mandelbrot set: Wolfgang Beye https://en.wikipedia.org/wiki/File:Mandel_zoom_08_satellite_antenna.jpg https://creativecommons.org/licenses/by-sa/2.5/deed.en;

(d) an abstract computer generated fractal: Diep Hoang Hai/Shutterstock.com

Figure 11: Russian dolls: demidoffaleks/Getty Images

Figure 12: Koch snowflake: António Miguel de Campos https://commons.wikimedia.org/wiki/File:Von_Koch_curve.gif

Figure 13: Sierpinski triangle: Beojan Stanislaus https://commons.wikimedia.org/wiki/File:Sierpinski_triangle.svg

Week 8

Figure 1: Eugene Wigner (1902–1995) © unknown

Figure 2: Neptune – a mathematical discovery: NASA/JPL/Voyager-ISS/Justin Cowart

Figure 3: Immanuel Kant (1724–1804): https://commons.wikimedia.org/wiki/File:Kant_-gemaelde_3.jpg

Figure 4: Sabine Hossfelder: File:Sabine Hossenfelder.jpg – Wikimedia Commons – Creative Commons — Attribution-ShareAlike 4.0 International — CC BY-SA 4.0

Figure 5: Kurt Gödel (1906–1978): https://www.flickr.com/photos/levanrami/24246848265 – public domain

Figure 6: What can we know? LEOcrafts/Getty Images

Audio/Video

Week 1

Video 4: The restricted three-body problem by Jim Belk: https://www.youtube.com/watch?v=jarcgP1rRWs Creative Commons — Attribution 3.0 Unported — CC BY 3.0

Week 2

Video 2: Voyage into the world: Daniel Dominguez/CERN

Week 4

Video 2: Timeline of our understanding © The Open University (2022) with thanks to the following:NASA/MSFC/Meteoroid Environment Office/Bill Cooke; NASA https://commons.wikimedia.org/wiki/File:COBE_cmb_fluctuations.png; G. Hüdepohl (atacamaphoto.com)/ESO https://www.eso.org/public/outreach/copyright/; Credit: NASA, ESA, A. Zitrin (California Institute of Technology), and J. Lotz, M. Mountain, A. Koekemoer, and the HFF Team (STScI) https://creativecommons.org/licenses/by/2.0/deed.en; Alfredo Garcia Saz/Alamy Stock Photo; NASA https://commons.wikimedia.org/wiki/File:Horn_Antenna-in_-Holmdel,_New_Jersey.jpeg; Credit: ESA/Hubble & NASA; https://commons.wikimedia.org/wiki/File:Milky_Way_galaxy1.jpg; ESO/S. Brunier https://creativecommons.org/licenses/by/4.0/; NASA Goddard Space Flight Center from Greenbelt, MD, USA https://creativecommons.org/licenses/by/2.0/

Video 3: The size of the universe © The Open University

Week 5

Video 7: Wonders of the Universe – Black Holes; BBC series: Wonders of the Universe. Episode title: Falling. BBC2 © BBC

Week 6

Video 5: Neuron networks viktorov.pro/Shutterstock

Video 6: How neurons communicate: Creative Commons — Attribution-NonCommercial 4.0 International — CC BY-NC 4.0. https://www.brainfacts.org/core-concepts/how-neurons-communicate BrainFacts.org

Video 8: Brain wiring: courtesy: National Institute of Mental Health (NIMH): https://www.nimh.nih.gov/

Video 9: The Rubber Hand Illusion, Horizon, BBC 2 © BBC

Video 11: Neural network simulation: Courtesy Denis Dmitriev: https://www.cybercontrols.org

Week 7

Video 4: Fractals in animation. Clip from BBC Two series from 2011. Marcus du Sautoy interviews Loren Carpenter. Collaboration with The Open University © BBC (2011)

Every effort has been made to contact copyright owners. If any have been inadvertently overlooked, the publishers will be pleased to make the necessary arrangements at the first opportunity.

Don’t miss out

If reading this text has inspired you to learn more, you may be interested in joining the millions of people who discover our free learning resources and qualifications by visiting The Open University – www.open.edu/openlearn/free-courses.