1

A Gentle Introduction

> Mathematics is the language with which God wrote the universe.[1]
>
> Logic is the foundation of the certainty of all the knowledge we acquire.[2]
>
> Mathematicians are like Frenchmen: whatever you say to them they translate into their own language, and forthwith it is something entirely different.[3]

1.1 What Is This Book About?

As I mentioned in the introduction, this is a book about numbers, but not just the numbers we are familiar with and use every day (such as 1, 2, 3, and fractions like $\frac{1}{2}$ and $\frac{3}{4}$). I want to introduce you to some rather less familiar types of numbers, which, whilst not something we as individuals use every day, nonetheless underlie much of the information and communications technology we take for granted. At the same time I hope to introduce you to mathematics as seen by mathematicians—which may be rather different, and perhaps more abstract, than what you thought of as mathematics at school. I am, of course, assuming that you would not call yourself a professional mathematician, but are nonetheless curious

1 The quote is attributed to Galileo Galilei, 1565–1642.

2 Leonhard Euler, 1707–1783. Note that 'Euler' is pronounced something like 'oiler', and not 'yooler' as I mistakenly thought for many years. Whilst on the subject of mispronounced names of mathematicians, apparently the surname of the great 19th century French mathematician Augustin-Louis Cauchy was routinely mispronounced as 'corky' by many British mathematicians.

3 Johann Wolfgang von Goethe, 1749–1832.

Mathematics for the Digital Systems Engineer: Essentials for Modern Cryptography, Computer Security and Communications Technology, First Edition. Chris J. Mitchell.

about the subject. If, for example, you have a degree in mathematics, you may find pretty much everything I have to say is routine, but I have nonetheless tried to stay on the right side of rigour; that is, I am hoping there are not too many errors!

Why is the title of this book *Mathematics for the digital systems engineer?* Well, as I just mentioned, although the numbers introduced in later chapters are not numbers as we are used to them, they are key to today's information-centric world. I find the whole subject fascinating, as I have since I was introduced to them over 50 years ago as an undergraduate, and I hope to share some of that fascination with you, and give you the tools to engage with the mathematics involved when working with and developing new systems.

This is perhaps not a typical textbook. I have been rather self-indulgent and allowing myself to go down all sorts of rabbit holes. It has made writing this book more fun for me, and I hope you enjoy some of these diversions. Apart from its intended role as an introduction to the mathematics of many modern digital systems, I would also hope that it might be something of interest to anyone thinking of studying mathematics at university, as it attempts to give an introduction to one of the subjects forming a key part of any mathematics degree.

1.2 Mathematics as Mathematicians See It

I would like to start by saying that what professional mathematicians do, i.e. inventing (or perhaps *discovering*) new mathematics, is a creative activity. That is, a goal of mathematical research is to state and prove new theorems or conjectures, or to find examples which disprove conjectures or show that some set of conditions can actually be met. The hope of every working mathematician is always that in some way they can extend the body of mathematics.

At this point, I would like to have a quick gripe about the use of the word 'creative' in relation to certain endeavours, as if no other activities are creative. Indeed, the term creative is often used in connection with the so-called 'creative arts', as if to deliberately exclude any other pursuits, including science and mathematics. I would like to make the case that while undoubtedly painting a picture or composing a piece of music are creative acts, so is the development of a new piece of mathematics.

However, at this point I want to slightly contradict myself! Whilst mathematics is clearly something completely created by humanity, it is important to observe that many—if not most—mathematicians, consciously or unconsciously, talk about mathematics as if it has always existed. For example, they (and that includes me) will talk about 'discovering' a theorem rather than creating a theorem. Indeed, there is a widespread—often unarticulated—feeling that the ideas we develop and the theorems we prove always existed, and it was just up to us to *discover* them. This has been discussed at length by many previous authors, and *platonism* is the

name usually given to this belief (although this is not the same as Platonism, the philosophy of Plato). Of course, when questioned about this seemingly bizarre belief, many mathematicians will row back somewhat and avow that mathematics is something created, but this doesn't quite explain why talk of *discovering* a theorem or proof is so widespread in the mathematics community.

The fact that mathematics *works*, i.e. that it is not self-contradictory and is so effective when applied to a very wide range of activities, adds to the feeling that it is universal and has an existence beyond its human creators.[4] Indeed, perhaps mathematics is a universal language and something that we might share with intelligent aliens—in the unlikely event we encounter any, even if they exist somewhere in the universe. Anyway, books could be written on this topic, and no doubt have, but I need to move on.

One of the great powers of mathematics comes from abstraction and generalisation. By abstraction I mean taking something concrete, for example addition over the whole numbers (the *integers*[5]), and abstracting away to the notion of a set of objects with an operation on them that has certain properties. In this example, if the set of properties is sufficiently rich, we can prove general results about the behaviour of our abstract objects, which will hold not only for the integers but also for any other structures which have the specified properties. Similarly, generalisation involves looking at a result in mathematics that holds in certain conditions and wondering whether somewhat similar results can be obtained if the conditions are weakened in some way, making the result more *general*, again in the hope that the result can then be applied elsewhere.

This leads us to what I regard as the core of mathematics: theorems and proofs. Theorems and lemmas are simply statements about the properties of certain mathematical structures. To be a theorem, and not simply a conjecture, there must be an accompanying proof. The proof will establish the truth of the theorem in a logical way. Until a theorem has a complete proof, it will not be accepted as true by the mathematical community.

So why do we need proofs? Many of the theorems and lemmas I will state and prove in this book will seem obvious or intuitively true. However, while mathematical intuition is hugely important, it is also very unreliable, and 'obvious' results can turn out to be false. This is the main reason so much effort was made

4 The observation that mathematics is surprisingly effective is something that has been widely explored; in particular, much discussion was triggered by the 1960 article 'The Unreasonable Effectiveness of Mathematics in the Natural Sciences' by the physicist Eugene Wigner.

5 At this point I can hear you, as a reader, complaining that the integers are not concrete—they are themselves an abstraction from the notion of a number of physical objects. However, from a mathematical point of view, the integers are concrete objects.

by mathematicians during the 18th and 19th centuries to put the mathematics of real and complex numbers on a firm foundation.

Of course, at this point I have to concede that we are all fallible; indeed, in the words of Alexander Pope 'To err is human'.[6] This means that sometimes 'proofs' are found to be erroneous—indeed, this can apply to proofs published in reputable journals, and errors may go undiscovered for quite a period of time.

The Four Colour Theorem (the '4CT') is a famous case in point.[7] This theorem asserts that any map on a plane (i.e. a flat surface) can be coloured in at most four colours, where countries sharing a border must not be coloured the same, although countries meeting in just a point can have the same colour. To make the problem well-defined, it is also necessary to disallow countries which consist of more than one disconnected region—this means that the result is not applicable to real world maps since there are many examples of countries which disobey this rule and which possess *exclaves*. For example, the Netherlands municipality of Baarle-Nassau is home to nearly 30 Belgian enclaves, known collectively as Baarle-Hertog, some of which possess Dutch nuclei, and the Musandam Governorate is part of Oman, despite being surrounded by the UAE (strictly speaking, the latter is only a semi-exclave, since it is bordered by a coastline with access to international waters). An example of nested exclaves is provided by Madha in Oman that it is completely surrounded by the UAE; moreover, the village of Nahwa that forms part of the UAE is completely surrounded by Madha.

A 'proof' of this theorem was published by Alfred Kempe in 1879, which was accepted at the time. However, in 1890, a flaw was found by Percy Heawood. So, at that point, it became the 'Four Colour Conjecture', and it didn't become the 4CT again until 1976, when Kenneth Appel and Wolfgang Haken announced their computer-assisted proof. They needed the assistance of a computer to complete their proof to check through 1482 individual cases.

Generally speaking, we rely on other mathematicians checking proofs during the refereeing process for a paper when submitted for publication, although this is clearly not infallible. More generally, any major results are likely to be looked at and checked by many mathematicians. Interestingly, the fact that the 1976 proof of the 4CT was computer-assisted made human checking almost impossible, and this gave rise to a lot of controversy at the time, although the result is today well-accepted. Indeed, a second proof has since been published (or should I say discovered?)—requiring a substantially smaller number of cases to be checked by computer.

---

6 The quote comes from the poem 'An Essay on Criticism'.

7 Readers interested in a detailed account of the history of the problem are referred to the excellent book *Four Colours Suffice*, by Robin Wilson (Penguin Books, 2002).

1.3 Theorems and Proofs

So if theorems are the bread and butter of mathematics, what do they look like? The statement of a theorem will typically start by saying what mathematical object (or objects) it applies to, such as the integers. It may also say what additional restrictions are in place—for example, a theorem about integers might only apply to certain specific integers, for example, prime numbers. The theorem will then list one or more statements which are true about the objects under consideration. This is all rather abstract, but the details of the statement of a theorem will vary a great deal depending on the subject matter.

A simple example is what is known as *Pythagoras' Theorem*, although, as is often the case for named theorems, Pythagoras was not the first to know about it. Apparently, the Pythagorean rule was well-known over a 1000 years before Pythagoras was born. The theorem goes as follows.

Theorem 1.1 (Pythagoras' Theorem) *For a planar right-angled triangle, with sides adjacent to the right angle having lengths a and b, and the other side (the hypotenuse) having length h:*

$$h^2 = a^2 + b^2.$$

The statement of this theorem first mentions the objects it is dealing with, namely right-angled triangles (embedded in the plane), and then sets up some notation. The theorem ends up with the main statement, i.e. that the length of the hypotenuse squared is equal to the sum of the squares of the other two sides. Having said that every theorem has a proof, even for this example I feel obliged to give one (I have chosen one of the many ways that are known for proving this theorem).

Proof. Take four copies of the triangle and arrange them in a square shape with rotational symmetry, so that:

- the right angles of the triangles form the corners of the square, and
- part-way along each side of the square, two copies of the triangle meet in a point.

The arrangement is shown in Figure 1.1.

As shown in the figure, this results in a square whose sides have length $a + b$. Moreover, there is a square in the middle with side length h. The areas of the four triangles and the smaller square in the middle (with side length h), when added together, must have the same area as the large square. Now the area of the large square is $(a + b)^2$ and the area of the smaller square is h^2. Each of the four triangles has area $\frac{ab}{2}$, since two such triangles can be joined together to make a rectangle with side lengths a and b.

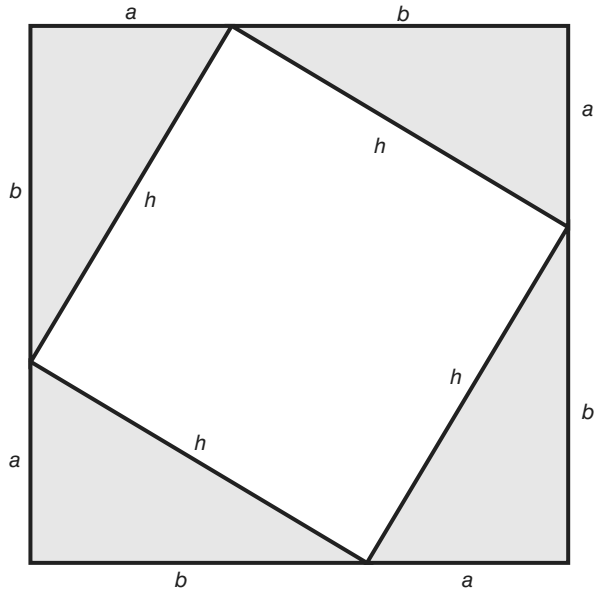

Figure 1.1 Proving Pythagoras' theorem.

Thus:

$$(a+b)^2 = h^2 + 4\frac{ab}{2} = h^2 + 2ab.$$

But $(a+b)^2 = a^2 + 2ab + b^2$, and so

$$a^2 + 2ab + b^2 = h^2 + 2ab;$$

and subtracting $2ab$ from both sides gives the desired result. □

The little box at the end of the previous line (the □ symbol) marks the end of the proof,[8] and can be read as meaning QED (short for the Latin phrase 'quod erat demonstrandum', or 'that which was to be demonstrated'). The alternative interpretation of QED as 'quite enough done' sums the idea up pretty well.

So now we know what a theorem looks like, I should also try to explain the difference between a theorem and a lemma. At the most fundamental level there is no difference, since they both involve the statement of the truth of something and must both be accompanied by a proof. It is often a matter of authorial taste as to what gets labelled a lemma and what gets called a theorem. In general, a theorem

8 This symbol, sometimes known as a *tombstone*, was apparently first used in this way in 1950 by the renowned Hungarian-born American mathematician Paul Halmos.

is something that is worth stating for its own sake, whereas a lemma is typically something whose only purpose is to help prove other lemmas or theorems.

When stating a theorem, or indeed within a proof of a theorem, it is common practice to use the phrase 'if and only if' (sometimes abbreviated to 'iff') for a two-way implication. The usual context is something like 'A is true if and only if B is true'. The meaning is simply that if A is true, then B is true *and also* that if B is true, then A is true.

There are a number of ways a theorem can be proved; indeed, logicians interested in the underlying rules of mathematics have identified a range of different *types* of proof. Probably the most obvious method of proof is to give a logical progression of arguments starting from the assumptions of the theorem that lead to the desired result. The proof I gave of Pythagoras theorem is of this type.

Another very common type of argument is the *reductio ad absurdum*. This involves assuming the opposite of the result that is to be proved, and then following a series of logical deductions to result in something that cannot possibly be true—typically something that contradicts one of the assumptions underlying the theorem. As a result, the statement must be true (since, if it cannot be false, it must be true[9]). Such proofs are often referred to as 'proofs by contradiction'.

A third form of argument, known as *induction*, only applies to theorems which assert the truth of a statement $\mathcal{P}(n)$ that is a function of the positive integer n. This method of proof is explained in Chapter 3.

Finally, for theorems that state that a certain mathematical object exists, for example, a theorem stating that there exist Pythagorean triples, i.e. sets of three integers a, b and h such that $a^2 + b^2 = h^2$, a proof need only exhibit an example. For the case of the Pythagorean triples theorem, it is simply necessary to give an example of a suitable triple (a, b, h), e.g. $(3, 4, 5)$.

There are other methods of proving theorems, for example, giving counterexamples to show that a claim is untrue, but the preceding list is good enough for this book.

Before moving on, I feel I should explain a little about how one might come up with a proof for a theorem. Often, when trying to prove a result, I might wrestle with the problem in my head until I think I have something—perhaps not a full proof, but maybe a step in the right direction. The acid test then comes when I try to write it down; sadly, often when I come to write down an idea, I find it doesn't actually work because I missed some point that only came to light when I worked through the details. So writing things down is an absolutely key step; actually, I believe this principle applies in many areas of study, not just in mathematics.

When I have actually proved something, the first proof I come up with may be rather long and inelegant—indeed, it may still be wrong, as even writing things

9 More formally, this principle is known as the *Law of Excluded Middle*, and the idea in some form goes back to Aristotle.

down doesn't always catch subtle errors! The final proof that gets included in an academic paper or book is typically not the first proof that was found but a polished version, i.e. a version which is as succinct and neat as possible. Indeed, for well-known theorems, such as those included in this book, the mathematical community has often had hundreds, if not thousands, of years to come up with the cleverest, most elegant proof.

This is not something that is always discussed as much as it should be; you could even call it a slightly dirty secret of mathematicians. When you are presented with a proof as a student of mathematics, they always look so much better than any proof you might come up with when asked to devise one for 'homework' or as an exercise. But, of course, you shouldn't feel bad about this—the same is true for most of us! So, when at various points in this book I ask you to try to prove a result, don't be disheartened if your proof is rather less elegant or simple than the one I present at the end of the book; remember that I've cheated and used a wide variety of established sources to find the proof I like best. Also, don't worry if your proof is different to mine; there are many different and equally valid ways to establish the same result.

In fact, finding a proof is often much simpler than working out what the theorem is that one should try to prove.[10] Often this involves looking at small examples or doing mathematical experiments. Of course, these days we can use computers to look at much larger examples than can be examined using pencil and paper, although there is something very illuminating in playing with examples 'by hand'. These examples and experiments may give you an idea of what might be true in general, which perhaps you can then try to establish through proof. But it is often a very messy and laborious process—unlike the polished theorem and proof you might see written down in a textbook.[11]

1.4 Abstract Algebra

I already briefly talked about the ideas of abstraction and generalisation. Abstract algebra, as its name suggests, is about investigating what happens when you throw away your preconceptions about what adding or multiplying together two objects means. The objects involved can be drawn from a finite or an infinite set, and the objects could be the familiar numbers, geometric objects, or more complex entities such as polynomials. As you proceed through the book, you will encounter a wide range of examples.

---

10 The following quote is attributed to the great German mathematician Bernhard Riemann: 'If only I had the theorems! Then I should find the proofs easily enough.'

11 As Paul Halmos wrote 'Mathematics is not a deductive science—that's a cliché. When you try to prove a theorem, you don't just list the hypotheses, and then start to reason. What you do is trial and error, experimentation, guesswork.'

At the end of the next chapter, I will give some examples of abstract properties we might expect for our generalised notions of addition or multiplication (or *operations* as I call them). The list given there is by no means complete, and again later in the book, I will introduce other examples, notably one governing how two operations interact.

1.5 What Do You Need to Know to Make Sense of This Book?

This is, of course, a question of key importance. I am assuming a very basic familiarity with mathematics that you may have learnt at school, in particular elementary ideas about whole numbers, addition, multiplication, subtraction, and exponentiation (i.e. raising one number to the power of another).

However, in other respects, I have tried to make the book self-contained, and where I introduce new ideas I give examples which I hope you will find helpful. At times I suspect you may need to take a little time to play with a concept before proceeding, and I would recommend working out your own examples using pen and paper if you have trouble understanding something. Don't worry, every mathematician (except perhaps those with planetary-sized intelligences) needs to scribble on a piece of paper every now and then to try to grasp a concept or check that something really does work.

1.6 Case Studies of Applications

I already mentioned in the preface that the main goal of this book is to provide an accessible introduction to topics such as groups and finite fields for the benefit of engineers and computer scientists who find themselves working on applications of these mathematical notions. Of course, I am hoping that the book will also be of interest to anyone interested in how key modern digital technologies actually work. The main emphasis is on the mathematics itself; however, in every chapter after the opening three introductory chapters, I provide case studies of applications of the mathematics that has been introduced. The main focus is on topics I am personally familiar with, in particular relating to the protection of digital communications, but I have tried to cover a range of application domains. The main case study topics are summarised as follows.

- In Chapter 4, I introduce two quite different applications, both resting on modular arithmetic. The first is the notion of check digits, as used to protect the integrity of numerical data against accidental errors. Second, I introduce both the idea of public key cryptography and the RSA public key cryptosystem, whose security rests on the difficulty of finding the prime factors of very

large numbers; RSA has been very widely used in many domains, notably in the protection of credit card transactions.

- Chapter 5, which introduces groups, contains a description of a different type of public key cryptosystem, namely Diffie–Hellman key exchange, which has been widely implemented using a range of underlying groups.
- In Chapter 6, where I introduce rings, I return to the problem of protecting data integrity and introduce a technology of huge practical importance over the last 60 years or so, namely block codes. I also introduce one key example of a class of block code, namely the Hamming codes.
- Chapter 7 introduces the idea of polynomials as mathematical objects in their own right. The key application explored here is the linear feedback shift register and the sequences such registers generate. Shift register sequences have been widely used in a range of digital communication domains, including in cryptography and spread spectrum communications.
- The summit of the mathematical mountain is reached in Chapter 8, where I introduce finite fields. The case study application described here again relates to cryptography, and in particular to what is known as elliptic curve cryptography.
- The last main chapter, Chapter 9, is a little different from the rest of the book in that it consists of a compendium of topics I could not bear to leave out. Whilst some of the topics do not have immediate application, the introduction to difference sets and finite geometry touches on a range of topics of practical interest, including cryptography, although the applications are not explored here.