CTYI "The Maths Experience" Handbook

Brian Tyrrell

Summer 2020

Contents

1 Introduction - What is Mathematics?

First and foremost, mathematics is *not* encompassed by the Junior or Leaving Certificate.

"Pure" mathematics, for the most part, is not concerned with solving complicated looking equations, fiddling around with prime numbers, or rehashing old formulas – like is portrayed in TV and film. Pure mathematics as a whole bridges the methodologies of science and philosophy. This doesn't mean that pure mathematics has no applications, but it does mean the object of principal concern is the *theory* behind an idea, and not its implementation or practise. Some of the world's leading universities say this about pure mathematics:

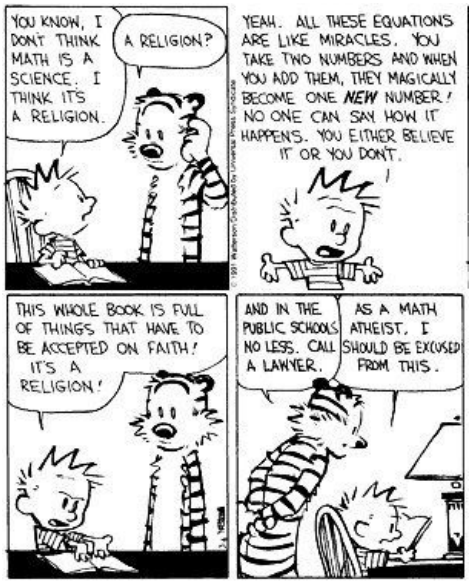
- (MIT) "Its purpose is to search for a deeper understanding and an expanded knowledge of mathematics itself."
- (Oxford) "Above all, mathematics is a logical subject, and you will need to think mathematically, arguing clearly and concisely as you solve problems."
- (Waterloo) "Mathematics is both an art and a science, and pure mathematics lies at its heart. Pure mathematics explores the boundary of mathematics and pure reason. It has been described as "that part of mathematical activity that is done without explicit or immediate consideration of direct application," although what is "pure" in one era often becomes applied later. Finance and cryptography are current examples of areas to which pure mathematics is applied in significant ways."

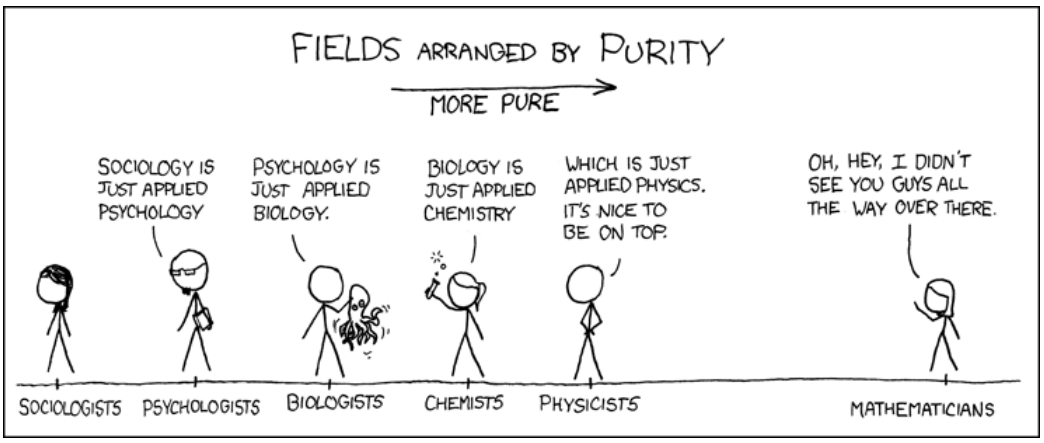
In many ways, mathematics is closer to philosophy than you might realise. In particular, research in one of its largest areas – abstract algebra – has often been compared to studying "applied philosophy", or "philosophy where the questions are extremely precise and technical".

There are many areas in mathematics studied by philosophers alongside mathematicians – set theory and logic being the prime examples. As well as this, mathematics can provide a window into philosophical arguments. Kurt Gödel, a famous logician, provided a *formal* (in the sense of using logical symbols) ontological argument – an argument for the existence of God.

Overleaf I've included some comic strips by writers who try to express this broader view of mathematics. Two of the most well known are:

- Saturday Morning Breakfast Cereal.
- xkcd.





So, what can you do with a degree in (pure) mathematics? Pretty much anything.

- **Academia.** You could continue your studies and research, gaining a PhD (also known as a "doctorate"). This qualifies you to join the worldwide academic community in mathematics, teach in universities, and research.

  Research in pure mathematics is just like research in science – except cheaper. Research is about discovering the solutions to unsolved problems, and as we shall see, mathematics has plenty of open questions.

- **Finance.** Studying pure mathematics is no barrier to a career applying mathematics. (In fact, it's often seen as a bonus: you're proven to have strong problem solving abilities, and the ability to learn hard things fast.)

- Most importantly, **there are other options**. Studying mathematics does not doom you to teach or work for banks. I know pure mathematicians who are now oceanographers, politicians, international relations experts, and have worked in various start-ups. Of course, major corporations in Ireland such as Google and Facebook employ hundreds of mathematicians too.

2 Set Theory

The source for this section is primarily "Introduction to University Mathematics" by Prof. Earl.

We shall begin our study of mathematics in its very foundations: set theory. Set theory is a branch of mathematical logic which studies sets – "collections of objects". Everything we consider in algebra and geometry is a set, so these are the most fundamental (and some would say, most important) objects to study.

Modern set theory began in 1874 with the work of Georg Cantor. His ideas, though controversial at the time, allowed mathematicians to rigorously understand the concept of "infinity". Furthering Cantor's work in the early years of the 20th century, Ernst Zermelo, Abraham Fraenkel, Bertrand Russell, and Thoralf Skolem produced the basic theory of sets and their properties (known as ZFC) that is still used to this day.

Sets are amongst the most primitive objects in mathematics, so primitive in fact that it is somewhat difficult to give a precise definition of what one means by a set – i.e. a definition which uses words with entirely unambiguous meanings. For example, here is a description due to Cantor:

> *By an "aggregate" [a set] we are to understand any collection into a whole M of definite and separate objects m of our intuition or our thought. These objects we call the "elements" of M.*

One might now ask exactly what one means by a "collection" or by "objects", but the point is that we all know intuitively what Cantor is talking about. Cantor's "aggregate" is what we call a set.

Notation 2.1.

1. Let S be a set. We then write $x \in S$ to denote that x is an element of S. That it is one of the "objects" in S. And we write $x \notin S$ to denote that x is not an element of S.

2. Let S and T be sets. We write $T \subseteq S$ to denote that whenever $x \in T$ then $x \in S$. That is, every element of T is an element of S. In this case T is said to be a subset of S.

At the same time, too liberal an understanding of what a "collection" means can lead to famous paradoxes.

Remark 2.2. (Russell's Paradox). Let

$$H = \{\text{sets } S : S \notin S\}.$$

That is, H is the collection of sets S which are *not elements of themselves*. This, at first glance, is an odd choice of set to consider but also currently seems a perfectly valid set for our consideration.

Most sets that we can think of seem to be in H. For example, $\mathbb{N}$ is in H, as the elements of $\mathbb{N}$ are single natural numbers, and no element is the set $\mathbb{N}$ itself. The problem arises when we ask the question: is $H \in H$?

There are two possibilities: either $H \in H$ or $H \notin H$. On the one hand, if $H \notin H$ then H meets the precise criterion for being in H and so $H \in H$ – a contradiction. On the other hand, if $H \in H$ then $H \notin H$ is false, and so H does not meet the criterion for being in H and hence $H \notin H$ – another contradiction. □

So we have a contradiction either way. A modern take on Russell's Paradox is that the set H is *inherently self-contradictory*. It would be akin to starting a proof with "let x be the smallest positive real number" or "let n be the largest natural number". There are no such numbers, so it is not surprising that contradictory or nonsensical proofs might result from such a beginning.

As previously mentioned, a modern "definition" of sets are given by the ZFC axioms. These are a list of mathematical statements outlining precisely how a set might be constructed – for example by taking unions or intersections of axiomatically assumed sets. Russell's set H is not constructible via the ZFC axioms and so simply would not be considered a set.

Question 2.3. Prove or disprove: there is a set of all sets. ■

Full details of these axioms are given in introductory set theory courses at university. For now, we will continue exploring set theory in this "naïve" way (that is, not exploring the ZFC axioms in detail) but be comforted in knowing that set theory recognises these "paradoxes" and has dealt with them.

Remark 2.4. It is unknown whether ZFC is "consistent", which is to say it is unknown whether there are paradoxes similar to Russell's that remain even

when we consider sets in this 'restricted', 'non-naïve' setting. However it is generally believed that ZFC is consistent – and if it is not, the problems are very minor and easily correctable. If ZFC were *inconsistent* in a very major, fundamental way, I think this would demonstrate a profound inconsistency in the thinking processes of humans, so in fact may never be discovered! It is also worth mentioning that it is in fact *impossible to prove the consistency of* ZFC. This is to say that, if it is true ZFC has no paradoxes, we cannot prove this fact. Kurt Gödel discovered this in the 1930's; the second of his two famous *Incompleteness Theorems*. We will discuss the first tomorrow. □

Let us now look at some examples.

Example 2.5. Let $A = \{1, 2, 3\}$. There are three elements of A namely 1, 2 and 3. There are 8 subsets of A namely

$$\{1, 2, 3\},\ \{1, 2\},\ \{1, 3\},\ \{2, 3\},\ \{1\},\ \{2\},\ \{3\},\ \emptyset,$$

where the last symbol $\emptyset$ denotes the empty set, the set with no elements. Note that the order in which a set's elements are listed is unimportant so that $\{1, 2, 3\} = \{1, 3, 2\}$ for example. □

Definition 2.6. Given a set A, its power set, denoted $\mathcal{P}(A)$, is the set of all subsets of A.

Example 2.7. With $A = \{1, 2, 3\}$ again, then

$$\mathcal{P}(A) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}.$$

Note that $1 \in A$, that $1 \notin \mathcal{P}(A)$, but that $\{1\} \in \mathcal{P}(A)$, the last being equivalent to writing $\{1\} \subseteq A$. That is, 1 is an element of A but the set $\{1\}$ is a subset of A. □

Also note the difference between 1 and $\{1\}$.

Question 2.8. How many elements in $\emptyset$? How many elements in $\{\emptyset\}$? ■

We have all already met certain important mathematical sets, though the following notation may well be new to you.

Definition 2.9.

1. We denote the set of natural numbers as $\mathbb{N}$. That is the set of non-negative whole numbers
$$\mathbb{N} = \{0, 1, 2, 3, \dots\}.$$
2. The set of integers, that is the set of whole numbers, is denoted $\mathbb{Z}$. The letter Z arises from the German word "zahlen" for "numbers". So
$$\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}.$$
3. The set of rational numbers (or just simply rationals) is denoted $\mathbb{Q}$. This is the set comprising all fractions where the numerator and denominator are both integers. So
$$\mathbb{Q} = \{\tfrac{m}{n} : m, n \in \mathbb{Z}\,,\, n \neq 0\}.$$
4. The set of real numbers, $\mathbb{R}$, are harder to define. You can think of the real numbers are being the "limit" of rational numbers: the real numbers are those numbers with a decimal expansion. This includes the rational numbers but also includes irrational numbers such as $\sqrt{2}$ and π.
5. The set of complex numbers $\mathbb{C}$ is the set of numbers
$$\mathbb{C} = \{a + bi : a, b \in \mathbb{R}\},$$
where $i = \sqrt{-1}$. They seem like a strange thing to consider ("imaginary" numbers? $\sqrt{-1}$?) however they arise very naturally in algebra.

   For example, over the real numbers, it is not true that every polynomial has a solution – e.g. $x^2 + 3 = 0$ has no solution in $\mathbb{R}$. However over $\mathbb{C}$, every polynomial has a solution. ($\mathbb{C}$ is *algebraically closed*.)

Note that
$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}\,.$$

2.1 Cardinality

Now, we'll talk about the sizes of sets. At first glance, it seems like this would be easy – sets are either finite (and have a specific number $n \in \mathbb{N}$ of elements) or they're infinite. It turns out, however, that one can say a lot more about an infinite set.

First, we'll need the idea of a *function*, also known as a *map*.

Definition 2.10. Let X and Y be sets. A function $f : X \to Y$ is an assignment of a value $f(x) \in Y$ for each $x \in X$. So for each $x \in X$, $f : x \mapsto f(x) \in Y$.

Functions must be *well-defined*, meaning for every $x \in X$, there is a *unique* $y \in Y$ such that $f(x) = y$ (i.e. $f : x \mapsto y$).

Example 2.11. Functions are incredibly general things. Here are some examples:

$$f : \mathbb{R} \to \mathbb{R} \quad \text{given by} \quad f(x) = x^2, \tag{2.1.1}$$
$$g : \mathbb{N} \to \mathbb{R} \quad \text{given by} \quad g(x) = x - 1, \tag{2.1.2}$$
$$h : \{1, 2, 3\} \to \{2, 4, 6\} \quad \text{given by} \quad h(x) = 2x, \tag{2.1.3}$$
$$j : \{1, 2, 3\} \to \{1, 2, 3\} \quad \text{given by} \quad j(1) = 3,\ j(2) = 2,\ j(3) = 1, \tag{2.1.4}$$
$$l : \mathcal{P}(\mathbb{N}) \setminus \{\emptyset\} \to \mathbb{N} \quad \text{given by} \quad l(A) = \min(A). \tag{2.1.5}$$

There are plenty of nonexamples too.

$$f : \mathbb{R} \to \mathbb{N} \quad \text{given by} \quad f(x) = \tfrac{x}{2}, \tag{2.1.6}$$
$$g : \mathbb{R} \to \mathbb{R} \quad \text{given by} \quad g(x) = \sqrt{x}, \tag{2.1.7}$$
$$h : \mathbb{R}_{\geq 0} \to \mathbb{R} \quad \text{given by} \quad h(x) = \sqrt{x}. \tag{2.1.8}$$

The first fails to be a function as f maps elements outside of $\mathbb{N}$ – e.g. $5 \in R$ but $f(5) = \frac{5}{2} \notin \mathbb{N}$.

The second fails to be a function as g is not defined everywhere – e.g. $-7 \in \mathbb{R}$ but $g(-7) = \sqrt{-7} \notin \mathbb{R}$.

The third fails to be a function because of its ambiguity. For, e.g., $3 \in \mathbb{R}$, there are technically two square roots – $\sqrt{3}$ and $-\sqrt{3}$. Both square to 3.

This last example is important: the problem is that an element $x \in \mathbb{R}_{\geq 0}$ can reasonably be sent two different places. We don't want functions to have this property.

A really nice property for a function to have is the following:

Definition 2.12. Let X, Y be sets. A function $f : X \to Y$ is a *bijection* (we also say "f is *bijective*") if for every $y \in Y$, there is a unique $x \in X$ such that $f(x) = y$ (i.e. $f : x \mapsto y$).

In this way, bijections describe an exact correspondence between the sets X and Y. One can look at this definition as demanding there exists an "inverse" to f.

Some examples:

$$f : \mathbb{Z} \to \mathbb{Z} \qquad \text{given by} \quad f(x) = x - 1, \tag{2.1.9}$$

$$g : \{1, 2, 3\} \to \{1, 2, 3\} \qquad \text{given by} \quad g(1) = 3,\ g(2) = 2,\ g(3) = 1, \tag{2.1.10}$$

$$h : \mathbb{Q} \to \mathbb{Q} \qquad \text{given by} \quad h(x) = 2x. \tag{2.1.11}$$

Some non-examples:

$$j : \mathbb{R} \to \mathbb{R} \qquad \text{given by} \quad j(x) = x^2, \tag{2.1.12}$$

$$l : \mathbb{N} \to \mathbb{R} \qquad \text{given by} \quad l(x) = x - 1. \tag{2.1.13}$$

Question 2.13. Prove that j and l are not bijections. ■

This idea of an "exact correspondance" between the sets X and Y is useful if we want to compare their sizes, or *cardinalities*. "Cardinality" is a fancy word for size – given a set X we wish to *rigorously* define $|X|$, the cardinality of X, to be the number of distinct elements in the set X. For finite sets this will not throw up any surprises – more surprising results will emerge when infinite sets are encountered later today.

Definition 2.14. Let $n \geq 1$ be a natural number and X be a set. We define the cardinality $|X|$ of X to be n if there exists a bijection from X to the set $\{1, 2, \ldots, n\}$ – the set with exactly n elements.

The cardinality of the empty set is defined to be 0.

Question 2.15. What is the cardinality of the set $\{1, 2, 3\}$? What about $\{10, 29, 33\}$?

Definition 2.16. A set X is said to be *finite* if its cardinality is some natural number $n \in \mathbb{N}$.

If there does not exist a bijection between between X and $\{1, 2, \ldots, n\}$, for all $n \geq 1$, then X is *infinite*. This is to say X has *infinite cardinality* if its cardinality is not a natural number.

Definition 2.17. Two sets X, Y are of the same cardinality, written $|X| = |Y|$, if there is a bijection between X and Y.

For example,

Theorem 2.18. $|\,\mathbb{N}\,| = |\,\mathbb{Z}\,|$; this is to say $\mathbb{N}$ and $\mathbb{Z}$ have the same "size".

(Note that both sets are infinite.)

Proof. I will describe a bijection between $\mathbb{Z}$ and $\mathbb{N}$.

Consider the function $f : \mathbb{Z} \to \mathbb{N}$ defined as follows:

- On non-negative elements $x \in \mathbb{Z}_{\geq 0}$, $f : x \mapsto 2x$. So $0 \mapsto 0$, $1 \mapsto 2$, $\ldots$, $17 \mapsto 34$, $\ldots$.
- On negative integers $x \in \mathbb{Z}_{<0}$, $f : x \mapsto -2x - 1$. Therefore $-1 \mapsto 1$, $-2 \mapsto 3$, $-3 \mapsto 5$, $\ldots$, $-50 \mapsto -2(-50) - 1 = 99$, $\ldots$.

Therefore for every $x \in \mathbb{Z}$, there is a unique $y \in \mathbb{N}$ such that $f : x \mapsto y$ (the y depends on whether $x \geq 0$ or $x < 0$). Also, for every $y \in \mathbb{N}$, there is a unique $x \in \mathbb{Z}$ such that $f : x \mapsto y$ (the x depends on whether y is even or odd). Therefore, by definition, $f : \mathbb{Z} \mapsto \mathbb{N}$ is a bijection and $|\mathbb{Z}| = |\mathbb{N}|$ as required. ■

What is this size? It's not 1, 2, 3, ... Let's give it a name.

Definition 2.19. Define $\aleph_0$ ("aleph zero" or "aleph naught") to be the cardinality of $\mathbb{N}$.

In some sense, it is an "infinite number".

Since this is a number, we might expect to be able to do some arithmetic with it.

Question 2.20. What is $\aleph_0 + 1$? What is $\aleph_0 - 1$? What is $\aleph_0 + \aleph_0$? What is $\aleph_0 \times \aleph_0$? ■

We can answer these questions by considering the famous "Hilbert Hotel".

So, what did we learn?

- $\aleph_0 + 1 =$
- $\aleph_0 + \aleph_0 =$
- $\aleph_0 \times \aleph_0 =$

It's not too hard to see that this is the "smallest infinite number" – anything *smaller* than $\mathbb{N}$ is finite. What about things "larger than $\mathbb{N}$"?

Maybe something we recognise is larger than $\mathbb{N}$. We know $|\mathbb{Z}| = |\mathbb{N}|$, but what about $\mathbb{Q}$?

Question 2.21. Is it the case that $|\mathbb{Q}| < |\mathbb{N}|$? Or $|\mathbb{Q}| > |\mathbb{N}|$? Or $|\mathbb{Q}| = |\mathbb{N}|$? ■

What about other sets?

Theorem 2.22. $|\mathcal{P}(\mathbb{N})| > |\mathbb{N}|$; this is to say there is no bijection between $\mathbb{N}$ and $\mathcal{P}(\mathbb{N})$.

Proof. This proof is originally due to Cantor. Suppose for the purpose of contradiction that there exists a bijection $f : \mathbb{N} \to \mathcal{P}(\mathbb{N})$. Consider the set $T = \{x \in \mathbb{N} : x \notin f(x)\}$. (This is indeed a set by the axioms of ZFC.) The set looks like this:

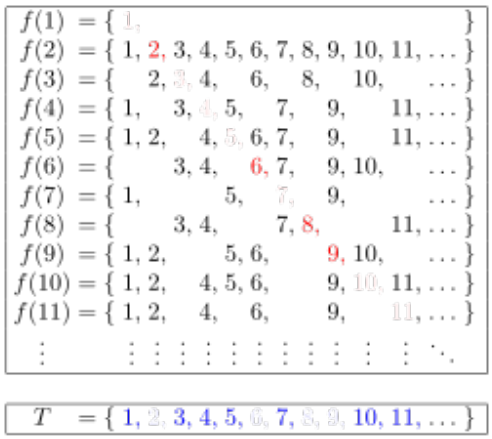


Figure 1: Image credit: Wikipedia.

By the definition of bijectivity, there exists $n \in \mathbb{N}$ such that $f(n) = T$. If $n \in T$, then $n \in f(n)$ hence $n \notin T$ by design. We reach a contradiction in this case. If $n \notin T$, then $n \notin f(n)$ and thus $n \in T$ by definition – another contradiction. We conclude such bijections cannot exist. ■

Therefore $\mathcal{P}(\mathbb{N})$ is infinite, but its size (or cardinality) is greater than $\mathbb{N}$. We call $\mathcal{P}(\mathbb{N})$ *uncountably infinite*.

The next big question is that of the real numbers. As it turns out:

Theorem 2.23. $\mathbb{R}$ is *uncountably* infinite.

Proof. To prove this we will show there is no bijection between $\mathbb{N}$ and $\mathbb{R}$. In fact, we will show there is no bijection between $\mathbb{N}$ and elements of the

form $0.x_1x_2x_3\ldots$ where $x_1, x_2, x_3, \cdots \in \{0, 1\}$. Let the set of such elements be S. This means $|\,\mathbb{N}\,| < |S|$ and hence $\mathbb{R}$ must be uncountably infinite.

To prove there is no bijection between $\mathbb{N}$ and S, we will use "Cantor's Diagonalisation Argument". It is a proof by contradiction as follows:

Assume there is a bijection $f : \mathbb{N} \to S$. This makes S 'countable', and in particular 'listable' – we can create a list of elements of S as follows: The first item, s_1, on the list is $f(0)$. The second item is $s_2 = f(1)$. The third item is $s_3 = f(2)$, ..., the nth item is $s_n = f(n-1)$, ... In this way we create a list $s_1, s_2, s_3, \ldots$ of *all* elements of S.

Let us define an element $s = 0.a_1a_2a_3\cdots \in S$ as follows. We may write $s_1 = 0.b_1b_2b_3\ldots$ where b_1 is 0 or 1. If $b_1 = 0$, set $a_1 = 1$. If $b_1 = 1$, then $a_1 = 0$. We may write $s_2 = 0.c_1c_2c_3\ldots$ where c_2 is either 0 or 1. If $c_2 = 0$, set $a_2 = 1$; if $c_2 = 1$, set $a_2 = 0$. We continue down the list in this fashion, creating an element $s \in S$ like so:

$$
\begin{array}{|ll|}
\hline
s_1 & = 0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\ldots \\
s_2 & = 1\,1\,1\,1\,1\,1\,1\,1\,1\,1\,1\ldots \\
s_3 & = 0\,1\,0\,1\,0\,1\,0\,1\,0\,1\,0\ldots \\
s_4 & = 1\,0\,1\,0\,1\,0\,1\,0\,1\,0\,1\ldots \\
s_5 & = 1\,1\,0\,1\,0\,1\,1\,0\,1\,0\,1\ldots \\
s_6 & = 0\,0\,1\,1\,0\,1\,1\,0\,1\,1\,0\ldots \\
s_7 & = 1\,0\,0\,0\,1\,0\,0\,0\,1\,0\,0\ldots \\
s_8 & = 0\,0\,1\,1\,0\,0\,1\,1\,0\,0\,1\ldots \\
s_9 & = 1\,1\,0\,0\,1\,1\,0\,0\,1\,1\,0\ldots \\
s_{10} & = 1\,1\,0\,1\,1\,1\,0\,0\,1\,0\,1\ldots \\
s_{11} & = 1\,1\,0\,1\,0\,1\,0\,0\,1\,0\,0\ldots \\
\vdots & \quad \vdots\;\vdots\;\vdots\;\vdots\;\vdots\;\vdots\;\vdots\;\vdots\;\vdots\;\vdots\;\ddots \\
\hline
\end{array}
$$

$$
\begin{array}{|ll|}
\hline
s & = 1\,0\,1\,1\,1\,0\,1\,0\,0\,1\,1\ldots \\
\hline
\end{array}
$$

Figure 2: Image credit: Wikipedia.

This process defines an element $s \in S$ that is *not* on our list. Indeed, if $s = s_m$ for some $m \in \mathbb{N}$, we see that in fact s and s_m differ at the mth decimal place, thus $s \neq s_m$. We conclude our list of elements is not complete

– the list doesn't contain every element of S. Therefore $|\mathbb{N}| < |S|$ and in particular, $\mathbb{R}$ is uncountably infinite, as required. ■

In fact, the two previous theorems are related – but in a way that is outside the scope of our course.

Theorem 2.24. $|\mathbb{R}| = |\mathcal{P}(\mathbb{N})|$. This size is written "$2^{\aleph_0}$". ■

This seems like a weird convention – writing $2^{\aleph_0}$. It should seem less weird after the following exercise:

Question 2.25. What is $|\mathcal{P}(\{1,2\})|$? What about $|\mathcal{P}(\{1,2,3\})|$ or $|\mathcal{P}(\{1,2,3,4\})|$?

What do you think $|\mathcal{P}(\{1,2,\ldots,n\})|$ is? ■

2.2 Final Comments

We mentioned previously that $\aleph_0$ is the smallest infinite number. Frequently, instead of "number", we say the word *cardinal* instead.

So 1 is a cardinal – it is the cardinality of a set with a single element. 2 is also a cardinal, as is 3, 4, ... These are the finite cardinals. The "next largest" cardinal is $\aleph_0$; this is the smallest infinite cardinal. However we have also seen that $\aleph_0 + 1 = \aleph_0$, and $\aleph_0 + \aleph_0 = \aleph_0$, and even $\aleph_0 \times \aleph_0 = \aleph_0$. So it's not entirely clear what the "next largest" cardinal after $\aleph_0$ is. Let's give it a name first.

Definition 2.26. Define $\aleph_1$ to be the smallest infinite cardinal larger than $\aleph_0$. (This is the "next largest" cardinal after $\aleph_0$.)

We know such cardinals exist, as the cardinal $2^{\aleph_0} > \aleph_0$ (so yes, there are bigger cardinals, and $\aleph_1$ is the smallest of these bigger cardinals!).

Unfortunately I don't have a simple example of a set of cardinality $\aleph_1$, and this is ultimately due to the following question:

Question 2.27. By definition, $\aleph_1 \leq 2^{\aleph_0}$ (as $2^{\aleph_0} > \aleph_0$, and $\aleph_1$ is the smallest cardinal bigger than $\aleph_0$). Is it the case that $\aleph_1 = 2^{\aleph_0}$? ■

This question – is $\aleph_1 = 2^{\aleph_0}$? – is known as the *Continuum Hypothesis* and was one of, if not the, largest questions in mathematics in the 20th century. It was first asked by Cantor in 1878 (who strongly believed it to be true), and it was first on David Hilbert's famous 23 open problems in mathematics presented to the International Congress of Mathematicians in Paris in 1900. It's proof (or disproof) was seen as one of the central questions of mathematics for nearly 100 years.

The resolution of this statement was suprising, and came in two parts. The ultimate answer is that this question – is $\aleph_1 = 2^{\aleph_0}$? – is *independent of* ZFC. *This is to say the Continuum Hypothesis cannot be disproven or proven.*

(Yes, there are statements in mathematics that we know cannot be proven or disproven – it is a provable fact that they cannot be proven or disproven. It is mindblowing and sounds crazy, but it's true.)

The first half of the proof ("cannot be disproven") is due to Kurt Gödel in 1940. The second half ("cannot be proven") is due to Paul Cohen in 1963/1964. Cohen received a Fields Medal[1] in 1966 for his work.

Question 2.28. Are the following sets finite, countably infinite, or uncountable (have cardinality larger than $\mathbb{N}$)?

1. The set of even natural numbers.
2. The set of prime numbers.
3. $\mathbb{C}$.
4. The set of all functions $f : \mathbb{R} \to \mathbb{R}$.
5. $\mathcal{P}(B)$, where

$$B = \{n \in \mathbb{N} \,:\, n \text{ has remainder 1 when divided by 12}\}.$$

Determine the cardinalities of the following sets.

[1]Mathematics has no Nobel prize – the highest prize in the subject is the Fields Medal. The 'Abel Prize' is a very close second.

1. The set of all functions $f : \mathbb{R} \to \mathbb{R}$ (I don't expect you to be able to do this rigorously).
2. The set of finite subsets of $\mathbb{N}$.
3. $(-\frac{\pi}{2}, \frac{\pi}{2}) = \{x \in \mathbb{R} : \ -\frac{\pi}{2} < x < \frac{\pi}{2}\}$.
4. $\mathbb{C}$.
5. $\{\{\emptyset\}\}$.

2.3 Euclidean Geometry

Instead of considering a complicated system of axioms (like ZFC) we will consider first the axioms of Euclidean geometry. The basics of Euclidean geometry are encapsulated by one of the most famous books ever written: Euclid's *Elements*. (See here for a complete, online and interactive version.)

The *Elements* is a mathematical treatise consisting of 13 books attributed to the ancient Greek mathematician Euclid in Alexandria, Egypt c. 300 BC. It is a collection of definitions, postulates, theorems, constructions, and mathematical proofs. The books cover plane (Books I–IV, VI) and solid (Books XI–XIII) Euclidean geometry and elementary number theory (Books V, VII–X). *Elements* is the oldest still existing large-scale deductive treatment of mathematics. It has proven instrumental in the development of logic and modern science, and its logical rigour was not surpassed until the 19th century.

The beauty of the *Elements* is in its simplicity; for instance, Euclid assumed five axioms about plane geometry (that is, geometry in two dimensions) and using logic and rigour deduced several books worth of theorems from these five assumptions. The five axioms are the following:

1. Between any two points we may draw a line;
2. Any line segment (part of a line) may be extended;
3. Given a point p and a length l, we may construct a circle with centre p and radius l;
4. All right angles are congruent (equal to each other);

5. (The parallel postulate) If a line segment intersects two straight lines forming two interior angles α, β on the same side that sum to less than two right angles, then the two lines, if extended indefinitely, meet on that side on which the angles sum to less than two right angles.

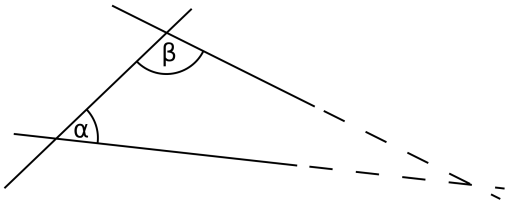


Figure 3: Image credit: Wikipedia.

This last axiom is a bit unwieldy, so commonly we use an equivalent formulation due to John Playfair:

5. (The parallel postulate) In a plane, given a line l and a point p not on it, exactly one line parallel to l can be drawn through p.

The first three axioms mean we are in a setting where all constructions must be done (and can be done) by a ruler and compass. The fourth allows us to compare (say) triangles that are of different scales, and develops the theory of congruence. The fifth is not only wordy, it is significantly less obvious than the first four. Because of this, for over two thousand years many attempts were made to prove the parallel postulate from the other four axioms. Eventually, however, it was proven by János Bolyai in 1831 that the parallel postulate is independent of the first four axioms: it cannot be proven or disproven from axioms 1–4. This is partly because there exist settings where the first four axioms are true, and the fifth is not; non-Euclidean geometry.

The standard example of non-Euclidean geometry is spherical geometry; on a sphere (as opposed to a plane) axioms 1–4 are true, however the parallel postulate is not.

Theorem 2.29. The parallel postulate fails in spherical geometry.

Proof. Consider the equator line l and the north pole p. Any line through p can be extended to intersect l. Therefore there is no line through p parallel to l. ■

This is an instance where there exists a line l and a point p and there are 0 parallel lines to l through p. What about when there are *more than 1* parallel lines through p? This is hyperbolic geometry.

In hyperbolic geometry, the "plane" we work in has a saddle shape, as opposed to a flat or spherical shape. This has many consequences, one of which is that parallel lines eventually diverge; that is, they get further and further apart. A consequence of this is that given a line l and p not on l, there are *infinitely many* lines through p parallel to l.

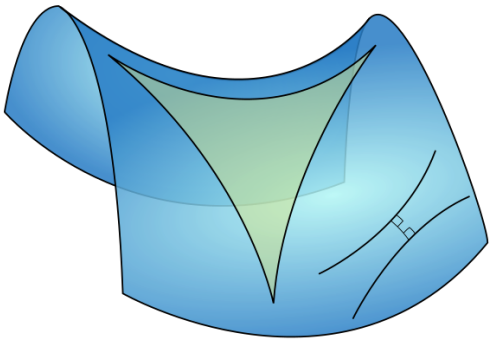


Figure 4: Image credit: Wikipedia.

Question 2.30. Prove the following results clearly stating your assumptions (so we may check if they follow from Euclid's 5 axioms).

1. The sum of the angles of a triangle is 180°.
2. (Thale's Theorem) If A, B, C are distinct points on a circle such that the line $\overline{AC}$ is a diameter, then the angle $\angle ABC$ is a right angle.
3. In any triangle the sum of any two sides is greater than the remaining one.

Finally: is it true that, without the parallel postulate, the sum of the angles of a triangle is 180°? Prove or give a counterexample.

3 Logic & Computability

We now need to introduce some logical notation and language to help formalise our notion of "proof".

Notation 3.1. Let P and Q denote logical statements – statements such as "$x \leq y$" or "for all $a \in \mathbb{R}$, $a^2 > 0$".

1. $P \implies Q$. This reads "P implies Q". This means that whenever the statement P is true then the statement Q is true. This implication may be *strict*, meaning that *it may be possible for Q to be true and P false.* For example, for all x, $x \geq 4 \implies x \geq 2$. However, when $x = 3$, $4 > 3 \geq 2$.

2. $P \iff Q$. This reads "P if and only if Q". This means P and Q are logically equivalent, i.e. P is true precisely when Q is true. Note that the *context* of the statement is an important part of its truth or falsity; in $\mathbb{N}$, $x > 2 \iff x^2 > 4$ is true, however this statement is false in $\mathbb{R}$ (where $x = -3$ is a counterexample).

3. $P \wedge Q$ represents "P and Q" and is true only when both P and Q are true.

4. $P \vee Q$ represents "P or Q" which holds when one of P or Q is true. (It also holds when both P and Q are true, i.e. $P \wedge Q \implies P \vee Q$.)

5. We write $\neg P$ for "not P" or the "negation of P". This is the statement that is true precisely when P is false.

We can detail exactly when these logical statements are true by the following *truth table*:

P	Q	$\neg P$	$P \wedge Q$	$P \vee Q$	$P \implies Q$	$P \iff Q$
T	T	F	T	T	T	T
T	F	F	F	T	F	F
F	T	T	F	T	T	F
F	F	T	F	F	T	T

We can use these tables to prove important identities.

Theorem 3.2. The statement $P \implies Q$ is logically equivalent to the statements $\neg P \vee Q$ and $\neg Q \implies \neg P$.

Proof. Examining the truth table:

P	Q	$\neg P$	$\neg Q$	$\neg P \vee Q$	$\neg Q \implies \neg P$	$P \implies Q$
T	T	F	F	T	T	T
T	F	F	T	F	F	F
F	T	T	F	T	T	T
F	F	T	T	T	T	T

we see the statements are all the same. ■

Definition 3.3. The statement $\neg Q \implies \neg P$ is known as the *contrapositive* of $P \implies Q$.

Now we move from "propositional" statements that have no context, to "predicate" statements that are true depending on their context.

Notation 3.4.

1. The symbol $\forall$ denotes "for all". So for example, $\forall x \in \mathbb{R}$, $x^2 \geq 0$ is true. This use of "for all" means that we have a family of statements, one for each $x \in \mathbb{R}$. It is good practice to make clear *what set is being varied over* – for example, $\forall x \in \mathbb{C}$, $x^2 \geq 0$ is false, with $x = i$ being a counter-example.

2. The symbol $\exists$ denotes "there exists". So for example $\exists x \in \mathbb{Q}$, $x^2 = 2$ is false, because no such rational x exists. But again, $\exists x \in \mathbb{R}$, $x^2 = 2$ is true – giving the example $x = \sqrt{2}$ is enough to prove the statement.

3. The symbols $\forall$ and $\exists$ are called *quantifiers* and are sometimes referred to as the *universal quantifier* and the *existential quantifier* respectively.

It is very important to note that the order of quantifiers is very important.

Example 3.5. Let S be the set of capital cities and T be the set of countries in the world. Let $P(x, y)$ be the statement "x is the capital of y". The statement

$$\forall y \in T \quad \exists x \in S \quad P(x, y)$$

is then true – it says every country has a capital city (and it doesn't really matter that some countries have arguably more than one capital). It is important here that the x is permitted to depend on the y *as the quantifier comes second*. So for $y =$ Ireland, there exists $x =$ Dublin, and for $y =$ France there exists $x =$ Paris.

However the statement

$$\exists x \in S \quad \forall y \in T \quad P(x, y)$$

is far from true. This time the existential quantifier comes first and this single capital city x *is required to be the capital of all countries* – there is no such city. □

Question 3.6. Translate the following mathematical statements into English: let $L(x, y)$ denote "x loves y" and let P be the set of people.

1. $\forall x \in P, \exists y \in P \quad L(x, y)$.
2. $\forall x \in P, \exists y \in P \quad (y \neq x \wedge L(x, y))$.
3. $\forall x \in P, \forall y \in P \quad (x = y \implies L(x, y))$.
4. $\exists x \in P, \forall y \in P \quad L(x, y)$.
5. $\exists x \in P, \exists y \in P \quad ((x \neq y \wedge L(x, y)) \vee (x = y \wedge \neg L(x, y)))$.

Translate the following English statements into logical ones:

1. All dogs hate all cats.
2. If cats hate dogs, then dogs love cats.
3. If a person loves a cat then that person hates all dogs.
4. Birds don't love cats or dogs, and love people exactly when people love them. □

Finally, I will introduce notation that precisely captures "truth":

Notation 3.7. Let $\models$ (spoken "turnstile" or "models") denote truth with context. This is to say the *context* appears on the left hand side of $\models$, while the statement true in this context appears on the right hand side.

For example,

$$\mathbb{R} \models \forall x (x^2 \geq 0).$$

This replaces the notation "$\forall x \in \mathbb{R}\,,\, x^2 \geq 0$".

Question 3.8. True or false?

1. $\mathbb{R} \models \forall x(x > 0 \vee x < 0)$.
2. $\mathbb{Z} \models \forall x(x^2 > 0 \vee x + x = x)$.
3. $\mathbb{N} \models \forall x, \exists y(x \times y = 1)$.
4. $\mathbb{Q} \models \forall x, \exists y(x \times y = 1)$.
5. $\mathbb{Q} \models \exists y, \forall x(x \times y = 1)$.
6. $\mathbb{C} \models \exists x(x^2 + x + 5 = 0)$. □

We have now seen exactly what it means for a statement to be true. However, results due to Gödel, Cohen and Bolyai (like the independence of certain statements from other axioms) tell us there is a conceptual difference between *absolute truth* and *provability*. Now we shall outline exactly what we mean by *provable*.

3.1 Proofs

The ideas here are quite complex, so I've had to omit many details. What follows is a rough sketch of the methodology; for a complete guide, see here (though "proof theory" is a subject of study in itself, and a detailed understanding is generally not needed elsewhere in mathematics).

Fundamentally, we will need to assume some axioms. These axioms, like Euclid's first four postulates, are designed to be self evident and unquestionably true. (The question of what we should consider to be an axiom, and more generally, what should be "unquestionably true", is a philosophical one. Mathematics starts with axioms and goes from there; philosophy provides the axioms.) Some examples of these axioms are:

- $\forall x, x = x$.
- For all statements α, β, γ:
  - $(\alpha \implies (\beta \implies \gamma)) \implies ((\alpha \implies \beta) \implies (\alpha \implies \gamma))$.
  - $(\alpha \implies \beta) \implies (\neg\beta \implies \neg\alpha)$.

Why some of these statements are considered axioms comes from their truth table. As we have seen previously, by truth tables,

$$(\alpha \implies \beta) \iff (\neg\beta \implies \neg\alpha),$$

so it seems reasonable to take "$(\alpha \implies \beta) \implies (\neg\beta \implies \neg\alpha)$" as an axiom.

Similarly, when we examine the truth table of

$$(\alpha \implies (\beta \implies \gamma)) \implies ((\alpha \implies \beta) \implies (\alpha \implies \gamma)),$$

we see that there is a "T" in every row – this is to say this statement is always true, i.e. a *tautology*.

Question 3.9. As an exercise, verify this: draw the truth table of $(\alpha \implies (\beta \implies \gamma)) \implies ((\alpha \implies \beta) \implies (\alpha \implies \gamma))$. □

So we will presuppose some axioms in our proof system. One other thing we must assume is a rule of logic known as *modus ponens* (Latin for "mode that by affirming affirms"). This is a *rule of inference* first explicitly described by the philosopher Theophrastus circa 300 BC. It is the following rule:

If P is assumed, and $P \implies Q$, then Q may be assumed.

Example 3.10. If today is Tuesday, then John will go to work. Today is Tuesday, therefore John will go to work.

If $x = 3$, then $x^2 = 9$. Assume $x = 3$. We conclude $x^2 = 9$. □

Definition 3.11. Let A be a set of assumptions (e.g. ZFC, or Euclid's postulates). A theorem T is *provable from* A (or *deducible from* A), written "$A \vdash T$" ("A deduces T"), if there exists a finite list $t_1, \ldots, t_n$ of statements such that

- $t_n = T$;
- Each t_i for $i = 1, 2, \ldots, n$ is either an axiom, an assumption from A, or follows from modus ponens:

  This is to say there exist statements "s", "$s \implies t_i$" such that $A \vdash s$, $A \vdash s \implies t_i$, and then $A \vdash t_i$ by modus ponens.

Writing proofs in this manner is typically long and quite complicated. Let us attempt to prove a relatively simple statement with full rigour.

Theorem 3.12. Let A be the set of Euclid's five postulates, and T the statement "given points A, B, there exists an equilateral triangle with side $\overline{AB}$".

Then $A \vdash T$.

Proof.

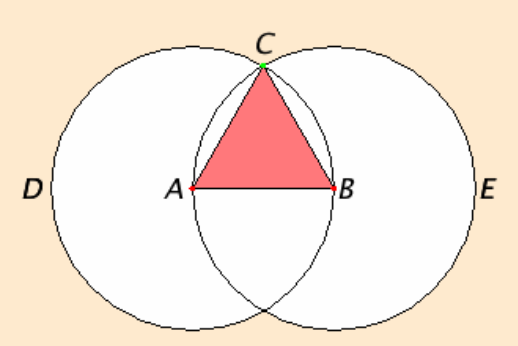


Figure 5: Image credit: David Joyce.

t_1: Postulate 1. Draw the line $\overline{AB}$.

t_2: Postulate 3. Draw the circle D with centre A and radius $\overline{AB}$.

t_3: Postulate 3. Draw the circle E with centre B and radius $\overline{AB}$.

t_4: Axiom. D and E intersect at 2 points, by construction. Label one of these points C.

t_5: Postulate 1. Draw the line $\overline{AC}$.

t_6: Postulate 1. Draw the line $\overline{BC}$.

t_7: Axiom. As $\overline{AB}$ and $\overline{AC}$ are radii of the circle D, $|\overline{AB}| = |\overline{AC}|$.

t_8: Axiom. As $\overline{AB}$ and $\overline{BC}$ are radii of the circle E, $|\overline{AB}| = |\overline{BC}|$.

t_9: Axiom. Therefore $|\overline{AB}| = |\overline{AC}|$ and $|\overline{AB}| = |\overline{BC}|$.

t_{10}: Axiom. For all x, y, z, if $x = y$ and $y = z$ then $x = z$.

t_{11}: Modus ponens (t_9, t_{10}). Therefore $|\overline{AC}| = |\overline{BC}|$.

T: We have concluded $|\overline{AB}| = |\overline{BC}| = |\overline{AC}|$ therefore triangle ABC is equilateral.

We have demonstrated the existence of an equilateral triangle with side $\overline{AB}$, as desired. ■

Example 3.13. Let A be the set of Euclid's five postulates. Then, for example, $A \vdash$ "the parallel postulate". This is a trivial example; other more complicated examples are

- $A \vdash$ "the sum of the angles of a triangle is 180°".
- $A \vdash$ "Thale's theorem".

These can be written as a "proof" in the sense of the above definition, however it would be long and complicated. □

As you can see, this is incredibly painful and intricate. Modern mathematics doesn't prove things this way – we give formal arguments but are not consistently referring back to basic axioms and rules of inference. However, defining "proof" in this rigorous and formulaic manner is necessary, and it allows computers to "prove" things (or at least verify long, complicated proofs).

Remark 3.14. There are philosophical objections to "non-surveyable proofs" (also known as "machine-verified proofs") – see here for further discussion.

Modern mathematics is also making increasing use of proof assistants and developing automated theorem proving. □

3.2 Final Remarks

So what is the connection between provability and truth? The connection is described by two theorems, the proof of either of which is outside the scope of this course.

Theorem 3.15. (Soundness). Let A be a collection of assumptions and T a theorem. If $A \vdash T$, then $A \models T$. ■

Loosely, "if something is provable, it is true" (in the context of A). To give an example: suppose N was the set of assumptions which define the natural numbers $\mathbb{N}$. The set N contains sentences such as "there is no element smaller than 0" and "if $x > 0$ then $x - 1$ is an element of $\mathbb{N}$".

The Soundness theorem tells us that if $N \vdash \exists x(x + x = 4)$ then there does exist a natural number n with the property that $n + n = 4$.

Maybe this is expected; provable things should be true! What about the other direction (the "converse")?

Theorem 3.16. (Completeness). Let A be a collection of assumptions and T a theorem. If $A \models T$, then $A \vdash T$. ■

To me, this says something very surprising. If a theorem is true, then *there exists a proof of it*. For example, since $\mathbb{N} \models (2 + 2 = 4)$, there is a formal proof (in the sense of *Theorem 3.12*) of that fact from N. The existence of such a proof is extremely non-obvious!

Where does that leave us with, say, our inability to prove or disprove the parallel postulate from Euclid's first four postulates? The above theorems tell us that, *in context*, provability and truth are one and the same. Therefore "our inability to prove or disprove the parallel postulate from Euclid's first four postulates" is equivalent to "there exists a context where Euclid's first four postulates are true, and the parallel postulate is *false*, and there exists a context where where Euclid's first four postulates are true, and the parallel postulate is *true*". We have already seen these contexts – in spherical geometry, postulates 1–4 are true but the parallel postulate is false. In plane geometry, all 5 postulates are true.

You can imagine, a similar idea of "different contexts" was used to prove the independence of the Continuum Hypothesis from ZFC.

3.3 Computability

So, we know when statements are true and when they are provable. However this is "true" and "provable" in the abstract sense – how can we *actually determine* if a given statement is true or false? To do this, we would need a finite set of instructions to follow that would, upon input of a statement, always output whether the statement is true or false (in a *finite amount of time*). What we need, is an *algorithm*.

The "concept" of an *algorithm* has been around since antiquity, however the "notion" of algorithm was not formalised until the early 1900's. Babylonian mathematicians circa 2500 BC were describing the use of algorithms, and this process continued among Egyptian, Greek, and Arabic mathematicians for thousands of years. In the 1930's it was, for the first time, explicitly set out the properties of "recursive" functions, and finally Alan Turing's definition in terms of *Turing machines* in 1936-1937 encapsulates the complete, modern notion of "algorithm" as we understand it today.

We'll begin with one of the most famous examples of an algorithm: Euclid's GCD algorithm.

Example 3.17. Euclid, in his *Elements*, sets out an explicit procedure for calculating the *greatest common divisor* (GCD) of two given natural numbers. The GCD of numbers n, m is the largest natural number g such that g divides both n and m without remainder. (For example, the GCD of $50, 72$ is 2, the GCD of $10, 20$ is 10, and the GCD of $17, 123$ is 1.)

The algorithm is as follows: take two natural numbers n, m where $n < m$. The GCD of n, m is actually the GCD of $m-n, n$. Since the numbers involved become smaller ($m - n, n$ is smaller than m, n), if we continue this process we eventually reach $0, a$. Then a is the GCD of n, m.

For example, let us consider the GCD of 252, 105. This is the GCD of (252-105 =) 147, 105. This is the GCD of (147-105=) 42, 105. This is the GCD of (105-42=) 63, 42. This is the GCD of (63-42=) 21, 42. This is the GCD of (42-21=)21, 21. This is the GCD of (21-21=) 0, 21. We conclude that 21 is the GCD of 252, 105.

Note that an algorithm is not a proof, and I have not proven *why* this algorithm works! □

We can see this is an algorithm, as it is given by a finite set of instructions and requires a finite amount of time to give an answer. Again, note that I have not made any comments on its efficiency, or how long it takes to run, or even how complicated the instructions are. For us, it suffices just that the algorithm *exists*.

Question 3.18. What other algorithms do you know? □

A *Turing machine* is a kind of computer which applies simple operations working with a limitless memory – the memory is an infinite ('paper') tape

divided into squares. Each square contains one symbol from a fixed finite set, the *tape alphabet* G. At any stage of a computation, the "machine" can read the symbol written on the tape square ("access its memory"), chose to replace the symbol with a different symbol, change "state", then choose to move left or right along the tape.

This is an illustration of a Turing machine – see the machine's "programming" displayed behind it.

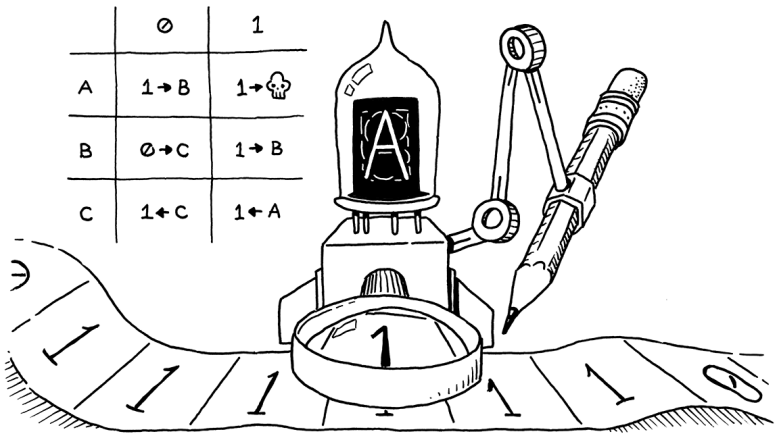


Figure 6: Image credit: Bob Nystrom.

Before we formally define Turing machines, let's take a look at this video.

Definition 3.19. A *Turing machine* is given by the following data:

- A *tape* divided into consecutive cells, each containing either a blank space or a symbol from a finite collection G of symbols. (The tape, in total, should have only finite many non-blank cells.)
- A *head* that can read/write symbols from G and move one space left or right.
- A finite set S of *states* of the Turing machine. Among these states is the unique *starting state* where the Turing machine is initialised, and a finite set of *halting states* where computation is ordered to stop.
- A finite *table of instructions*, of the following form:

  "Given the machine is in state s, and reads symbol g on the tape, do the following: replace g with symbol g', change state s to s', and move the head left/right."

The following is an example of a simple calculation that can be achieved by a Turing machine:

Example 3.20. Let $G = \{1, 0, B\}$ (where "B" represents a blank space), $S = \{s_1, s_2, s_3\}$ be three states with initial state s_1 and halting state s_3. Finally, set the following instructions:

$$(s_1, 1, B, R, s_2), \quad (s_2, 1, B, R, s_1), \quad (s_1, B, 0, R, s_3), \quad (s_2, B, 1, R, s_3).$$

What does this Turing machine do?[2] □

Examples become complicated quickly. There are multiple resources online (see here) that allow us to visualise more complicated Turing machines.

We see we can quickly put together complicated machines that can, say, multiply any given numbers[3]. This leads us to the following definition:

Definition 3.21. An *algorithm* is the table of instructions of some Turing machine. A "problem" can be *solved algorithmically* if it can be solved by a Turing machine.

Any problem solvable by a Turing machine is *decidable*.

Example 3.22. There are plenty examples of decidable problems:

1. Determining the GCD of two natural numbers.
2. Determine if a given natural number n is prime.
3. Determining the solutions to any quadratic equation $ax^2 + bx + c = 0$ over $\mathbb{R}$.
4. Determining whether two knots are equivalent.

   (I'm being rather vague on this one, as to formally state the problem we require more advanced mathematics. See here for the statement of the problem.) □

Of course, there are *undecidable* problems.

[2]This Turing machine takes as input a list of 1's, and returns 0 if the list has even length, and 1 if the list has odd length.

[3]In the previous link, one can first use a Turing machine to convert any two given numbers into binary, then use the binary multiplication Turing machine.

Definition 3.23. A problem is *undecidable* if there does not exist a Turing machine to solve it.

For these problems one is required to mathematically prove no sufficient Turing machines exist – it is impossible "for a computer" to solve this problem, as there is something intrinsically difficult about the problem making it impossible to solve in a finite amount of time with finite instructions. (This is where my interests lie, so I am ultimately biased.)

Example 3.24. There are also plenty of examples of undecidable problems.

1. Given the description of an arbitrary computer program and a finite input, decide whether the program finishes running or will run forever.

   This is one of the most famous examples of an undecidable problem, and is known as the "Halting Problem". It was proven to be undecidable by Alan Turing in 1936.

2. Determining whether a player has a winning strategy in a game of Magic: The Gathering.

   This is quite a new result (see here for the proof).

3. In Conway's *Game of Life*, determine whether given an initial pattern and another pattern, can the latter pattern ever appear from the initial one.

4. Given an equation in any number of variables with coefficients in $\mathbb{Z}$ (e.g. "$x^2 + 2yz - 17w^3 = 44$"), determine whether or not the equation has solutions in $\mathbb{Z}$.

   Note that this problem isn't asking what the solutions are, just to determine whether or not solutions exist. (This is known as "Hilbert's Tenth Problem", as it was number ten on his list of 23 problems presented to the International Congress of Mathematicians in 1900. The proof of this undecidability was completed by Yuri Matiyasevich in 1970 building on earlier work by Julia Robinson, Hilary Putnam and Martin Davis.)

Remark 3.25. Offhand remark: by the definition of a Turing machine, there exists only *countably many different algorithms*. However, there are uncountably many "problems" (e.g. there are uncountably many subsets of $\mathbb{N}$). Therefore most problems are undecidable! □

Bjorn Poonen has a very interesting paper on (mainly) undecidable problems – but there are some decidable ones thrown in there, along which there are open problems; problems which no one knows yet whether they are decidable or not.

There is also an overlap between "decidability" and game theory; there exist two player games for which it is undecidable whether player 1 or player 2 has a winning strategy. For example, the following problem is in fact open:

Given finitely many chess pieces on an arbitrarily large edgeless board, can White force checkmate?

Question 3.26. Research and present the following:

1. A Turing machine doing an interesting thing.
2. An interesting decidable problem.
3. An interesting undecidable problem.

3.4 Final Remarks

Turing machines may seem complicated, however their definition is extremely formulaic. In fact, how a Turing machine operates boils down to its finite table of instructions – remember, these were finitely many sequences/'tuples' of the form "$(s_1, 1, B, R, s_2)$". Let us suppose $G = \{1, 0, B\}$, i.e. that the Turing machine just deals with 1's and 0's[4]. You could maybe convince yourself that it is possible to represent "$(s_1, 1, B, R, s_2)$" as a sequence of 1's and 0's; maybe I decide "R" is represented by 00000 and "L" by 11111. Maybe "B" is 10001, s_1 is 11001, and s_2 is 11011. Then:

$$\begin{aligned}(s_1, 1, B, R, s_2) &= (11001, 1, 10001, 00000, 11011)\\ &= 11001\ 010\ 1\ 010\ 10001\ 010\ 00000\ 010\ 11011.\end{aligned}$$

It therefore seems reasonable one could represent the finite table of instructions, and hence the *entire Turing machine*, as one big long list of 1's and 0's.

[4]One can convince oneself that in fact any collection of symbols can be represented by long sequences of 1's and 0's.

Definition 3.27. An *encoding* of a Turing machine is a sequence of 1's and 0's arising in the manner described above.

Since I have (more or less) given an algorithm for constructing the encoding of a Turing machine, notice that this algorithm can be reversed: given a Turing machine encoding, from it I could pull the details of the Turing machine's instructions (assuming I know how symbols are coded!). E.g.

$$\begin{aligned} 11001\ 010\ 1\ 010\ 10001\ 010\ 00000\ 010\ 11011 &= (11001, 1, 10001, 00000, 11011) \\ &= (s_1, 1, B, R, s_2) \end{aligned}$$

This is to say that there is an *algorithm* which, on input a Turing machine encoding, can effectively produce the Turing machine being encoded. Therefore there is a *Turing machine* which takes as input a Turing machine encoding, and "outputs" the encoded Turing machine, ready for use. This brings us to the following definition:

Definition 3.28. A *Universal Turing Machine* is a Turing machine that can simulate any arbitrary Turing machine on any arbitrary input.

How this Turing machine would work is based on the idea of encoding; to simulate a Turing machine M on input i, I simply need to encode M into a string of 1's and 0's, then add i to the end of the string. This is my input to the Universal Turing machine, which 'decodes' my input into M and i, then runs M with input i, giving the final output.

Remark 3.29. This definition (given by Alan Turing in 1936/1937) encapsulates the idea of a modern, programmable computer, and took nearly ten years to physically implement.

Over the subsequent years, mathematicians have set about constructing Universal Turing machines – in particular, there is an interest around finding the *smallest/simplest* Universal Turing Machine. The smallest known Universal Turing Machine was discovered by Yurii Rogozhin in 1996, and has 4 states, 6 symbols, and 22 instruction tuples. See here for his paper. □

Along these lines is the idea of *Turing completeness*.

Definition 3.30. A "computational system" is *Turing complete* if it can simulate a Universal Turing machine; this is to say it can simulate any arbitrary Turing machine.

For example, the standard programming languages such as C, Python, Java, etc. are Turing complete.

Of course, all modern programmable computers are Turing complete too.

There is a fantastic subsection of the Turing completeness Wikipedia page called "Unintentional Turing completeness" – some computer programs, games etc. are so complex they can unintentionally simulate any Turing machine. Some examples are:

- Minecraft (see here and here for example videos).
- Minesweeper (see here for a paper).
- Magic: The Gathering (see here for the paper and here for an article).
- Conway's Game of Life (see here).
- MS Powerpoint (see here – though the claim needs further verification!).

These ideas of encoding Turing machines link further to the notion of undecidability. One key result proven by Alan Turing in 1936 is the undecidability of the *Halting Problem*: Given the description of an arbitrary computer program and a finite input, decide whether the program finishes running or will run forever.

Remark 3.31. We see immediately, for example, that *determining the winner in a game of Magic: The Gathering is undecidable*, as we can encode a Universal Turing Machine in some game of *Magic: The Gathering* – hence to determine a winner algorithmically, this requires the Halting Problem to be decidable; a contradiction. ■

Theorem 3.32. The Halting Problem is undecidable.

Proof. This is an argument similar to Cantor's Diagonalisation argument, or even to Russell's paradox. (It's very mind-bendy and intricate, so be warned.)

Suppose for the purpose of contradiction there exists a Turing machine $H(T, i)$ which takes as input a Turing machine T and input i. H returns 1 if T halts on i, and returns 0 if T doesn't halt on i (i.e. loops forever).

Recall we may encode a Turing machine into a string of 1's and 0's. Define a new Turing machine M according to the following procedure:

Given input i:

- If i does not encode a Turing machine, then return 1.
- Otherwise, run $H(i, i)$ (which computes whether Turing machine i running on input i halts or not). If 1 is returned by H, M loops forever. If 0 is returned by H, M returns 0.

Now encode Turing machine M into a string of 1's and 0's; call this string s. What happens when M is given input s?

1. Suppose M halts on s. Then necessarily M returns 0, which means $H(s, s)$ returns 0, which means M does not halt on s – a contradiction.
2. Suppose M loops forever and does not halt on s. Necessarily $H(s, s)$ must have returned 1, which is to say M on input s halts. This is a contradiction.

In either case, we reach a contradiction, so H must not exist. Therefore the Halting Problem is undecidable, as required. ■

Question 3.33. Pick an "unintentionally Turing complete" computational system and read up on how a Universal/arbitrary Turing machine can be encoded. □

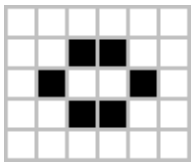
4 Casual Topics

4.1 Conway's Game of Life

We've mentioned Conway's *Game of Life* several times now, so let us explore it in more detail. (The source for this section is mainly the Wikipedia page.)

Definition 4.1. A *cellular automaton* is a two dimensional infinite grid of *cells*, each of which is in some *state* (of which there are finitely many), which collectively changes state at regular time intervals according to a fixed, finite set of rules.

The *Game of Life* is a cellular automaton devised by the British mathematician John Conway in 1970. It is a zero-player game, meaning that its evolution is determined by its initial state, requiring no further input.



One interacts with the *Game of Life* by creating an initial configuration and observing how it evolves, according to the following rules:

- There are two states; *alive* or *dead* (sometimes called *populated* and *unpopulated*, respectively).
- Every cell interacts with its eight *neighbours*, which are the cells that are horizontally, vertically, or diagonally adjacent.
- At each step in time, the following transitions occur:
  - Any live cell with fewer than two live neighbours dies, as if by underpopulation.
  - Any live cell with two or three live neighbours lives on to the next generation.
  - Any live cell with more than three live neighbours dies, as if by overpopulation.
  - Any dead cell with exactly three live neighbours becomes a live cell, as if by reproduction.

Let's watch this Youtube video with the man himself, explaining the rules and talking about the game in general.

At the end of the video, Conway mentions something we've come across before: in the *Game of Life*, it is undecidable to determine whether, given an initial pattern and another pattern, if the latter pattern can ever appear from the initial one.

- Let's have a look at some examples of patterns.
- See here for the Gosper glider gun.
- Here are some basic example patterns.
- Here is a pretty exhaustive lexicon of configurations.

One interesting question answered only recently is *self-replication* – does there exist a finite pattern which creates copies of itself? To answer this, we need some terminology.

Definition 4.2. A *spaceship* is a finite pattern that reappears (without additions or losses) after a number of generations and displaced by a non-zero amount.

The simplest example of a spaceship is a *glider*.

In 2010 Andrew Wade announced a *self-constructing* pattern, dubbed "Gemini", that creates a copy of itself while destroying its parent. This pattern replicates in 34 million generations, and uses an instruction tape made of gliders oscillating between two stable configurations made of "Chapman–Greene construction arms". These, in turn, create new copies of the pattern, and destroy the previous copy. Gemini is also a spaceship, and is the first spaceship constructed in the *Game of Life* that is an *oblique* spaceship, which is a spaceship that moves neither purely orthogonally nor purely diagonally. In 2015, diagonal-moving versions of Gemini were built.

In 2013, Dave Greene built the first *replicator* in the *Game of Life* that creates a complete copy of itself, including the instruction tape (some more details can be found here, including a talk by Dave Green here).

Question 4.3. Check out this Youtube video displaying some interesting patterns.

There is also the concept of "metacells"; *Game of Life* configurations which emulate the *Game of Life* on a larger scale. See this Youtube video for an example of "meta" glider guns on OTCA metapixels.

Finally, using the lexicon here, find me something interesting! □

4.2 The Three Jugs Problem

We will now move in a different direction by considering the *Three Jugs Problem* (though this is a logic and geometry puzzle, ultimately, so it is connected to our previous discussions). The source for this discussion is Cut the Knot!.

According to one story, Siméon Poisson, one of the greatest mathematicians of the 19th century, owed his interest in mathematics to a chance encounter with the following simple problem:

> Two friends who have an eight litre jug of water wish to share it evenly. They also have two empty jars, one holding five litres, the other three. How can they each measure exactly 4 litres of water?

In the pop-maths book "Mathematical Recreations and Essays" by Ball & Coxeter, the problem appears with the remark that "the solution presents no difficulty". Is this the case?

Question 4.4. Solve the Three Jugs Problem. □

This problem precedes another with four jugs of capacities 5, 11, 13, and 24 litres for which a solution "can be worked out only by trial". The problem is presented by the slightly more exciting narrative:

> Three men robbed a gentleman of a vase, containing 24 kilos of spice. Whilst running away they met a glass seller, from whom they purchased three vessels. Upon reaching a place of safety they wished to divide the booty, but found that their vessels could only hold 5, 11, and 13 kilos respectively. How could they divide the spice into equal portions?

Question 4.5. Solve the Four Jugs Problem. □

If nothing else, such problems wrap up a meaningful counting exercise that can be handed out to children in early grades. But there is also some worth-

while mathematics involved that was mostly overlooked by teachers and students alike.

Let's return to the original problem and tackle it again, this time in more abstract terms. First, I need to introduce *modular arithmetic*.

Definition 4.6. *Modular arithmetic* is a system of arithmetic for integers, where numbers "wrap around" to 0 after reaching a certain value.

We see an example of modular arithmetic everyday in (analogue) clocks – these are *modulo 12*. This is to say once 12 o'clock is reached, the numbers "begin again" proceeding to 1 o'clock, 2 o'clock, etc. The *arithmetic* (adding, subtracting, multiplying, etc.) of these numbers is again something familiar; 4 hours after 11 o'clock is 3 o'clock – that is to say $11 + 4 = 3$ *modulo 12*.

So how do we compute modular arithmetic? In the case of clocks, as we saw above, we take the *remainder* after dividing by 12. That is:

$$11 + 4 = 15 \implies 15 \div 12 = 1, \text{ remainder } 3 \implies 11 + 4 = 3 \text{ modulo } 12.$$

What about $3 + 162$ modulo 12?

$$3+162 = 165 \implies 165 \div 12 = 13, \text{ remainder } 9 \implies 3+162 = 9 \text{ modulo } 12.$$

What about $2 + 12$ modulo 12?

$$2 + 12 = 14 \implies 14 \div 12 = 1, \text{ remainder } 2 \implies 2 + 12 = 2 \text{ modulo } 12.$$

So notice in particular that $12 = 0$ modulo 12.

Question 4.7. Now you have the basics modulo 12, consider modular arithmetic with other values. Compute:

1. 5 modulo 3;
2. $15 + 7$ modulo 9;
3. 23×65 modulo 4. □

Now let's return to the original jug problem.

Label the jugs A, B, C in the increasing order of their capacities. Let's agree to use the same letters for the capacities themselves. Let x, y, z denote the

quantities of water in the jugs. In particular, $x + y + z = C$. A typical state – distribution of water – of the puzzle is described by a triple (x, y, z). For the original problem, the initial state is thus $(0, 0, C)$ with $C = 8$; the final state should be $(0, 4, 4)$.

Theorem 4.8. Let $C = A + B$, where A and B are mutually prime (i.e. their GCD is 1). Then any quantity Q with $0 \leq Q \leq C$ can be measured with the three jugs A, B and C.

Proof. Start with $(0, 0, C)$, and pour from C to A and then from A to B to obtain $(0, A, C - A)$. This is the first basic step that must be repeated until B becomes full: eventually, the state reached is $(r, B, C - Aq)$, where $q, r \in \mathbb{N}$ and $q, r > 0$. (q is the number of times the above step is performed, and r is the "remainder"; notice $B = Aq - r$.)

At this point, pour from B to C and from A to B: we obtain the state $(0, r, C - Aq + B)$, which by algebra is just $(0, r, 2C - A(q + 1))$. This the second step. Follow with the first step until B becomes full, after which apply the secondary step, and so on.

Modulo C, the third vessel will successively contain the quantities 0, $-A$, $-2A$, $-3A$, ... as q increases. Since A and B have been assumed to be mutually prime, so are A and C (i.e. the GCD of A and $A + B$ is the same as the GCD of A and B).

Because of this, all the quantities 0, $-A$, $-2A$, $-3A$, ..., $-A(C - 1)$ are *different* modulo C. Think about it:

$$\begin{aligned} \text{if } -2A = -5A \text{ modulo } C &\implies -2A + 5A = 0 \text{ modulo } C \\ &\implies 3A = 0 \text{ modulo } C \\ &\implies C \text{ divides } A, \end{aligned}$$

a contradiction.

So 0, $-A$, $-2A$, $-3A$, ..., $-A(C - 1)$ are all different modulo C. Therefore all quantities in the third jug are always different, according to this algorithm. There are C different quantities, meaning the algorithm gives us the quantities 1 litre, 2 litres, ..., C litres – though not in order! ■

The problem and the proof have a surprising geometric interpretation in terms of "triangular coordinates". We are most familiar with "Cartesian" coordinates – coordinates on the two dimensional plane – however there are several systems of coordinates in which vertices and sides of a triangle are

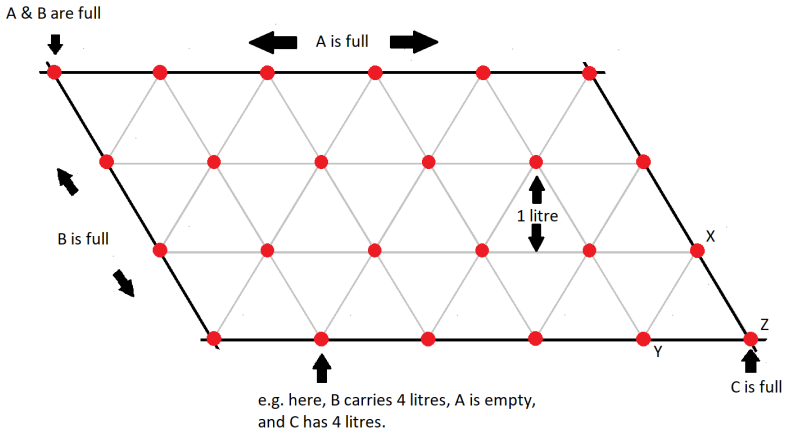
treated in an equal manner. The most important are the *trilinear* coordinates.

Definition 4.9. For a point P in the plane of $\triangle XYZ$, the triple of its (*signed*) distances to the sides $\overline{YZ}$, $\overline{XZ}$, and $\overline{XY}$ is called *trilinear coordinates of P (with respect to $\triangle XYZ$.)*

(The distances are *signed* such that, for example, the distance to $\overline{XY}$ is positive or negative depending on whether P is located on the same or different side of $\overline{XY}$ as vertex Z.)

The description of the Three Jugs Problem as a triple of quantities (x, y, z) fits nicely with trilinear coordinates. Draw an equilateral triangle XYZ and let the vertices have trilinear coordinates $(1, 0, 0)$, $(0, C, 0)$, and $(0, 0, C)$ – this is to say the distance between $\overline{XY}$ and Z is 1, the distance between $\overline{YZ}$ and X is 1, and the distance between $\overline{XZ}$ and Y is 1. The sides $\overline{XY}$, $\overline{YZ}$, and $\overline{XZ}$ are defined by $z = 0$, $x = 0$, and $y = 0$, respectively.

Consider the triangular grid formed of lines parallel to $x = 0$, $y = 0$, $z = 0$. The vertices of the triangles inside a parallelogram correspond to the integers Q with $0 \leq Q \leq C$ that solve the Three Jugs Problem.



Only the points on the boundary of that parallelogram could be attained as a result of valid puzzle moves. The "first basic step" corresponds to an inverted "V" path with one side parallel to $\overline{XZ}$ (pour from C to A), the other to $\overline{XY}$ (pour from A to B). See the left image below:

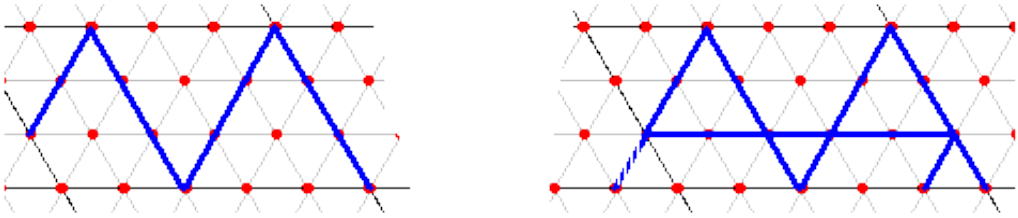


Figure 7: Image credit: Cut the Knot!

The jug B is full at the points of the "western" side of the parallelogram. Close to that side, the left leg of the inverted "V" may not reach the bottom line of the parallelogram. In which case, a secondary move must be made: first parallel to the line $\overline{YZ}$ to the "eastern" side of the parallelogram (pour B to C), and then to the bottom side (pour from A to B). See the right image above.

We are left with A being empty, B containing 1 litre, and C containing the remaining seven litres. However, since we are only interested in the *modular arithmetic*, we may overlook the need for secondary moves on the western side of the parallelogram and keep applying only the basic first step.

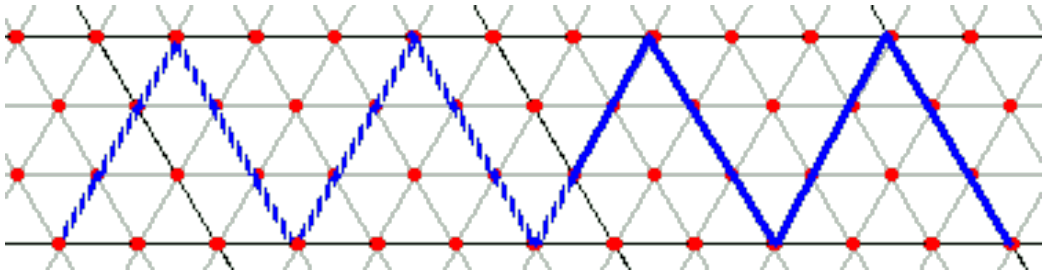


Figure 8: Image credit: Cut the Knot!

Finally, as we can see, since A and B are mutually prime (recall this means the GCD of A and B is 1) all paths parallel to $\overline{XY}$ will be taken, meaning every red point in the parallelogram will be covered, hence every integer value Q between 0 and C can be constructed.

The condition "$A + B = C$" serves a double purpose. First, together with the relative primality of A and B, it insures that a*ll three capacities share no common factor*, save 1. If this were not the case, the quantities that could be measured with three vessels of the specified capacities would *share their*

common factor. E.g., consider vessels of capacities 2 litres, 4 litres, and 6 litres – can you ever isolate 1 litre? Or 3?

For the problem to be solvable in general the mutual primality of all three capacities is a necessary condition. However, this isn't the only condition we require; anomalies also arise when the three jugs are rather big and $A + B \neq C$.

Question 4.10. Consider jugs of capacity 6, 7, and 8 litres. Can one isolate 5 litres? If not, why? □

This completes the analysis of existence of a solution to the Three Jugs problem. That of the Two and Four Jugs problems is left as an enticement for the future Poissons ...

Question 4.11. Here is the Two Jugs problem:

A farm hand was sent to a nearby pond to fetch 8 litres of water. He was given two jugs – one 11, the other 6 litres. How can he measure the requested amount of water?

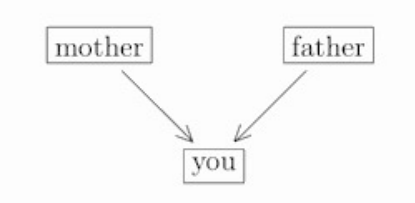
See if you can solve it using the "geometric" analysis above as well. □

4.3 Family Trees

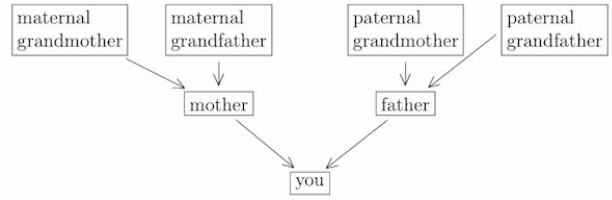
Tomorrow we'll be looking at graph theory, so today, we'll get a taster. (All images in this section, as well as the source material, belong to +plus magazine.)

Keeping track of family relations can be difficult. If Edna marries your mother's uncle Charlie, what should you call her? If your father's cousin's daughter just had a baby boy, how should you two be introduced? Who is your "great great aunt", and how can you find your "first cousin twice removed"? Fortunately, a bit of graph theory can clarify who should be called what, and why – and even measure the degree of genetic similarity between different relatives.

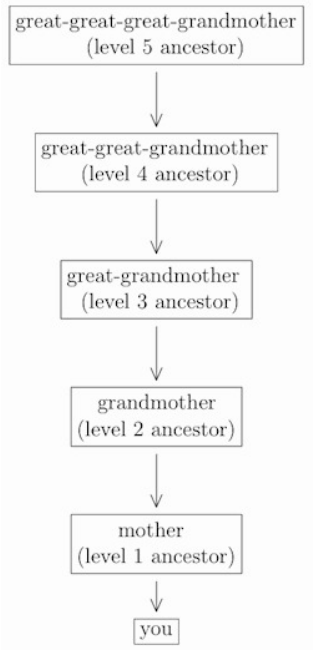
To begin at the beginning (well, your beginning, anyway), you had two parents, a mother and father:



Continuing backwards, they each had two parents, giving you a total of four grandparents:



Going back still further, each of your ancestors in turn had two parents, indicated by prefixing an extra "great" each time. For example, your maternal lineage is:



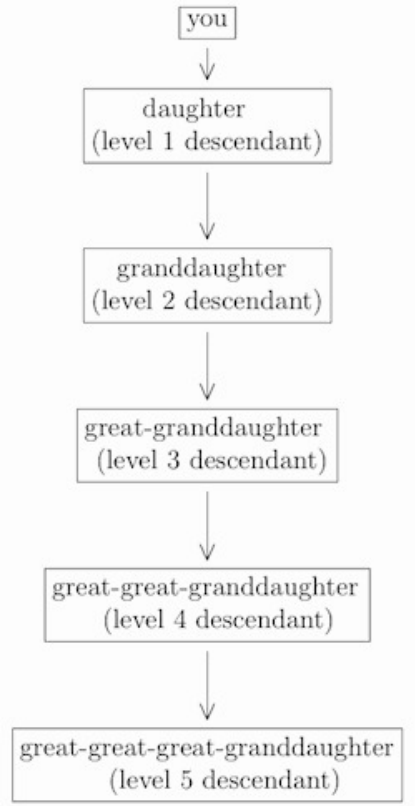
. . . and so on (and similarly for "fathers" instead of "mothers" at any level).

Since each ancestor has two parents (one mother and one father), you have a total of $2n$ ancestors at level n: two parents, four grandparents, eight great-grandparents, sixteen great-great-grandparents, and so on. Summing up, you have a total of $2 + 2^2 + 2^3 + \cdots + 2^n = 2^{n+1} - 2$ ancestors of level n or lower.

Question 4.12. Prove by induction: $2 + 2^2 + 2^3 + \cdots + 2^n = 2^{n+1} - 2$. □

For example, your total number of parents and grandparents and great-grandparents combined is $2^{3+1} - 2 = 16 - 2 = 14$. In short, your ancestors form a *perfect binary tree*.

What about descendants? If you have children yourself, then their children are your grandchildren, and your grand-children's children are your great-grandchildren, and so on:

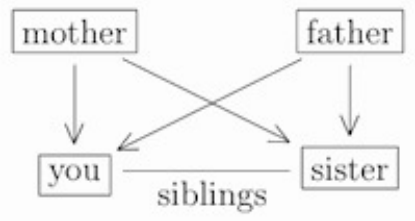


(and similarly for "son" instead of "daughter" at any level).

Unlike with ancestors, clearly there is no simple formula for your number of descendants. Rather, you have to count up all of your children, and all of their children, and so on. For example, even if you have five children, it is

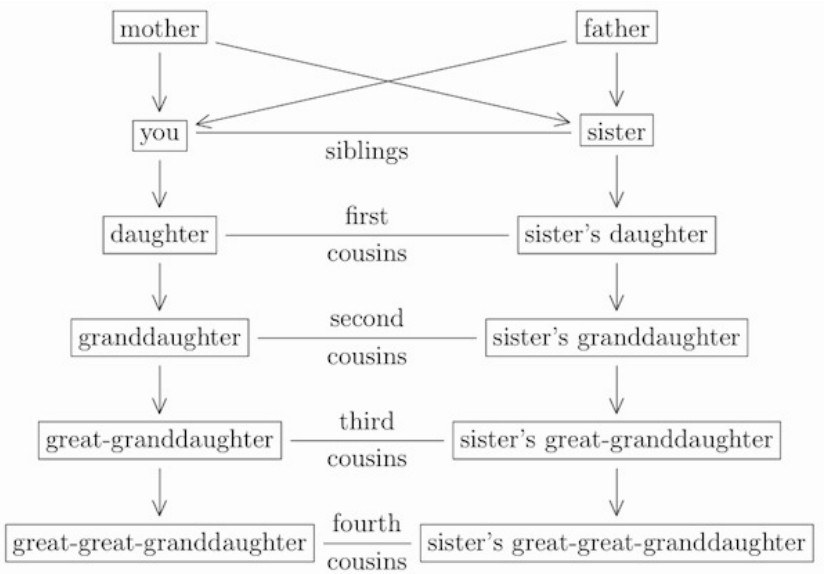
possible that none of them will have children of their own, in which case your number of grandchildren will be zero. On the other hand, if they each have five children of their own, then you will have twenty-five grandchildren – a lot more.

When people have more than one child, this fattens the family tree, creating new relationships like sister and niece and great-aunt and more. For starters, if your parents have additional children besides you, then they are of course your siblings, that is your sisters and brothers:



(Here, and throughout, relationships to "you" are written within the boxes, and relationships between other pairs of individuals are indicated by connecting lines.)

If you and your siblings each have children, then those children are first-cousins of each other. Then, if the two first-cousins each have children, then those children are second-cousins of each other; and their children are third-cousins, and so on:

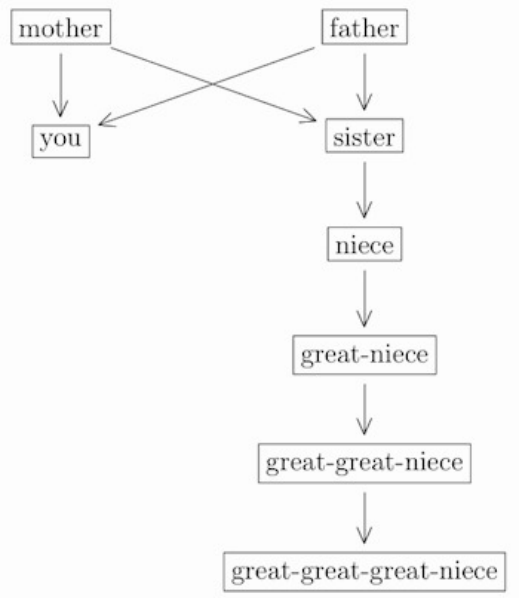


(and similarly for "son" instead of "daughter" at any level).

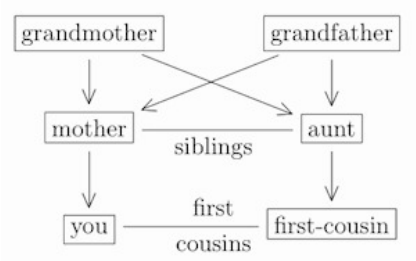
In general, n-level cousins share two $(n+1)$-level ancestors (but no n-level ancestors). Thus, first-cousins share two grandparents (but no parents), and second-cousins share two great-grandparents (but no grandparents), and so on.

It follows that if A and B are n-level cousins, then A's child and B's child are $(n+1)$-level cousins. Thus, children of first-cousins are second-cousins, and children of second-cousins are third-cousins, and so on. In fact, if we regard siblings as "0-level cousins", then this reasoning applies to siblings too: children of 0-level cousins (i.e. siblings) are themselves first-cousins (i.e. 1-level cousins).

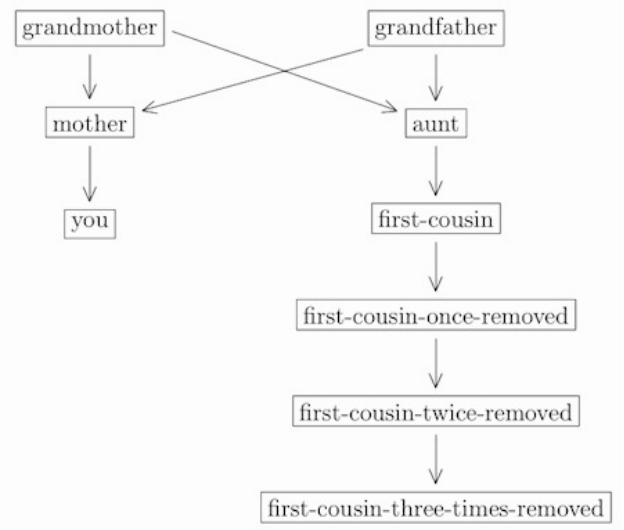
Finally, your sibling's child is your niece (or nephew, if male), and their child is your great-niece (or great-nephew), and so on:



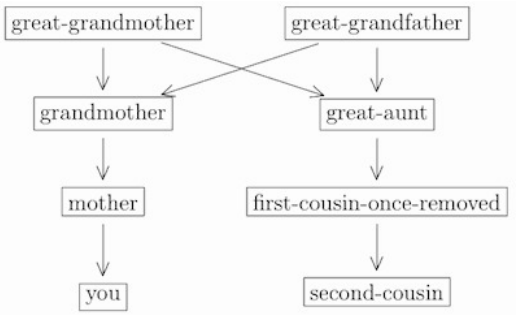
(and similarly for "nephew" instead of "niece" at any level). So now we know where your descendants' cousins come from. To see where *your* cousins come from, we have to move up to your parents' level. Your parents' siblings are your aunts and uncles, and their children are your first-cousins (since you and they share the same grandparents, but not the same parents):



If your cousins have children, then what are they to you? Well, children of your first-cousin are called your "first-cousins-once-removed", and their children are your "first-cousins-twice-removed", and so on:

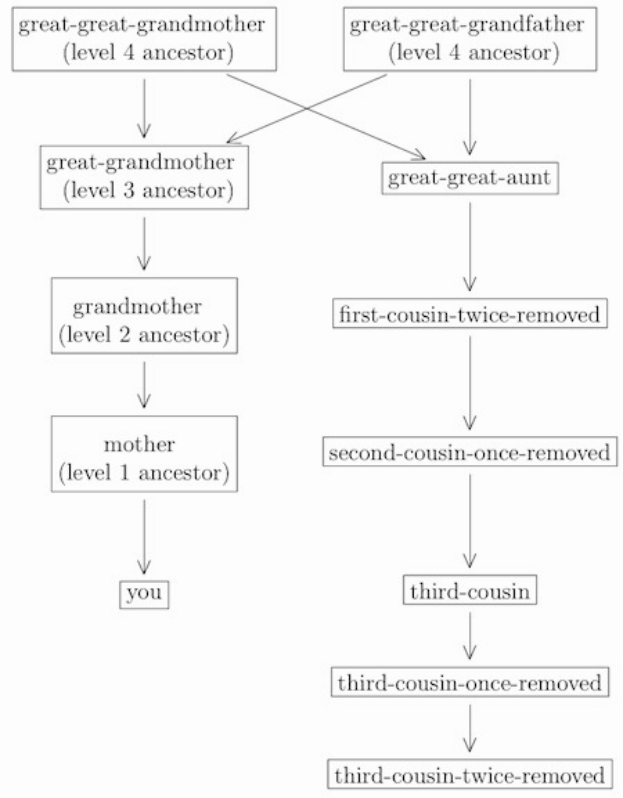


To see where your second-cousins come from, we have to move one more level up. Your grandparents' siblings are your great-aunts and great-uncles. So their children (i.e. your parents' cousins) are your first-cousins-once-removed. And their children are your second-cousins:



The same pattern continues upwards for all earlier generations. Once again, your nth cousins share your $(n+1)$-level ancestors, but not your nth-level ancestors. Siblings of your nth-level ancestors are your great-...-great aunts and great-...-great uncles, where "great" is repeated $n-1$ times. Furthermore, the nth cousins of your mth-level ancestors, and also the mth-level descendants of your nth cousins, are your nth cousins m times removed.

For example, with $n = 3$ and $m = 2$, this says that your *grandparents' third-cousins* are your third-cousins-twice-removed, and your *third-cousins' grandchildren* are also your third-cousins-twice-removed. Tracing back to $n = 3$ gives:



In this diagram, your third-cousin ($n = 3$) shares two of your great-great-grandparents (level $n+1 = 4$ ancestors) but none of your great-grandparents (level $n = 3$ ancestors). Your great-great-aunt is a sibling of your great-grandmother ($n = 3$). Your second-cousin-once-removed achieved that designation by being the second cousin ($n = 2$) of your mother (level $m = 1$ ancestor), while your third-cousin-once-removed achieved that designation by being the daughter (level $m = 1$ descendant) of your third cousin ($n = 3$).

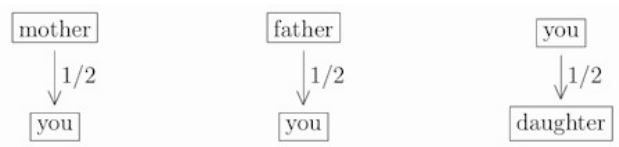
Tricky, right?

One of the reasons we care about family trees is because of a sense that certain family relations are "more related" to us, and should be assisted, protected and loved on that basis. This attitude presumably has an evolutionary basis: our genes survived through the ages because our ancestors made efforts to help them survive by caring not only for themselves, but for their *close* relatives too.

This raises the question of just how similar our relatives' genes are to our own. Well, first of all, about 99.9% of our genetic material is common to all humans (yes, even your girlfriend/boyfriend/partner), and indeed is what makes us human. Furthermore, some people may share other genes with us just by chance; for example, if I meet a stranger whose eyes are blue just like mine are, that does not necessarily establish that we are close relatives. In addition, there is lots of randomness in how genes are passed on (each individual gets half of their genetic material from their mother and half from their father, but which bits come from which parent is chosen at random and cannot be predicted), so we cannot draw precise conclusions with certainty.

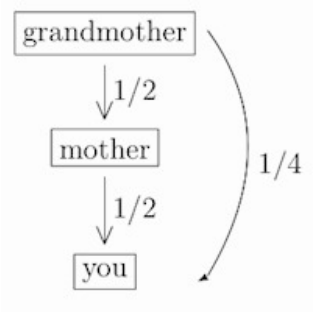
To deal with all of this, we assign to each pair of individuals a *relatedness coefficient* which represents the expected fraction (that is, the fraction *on average*) of their genes which are forced to be identical by virtue of their family relationship. This approach averages out all of the randomness, while focusing on genetic similarities specifically due to family connections.

According to this definition, strangers have a relatedness of 0 (the smallest possible value). By contrast, your relatedness with yourself is 1 (the largest possible value). Other relatedness coefficients fall between these two extremes. For example, your relatedness with your mother is $\frac{1}{2}$, since you obtain half of your genetic material from her. And your relatedness with your father is also $\frac{1}{2}$. By the same reasoning, your relatedness with your child is again $\frac{1}{2}$. So far so good:

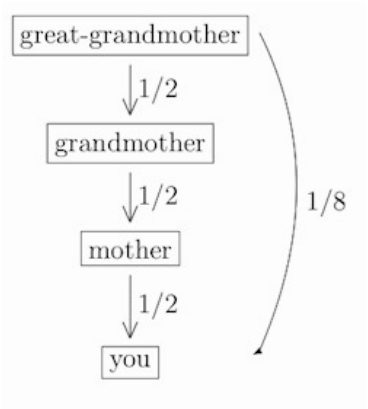


Next consider your maternal grandmother. She gave half of her genes to your mother, and then your mother gave half of her genes to you. It is possible that the half you took is exactly the same as the half your grandmother

gave. It is also possible that the half you took has no overlap at all with the half your grandmother gave. But on average, that is, in *expectation*, exactly half of the genetic material you took from your mother originated from your maternal grandmother. So, your relatedness coefficient with your grandmother is one-half of one-half, that is, $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$:



Continuing up the tree, your relatedness with your great-grandmother is one-half of one-half of one-half, that is $\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$:

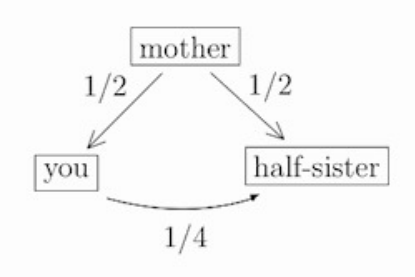


and similarly for "father" instead of "mother" at any level). In general, your relatedness coefficient with your level-n ancestor is $\frac{1}{2^n}$.

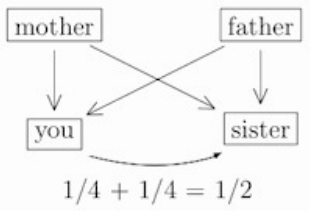
By the same reasoning, your relatedness coefficient with your level-n descendant is also $\frac{1}{2^n}$. So, for example, your relatedness coefficient with your daughter is $\frac{1}{2}$; with your granddaughter is $\frac{1}{4}$; and with your great-granddaughter is $\frac{1}{8}$ (and similarly for "son" instead of "daughter").

For siblings, the situation is a little bit more complex. Consider first the case of two half-siblings (half-sisters or half-brothers), that is, people who share just one parent. Since they each got half of their genetic material from that

one shared parent, their relatedness coefficient is one-half of one-half, that is $\frac{1}{4}$:

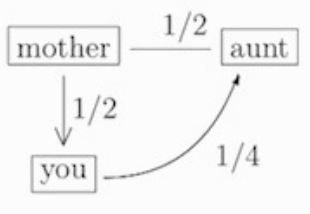


Regular (full) siblings similarly share $\frac{1}{4}$ of their genetic material through their mother, but also share $\frac{1}{4}$ of their genetic material through their father. This gives a total relatedness coefficient of $\frac{1}{4} + \frac{1}{4} = \frac{1}{2}$:



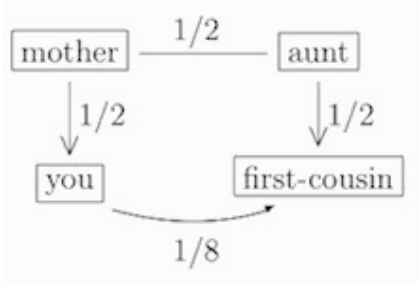
(One special case is identical twins, who we take to have identical genes and thus a relatedness coefficient of 1. But fraternal twins have relatedness coefficient $\frac{1}{2}$, just like other siblings.)

Continuing onward, since your mother and aunt are siblings, they have relatedness coefficient $\frac{1}{2}$. Meanwhile, you and your mother have relatedness coefficient $\frac{1}{2}$. Putting this together, you and your aunt have relatedness coefficient $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$:

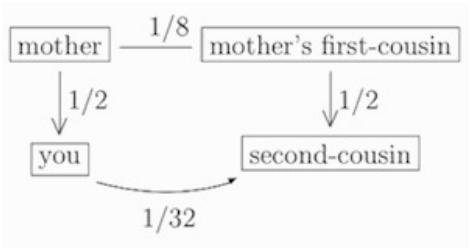


(and similarly with "aunt" replaced by "uncle"). And, your relatedness coefficient with your niece or nephew is also $\frac{1}{4}$.

Then, since your first-cousin has relatedness coefficient $\frac{1}{2}$ with your aunt, who in turn has relatedness coefficient $\frac{1}{4}$ with you, it follows that you and your first-cousin share relatedness coefficient $\frac{1}{8}$:



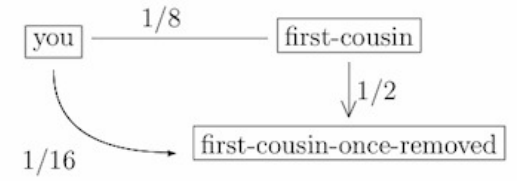
Now, since your mother and her first-cousin have relatedness coefficient $\frac{1}{8}$, and since you have relatedness coefficient $\frac{1}{2}$ with your mother, and since your mother's first-cousin has relatedness coefficient $\frac{1}{2}$ with her own child (who is your second-cousin), it follows that your relatedness coefficient with your second-cousin is $\frac{1}{2} \times \frac{1}{8} \times \frac{1}{2} = \frac{1}{32}$:



In general, switching to level-n cousins from level-$(n-1)$ cousins introduces two new factors of $\frac{1}{2}$. Since $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$, this means that your relatedness coefficient with your level-n cousin is always $\frac{1}{4}$ times your relatedness coefficient with your level-$(n-1)$ cousin.

It follows that your relatedness coefficient with your level-n cousin is equal to $\frac{1}{2^{2n+1}}$. So, your relatedness coefficient with your first cousin is $\frac{1}{8}$; with your second cousin is $\frac{1}{32}$; with your third cousin is $\frac{1}{128}$; and so on.

What about first-cousins-once-removed, and all of that? Well, since you and your first-cousin have relatedness $\frac{1}{8}$, and since your first-cousin and their child (your first-cousin- once-removed) have relatedness $\frac{1}{2}$, it follows that you and your first-cousin-once-removed have relatedness coefficient $\frac{1}{8} \times \frac{1}{2} = \frac{1}{16}$:



The pattern continues, with each new "removed" introducing an extra factor of $\frac{1}{2}$ into the product. It follows that your relatedness coefficient with your nth cousin, m times removed, is equal to $\frac{1}{2^{2n+m+1}}$. For example, your relatedness coefficient with your third cousin ($n = 3$) twice removed ($m = 2$) is equal to $\frac{1}{2^{6+2+1}} = \frac{1}{2^9} = \frac{1}{512}$ – not very close at all!

We can summarise the relatedness coefficients of various relationships in a table:

Relationship to you	**relatedness coefficient**
yourself	1
identical twin	1
parent, child	1/2
grandparent, grandchild	1/4
great-grandparent, great-grandchild	1/8
n^{th} level ancestor or descendant	$1/2^n$
sibling (sister or brother)	1/2
half-sibling	1/4
aunt, uncle	1/4
niece, nephew	1/4
great-aunt, great-uncle	1/8
great-niece, great-nephew	1/8
first-cousin	1/8
first-cousin-once-removed	1/16
second-cousin	1/32
second-cousin-once-removed	1/64
third-cousin	1/128
n^{th} cousin	$1/2^{2n+1}$
n^{th} cousin, m times removed	$1/2^{2n+m+1}$
stranger	0

This table can be thought of as indicating your level of evolutionary imperative to protect and assist your various relatives. That perspective was nicely summarised by the early evolutionary biologist J.B.S. Haldane; when he was asked if he would give his life to save a drowning brother, and replied "No,

but I would to save two brothers or eight cousins." He was merely observing that $2 \times \frac{1}{2} = 8 \times \frac{1}{8} = 1$, i.e. that two brothers, or eight cousins, are each "equal" (in evolutionary terms) to one copy of yourself.

There is an ancient Bedouin Arab saying: "I against my brother, my brothers and me against my cousins, then my cousins and I against strangers". Well, in the context of relatedness coefficients, it corresponds to the observation that your relatedness coefficient is higher with yourself (1) than with your brother ($\frac{1}{2}$), higher with your brother ($\frac{1}{2}$) than with your first-cousin ($\frac{1}{8}$), and higher with your first-cousin ($\frac{1}{8}$) than with a stranger (0). That is:

$$1 > \frac{1}{2} > \frac{1}{8} > 0.$$

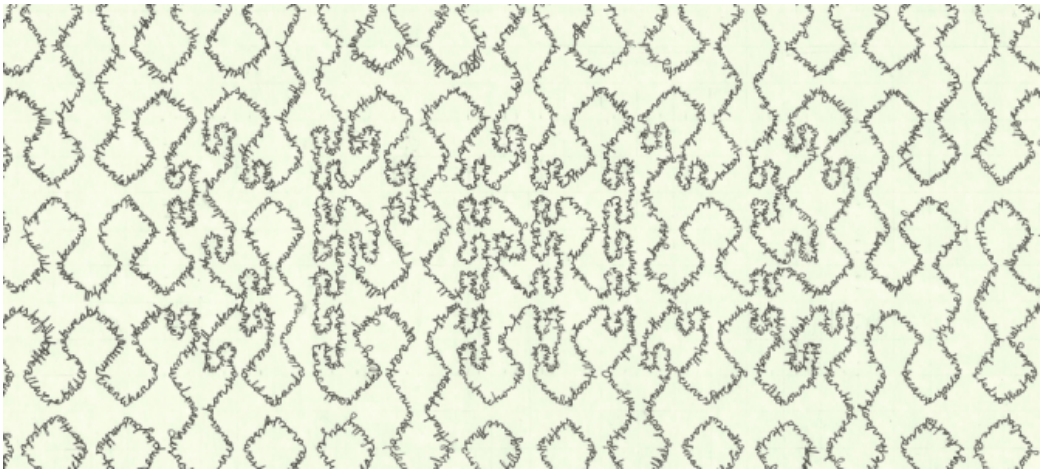
Remark 4.13. Of course, this model does not take into account all possible relationships. One can expand the model by considering spouses, adoption, in-laws, step parents, etc. □

4.4 Space-Filling Curves

We will continue our exploration of geometry, and dip into the study of dimension, by reading an article on space-filling curves. (*The following has been adapted from this article by Andrew Stacey, with whom all subsequent images originate.*)

A few years back, the students and staff at my school were set an art challenge over the summer break. Those who chose to take part were given a blank postcard and asked the question: What can you do with this space?

I feel that the obvious answer to this question is "fill it". But in mathematics, the combination of the words space and fill links to a very specific concept: space-filling curves. In this article, I'm going to explain what these are, and how this leads to my entry (shown below).



There are many variations on the theme of space-filling curves, but the original and simplest version is: *a continuous curve that passes through every point in the unit square.* Here *continuous* means it can be drawn without lifting your pencil (sharp corners are allowed, too) – though of course there is a more rigorous mathematical definition as well.

The first space-filling curve was designed by Giuseppe Peano in 1890. Surprisingly, his paper has no pictures. A year later, based on Peano's work, David Hilbert came up with a slightly different construction and included pictures. Hilbert curves seem to be the more popular of the two, with about 20 times more search hits on the internet. You might conclude that pictures are important in mathematics!

The reason why Peano and Hilbert (and others) were interested in these curves was because of Cantor. Not long before Peano published his paper, Cantor had started on his exploration of infinity and established that the unit interval $[0, 1]$ and the unit square $[0, 1] \times [0, 1]$ had the same quantity of points – to us, this is equivalent to knowing there exists a bijection between $\mathbb{R}$ and the plane $\mathbb{R}^2$. This result is what caused Cantor's famous utterance "I see it, but I don't believe it."

In particular, this meant that there was a function from the unit interval to the unit square that hit every single point. It was established that there couldn't be a continuous *bijection* between the unit interval and unit square, but the question remained as to whether it was possible to create a *surjective* function that was continuous. (That is, a function that visits every point in the unit square but possibly more than once. This is weaker than the "exact correspondance" of a bijection.) This is precisely the question "does there exist a curve completely filling the unit square that can be drawn without

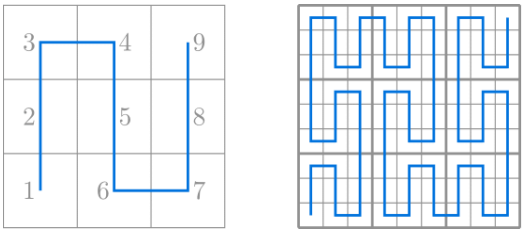
lifting one's pencil"? Peano's paper answers this with a definitive "yes".

Interestingly, Peano notes that his function is not *differentiable* - it has corners and "sharp" points. It would later be shown that no such differentiable function could exist.

My interpretation of what Peano and Hilbert were doing is that they were experimenting with mathematics. Cantor's ideas were spreading and they wanted to understand them better. To do so, they made examples testing the ideas, pushing them to their limits to see if they would break.

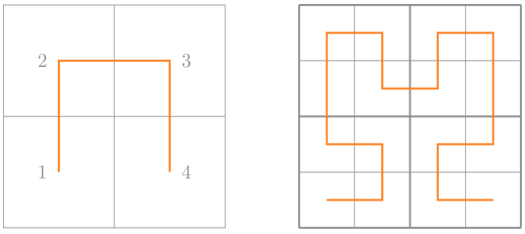
Of the several variants of space-filling curves out there, I'm going to focus on Peano curves. I'll explain why once I've shown what they are. A space-filling curve is actually very easy to draw; it is defining them mathematically that takes a little more work.

The modern way to construct a space-filling curve is as a limit of a family of curves that are more straightforward to define. This is not how Peano originally constructed his curve; it is more akin to how Hilbert constructed his curve, but the curves I will initially define are named after Peano. The construction of Peano's curve that I know best starts by dividing the unit square into 9 smaller squares and joining their centres in a specific order, as I've illustrated here on the left:



The second iteration splits each of the 9 squares into 9 smaller squares and repeats the pattern in each smaller grid, with some rotations and reflections so that they join up. There are some options as to how to choose the rotations. The one I've used is shown next to the first iteration above. This process can be repeated and the Peano curve is what you get by doing this "infinitely many times". Formally, it is the limit of this family.

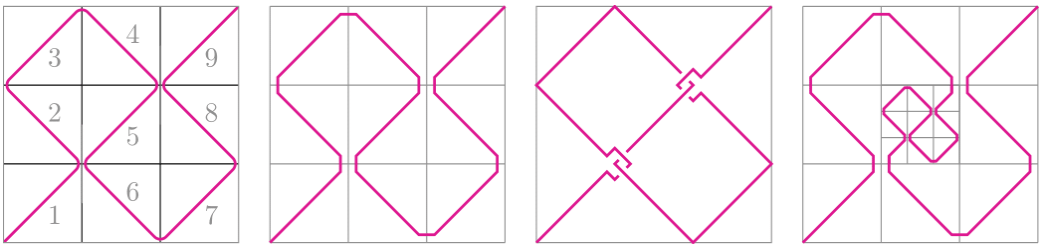
The Hilbert curve is constructed in a similar fashion except that the unit square is divided initially into 4 smaller squares, which are then each further divided into 4 and so on. The first two iterations are shown below:



I suspect that the real reason Hilbert curves are more well-known is because of the superiority of this division-by-two strategy and not due to Hilbert drawing pictures.

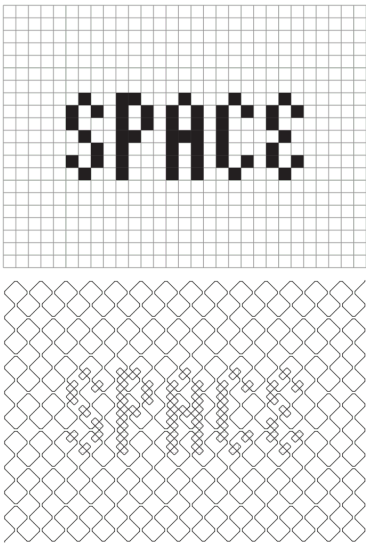
So, how do you adapt these ideas to fill a postcard, and outline the word "SPACE"?

The key insight was that the path taken from the centre of one square to the next was relatively unimportant. So long as the route from one centre to the next doesn't stray outside the two squares themselves, the limit will still be *the* Peano curve. This is because as the squares get smaller, the distance between the original construction and a variation will be no bigger than the width of the squares – and this gets vanishingly small. So rather than joining the centres with straight lines, we travel from one corner to the opposite:



(As the curve now meets itself, it can be somewhat tricky to trace its actual path. To combat that, I've squared-off the corners in the second image. An alternative is to tie a little knot at each corner – the third image.)

With this basic pattern it is possible to replace an individual square by a copy of itself (suitably scaled) without disturbing the rest of the curve, as shown in the rightmost picture above. This can also be viewed as a type of path replacement where a single diagonal line is replaced by the crinkled line, and then each segment of that is replaced by a copy of itself, and so on. Using this I could make a picture: start with a grid of squares of suitable dimensions. Create a black-and-white picture by colouring in some squares, then convert that to a Peano curve by using the *next level* for the darker squares.



Finally, I turned the curve into a text written in joined writing.

This technique is interesting as, in the limit, it produces *the* space-filling Peano curve. Therefore by careful choices, it is possible to find a family of curves starting with just a diagonal line, ending with the full Peano curve, and which makes a picture of Giuseppe Peano himself somewhere along the way.

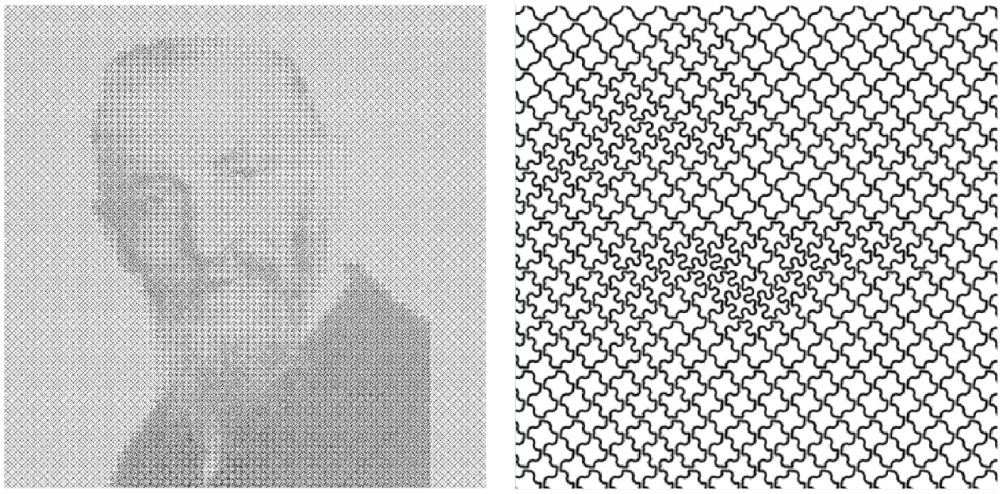


Figure 9: Giuseppe Peano to the left; close up detail of his left eye to the right.

There is a lot of literature on space-filling curves, ranging from viewing them as art forms, as I have done here, through hearing them as music, to practical applications in the field of image analysis. From the historical perspective, the original papers of Cantor, Peano, and Hilbert can be found online (albeit in their original languages!) and there is a chapter on them in the book *Curves for the Mathematically Curious*.

Question 4.14. This article (unedited by me) was originally published in *chalkdust magazine* issue 11. Some other articles from this issue:

1. Adopt a polyhedron.
2. Automated Joke Generation. □

4.5 Guest speaker: Dr. Sylvy Anscombe

Sylvy received her PhD in Mathematics from the University of Oxford in 2013. She researches algebra, number theory, and mathematical logic. Today she will introduce a definition of "dimension" that is useful for measuring shapes that don't fit into our usual way of thinking, such as *fractals*.

Sylvy is a coauthor of *Unmasked: the Science of Superheros*.

5 The Game of Hex

We will now discuss some of the mathematics behind the game 'Hex', invented independently by mathematicians Piet Hein and John Nash. John Nash is famous in particular for his fundamental contributions to game theory, which (co-)won him the 1994 Nobel Prize in Economics. The game is traditionally played on an 11×11 rhombus composed of hexagonal compartments. Players shade in one hexagonal square per turn (no skipping turns) and the winner is the first player who manages to connect their opposite sides in an unbroken chain. See below for a game of Hex won by the blue player.

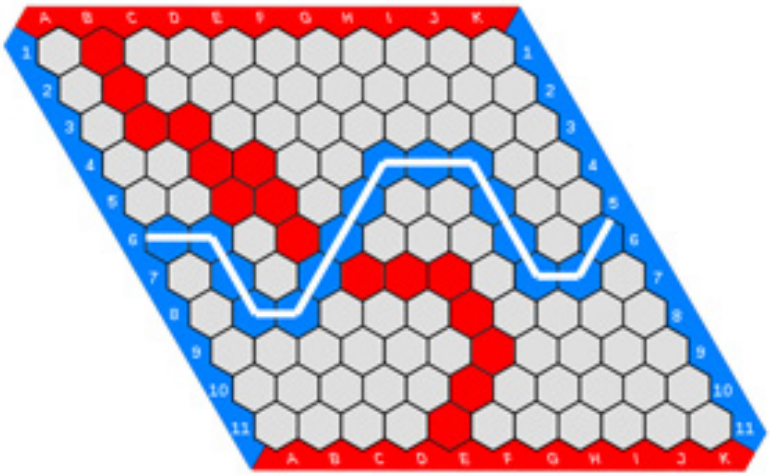
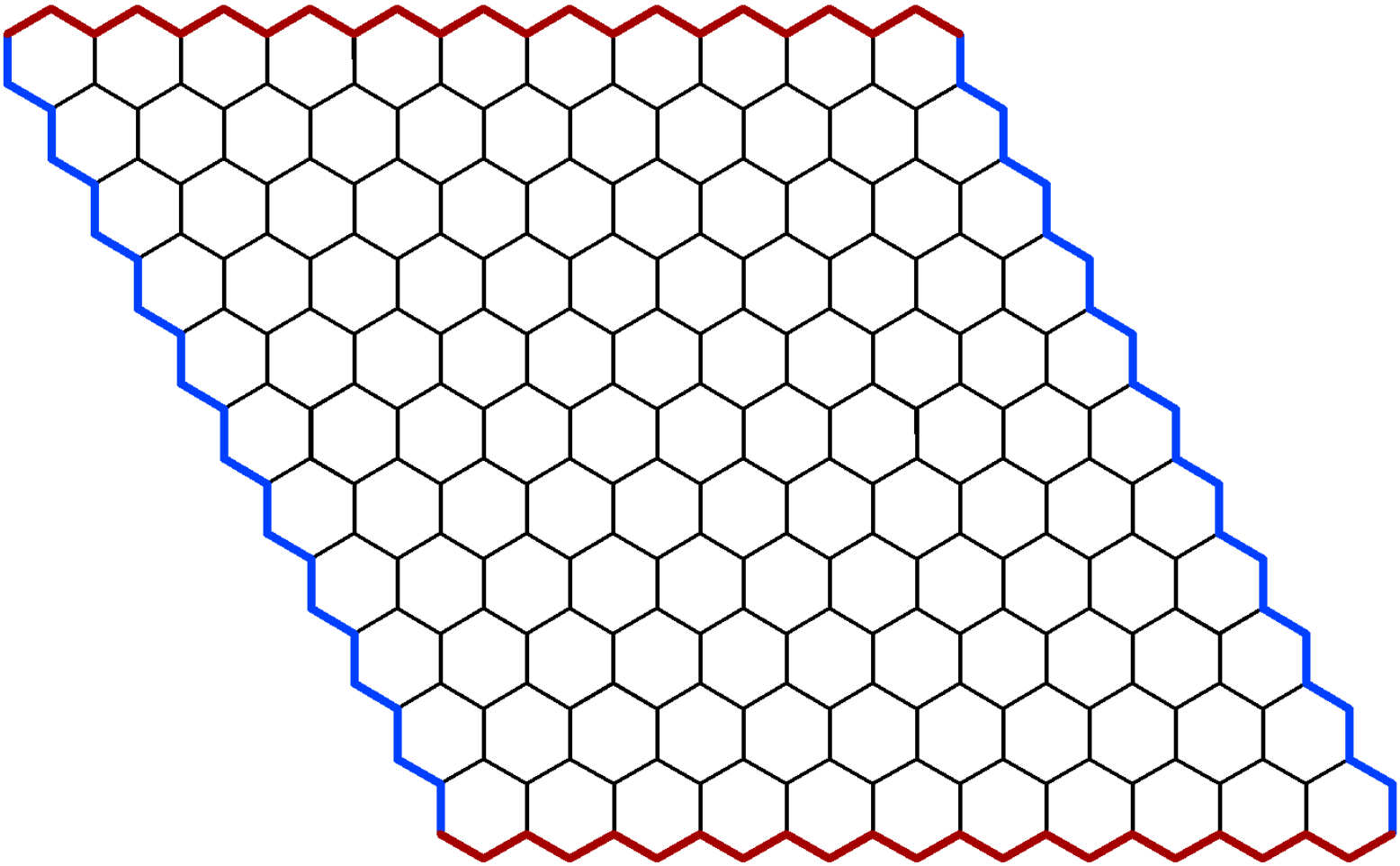


Figure 10: Image credit: Wikipedia.

Without knowing anything about strategy or proofs, I want the class to split into pairs of students and play against each other using either the board on the next page or this link (see here for online play against a computer). In game theory, when discussing strategy, we assume the players are both *intelligent* (they can determine their best option between strategies) and *rational* (they choose the strategy which maximises their payoff). So play to the best of your ability, think your moves through and play to win!

Question 5.1. Did you discover any interesting or advantageous strategies? What are they? □



Here is a very good strategy guide, if anyone would like to do some independent reading.

Now let's talk about some results regarding this game:

1. Hex is a *determined game* - that is, it can *never* end in a tie. There is always going to be a winner and loser. This was proven by Nash c. 1949 in the "Hex Theorem" which we will reprove ourselves shortly.

   Fun fact: this result is *equivalent* to a result known as the "Brower Fixed Point Theorem" in topology, which says that any continuous function $f : [0,1]^2 \to [0,1]^2$ has a fixed point (i.e. there exists $x_0 \in [0,1]^2$ such that $f(x_0) = x_0$).

2. In Hex, having an extra piece on the board is always an advantage, never a handicap - even if the piece is placed randomly on the board. We'll prove this result shortly too.

3. The first player (first person to make a move) has a winning strategy. That is, there exists a strategy the first player can play that guarantees them to win. This was prove by Nash in 1952 and we will reprove it ourselves shortly. Note that Nash's proof was not a constructive one - he proved a winning strategy **existed**, but we don't automatically known **what that strategy is**.

4. Related to the previous point, in 2002 the first explicit winning strategy on a 7×7 board was described. In the 2000s, by using brute force search computer algorithms, Hex boards up to size 9×9 (as of 2016) have been completely solved. The 11×11 board is (as of right now) still unsolved. This means you should never face an opponent on a 9×9 board or smaller - they might know the winning strategy from the get go!

5. There is a "Hex Uniqueness Theorem" that says it is impossible for any Hex Board to be coloured in such a way as to satisfy winning conditions for more than one player. We will also prove this ourselves.

   When combined with the *Hex Theorem*, this tells us that in Hex there is *always exactly one winner and loser*.

6. Hex is a finite, perfect information game. 'Perfect Information' means all players have all the knowledge possible of all the previous moves, all the moves that they and their opponent could make, and all of the possible consequences of any move. Chess is another example of a game

with perfect information as each player can see all of the pieces on the board at all times.

7. According to the Wikipedia page on Hex, in 11×11 Hex there are approximately 2.4×10^{56} possible legal positions; this compares to 4.6×10^{46} legal positions in chess.

   There are also various variants of Hex possible.

Let's tackle the third point first. We will need the following general result on games:

Theorem 5.2. (Zermelo). In a finite two-person game of perfect information in which the players move alternately and in which chance does not affect the decision making process, if the game cannot end in a draw, then one of the two players must have a winning strategy. ■

The proof of this is very complicated and outside the scope of what we can cover in this course. For now, you'll just have to believe this result. This was proven in 1913 by Ernst Zermelo – the same Zermelo from ZFC.

Theorem 5.3. The first player in Hex on a board of any size has a winning strategy.

This is a *reductio ad absurdum* (proof by contradiction) existence proof attributed to John Nash. Such a proof gives no indication of a correct strategy for play. The proof is common to a number of games including Hex, and has come to be called the "strategy-stealing" argument. Here is a highly condensed informal statement of the proof:

Proof.

1. Either the first or second player must win (by the first point), therefore there must be a winning strategy for either the first or second player (by Zermelo's Theorem).

2. Let us assume that the second player has a winning strategy (proof by contradiction, remember).

3. The first player can now adopt the following defence: She makes an arbitrary move. By point # 2, she is not at a disadvantage because of this. Thereafter she plays the *winning second player strategy assumed above*. If in playing this strategy, she is required to play on the cell

where an arbitrary move was made, she makes another arbitrary move. In this way she plays the winning strategy with one extra piece always on the board (which, again, is not a handicap). Moral of the story: the first player 'becomes' the second player.

4. Therefore the first player can win, and the second player can win – a contradiction to the Hex Uniqueness theorem.

5. Because we have now contradicted our assumption that there is a winning strategy for the second player, we are forced to drop this assumption.

6. Consequently, there must be a winning strategy for the first player, otherwise we reach the above contradiction.

■

Comments on this:

- What is keeping player B from stealing the strategy back? As in, what if B plays another bogus move, and we're back at A being the effective first player? We are assuming perfect play, so if A's winning strategy beats B when B is playing the best she can, that strategy will also beat B when she is playing less than perfectly.

  Once the strategy is stolen, it cannot be lost by the stealer!

- Given this, we would then wonder, how can A steal the strategy in the first place? Player B had the winning strategy, so if A started out playing a random move, then surely B can just keep playing the winning strategy and beat A, right? This is the heart of the contradiction and our proof. Given that A played a random move and has effectively made B the first player, B is no longer in the position to use the winning strategy even though our assumption would imply that B can just keep playing perfectly and beat A. Hence, our assumption that B had the winning strategy must have been flawed.

- This strategy-stealing argument can be applied to any other symmetric game where having an extra move or game piece on the board can never hurt you, and there can only be one winner – e.g. tic-tac-toe.

Now, the second point:

Theorem 5.4. Having extra pieces of your own colour lying on the board cannot hurt you.

Proof. A short and sweet argument is as follows: Suppose that there is an extra piece at position x on the board. If x is part of your strategy, then on the turn when you should be playing at position x, you could instead lay down another piece somewhere else. If there is nowhere else to place your piece, the board is full and the game ended at the previous player's turn. If x is *not* part of your strategy, then you would not care that it is occupied. In either case, your strategy is unaffected so an extra piece of your own colour on the board has not hurt you, as required. ■

Finally, the first point. To prove the *Hex Theorem* we will first need to learn some graph theory first.

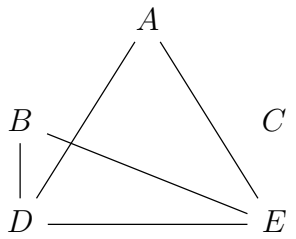
5.1 Graph Theory

Graph theory is a huge area of mathematics, and we will really not do it justice by dipping in and introducing a few particular ideas. For a more complete picture of graph theory, see e.g. here.

Definition 5.5. A *graph* is a mathematical structure consisting of *vertices* and *edges*.

A graph is written as a pair (V, E) where $V = \{v_1, v_2 \dots\}$ is a (possibly infinite) set of points, known as vertices, and $E = \{(v_i, v_j), \dots\}$ is itself a set of pairs, where (v_i, v_{ij}) denotes that vertex v_i is connected to vertex v_j by a line.

Graphs are *undirected*, i.e. $(v_i, v_j) = (v_j, v_i)$. An example of a graph is the following:



Here, $V = \{A, B.C, D, E\}$ and

$$E = \{(A, B), (A, D), (A, E), (B, D), (B, E), (D, E)\}.$$

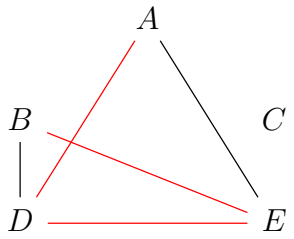
Here, C is called *isolated*. Another way to phrase this would be in terms of the *degree* of C:

Definition 5.6. Let (V, E) be a graph. The degree "$\deg v$" of a vertex $v \in V$ is defined as the number of edges of the graph that are *incident* to v, i.e. the number of edge with v as one of their endpoints.

So $\deg C = 0$, $\deg A = 2$, $\deg D = 3$, etc. Degree is fundamentally connected to what *walks* may be taken through a graph.

Definition 5.7. A *walk* is a sequence of consecutive edges which joins a sequence of vertices.

A walk is quite literally a "walk" through the graph. For example, in red is a walk:

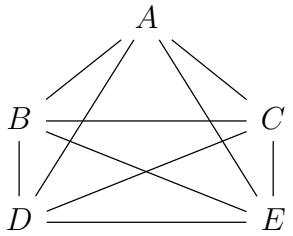


This walk begins at (say) A, visits D, then E, then B. We write this walk by smashing all the vertices together: $ADEB$. The only condition on a walk is that it is "connected", in some sense – in this graph, there cannot be a walk including C, and there cannot be a walk ABD.

Of course, these are very general objects – if we are to get some sort of classification or control over them, we need to think of more specific walks. That is what this next definition does:

Definition 5.8. A *trail* is a walk in which all edges are distinct. A *path* is a trail in which all vertices are distinct. (Therefore in paths, all vertices and edges are distinct.)

Let's discuss this with the following "complete" graph:

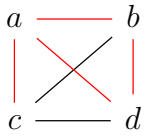


When remembering the terminology, keep this mnemonic in mind:

- A TRAIL TRAVERSES each EDGE once.
- A PATH PASSES THROUGH each POINT (i.e. vertex) once.

It's easy to think of examples of walks that are trails but not paths. Consider the following walk on k_4:

$abdac$



This is clearly not a path, but is a trail. What about the other way around?

Theorem 5.9. All paths are trails.

Proof. We will prove this by contradiction. Assume there is a path $p = v_{i_0} v_{i_1} \dots v_{i_n}$ which is not a trail. Then it must have a repeated edge (a, b):

$$p = v_{i_0} v_{i_1} \dots ab \dots ab \dots v_{i_n}$$

or

$$p = v_{i_0} v_{i_1} \dots ab \dots ba \dots v_{i_n}$$

Suppose the repeated edges fall on vertices v_w, v_x, v_y, v_z respectively. There are two cases:

1. $x = y$, e.g. something like $v_0 aba v_2$. In this case $v_w = v_z$ or $v_y = v_z$, so a vertex is repeated in the path.

2. $x \neq y$, e.g something like v_0abv_2ab. Again, at least one vertex is repeated in the path.

It boils down to the fact that no matter what, a vertex is repeated; a contradiction to the fact p is a path. We conclude the assumption is false; all paths are trails. ■

What about those special trails or paths which "loop back" to the start?

Definition 5.10. A walk is *closed* if its starting vertex is the same as its ending vertex (i.e. it loops back to the start).

A *circuit* is a closed trail – that is, a walk with no repeated edges that returns to its starting point.

A *cycle* (or *simple circuit*) is defined as a closed trail where no other vertices are repeated apart from the start/end vertex (so in this way, it's like a "closed path").

One can use the following mnemonic:
CIRCUIT = CLOSED TRAIL.

Finally, one last question: what about those walks that use *every* edge of the graph?

Definition 5.11. A walk is *Eulerian* if it traverses every edge. (These aren't particularly special.)

An *Eulerian trail* in a graph is a trail that traverses every edge of the graph. As it is a trail it can only traverse edges once. Thus an Eulerian trail is a walk traversing every edge *exactly* once.

An *Eulerian circuit* in a graph is a circuit that traverses every edge of the graph. (The difference between an Eulerian *trail* and an Eulerian *circuit* is that an Eulerian circuit is closed – that's it!)

Use the following mnemonic:
EULERIAN TRAIL = EVERY EDGE is TRAVERSED.
EULERIAN CIRCUITS ARE CLOSED.

I promise that is the end of the definitions.

All of this material was developed by Leonhard Euler in the 1730's, and (one of) his big result(s) was the following:

Theorem 5.12. Let v be a vertex of the graph. Given any circuit in the graph, the number of edges incident to v traversed by that circuit is even. ■

This statement should make intuitive sense, once we look at a few examples of circuits.

A consequence of that theorem is the following one:

Theorem 5.13. If a graph has an *Eulerian circuit*, then the degree of *every* vertex of that graph must be even. ■

Question 5.14. Consider the famous "Seven Bridges of Königsberg" Problem:

The city of Königsberg in Prussia was set on both sides of the Pregel River, and included two large islands – Kneiphof and Lomse – which were connected to each other, or to the two mainland portions of the city, by seven bridges (see picture). The problem is to devise a stroll through the city that would cross each of those bridges *once and only once*.

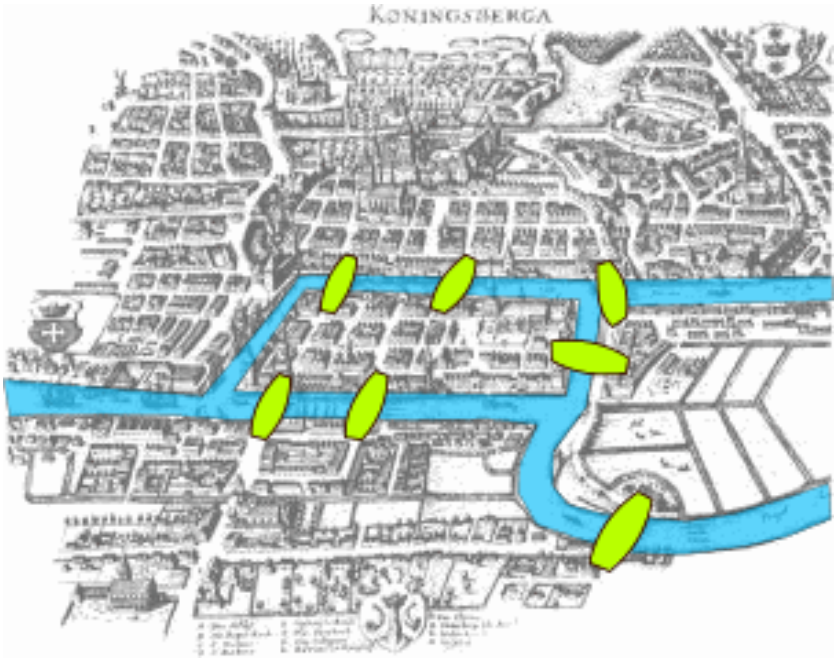


Figure 11: Image credit: Wikipedia.

Can you solve this problem?

HINT: start with the easier problem of a stroll that begins and ends in the same location. □

5.2 Applying Graph Theory

Now let's return to the context of Hex, and the Hex theorem. To prove it, we will need the following little theorem ("lemma"):

Lemma 5.15. A finite graph whose vertices have degree at most two is the union of disjoint subgraphs, each of which is either (i) an isolated vertex, (ii) a cycle, or, (iii) a path.

This is proven by *induction*: one proves a statement $P(n)$ about natural numbers $n \geq n_0$ by induction by proving

$$P(n_0) \wedge \forall n \geq n_0 (P(n) \implies P(n+1)).$$

Typically, $n_0 = 0$ or 1. Try to prove the following to practise proofs by induction:

Question 5.16. Prove for all $n \in \mathbb{N}$, $n \geq 1$, that $1 + 2 + \cdots + n = \frac{n(n+1)}{2}$. □

Proof of the lemma.
We induct on the number of edges in a graph. Consider a graph G with n vertices. Each vertex can have degree at most two, so G can have at most n edges.

For simplicity, we denote a graph with k edges as G_k. In the base case, G_0, all the vertices are isolated. When a graph has $n+1$ edges, we randomly choose an edge to remove; call the edge (u, v). Then the vertices u and v now have degree at most 1 since they had degree at most two before we removed edge (u, v). Therefore u and v cannot be on any cycles. By assumption, G_n is the union of disjoint isolated vertices, cycles, and paths.

We now add (u, v) back into the graph. The subgraphs that were disjoint from u and v in G_n are unchanged by the addition of (u, v), and the vertices u and v are now either on the same path or cycle. Therefore, G_{n+1} is also

the union of disjoint isolated vertices, cycles, and paths. Hence the lemma is true for *all* G_k with $0 \leq k \leq N$; in particular this is true for G_N, as required. ■

How is this lemma useful?

Next step: For simplicity, we substitute the coloured tiles in the game with x's and o's. We represent the game board as a graph $G = (V, E)$, with a set of vertices V and a set of edges E. **Each vertex of a hexagonal board space** (a corner of the hexagonal tile) is a vertex in V, and each side of a hexagonal board space is an edge in E. We create four additional vertices, one connected to each of the four corners of the core graph; call these new vertices u_1, u_2, u_3, and u_4 and the edges that connect them to the core graph e_1, e_2, e_3, and e_4. An X-face is either a tile marked with an x or one of the regions marked X or X'. Similarly, O-face is either a tile marked with an o or one of the regions marked O or O'. Hence, the edges e_1, e_2, e_3, and e_4 lie between an X-face and an O-face since the regions X, X', O', and O are considered 'faces' as well. Let's draw a picture to illustrate this:

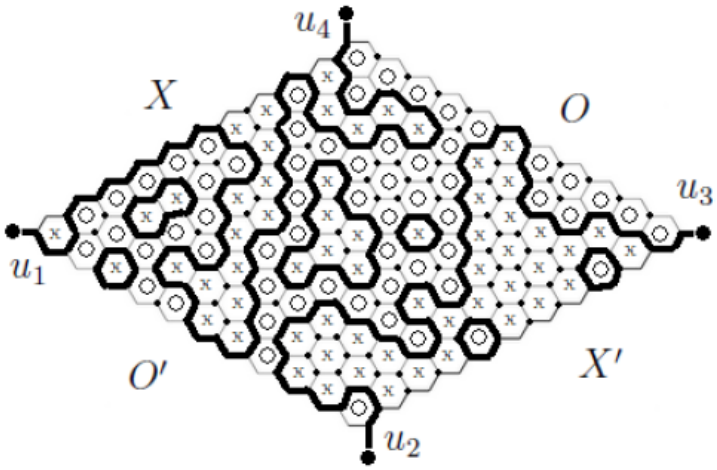


Figure 12: Image credit: SP.268: The Mathematics of Toys and Games

Theorem 5.17. (Hex Theorem). If every tile of the Hex board is marked either x or o, then there is either an x-path connecting regions X and X' or an o-path connecting regions O and O'.

Proof.
First, we construct a subgraph $G' = (V, E')$ of G, with the same vertices but a subset of the edges. We define an edge to belong to E' only if it lies

between a X-face and an O-face (figure 12). Therefore, e_1, e_2, e_3, and e_4 belong to E'. Note the vertices u_1, u_2, u_3, and u_4 each have degree one.

If all three hexagons around a vertex are marked the same, then the vertex is isolated in G' and has degree zero. If a vertex is surrounded by two hexagons of one pattern and one hexagon of the other pattern, then that vertex has two incident edges. Hence, each vertex in the core graph has degree either zero or two. Since G' has vertices with degree at most two, by the lemma, G' is a union of disjoint subgraphs, each of which are isolated vertices, cycles, or paths. Each of the vertices u_1, u_2, u_3, and u_4 are ends of some *path* because they have degree one. The disjointness of subgraphs in G' ensures that these paths do not cycle (loop back to the starting node). Therefore, there exist two simple paths in G', each connecting two of u_1, u_2, u_3, and u_4. Note that these paths cannot overlap as at an 'overlap' node, its degree would be 4, a contradiction to what we proved at the start of this paragraph.

Although the winner depends on the orientation of the paths, the paths do trace out a winning chain of hexagons. Therefore, for any arbitrary configuration of the Hex board, a winning path for one of the players exists, as required. ■

I have some notes regarding the equivalence of the Hex Theorem and the Brower Fixed Point Theorem, however we probably won't get a chance to cover the proof in class as we would need to cover a lot of topology first. The notes can be found here and they are based off this paper.

With regards to the fifth bullet point, the Hex Uniqueness Theorem, the proof we will discuss comes from this paper. According to the paper, it is "inspired by David Berman's inductive proof of the fact that Hex always has a winner" – what we saw above, but elaborating on and formalising one small throwaway remark we made. The proof proceeds as follows:

Theorem 5.18. (The Hex Uniqueness Theorem). It is impossible for any Hex Board to be coloured in such a way as to satisfy winning conditions for more than one player.

Proof.
We shall prove this via a "double induction" for $n \times m$ boards.

The (first serious) base case regards $n = 2$ or $m = 2$. It is easily demonstrable that Hex Uniqueness holds for any $2 \times m$, $n \times 2$, and smaller-dimensional boards.

We therefore assume it true for all $i \times j$ boards, with $i < n$ and $j < m$, $i = n$ and $j < m$, or $i < n$ and $j = m$. Furthermore, for the purpose of contraction, imagine an $n \times m$ board $H(n, m)$ coloured such that both Black and White have won. Each player therefore has a winning path connecting opposite sides of the board. More specifically, each player has a *minimal* path, which we define as a winning path M contained in a given winning path such that M contains precisely one hexagon adjacent to each necessary edge; these hexagons in turn are each adjacent to precisely one other hexagon on M, and all other hexagons contained in M border precisely two other component hexagons.

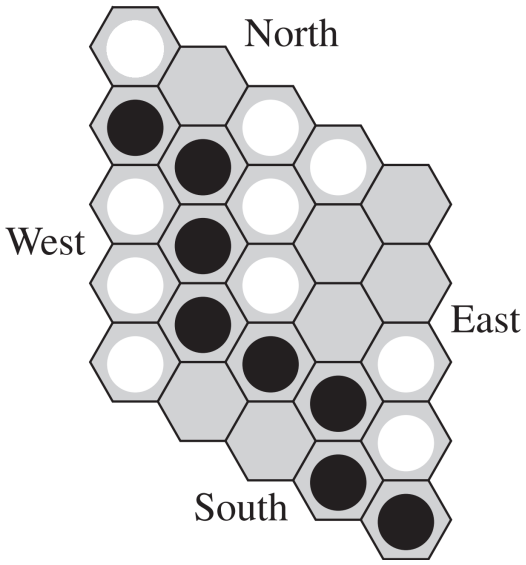


Figure 13: Image credit: modified from Samuel Clowes Huneke.

We leave it to the reader to certify that such a minimal path is indeed contained in any winning path. Suppose Black wishes to connect East to West, and White wishes to connect North to South.

First, consider Black. Because she has a path from East to West, we can remove the nth column from the board and Black will retain a winning path. However, by our above assumption, there can be only one winner on this new $(n-1) \times m$ board. Hence, White's minimal path from North to South on $H(n, m)$ must contain a hexagon in the nth column. We follow the same argument to show that White's minimal path must contain a hexagon in the first column too. Hence, on the $(n-2) \times m$ board created by removing the first and last columns, White retains a path P connecting *East and West*. Note that none of the hexagons contained in P may be in the first or final

rows; this would contradict the fact that the path we are considering for White is minimal.

Now run the same argument from White's perspective: remove the first row of $H(n, m)$. Because White retains a winning path on the new $n \times (m-1)$ board, Black cannot win on it, meaning that Black's original minimal path must contain a hexagon in the first row. Similarly, Black's minimal path must contain a hexagon in the mth row. Thus, by the same argument as above, on the $n \times (m-2)$ board created by removing the first and last rows, Black has a path connecting *North and South*, no hexagon of which can be contained in the first or final columns.

We now remove the first column, the first row, the nth column, and the mth row to create an $(n-2) \times (m-2)$ board. Note that White has a path connecting East and West and Black a path connecting North and South. Imagine that all black tiles are white and all white tiles black. Then, winning conditions would be satisfied for both players on this $(n-2) \times (m-2)$ Hex board, contradicting our inductive assumption for smaller-dimensional boards. Hence, by induction, no colouring exists for any Hex board that satisfies winning conditions for more than one player, as required. ■

5.3 Final Remarks

The Hex Uniqueness theorem may also be proven using the *Four Colour Theorem*, a famous result also in graph theory.

The *Four Colour Theorem* states that, given any separation of a plane into bordering regions, producing a figure called a "map", no more than four colours are required to colour the regions of the map so that no two adjacent regions have the same colour. Here, "adjacent" means that two regions share a common boundary curve segment, not merely a corner where three or more regions meet.

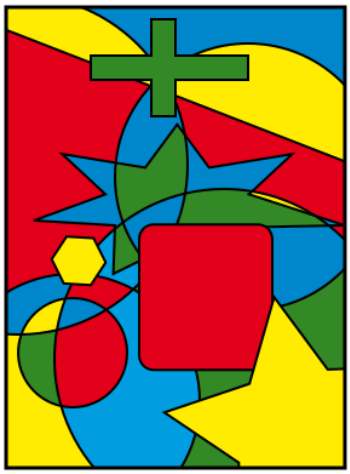


Figure 14: Image credit: Wikipedia.

The conjecture was first stated just over 150 years ago and finally proved conclusively in 1976. It is an outstanding example of how old ideas combine with new discoveries and techniques in different fields of mathematics to provide new approaches to a problem. It is also an example of how an apparently simple problem was thought to be "solved" but then became more complex, and it is the first spectacular example where *a computer was involved in proving a mathematical theorem.*

The conjecture that any map could be coloured using only four colours first appeared in a letter from Augustus De Morgan, first professor of mathematics at the new University College London, to his friend William Rowan Hamilton, the famous Irish mathematician, in 1852. It had been suggested to De Morgan by one of his students, Frederik Guthrie, on behalf of his elder brother Francis, who was trying to colour a map of England (Francis later became professor of mathematics at the University of Cape Town).

The problem, so simply described, but so tantalisingly difficult to prove, caught the imagination of many mathematicians at the time. In the late 1860s De Morgan even took the problem and his unfinished proof to America where among others, Benjamin Peirce (a famous mathematician and astronomer) became interested in it as a way to develop their own logical methods.

De Morgan used the fact that in a map with four regions, each touching the other three, one of them is completely enclosed by the others. Since he could not find a way of proving this, he used it as an axiom.

Figure 15: Image credit: NRICH.

In 1878 Arthur Cayley at a meeting of the London Mathematical Society asked whether anyone had found a solution for De Morgan's original question, but although there had been some interest, no one had made any significant progress. Cayley became interested in the problem and in 1879 published a short paper (*On the colouring of maps*) where he explained some of the difficulties in attempting a proof and made some important contributions to the way the problem was approached. His question that, "if a particular map is already successfully coloured with four colours, and we add another area, can we still keep the same colouring?" began another line of enquiry which led to the application of mathematical induction to the problem. However it turned out that no, the same colouring cannot be kept.

The best route forward in these situations is often to attempt an easier problem in the hopes that the proof of the easy problem will tell you something about how to tackle the difficult problem. First, the following result was shown:

Theorem 5.19. Every map has at least one country with five or fewer neighbours. ■

In 1879, Alfred Kempe started from the "five neighbours property" and developed a procedure known as the method of "Kempe Chains" to find a proof of the Four Colour Theorem. He published this proof in the American Journal of Mathematics. He found two simpler versions that were published in the next year, and his proof stood for ten years before Percy Heawood showed there was an important error in the proof-method that Kempe had used. Although Heawood found a major flaw in Kempe's proof method in 1890, the method itself was still important and useful. Heawood was unable to go on to prove the four colour theorem, but he made a significant breakthrough and proved conclusively that all maps could be coloured with *five colours*.

By 1900, mathematicians knew that a graph can be constructed from any map using the powerful concept of *duality*. In the dual, the regions are

represented by vertices and two vertices are joined by an edge if the regions are adjacent. In these graphs, the Four Colour Conjecture now asks if the vertices of the graph can be coloured with 4 colours so that no two adjacent vertices are the same colour.

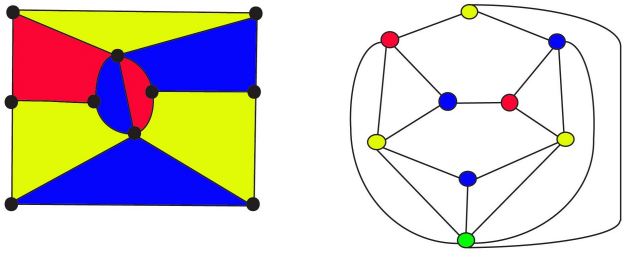


Figure 16: Image credit: NRICH.

During the first half of the twentieth century, mathematicians focused on modifying these kinds of techniques to reduce complicated maps to special cases which could be identified and classified, to investigate their particular properties and developed the idea of a "minimal set" of map configurations that needed to be tested.

In the first instance, the set was thought to contain nearly 9,000 members – which was an enormous task – and so the mathematicians turned to computer techniques to write algorithms that could do the testing for them. The algorithms used modified versions of Kempe's original idea of chains together with other techniques to reduce the number of members of the minimal set.

After collaborating with John Koch on the problem of reducibility, in 1976 at the University of Illinois, Kenneth Appel and Wolfgang Haken eventually reduced the testing problem to an unavoidable set with 1,936 configurations, and a complete solution to the Four Colour Conjecture was achieved. This problem of checking the reducibility of the maps one by one was double checked with different programs and different computers. Their proof showed that at least one map with the smallest possible number of regions requiring five colours cannot exist.

However, as the proof was done with the aid of a computer, there was an immediate outcry. Many mathematicians and philosophers claimed that the proof was not legitimate. Some said that proofs should only be "proved" by people, not machines, while others, of a more practical mind questioned the reliability of both the algorithms and the ability of the machines to carry them out without error. Whatever the opinions expressed, the situation

produced a serious discussion about the nature of proof which still continues today.

Nowadays, it is accepted that the Four Colour Theorem has been proven – but the larger philosophical questions of "what constitutes a proof?" and "can a computer produce a proof independent of a human mind?" remain.

Question 5.20.

- Try your hand at four-colouring here.
- Have a look at the Wikipedia page for some more basic information about the history, proof, and generalisations.
- See this thesis by Oscar Leward for a full, modern exposition of the proof.

6 University Mathematics 101

6.1 Guest speaker: Soinbhe Nic Dhonncha

Soinbhe is a mathematics PhD student at the University of Manchester studying the intersection of algebra, logic, and "category theory". She will introduce to us the basics of *algebraic topology* and explain why a coffee cup is the same as a doughnut.

6.2 An Introduction to Structure

To finish, I will briefly introduce some of the weird and wonderful mathematical structures you would encounter at the beginning of a university-level mathematics degree.

- We have already seen *groups* this morning. Recall: a group G is a set with a binary operation $- * - : G \times G \to G$ and distinguished element $e \in G$, such that
  - $\forall a, b, c \in G$, $(a * b) * c = a * (b * c)$.
  - $\forall a \in G$, $a * e = e * a = a$. (e is the *identity.*)
  - $\forall a \in G$, $\exists b \in G$ such that $a * b = b * a = e$. (b is the *inverse* of a.)

  The common example given is $(\mathbb{Z}, +, 0)$.
- We will discuss *rings* and *fields*, which incorporate more operations.
- We can then discuss *varieties* and *(real) manifolds*. Finally I will introduce a structure with two sets and two binary operations: a *vector space*.
- We will also discuss the maps between these objects. We will see the notions of *isomorphism*, *continuous*, and a *metric space*.
- The format of these discussions will be a definition (or two) followed by a discussion on some examples and non-examples. Some basic properties of each may be menitoned, however we will not go into details or prove anything.

6.3 The Q & A

I'd like to spend out last thirty minutes discussing your experiences with mathematics; what you like, what you don't like, and what you'd like to know. No question too big or small.