The Bright Child's Book of Numbers

William C. Schulz

March 18, 2020

Transgalactic Publishing Company
Flagstaff, Vienna, Cosmopolis

© 2018 by William C. Schulz

Every creator painfully experiences the chasm between his inner vision and its ultimate expression. Isaac Bashevis Singer

To my daughters Alexia and Danae and my grandchildren Jack and Klara: all a source of interest, pride and joy.

Contents

Introduction

WHO IS THIS BOOK FOR?

This is the little book that I wanted to read when I was a child. It is written for the reasonably bright child who is also curious. Such a child sits in school and listens to the teacher present the rules of numbers, without much effort to explain where the numbers come from and why the rules work. This is part of the reason for this book. I explain what kinds of numbers there are, how they came to be, what stupid ideas were prominent in history about them and how some of the ideas continue to plague teaching and learning. I write in what I hope is a colloquial style that a reasonably bright child can cope with and enjoy. This is a math book and it takes some thinking and some effort, so in addition to intelligence and curiosity the child must be willing to put forth a little effort, but I hope I am not going to strain anyone.

If you are a girl, people may discourage you from reading this book. These are people who think women should spend their lives washing dishes, canning fruit and changing diapers. There is nothing wrong with these activities but this book will give you a glimpse of ways to enrich your life with mathematics or some other science, and whatever you learn some of it will get passed on to your children. Domestic chores are necessary but the good life has room for other things, and if this book encourages you to seek them out one of my goals will have been accomplished.

This book in its initial stages talks about things that are well known to most kids by the 6th grade; whole numbers, signed whole numbers (integers), fractions and decimals (real numbers). I also talk about complex numbers, or to use the common slur, imaginary numbers. These are important for many purposes nowadays in science and technology, but there is another reason. From experience I know that past the age of 14 it is difficult psychologically to introduce new numbers. The experience can be frightening and troublesome for some. A child who accepted fractions and decimals without the least qualm, on encountering complex numbers at age 16, may be initially tripped up, and some students may be able to operate with them while simultaneously maintaining a disbelief in their existence. A few years using them will blunt this effect but we can much more easily ward it off by simply teaching them earlier along with

their representation in the plane. I learned them at about 12 and they seemed to me a natural extension of the real numbers. No mental trauma at all.

This book is very systematic. We first introduce the various numbers in an algebraic context, with virtually no geometry. Later when the algebra is comfortable we then explain numbers in a geometric context, which is much more difficult than it at first appears. The later sections talk about various ideas of great interest and little difficulty that are not now a part of the ordinary curriculum. Many of these ideas have some uses and show there is much fun in mathematics beyond the stuff in the standard curriculum. The hope is that they may encourage students to learn more about these ideas when they become available and of course I hope to suggest to some of the children to later consider careers in mathematics and science or related areas.

Some of this book is challenging but the book need not be read in order; I have been careful about this. If the reader gets to a point where she is no longer interested, skip to another part. It will be perfectly comprehensible at the beginning. Not everyone needs to know everything in the book and it's better to skip to another part than to flounder in things no longer interesting. We are all different and things that greatly interest some people will have zero or less) interest for others. Make the book your own by reading the parts of it you like. You can always come back later when you may want to take another look.

There are many fine books which present mathematics in a cultural context, as a part of general education. Some of these books present a good deal of philosophy along with the mathematics. This is not one of those books. This book is about mathematics itself. It may be elementary mathematics, (though some would claim I go beyond elementary things,) but it is still mathematics. I do from time to time present some of the history of the mathematics being discussed, but this is not an end in itself; it is for amusement and it is not systematic. For a systematic cultural and philosophical perspective on numbers I suggest the book WHAT IS NUMBER by Robert Tubbs.

Let me say a few words about what this book is *not*. It is not a book that will help you do your taxes or realign your sewer pipe. It will not get you to Calculus any faster; the book is virtually Calculus free. However, I can promise you that if you read this book you will be in a lot better position than many other students when you get to college math because you will have learned some mathematical ways of thinking and will have seen some things already that will be new to your fellow students. You will occasionally recognize old mathematical friends and you will not be nearly as prone to misunderstand what the teacher is telling you. And if I am at all successful in the project, your math classes will seem like friendly places rather than hostile territory. Of course not everything in a math course can be fun, but maybe you will have learned that if you persevere through the less amusing stuff you will then be in a position where math again becomes fun and you'll enjoy sense of achievement that comes with being able to do difficult things that have frightened off many an otherwise brave-hearted adventurer. Also, knowledge, like magic tricks or a never-fail ambrosia salad, is handy at parties and eventually generates a larger

paycheck.

One last word. Although this book does not cover any statistics and probability, never miss a chance to learn them. They are handy in every walk of life. And a note of caution: take them in the *math/stat department*; don't be fooled by classes with names like Economic Statistics, which will give you all of the fog and little of the clarity of the subject, and thus much less of the utility.

A NOTE ON PROBLEMS

This book contains a fair number of problems. Problems are keyed to section numbers; the problems for section 4.2 are found at the end of Chapter 4 and labled Sec 4.2. A few sections have no problems.

You are encouraged to do as many of the problems as you can. Mathematics is not really a spectator sport and you cannot get the full benefit of studying if you do not do at least some problems. The benefits of solving the problems are many; just solving a few problems gives you a feeling for the subject that can't be gained by just reading. The investment of time you make in doing problems pays big dividends in understanding. Also, the author hopes you will have fun doing the problems.

Chapter 1

NUMBERS WITHOUT GEOMETRY

1.1 Introduction

This[1] book is about numbers. Naturally it starts with the natural numbers $1, 2, 3, \ldots$. Natural numbers have a variety of names, whole numbers, counting numbers, positive integers, etc. Each name emphasizes a slightly different aspect. And, as we will find several times in our story, some of the names merely reflect prejudices that at one time or another people have had. Here, natural numbers reflects a prejudice that these numbers are somehow more "natural" than other numbers. Reflecting this, there is a famous quote from the mathematician Kronecker saying "The natural numbers are the work of God; all else is the work of man." Although philosophers tend to see it this way, it is a concept hard to defend rationally. We will point out these odd names at the appropriate times and discuss why the numbers are so called. In the present day, these names are sometimes misleading, but some of them we are probably stuck with. Others can perhaps be annihilated and I will do my best. Some names, while annoying in English, actually make sense if you know what sense of the word is meant, irrational number means number that is NOT a ratio, rather than number that makes bad decisions.

In this book we will introduce many types of numbers. We begin with the positive integers $\{1, 2, 3, ...\}$ and then the negative and rational numbers (fractions and the decimals that repeat: .142857142857142857.....). These do not suffice even for elementary purposes; if a circle has radius 1 then its circumference has length 2π which is not a rational number (that is, it is not a fraction. Hence we move along to the so-called real numbers, and this is where the average person stops. With the real numbers, that is all decimal numbers like 32.31784567753...., (code letter $\mathbb{R}$), the average person can count and measure everything she can see in the world around her. Even at this point though we can see how much richer mathematics is with $\mathbb{R}$ than with just the positive integers $\{1, 2, 3, ...\}$. The world's greatest logician, Kurt Gödel, (Austrian, 1906-1978), once remarked that each time the concept of number was generalized mathematics was greatly enriched. Hopefully you will see that in this book.

Returning to the natural numbers, some are more natural than others. We know this because only relatively recently (last 2500 years) have we had methods to express large numbers. Many popular works cite an aboriginal tribe (never explicitly identified) which got along fine with $1, 2, 3,$ many. Now that life has become more complicated we need more numbers than this. The ancestor of most of the languages of Western Eurasia, called Indo-European, had names for numbers up to a thousand, which then could be compounded to make numbers up to 999,999 after which a new word, million, is necessary but was not in the language[2]. Since the speakers of Indo-European were mostly herdsman, it was not necessary to go much over ten or twenty thousand. Eventually we came to realize that natural numbers naturally go on forever. The first person to

[1] Start 14 Jan 2018, current March 18, 2020

[2] This explains why the numbers in most of the European languges vaguely resemble one another; the words were descended from a common source, Indo-European

come up with an efficient means to write really large numbers was Archimedes (approximately 287-212 BCE), one of the greatest mathematicians in human history, but we will not go into the details of his methods.

The natural numbers have a grip on the human mind which seems qualitatively different from the other numbers (like fractions or decimals) that we will look at. For example only the natural numbers have been used in fortune telling or numerology. An example of numerology is assigning numbers to letters in some way (very natural for Greeks) and then, say, adding up the values of the letters in your name to tell your future. The value society places on this activity has varied considerably with time, and is now at low ebb, but, like Astrology, the interest never completely disappears.

Because of this mysterious grip on the human mind, properties of the natural numbers have been studied far beyond any practical need, and this subject, called number theory, has been vastly developed, often with surprising tools. For example, any even number greater than 4 seems to be the sum of two prime numbers, ($6 = 3 + 3$, $8 = 5 + 3$, $10 = 5 + 5$, $12 = 7 + 5$, etc.). We have never been able to prove this although we have been trying for almost 300 years. We will look at number theory in some detail in Chapter 6.

1.2 Exponents

We will begin with something easy partly because we will need the material later and partly because there is some interesting methodology involved. This means exponents will give us some insight into how mathematicians think. Note: for this section a should be thought of as a *positive* integer but in fact the methods will work for any positive real number. (We will discuss real numbers in chapter 2)[3]

If a is the length of the side of a square, then the area is $aa = a^2$. (This is the origin of the term "squared".) For a cube with side a the volume is $aaa = a^3$, pronounced a cubed. Notation like a^2 and a^3 was invented in the 1400s to simplify algebraic expressions. Now note that

$$a^2 \cdot a^3 = aa \cdot aaa = a^{2+3} = a^5$$

so the rule is to multiply expressions with exponents (and the same base a) add the exponents; this can be expressed more generally as

$$a^m \cdot a^n = a^{m+n} \qquad m, n \text{ positive integers}$$

Since, using this rule, $a^1 a^2 = a^{1+2} = a^3$ we see that $a^1 = a$ which is completely reasonable. Similarly it is easy to show that

$$(a^m)^n = a^{mn}$$

Just write out $(a^2)^3 = (aa)(aa)(aa)$ and you can see why.

[3]The name *real* number is one of those historically odd names for numbers but there is no hope of getting rid of it and perhaps it's not that bad.

Now we would like to know what a^0 means. From our original viewpoint it makes no sense. Many people's intuition suggests that it ought to be 0 but in mathematics one must be a little careful of jumping to the "obvious" answer too quickly. One should at least have some sort of reason rather than just gut feeling. Gut feeling is important but not reliable. One of the most reliable methods to answer questions of this sort is to try to preserve some law or calculational rule. This is usually a good guide. So let us take the law $a^m \cdot a^n = a^{m+n}$ and use it with $m = 0$ and $n = 1$. Then we have

$$a^0 a^1 = a^{0+1} = a^1 = 1 \cdot a^1$$

Now, *provided* $a^1 = a$ *is not* 0 we can divide it out and get

$$a^0 = 1 \qquad \text{for } a \neq 0$$

and that solves our problem. This leaves us with a slight problem. $0^m = 0$ for positive integer m, and $a^0 = 1$ for positive a. This leaves 0^0 in a kind of limbo. The sad fact is that there is no *possible* definition for 0^0 that preserves all the laws and so we must leave it undefined, which simply means that we can't use it. Which ever way we define it, some law will fail, and since in mathematics we are very law abiding we just declare 0^0 to have no usable value and forbid its use. WARNING: Various calculators and computer algebra packages[4] *may* give a value for 0^0. This is just to simplify the programming and *should not be taken as an indication that in mathematics* 0^0 *has a value.*

Well that was easy. Now what does a^{-1} mean. Well now we know the trick.

$$a^{-1}a = a^{-1}a^1 = a^{(-1)+1} = a^0 = 1 \qquad \text{for } a \neq 0$$

so we have

$$a^{-1} = \frac{1}{a}$$

and then

$$a^{-m} = a^{m\cdot(-1)} = (a^m)^{-1} = \frac{1}{a^m}$$

using a different law for a change.

Finally we want $a^{\frac{1}{3}}$. This is pretty easy; use the law $(a^m)^n = a^{mn}$

$$(a^{\frac{1}{3}})^3 = a^{\frac{1}{3}\cdot 3} = a^1 = a$$

Since the cube of $a^{\frac{1}{3}}$ is a, $a^{\frac{1}{3}}$ must be the cube root of a. Similarly $a^{\frac{1}{2}}$ is the square root of a and more generally

$$a^{\frac{m}{n}} = (\sqrt[n]{a})^m = \sqrt[n]{a^m}$$

[4] A computer algebra package is a program for use on a computer that does algebra. Mathematica© and Matlab© are popular and there are many more.

This leaves open the question of a^π where $\pi = 3.1415926535\ldots$ but we could get an approximation (indicated by $\approx$) by using $\pi \approx 22/7$ and so

$$a^\pi \approx (\sqrt[7]{a})^{22}$$

To do better one could use a better approximation for π, but this is a very clumsy method. Far better is to use logarithms, which sadly is not one of our topics. A scientific calculator can do these things (it uses logarithms) with no fuss at all.

1.3. A Little Algebra

Our basic plan is first to study numbers from an algebraic point of view. This point of view starts with the natural numbers $\{1, 2, 3, \ldots\}$ and looks at what numbers need to be created (or discovered, depending on your philosophical point of view) in order to create a system in which all algebraic processes are possible. So we must take a few moments to ask what *are* the algebraic processes and what are the basic laws governing them. From the modern point of view, it is essential to single out some subset of the laws of algebra which we regard as *basic*, and then show how the other laws can be derived from the basic ones.

How do we describe the basic laws. The Babylonians, the first culture that worried about such things, did it by example. The teacher writes on the big clay tablet in front of the class (using a dot for multiplication)

$$2 \cdot 3 = 3 \cdot 2$$

and tells the class that the order doesn't matter; 2 can come first and then 3, or 3 can come first and then 2; the answer is the same, 6. The students copy it all on their small clay tablets. The teacher then asks, is their anything special about 2 and 3 or would this work for any two numbers? The students say, nothing special; it'll always work. Good, says the teacher; you have now learned the commutative law of multiplication. This is teaching by example and it is a great method, which we use very often today. It is sometimes called student see, student do. But there are drawbacks.

If we wish to derive more complicated laws from simpler ones, it can become difficult for the novice to keep in mind that the numbers represent *any* numbers. Also, it can get quite confusing. Also confusing is what is called rhetorical algebra, where you say the law in a sentence:

if a first number is multiplied by a second number, the result is the same

as when the second number is multiplied by the first number

Clearly this says the same thing as $2 \cdot 3 = 3 \cdot 2$ but it is cumbersome to say and think about, and also to write. So a process of abbreviation set in, where some words were abbreviated and some words were represented by symbols.

The process was taken up seriously by the medieval 'Arabs[5] and after six or so centuries and a change of country and language we come, about 1500, to

$$r \cdot s = s \cdot r$$

where r and s come to mean any natural numbers you choose them to be and the dot means multiplication. Whatever natural numbers r and s you choose, the equation remains true. The takeaway here is that $r \cdot s = s \cdot r$ is the result of slow substitution of symbols for words and numbers to boil it down to the basic meaning we are trying to capture. The long process was exceedingly cumbersome and Darwinian, in the sense that many things were tried and the things that worked were retained and the others pitched overboard. So there are really two takeaways here

1. Numbers are replaced by letters to emphasize the generality.

2. The notation with its rules suggests what to do; it often thinks for itself.

That is the wonder of algebra. It is remarkable that such an efficient system could be developed by a process of abbreviating fixed phrases in 'Arabic or Italian.

Besides representing any natural number, letters have a second function in algebra which we now discuss. Consider the following two equations

$$\begin{aligned} x + x &= 2x \\ x + 2 &= 5 \end{aligned}$$

The first is true no matter what natural number is substituted for x. The second is true for only *one* number which is 3. We generally express this by $x = 3$. A large part of algebra consists of finding all the numbers which, when substituted for the letter, make the equation true. A little experience makes it possible for a student to tell the two types of equation apart though it can be puzzling for the beginner. Beware!

The equation (x^2 just means $x \cdot x$)

$$x^2 - 5x + 6 = 0$$

is true for exactly two numbers; $x = 2$ and $x = 3$. We say that 2 and 3 *satisfy* the equation or that 2 and 3 are *solutions* of the equation. We will not take up the question of *finding* the solutions at this time.

The solution of equations can be done for two reasons; because it's fun or because it has some real world application. If it is a real world application, then it always comes from a "word problem". Thus $x + 2 = 5$ comes from, or is an abbreviation for: A number is added to 2 and the result is 5; find the number. The solution consists of two steps; 1) Translate the problem into the equation and 2) Solve the equation. Step 1 is often much harder than step 2, but step

[5] the ' before the A in 'Arabic stands for the Semetic letter ayn. It is a sort of weak growl.

1 is what they pay you for. Nowadays step 2 can often be handed over to a machine.

Before attempting to solve an equation it is nice to know if a solution exists. Consider the following two equations, where we are looking for natural numbers in $\{1, 2, 3, 4, \ldots\}$ that solve the equations.

$$\begin{aligned} x + 2 &= 5 \\ x + 2 &= 2 \end{aligned}$$

The first has a solution, 3; the second does not. No natural number can be substituted for x in the second equation and have it come out true. Of course you are thinking that the solution is 0, but whoever set up the system forgot to begin the natural numbers with 0. That was dumb. We all know how important 0 is to get on in life. If your bank balance is 0 things are bad and probably going to get worse. In fact, 0 is so important that we almost surely should put it in with the natural numbers to get a bigger set [6]

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

This set is called the *non-negative integers*. If the name seems unclear it won't be in a few minutes. Note the extra bar in the $\mathbb{N}$. All the letters that refer to fixed kinds of numbers have an extra stroke like this. The natural numbers are then $\mathbb{N}^+ = \{1, 2, 3, \ldots\}$, the $+$ to the upper right in the $\mathbb{N}$ indicating the members of the set are *positive*. Now that we have added 0 to the system let us review why we did it. It was to provide us with a solution to $x + 2 = 2$, because we hate it when an equation has no solution. There is nothing special about 2 here. In fact, we require

$$x + 0 = x$$

for any x in $\mathbb{N}$. This is the way we express the general law: if 0 is added to any non-negative integer the result is the non-negative integer. You can usually tell the general laws from the equations pleading for solutions by the phrase "for any x in $\mathbb{N}$". Such "for any" or "for all" phrases are characteristic of general laws. And *you* should use them too.

Did we invent 0 or did we discover it? Do we have a right to throw in new numbers when we think we need them? Does 0 have a similar meaning to the other numbers? The first two questions you can ask your philosophy professor. No two will give the same answers. Make sure you have time to listen to their complete answers which may take an hour or two. But for the third question, think of 2 as describing the contents of a box with one red and one blue marble. Then 0 describes the contents of an empty box. That seems fair, doesn't it?

Now that we have disposed of $x + 2 = 2$ with solution 0, it seems natural to continue on with $x + 2 = 0$. There is of course no solution in $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$. However, once again real life rather demands a solution. Here the real life situation is assets and debts. Even the ancient Babylonians, enthusiastic bankers,

[6] In fact, some mathematicians do put 0 at the beginning of the natural numbers.

had to come up with a solution. In medieval Italy, the bankers wrote the debts in red ink and the assets in black ink. We still say that a person with debts is "in the red". The mathematicians eventually discovered that a good description of the solution of $x + 2 = 0$ is $x = -2$. So to make the bankers happy we must add a $-x$ to the system for each x in $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$ and to keep the system working we must have $0 = -0$ so the system becomes

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

Here the letter $\mathbb{Z}$ stands for the German word Zahl because the system was set up by Germans and Zahl means number. In English this system is called the *Integers*, and we have positive integers, negative integers, non-negative integers, etc.

Now a wonderful thing happens. Absolutely any equation

$$x + a = b$$

where a and b are any Integers has a solution which is an integer and moreover the solution is unique; only one integer x is a solution to the equation, no matter what you choose for a and b. For example $x+17 = -12$ has the unique solution $x = -29$. (We will discuss how to find the solution later, but you can figure it out now if you think in terms of assets and debts.)

We will now take a step back and look at what we have done from the point of view of a mathematician. We have solved all problems of the form $x + a = b$ where a and b are integers; we know the solution exists, we know that it is unique, and we have a way of finding it; indeed it is $x = b + (-a)$, which we write for convenience as $x = b - a$. What more could one ask? When a mathematician has got all these things under control he is happy, for a while. But notice he has killed off the problem; there is nothing more to do with this problem. He cannot make further progress with it because there is no further progress to make; it done, finished, complete, and of no further interest, except of course to the person who wishes to apply this knowledge to solve applied problems (usually for money). Not our concern.

So the next thing the mathematician must do is see if he can modify the problem to a new problem which has not yet been solved, so he again has something to do. That is the nature of mathematics; when a problem is solved, find a new problem. There is neither fun nor glory, although there may be money, in working with problems that are already solved. So the mathematician says to himself, how can I modify a problem like $x+a = b$ to get a new interesting problem.

This is pretty easy; one puts in a multiplier for x so the new problem is

$$c \cdot x + a = b$$

In practice the dot is often omitted. I will omit the dot in cases where this is the usual mathematical practice but will use it whenever I think it clarifies the

situation. (The great logician Alfred Tarski referred to this as the "Method of systematic forgetting".) So we now write the problem as

$$cx + a = b$$

where, as before, c and x are multiplied. (Algebra is very good at these sorts of abbreviations, since it came from words by abbreviations.) Let us look at an example:

$$3x + 7 = 22$$

We must move the 7 from the left side of the equation to the right. We do this by adding the additive inverse of 7, which is -7 to both sides of the equation getting

$$\begin{aligned} 3x + 7 + (-7) &= 22 + (-7) = 22 - 7 \\ 3x + 0 &= 15 \end{aligned}$$

So when the dust settles we have

$$3x = 15$$

Well if $3x = 15$ then we must have $x = 5$. The exact details of getting the 5 we will discover later. Everything seems to be going well, but appearances are deceiving sometimes and if we modify the problem just a tiny bit we have disaster. Consider

$$3x + 6 = 22$$

As before,

$$3x = 22 - 6 = 16$$

However, we must remember that the numbers we are working with here are *Integers* $\mathbb{Z} = \{\ldots -4, -3, -2, 0, 1, 2, 3, \ldots\}$ and there is no solution. That is, there is no integer which, when multiplied by 5 gives 16. This is often restated as 16 is not divisible by 5. (Divisibility is an interesting and important idea but we don't want to stop here and talk about it. We will examine it later.) Now many of you are are screaming, at least inside your head, that the answer is a fraction. This is correct, but fractions are a new kind of number which we must invent (or discover) so we can solve our new problem.

Note carefully the similarity to what we did for 0 and negative numbers. We had equations we could not solve. We needed new numbers to be solutions and so we wrote them down. The point is that when an equation is not solvable we should consider inventing (or discovering) some new numbers to make it solvable. The new numbers we need now are called *Rational Numbers* because they are *ratios* of integers. Ordinary people call them *fractions*; mathematicians who are seldom ordinary, call them rational numbers, which I again emphasize means that they are ratios of integers. The code letter for the rational numbers is $\mathbb{Q}$, which you can think of standing for Quotient, although it comes from the

long German word Quotientenkörper[7]. Numbers in $\mathbb{Q}$ are written as one integer over another provided the integer on the bottom (the denominator) is not 0.

$$\frac{1}{2} \quad \frac{2}{3} \quad \frac{-7}{16} \quad \frac{7}{14} \quad \frac{20}{5} \quad \frac{0}{7} \quad \frac{8}{1} \quad \frac{4}{6}$$

Notice not all these fractions are different. Do you know how to tell if two fractions are the same? Most people have no idea, but I am going to tell you the secret. The definition is

$$\frac{a}{b} = \frac{c}{d} \quad \text{if and only if} \quad ad = bc$$

For example $\frac{91}{247} = \frac{119}{323}$ because $91 \cdot 323 = 29393 = 247 \cdot 119$. Sometimes teachers will tell you that you can check if fractions are equal by reducing both of them to lowest terms but this is often quite difficult unless you know further secrets. Just try and reduce these two fractions above to lowest terms.

We'll put off adding fractions for awhile, but multiplication is easy. The definition is

$$\frac{a}{b} \cdot \frac{r}{s} = \frac{ar}{bs}$$

Suppose we multiply $\frac{a}{b}$ by $\frac{s}{s}$ where in the second fraction the top and bottom are equal. Then we have

$$\frac{a}{b} \cdot \frac{s}{s} = \frac{as}{bs}$$

Now we check $\frac{a}{b}$ and $\frac{as}{bs}$ for equality. They are equal if and only if $a \cdot bs = b \cdot as$ which is true. Hence for any non-zero s, the fraction $\frac{s}{s}$ does not change a fraction when you multiply by it, and hence it must be 1. Similarly $\frac{2s}{s}$ works like 2 and the same for any integer. Hence the rational numbers $\mathbb{Q}$ contain a copy of the integers. It is standard in mathematics to not make any distinction between objects with identical properties, so that we can write

$$2 = \frac{2s}{s} \quad \text{for any integer} \quad s \neq 0$$

We'll look at one more thing. Suppose $r = da$ and $s = db$. Then

$$\frac{r}{s} = \frac{da}{db} = \frac{d}{d} \cdot \frac{a}{b} = 1 \cdot \frac{a}{b} = \frac{a}{b}$$

and now you know why reducing fractions works. Did you ever wonder why it worked? It's a matter of multiplying by 1 where 1 is written in a clever manner. There are a lot of circumstances where writing 1 in a clever manner is very useful.

Notice that when we multiply a fraction by an integer we use the identification of integer a with $\frac{a}{1}$ so that

$$a \cdot \frac{r}{s} = \frac{a}{1} \cdot \frac{r}{s} = \frac{ar}{1s} = \frac{ar}{s}$$

[7]To make the German sound ö pronounce err like you were a Kennedy.

We are now in a position to solve the equation that started our excursion into rational numbers (fractions). Recall that it was

$$3x = 16$$

and the solution is

$$x = \frac{16}{3}$$

because

$$3 \cdot \frac{16}{3} = \frac{3}{1} \cdot \frac{16}{3} = \frac{16 \cdot 3}{1 \cdot 3} = \frac{16}{1} = 16$$

Naturally we don't go through all these steps when we check our solution but I wrote them all out so you could see what lies behind the calculation. The process is exactly the same for any other equation of this type:

The solution of $ax = b$ for integers b and $a \neq 0$ is $x = \frac{b}{a}$.

In fact we can do better. We can solve equations like this with fraction coefficients:

The solution of $\frac{r}{s} \cdot x = \frac{a}{b}$ for integers a and $r, s, b \neq 0$ is $x = \frac{s}{r} \cdot \frac{a}{b}$.

I will let you check the solution yourself as I did for $3x = 16$. It will give you a chance to practice with letters and show that you have gotten the ideas under control. Algebra requires courage more than anything else. Intelligence is helpful; courage is essential. Write down what you know and try and get to where you want, modeling on what I did. In the end you want to get to $\frac{a}{b}$.

So again we have solved a whole class of equations and we know that if the coefficients are rational numbers so is the solution. We say the rational numbers are *closed* under the solution of such equations. Closure is an important idea and we will have other examples.

A word now about division, which is a concept with fairly limited use in the mathematical world. If we want to divide 18 by 6, it means we must find a number so that

$$6x = 18$$

We know the answer is

$$\frac{18}{6} = \frac{3 \cdot 6}{1 \cdot 6} = \frac{3}{1} = 3$$

Now let's divide the fraction $\frac{r}{s}$ by $\frac{a}{b}$ (where $a, b, s \neq 0$ of course) which means we solve the equation

$$\frac{a}{b} \cdot x = \frac{r}{s}$$

We already know the answer $x = \frac{b}{a}\frac{r}{s}$. How can we express this easily in words. Look at it this way: We want to simplify

$$\frac{\frac{r}{s}}{\frac{a}{b}} \quad \text{to get} \quad \frac{b}{a} \cdot \frac{r}{s}$$

So the rule is to turn the denominator fraction $\frac{a}{b}$ upside down and multiply the numerator fraction $\frac{r}{s}$ by it, or even shorter *to divide fractions invert and*

multiply. Careful to note which one to invert. We have now explained why this old rule, which you may already know, works. In case you forget the rule, there is another method that has advantages. You multiply the *compound fraction* (fraction made of fractions) on top and bottom by something both denominators divide; in this case bs. The little fractions then disappear:

$$\frac{\frac{r}{s}}{\frac{a}{b}} \cdot \frac{bs}{bs} = \frac{rb}{sa}$$

So now we come to the basic laws of algebra. Here we could be talking about integers, but another wonder of algebra is that these same laws are valid over vast areas of mathematics where the letters refer to quite different things than natural numbers. It took time to figure out what laws held for many systems and which were special to a few systems but by the early 1900s we knew this. Here are the most basic laws and their code letters.

RING

$A1$	$a+(b+c)=(a+b)+c$	Associative law	$M1$	$a\cdot(b\cdot c)=(a\cdot b)\cdot c$
$A2$	$a+0=0+a$	Identity law	$M2$	$a\cdot 1=1\cdot a$
$A3$	$a+(-a)=(-a)+a=0$	Inverse law		
$A4$	$a+b=b+a$	Commutative law		
$D1$	$a(b+c)=ab+ac$	Distributive law		
$D2$	$(b+c)a=ba+ca$	Distributive law		

For ease of reference I have abbreviated the above laws. A1 should read: For all a, b, c, $a+(b+c)=(a+b)+c$. A2 should read There exists an element 0 ..... A3 should read For any a there exists a $-a$ ..... Similar comments should be added to the other laws and similarly for future systems.

A system of things in which addition and multiplication are defined and these laws hold is called a *Ring*. You know two rings; the integers $\mathbb{Z}$ and the rational numbers $\mathbb{Q}$. The natural numbers $\{1, 2, 3, \ldots\}$ are not a ring because of A2 and A3. The non-negative numbers $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$ are not a ring because, although A2 is now OK, A3 still doesn't work.

There are two kinds of rings, commutative and non commutative. Note that when we talk about the commutative law it is always the commutative law of *Multiplication*. The commutative law of *Addtion* we always have in a ring. In commutative rings there is an additional requirement, the commutative law of multiplication M4 which is $ab = ba$. You might think that this a a very natural requirement but in fact it is highly restrictive; the vast majority of rings are *not* commutative although many very important rings are, including all the ones you probably know. But remember, in most things in life order matters: put on your socks then your shoes versus putting on your shoes then your socks. If we want to model real world things with our mathematics then we need to have non-commutative multiplication. However, most of the systems in this book (but not matrices) will be commutative. The rules then are

COMMUTATIVE RING

$A1$	$a+(b+c)=(a+b)+c$	Associative law	$M1$	$a\cdot(b\cdot c)=(a\cdot b)\cdot c$
$A2$	$a+0=0+a$	Identity law	$M2$	$a\cdot 1=1\cdot a$
$A3$	$a+(-a)=(-a)+a=0$	Inverse law		
$A4$	$a+b=b+a$	Commutative law	$M4$	$ab=ba$ $\leftarrow$ **new**
$D1$	$a(b+c)=ab+ac$	Distributive law		
$D2$	$(b+c)a=ba+ca$	Distributive law		

The two rings we already discussed, $\mathbb{Z}$ and $\mathbb{Q}$, are both commutative rings, so that all the laws above are true for both rings. Now let us suppose that from these laws we derive other laws, for example $a\cdot 0=0$ for any a in the ring. Then this will be true in any *ring*; we don't have to prove it again and again whenever we find a new ring. It will automatically be true. Once we show the basic laws A1-A4, M1,M2,M4,D1,D2 are true in some system with addition and multiplication, then any law we derive from these will also be true.

Notice that most of the basic laws are things you never think about, like A1 or A4. The ones you need to learn really well are D1 and D2. Since the ring is commutative, we really only need one of D1 and D2, we can get one from the other by use of M4, but I put them both in for a kind of symmetry. Here is the real point though; your grade in your algebra course, when you get to it, will be determined by how well you control D1 and D2. They are the tricky part of algebra. Notice also that all the other laws concern either addition or multiplication, but not both. Only D1 and D2 connect addition and multiplication and thus it is obvious that they must be truly important. So remember, control the distributive laws well, get A. Control them half the time; get C. Fail to control them, FLUNK.

In modern mathematics the basic laws of a commutative ring are called the *Axioms*[8] of a commutative ring. Here is what Axioms[9] do in modern mathematics. The axioms (basic laws) cut out from all possible mathematical systems with addition and multiplication exactly those systems which are commutative rings. It's like cutting out the cows from a herd that have a particular brand. And since $a\cdot 0=0$ can be derived from the basic laws, it will be true in *every* ring; we need never check it once we know the basic laws are true so that the system is indeed a ring. It will automatically be true.

The laws that we derive from the basic laws (axioms) are called *theorems*[10]. Thus $a(b+c)=ab+ac$ is an axiom and $a\cdot 0=0$ is a theorem. The first is part of the *definition* of a commutative ring and the second is a mere consequence. Although it is heavy for the projected audience of this book I am now going to prove that $a\cdot 0=0$ from the basic laws (axioms). (Note that I never use M4 so that the results are true in *any* ring, not just commutative rings.) The following series of steps does the job.

Theorem $a\cdot 0=0$

[8] pronounced ak-see-um

[9] The Greeks, Euclid for example, thought of Axioms as *self-evident truths*. We no longer think this way and we suspect there are no self evident truths except those that don't have any content, like definitions or $a=a$.

[10] pronounced theer-um. The 'o' is very rarely pronounced.

Proof

$$\begin{array}{rcll} a(0+0) & = & a0+a0 & \text{Axiom D1 with } b=c=0 \\ a0 & = & a0+a0 & \text{Since 0+0=0 by A2} \\ -a0+a0 & = & -a0+(a0+a0) & \text{Adding equals to equals, A3} \\ 0 & = & (-a0+a0)+a0 & \text{A3 and A1} \\ 0 & = & 0+a0 & \text{A3} \\ 0 & = & a0 & \text{A2} \end{array}$$

Don't worry if you didn't completely understand it, but if you did understand it and more importantly enjoyed it you may be qualified for a mathematical career. Pure mathematicians spend most of their time thinking up and proving theorems, which by the way are two completely different processes.

Before we move on, let me show you one more little derivation that explains something you think is obvious.

$$a+a+a = 1a+1a+1a = (1+1+1)a = 3a$$

Notice how something like $a+a+a=3a$ is totally dependent on the distributive law. And you thought it was obvious!

Also important is that there is just *one* element that is an additive inverse for a. For if b is an additive inverses for a then $a+b=0$. So we have (name the law for each step as I did above)

$$\begin{array}{rcll} a+b & = & 0 & \text{given} \\ -a+(a+b) & = & -a+0 & \\ (-a+a)+b & = & -a & \\ 0+b & = & -a & \\ b & = & -a & \end{array}$$

Since this result was derived from the ring laws, it will be true in any ring, and thus is true for $\mathbb{Z}$ and $\mathbb{Q}$.

Now watch me use this result to prove

$$-(-a) = a$$

We note that

$$\begin{array}{rcl} -(-a)+(-a) & = & 0 \\ a+(-a) & = & 0 \end{array}$$

But there is only *one* element which, when added to $-a$ gives 0 by the above uniqueness result, so we must have $-(-a)=a$. This is a typical use of a uniqueness result and often gives very quick proofs of results much more complicated to do in other ways.

There are a couple of other pieces of elementary algebra that it is convenient to prove now. For example

$$-a = (-1)\cdot a$$

Again we use the uniqueness result.

$$\begin{aligned} -a + a &= 0 \\ (-1) \cdot a + a &= (-1) \cdot a + 1 \cdot a \\ &= (-1 + 1) \cdot a \\ &= 0 \cdot a \\ &= 0 \end{aligned}$$

using the ring laws and a previous theorem. But by uniqueness only one element when added to a gives 0. Since both $-a$ and $(-1) \cdot a$ have this property, they must be equal. (Notice that the second step is *not* a consequence of the first step. In the second step I start off again. It would be nice if I had warned you I was doing that, but mathematicians rarely are that polite. So it's important to keep reading for a while when you don't understand a step; it may become clearer later. I hope this proof is now clearer to you.)

Next we show the important case of eliminating parentheses. The rule is

$$-(a + b) = -a + (-b) = -a - b$$

Recall that $c - d$ is just an abbreviation for $c + (-d)$. Here's the little proof, where I am leaving out many of the dots when multiplication is obvious:

$$-(a + b) = (-1)(a + b) = (-1)a + (-1)b = -a + (-b) = -a - b$$

Now we want to add fractions. This is so complicated that many adults cannot do it. They have forgotten the critical trick, which is the *common denominator*. So let us do an example that takes a little effort:

$$\frac{8}{15} + \frac{9}{20}$$

The first step is to find some non-negative integer that 15 and 20 both divide. You can always use $15 \cdot 20 = 300$ but often there is a smaller number which makes the calculations easier. In this case it is 60. $60 = 20 \cdot 3$ and $60 = 15 \cdot 4$. Knowing this we have (note which laws are being used)

$$\begin{aligned} \frac{8}{15} + \frac{9}{20} &= \frac{8}{15} \cdot 1 + \frac{9}{20} \cdot 1 \\ &= \frac{8}{15} \cdot \frac{4}{4} + \frac{9}{20} \cdot \frac{3}{3} \\ &= \frac{32}{60} + \frac{27}{60} \\ &= \frac{32}{1} \cdot \frac{1}{60} + \frac{27}{1} \cdot \frac{1}{60} \end{aligned}$$

Note in the second line how I have written 1 in the intelligent manners $\frac{4}{4}$ and $\frac{3}{3}$. Note also that I am striving to get the last line where there is the common

factor in both terms $\frac{1}{60}$. This is the essence of the trick. We can now use the distributive law and a little rewriting to get

$$\begin{aligned} \frac{8}{15}+\frac{9}{20} &= (32+27)\cdot\frac{1}{60} \\ &= 59\cdot\frac{1}{60} = \frac{59}{1}\cdot\frac{1}{60} = \frac{59}{60} \end{aligned}$$

Note that in the calculation I have mucked around with the fractions until I get to a situation where I can add integers. This gets me to the final result after some more fraction manipulation. Naturally I do not want you to go through all these steps when you add fractions. The point is to see why the method works; it consists of setting things up so you can use the distributive law (which the teacher didn't mention, and probably didn't know). The distributive law lies at the root of everything except the most simple minded manipulations of numbers. For example it underlies most simplifications of algebraic expressions. Always keep it in mind.

Notice that the ring laws have nothing to say about cancellation. It is definitely not true in ring or a commutative ring that if you have $ab = ac$ and $a \neq 0$ then you can then say $b = c$. This is called the cancellation property and it fails in most rings, but works in $\mathbb{Z}$ and $\mathbb{Q}$. In the next section of the book we will see rings where it fails. If it is true in a ring, the ring is called an *integral domain* where usually the ring is assumed commutative. You do not have to learn this now.

More important for us is that the ring laws have nothing to say about division. In higher mathematics division is not talked about as much as in elementary mathematics; the idea division is replaced by multiplication by the multiplicative inverse which we now discuss.

In a ring, an element a is called a *unit* if it has a multiplicative inverse. The notation for the multiplicative inverse of $a \neq 0$ is a^{-1}. It looks like and exponent and in some circumstances acts like one, but you should look at it as just notation now. In equations

$$\text{if } a \neq 0 \text{ and } a \text{ is a unit then } a \cdot a^{-1} = a^{-1} \cdot a = 1$$

A ring may have many units or very few. The units of $\mathbb{Z}$ are just 1 and -1, and each is its own inverse. On the other hand in $\mathbb{Q}$ every element that could be a unit is a unit. For example $(\frac{2}{3})^{-1}$ is $\frac{3}{2}$ since

$$\frac{3}{2}\cdot\frac{2}{3} = \frac{6}{6} = 1$$

and the same goes for any other non-zero rational number (fraction). Thus in $\mathbb{Q}$ all the non-zero elements are units. This is an important property and so we have a name, actually two names, for it. We add into the ring axioms one more,

$M3$ if $a \neq 0$ then there is an element a^{-1} so that $a \cdot a^{-1} = a^{-1} \cdot a = 1$

If M3 is true in a commutative ring, then the ring is called a *field*. Rings, even commutative rings, are common; there are lots of them. Fields are scarce and they fall into a few types. The complete set of field laws, where M3 is abbreviated for lack of space, is:

FIELD

$A1$	$a+(b+c)=(a+b)+c$	Associative law	$M1$	$a\cdot(b\cdot c)=(a\cdot b)\cdot c$
$A2$	$a+0=0+a$	Identity law	$M2$	$a\cdot 1=1\cdot a$
$A3$	$a+(-a)=(-a)+a=0$	Inverse law	$M3$	if $a\neq 0$ then $a\cdot a^{-1}=a^{-1}\cdot a=1$
$A4$	$a+b=b+a$	Commutative law	$M4$	$ab=ba$
$D1$	$a(b+c)=ab+ac$	Distributive law		
$D2$	$(b+c)a=ba+ca$	Distributive law		

Since we have M4 in a field, only one of the distributive laws is actually necessary. However, I keep both to contrast with the following

DIVISION RING

$A1$	$a+(b+c)=(a+b)+c$	Associative law	$M1$	$a\cdot(b\cdot c)=(a\cdot b)\cdot c$
$A2$	$a+0=0+a$	Identity law	$M2$	$a\cdot 1=1\cdot a$
$A3$	$a+(-a)=(-a)+a=0$	Inverse law	$M3$	if $a\neq 0$ then $a\cdot a^{-1}=a^{-1}\cdot a=1$
$A4$	$a+b=b+a$	Commutative law		
$D1$	$a(b+c)=ab+ac$	Distributive law		
$D2$	$(b+c)a=ba+ca$	Distributive law		

Note that M4 is gone; commutativity is not assumed. We will meet a very important division ring later, the Quaternions. Division rings are still not common but there are many more of them than there are fields. Even without commutativity their theory is still pretty civilized, in contrast to rings or commutative rings which can be very wild in their behavior. Division rings which are definitely *not* commutative are sometimes called *Skew Fields*. We will not use this term.

In both fields and division rings every non-zero element is a unit.

Also note that every field is a division ring because the lack of M4 does *not* mean that $ab\neq ba$, it means ab does not have to be ba although it might be. Try and get this subtle difference clear in your head because almost all mathematics works this way. We have found that trying to do it any other way makes things *worse*. (Notice that the skew field is an exception to the usual practice.)

We will now prove that in a division ring (and therefore in a field) we have cancellation. Let $ab=ac$ and $a\neq 0$. Then (try to give reasons from the axioms of a division ring)

$$\begin{aligned} ab &= ac && \text{given}\\ a^{-1}(ab) &= a^{-1}(ac) && \text{remember } a\neq 0\\ (a^{-1}a)b &= (a^{-1}a)c\\ 1\cdot b &= 1\cdot c\\ b &= c \end{aligned}$$

The last piece of high school algebra that we will prove is that in a division ring if $ab=0$ then $a=0$ or $b=0$. This is important in solving equations. It is

more convenient to do this in the logically equivalent form

$$\text{if } a \neq 0 \text{ and } ab = 0 \text{ then } b = 0$$

We use the cancellation law we just proved.

$$\begin{aligned} ab &= 0 \\ ab &= a \cdot 0 \quad \text{old theorem } a \cdot 0 = 0 \\ b &= 0 \qquad \text{by cancellation} \end{aligned}$$

I want to mention a big difference between Division Rings and Fields. In a division ring, an equation like $x^2 - 1 = 0$ may have any number, even infinitely many, solutions. We will see this with Quaternions. But in a field, where multiplication is commutative, $x^2 - 1 = 0$ may have no more than 2 solutions, $x^3 - x - 1 = 0$ may have no more than 3 solutions, and so on. The degree of the equation gives the upper limit on the number of solutions. So you see, the commutative law is a powerful restraint.

We now have most of the theoretical development of the first year of high school algebra, and all you now need to pass the course is 50 hours of practice. High school algebra is essentially the algebra of fields.

In the next section we will study some small examples of rings and fields.

You may have wondered about the absence of decimals. Decimals are best understood with some geometric help and therefore won't show up until chapter 2 where we study numbers with the help of geometry.

1.4 Modular Rings and Fields

It is very amusing to look at some small examples of rings and fields. We will look first at a small ring; the ring modulo 6. These small rings have extremely important applications in Number Theory. Three examples of number theory, which I mention so you know what the subject is about, are below. A prime number is a number which has only 1 and itself as positive integers divisors. (This is not the proper definition but will do for the moment.)

Theorem If a prime number p has remainder 1 when divided by 4 then there are two natural numbers (uniquely determined up to order) a and b for which $p = a^2 + b^2$.

Examples; $17 = 1^2 + 4^2$, $41 = 5^2 + 4^2$, $101 = 1^2 + 10^2$, $89 = 5^2 + 8^2$

Theorem Any natural number is the sum of the squares of four natural numbers: $n = a^2 + b^2 + c^2 + d^2$.

Examples: $1 = 1^2+0^2+0^2+0^2$, $7 = 1^2+1^2+1^2+2^2$, $35 = 5^2+3^2+1^2+0^2$

Conjecture Any even number greater than 4 is the sum of two primes.

Examples: $6 = 3+3$, $12 = 7+5$, $30 = 23+7$, $50 = 37+13$, $80 = 73+7$

The first two are moderately difficult to prove. The third we have been trying to prove for more than 276 years. So far, no luck. It is called Goldbach's

conjecture. When a new computer comes on line, this conjecture is always tested out as far as the computer can handle.

One of the most important tools in proving theorems in number theory like the above is modular arithmetic. In the system modulo 6 (abbreviated mod 6) all integers are replaced by their remainders when divided by 6. So if $a = 6q + r$ then we write $a \equiv r \pmod 6$[11]. Thus the different integers mod 6 are 0,1,2,3,4,5. We replace 6 by 0, 7 by 1, 8 by 2 etc. So we have $2 \cdot 3 \equiv 0$, $3 \cdot 5 \equiv 3$ since $15 = 2 \cdot 6 + 3 \equiv 3$ etc. Now we can make addition and multiplication tables for the mod 6 system, which are.

+	0	1	2	3	4	5
0	0	1	2	3	4	5
1	1	2	3	4	5	0
2	2	3	4	5	0	1
3	3	4	5	0	1	2
4	4	5	0	1	2	3
5	5	0	1	2	3	4

×	0	1	2	3	4	5
0	0	0	0	0	0	0
1	0	1	2	3	4	5
2	0	2	4	0	2	4
3	0	3	0	3	0	3
4	0	4	2	0	4	2
5	0	5	4	3	2	1

There are many interesting things to see here. First and foremost is that this system is a commutative ring. It's interesting that that commutativity is visible in the times table by the symmetry around the main diagonal (upper right to lower left). Most of the laws are automatically true because they are true for integers and so they remain true when we replace the integers by their remainders. This may not seem totally obvious to you but it would take too long to go into in detail, so take my word for it. Notice that $2 \cdot 3 \equiv 0$ so that the system is *not* a field. (Remember in a field if $ab = 0$ then $a = 0$ or $b = 0$.) When a ring is not a field one of the things we worry about is which elements are units (that is which elements have multiplicative inverses)? To find the units, look for 1's in the body of the multiplication table. There are two. $1 \cdot 1 \equiv 1$ and $5 \cdot 5 \equiv 1$. Thus 1 and 5 are units, and each is its own inverse. Note also that the elements which are *not* units have 0's in their rows and columns (besides the predictable ones at the top and left side). It is possible to predict which elements are units, but 6 is not a big enough number to make this clear. If you make a table for mod 12 you can probably figure it out.

Similarly we can look at the addition table and find the additive inverses for each element. Since this is a ring, each element will have just one additive inverse. Look for the 0's, For example the additive inverse of 4 is 2 since $4+2 \equiv 0$. Here is a little table of additive inverses

element	x	0	1	2	3	4	5
additive inverse	-x	0	5	4	3	2	1

It is easy to see the pattern here. Note that since $1 + 5 \equiv 0$ we must have $-1 \equiv 5$, which explains why 5 is a unit, since -1 will always be a unit in a ring.

Suppose now we form similar tables for a number m which is not a prime. Since m is not a prime there will be two numbers a and b that are less than m for which $m = ab$. In that case $ab \equiv 0 \bmod m$ and so the multiplication table

[11] This definition is inconvenient when negative numbers are involved, and it is much easier to use $a \equiv r \pmod 6$ if and only if 6 divides $r - a$.

will have 0 for a on the left and b on top. Since the multiplication table has a 0, the system cannot be a field.

Now suppose that m *is* a positive prime integer. Let's try it for $m = 7$; that is we will make the addition and multiplication tables mod 7. This will turn out quite different. The numbers in the system are now 0,1,2,3,4,5,6 and, for example, $5 \cdot 6 = 30 = 4 \cdot 7 + 2$ so the remainder is 2 and $5 \cdot 6 \equiv 2$. If you look back at the six addition table you can see that addition tables can be written down easily because of the obvious pattern. It is quite different with multiplication tables. At any rate they are.

+	0	1	2	3	4	5	6
0	0	1	2	3	4	5	6
1	1	2	3	4	5	6	0
2	2	3	4	5	6	0	1
3	3	4	5	6	0	1	2
4	4	5	6	0	1	2	3
5	5	6	0	1	2	3	4
6	6	0	1	2	3	4	5

×	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	1	2	3	4	5	6
2	0	2	4	6	1	3	5
3	0	3	6	2	5	1	4
4	0	4	1	5	2	6	3
5	0	5	3	1	6	4	2
6	0	6	5	4	3	2	1

Like all good mathematicians the first thing we check is the units. It turns our that every non-zero element is a unit. Here is the table:

element	x	1	2	3	4	5	6
multiplicative inverse	x^{-1}	1	4	5	2	3	6

Thus we see that the integers mod 7 form a field. It turns out that this is true for every prime number, and as we saw it cannot be true for any non-prime numbers. These little fields are called *prime fields*. It turns out that there is also a field for every power of a prime p^n. And that is all the finite fields that there are. To reiterate, there is exactly *one* finite field with p^n elements, $n \geq 1$, and mathematicians know how to construct them all. This is the kind of result that mathematicians always hope for; a complete list of all the objects of a certain type, and their structure.[12]

We should probably note the smallest field, the field mod 2, which has addtion and multiplication tables

+	0	1
0	0	1
1	1	0

×	0	1
0	0	0
1	0	1

Here is another property of finite fields that is interesting and useful, but too difficult to prove in this book. Every finite field with more than 2 elements has at least two (and maybe more) elements r so that every non-zero element in the field is a power of r. In the case $p = 7$ we can take $r = 5$. then we have, for example, $5^2 = 25 \equiv 4$ and $5^4 = (5^2)^2 \equiv 4^2 = 16 \equiv 2$, etc. Here is the table.

exponent	n	1	2	3	4	5	6
power of 5	r^n	5	4	6	2	3	1

[12]The field with p^n elements is *not* the ring of integers mod p^n.

An element in a system whose powers give all the non-zero elements in the system is said to be a *generator* or *primitive root* so 5 is a primitive root mod 7. A system made from powers of a fixed element is called *cyclic*. Cyclic systems are fairly simple.

If we rearrange the multiplication table in the order of the elements as powers of 5 we get

$\times$	5	4	6	2	3	1
5	4	6	2	3	1	5
4	6	2	3	1	5	4
6	2	3	1	5	4	6
2	3	1	5	4	6	2
3	1	5	4	6	2	3
1	5	4	6	2	3	1

which is characteristic of cyclic systems.

Here we see the same pattern as we see in addition tables. This is related to the fact that $5^6 \equiv 1$ and so the if we watch the exponents they work like a mod 6 system and the table resembles a mod 6 addition table. This is useful for many purposes.

Although we understand almost everything about finite fields, at least in theory, some things remain elusive. For example, although we know exactly how many primitive roots a mod p system has, we have never been able to find a way to produce one except by trial and error. Since this is boring and time consuming their are big books that list the primitive roots for all the primes up to some big one, depending on the size of the book, and computer algebra programs like Mathamatica have memorized the primitive roots up to some really big prime. But note, they don't calculate the primitive roots, they memorize them. The computer terminology for memorization is *lookup table* which is put into the program as a big list.

(Historical Note: The terminology of computer science is usually English based and often quite colloquial, like "lookup table" or "throughput time". This is because most of the developers of computers in the early days were young, and felt no need to use the Latin terminology so common elsewhere in science in the English speaking world. So we have lookup table instead of tabula sequorensa.)

Since the integers mod p are fields, almost all the stuff you learn in high school algebra will work for them, if you keep your wits about you. You must remember that dividing by a has to be replaced by multiplication by a^{-1} and $\sqrt{a}$ must be found by looking at the multiplication table. For example mod 7, we have $3^2 \equiv 2$ so $\sqrt{2} \equiv \pm 3 \bmod 7$. Recall that $-3 \equiv 4 \bmod 7$ so $\sqrt{2} \equiv 3, 4 \bmod 7$. Note there is no way to select between 3 and 4 as the "official" square root the way we choose the positive one when working with integers. There is no good way to *order* the elements in a finite field. The diagonal of the multiplication table has only 1,4,2 on it so these are the only elements which have square roots. The elements 3,5,6 have no square roots, but we could create them, as shown in the next section, by making the system bigger.

Before moving on I want to show how use of finite rings and fields can have applications in number theory. There are many applications in many areas but this one is easy. Recall that I mentioned that primes $p \equiv 1 \bmod 4$ are the sum of two squares of integers. What about the $p \equiv 3 \bmod 4$, the other possibility (except for $2 = 1^2 + 1^2$). We look at the problem mod 4. The squares mod 4 are $0 \equiv 0^2,\ 1 \equiv 1^2,\ 0 \equiv 2^2,\ 1 \equiv 3^2$ so the squares are 0,1. Hence $a^2 \equiv 0, 1$ and $b^2 \equiv 0, 1 \bmod 4$. Thus if p is the sum of two squares, $p = a^2 + b^2$, using all possibilities we get $p \equiv 0, 1, 2 \bmod 4$. Hence if $p \equiv 3 \bmod 4$ it *cannot* be the sum of two squares. Isn't that neat?

1.5 Algebraic Numbers

In this section we will study the fields associated with $\sqrt{2}$, $\sqrt{-3}$ and $\sqrt{-1}$. These numbers arise from equations. You may recall in the previous section we invented 0 so we could solve $x + 2 = 2$. In a similar way we invent $\sqrt{2}$ so we can solve $x^2 - 2 = 0$. The first thing to realize is that $\sqrt{2}$ is not a rational number. This was one of the greatest discoveries of ancient Greek mathematics, and had catastrophic consequences. I will explain this after I prove the fact. This is done by contradiction. I assume that $\sqrt{2}$ is a rational number and then I reach a contradiction. Since mathematics is contradiction free, I must have assumed something false; namely I assumed that $\sqrt{2}$ is a rational number. This kind of proof is usually called proof by contradiction[13].

First we remember that if a fraction is reduced (often called *lowest terms*) it cannot have both numerator and denominator even, since then we could reduce it. So assume $\sqrt{2} = \frac{r}{s}$ where we reduce the fraction so that r and s are not both even. Also recall that the square of an even number is even and the square of an odd number is odd. So we have

$$\begin{aligned} \sqrt{2} &= \frac{r}{s} \qquad r \text{ and } s \text{ not both even} \\ 2 &= \frac{r^2}{s^2} \\ 2s^2 &= r^2 \end{aligned}$$

so we know that r^2 is even. But the squares of odd numbers are odd so r cannot be odd so we know r is even. Since r is even, $r = 2t$ for some integer t. Hence we have

$$\begin{aligned} 2s^2 &= r^2 \\ 2s^2 &= (2t)^2 = (2t) \cdot (2t) = 4t^2 = 2(2t^2) \\ s^2 &= 2t^2 \end{aligned}$$

By the same argument as above, s is even, so we have both r and s are even,

[13] A small subset of mathematicians, called Intuitionists and largely of Dutch origin, does not accept all proofs by contradiction, although this one might be OK.

which is a contradiction. Hence

$$\sqrt{2} \neq \frac{r}{s} \quad \text{for any integers } r \text{ and } s$$

and hence $\sqrt{2}$ is not a rational number. This is often expressed by "$\sqrt{2}$ is irrational", but care must be taken that the listener or reader understands the meaning of "irrational" (= not rational = not a fraction) in this context.

Up to the time when some brother[14] in the Pythagorean brotherhood discovered this, the Greeks were happily building mathematics on the natural numbers and their ratios the rational numbers (fractions). But if a square has side of length 1, the Pythagorean theorem tells us that the diagonal has length $\sqrt{2}$. This shows that geometry throws up "numbers" that are irrational, that is they are not rational numbers, not ratios of integers. The Greek reaction to this was to largely throw numbers overboard and reconstruct mathematics in terms of line segments. This had the effect of making mathematics much much harder, and divorcing it from applications. So the mathematics used in astronomy, which Ptolemy had taken over from the Babylonians, was no longer regarded by the Greeks as *real* mathematics. This had very negative effects on the development of science, probably slowing it considerably. It had negative effects on some kinds of mathematics too. For example, every positive integer greater than 1 can be factored into a product of prime numbers in essentially only one way. Euclid restructured this so it was stated in terms of line segments. This did not contribute to the ease of learning the theorem, or using it, or the ease of proving it, as you might well imagine.

Eudoxos was the mathematician who provided the fully formed theory of real irrational numbers masquerading as line segments, and this became part of Euclid's treatment of geometry (Book V). This is also where Euclid really start to get difficult.

Eudoxos' method can be put in a modern mathematical setting and this was done by Richard Dedekind in 1858. The basic concept is called the Dedekind cut. It is one of two methods to attack the problem of irrational numbers and is based on the *order relation* of rational numbers. The other was originated by the Babylonians and is essentially the method of decimal expansions which is based on *approximation*. (The Babylonians used 60 instead of ten for the expansion, a trivial difference.) Babylonian methods were used for 20 Centuries in Astronomy before Georg Cantor repatriated them to mathematics in the 1850s as a method of defining irrational numbers. The basic item of the construction is the Cauchy sequence. We will talk much more about this in Chapter 2.

In this section we want to look at a different method of constructing a solution to $x^2 - 2 = 0$ and other similar equations. The basic idea of the method can be used to construct solutions for any algebraic equation with integer coefficients, and thus we are lead to the concept of algebraic number.

[14] or sister. Almost all the founders of religious movements restrict membership to men. Pythagoras and Buddha were rare exceptions to this pernicious rule.

Def An *Algebraic Number* θ[15] is a solution to an algebraic equation

$$a_n x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_0 = 0$$

where the a_i are ordinary integers for $i = 0, 1, \ldots, n$ and $a_n \neq 0$. If n is as small as possible for an equation having θ as a solution, it is called the characteristic equation of θ.

For example $\sqrt{2}/3$ is an algebraic number because it satisfies (check!) $9x^2 - 2 = 0$. This is the characteristic equation for $\sqrt{2}/3$ since $\sqrt{2}/3$ could not satisfy a linear equation with integer coefficients.

Def An *Algebraic integer* is an algebraic number which satisfies an equation with ordinary integer coefficients where the first coefficient a_n is 1, so the characteristic equation would look like

$$x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_0 = 0$$

where all the a_i are ordinary integers.

For example, $\sqrt{2}/3$ is not an algebraic integer but $\sqrt{2}$ is an algebraic integer because it satisfies $x^2 - 2 = 0$ where the coefficient of x^2 is 1.

There is an extensive theory of algebraic numbers and algebraic integers and many problems have not been solved. For our purposes we are only interested in quadratic equations

$$a_2 x^2 + a_1 x + a_0 = 0$$

which we will write for convenience

$$ax^2 + bx + c = 0$$

The basic tool for such equations (and a great many other things; memorize) is the factorization

$$x^2 - d^2 = (x - d)(x + d)$$

To prove this, just multiply out the right side using the distributive and commutative laws:

$$(x-d)(x+d) = x(x+d) - d(x+d) = x^2 + xd - dx - d^2 = x^2 + dx - dx - d^2 = x^2 - d^2$$

Note that we had to use the commutative law, so this formula is true in any *commutative* ring, but need not be true if the ring is not commutative. A second useful formula valid in any commutative ring is $(x+e)^2 = x^2 + 2ex + e^2$. See if you can provide justifications for the following steps

$$\begin{aligned} (x+e)^2 &= (x+e)(x+e) = x(x+e) + e(x+e) = x^2 + xe + ex + e^2 \\ &= x^2 + ex + ex + e^2 = x^2 + 2ex + e^2 \end{aligned}$$

[15] θ rhymes with data when pronounced dayta. Theeta is also an acceptable pronunciation, but less common.

Using these we can find a formula to solve *any* quadratic equation. The derivation takes a little courage because it looks horrible, but you can skim it if it scares you. [16] We start with the standard quadratic equation $ax^2+bx+c=0$ where of course $a \neq 0$ since otherwise it would not be quadratic. First we factor ax^2+bx+c. Ready? Go!

$$\begin{aligned}
ax^2+bx+c &= a\Big(x^2+\frac{b}{a}+\frac{c}{a}\Big) \\
&= a\Big[\Big(x^2+2\frac{b}{2a}+\frac{b^2}{4a^2}\Big)+\frac{c}{a}-\frac{b^2}{4a^2}\Big] \\
&= a\Big[\Big(x+\frac{b}{2a}\Big)^2-\frac{b^2-4ac}{4a^2}\Big] \\
&= a\Big[\Big(x+\frac{b}{2a}\Big)^2-\Big(\frac{\sqrt{b^2-4ac}}{2a}\Big)^2\Big] \\
&= a\Big[\Big(x+\frac{b}{2a}-\frac{\sqrt{b^2-4ac}}{2a}\Big)\Big(x+\frac{b}{2a}+\frac{\sqrt{b^2-4ac}}{2a}\Big)\Big] \\
&= a\Big[\Big(x-\frac{-b+\sqrt{b^2-4ac}}{2a}\Big)\Big(x-\frac{-b-\sqrt{b^2-4ac}}{2a}\Big)\Big]
\end{aligned}$$

Thus we have

$$ax^2+bx+c=0 \text{ is equivalent to } a\Big[\Big(x-\frac{-b+\sqrt{b^2-4ac}}{2a}\Big)\Big(x-\frac{-b-\sqrt{b^2-4ac}}{2a}\Big)\Big]=0.$$

Now this is not that helpful in a general commutative ring but in an integral domain[17] or field we have if $ab=0$ then $a=0$ or $b=0$. Hence in an integral domain or field the only possible solutions θ of $ax^2+bx+c=0$ will be

$$\theta=\frac{-b+\sqrt{b^2-4ac}}{2a} \quad \text{or} \quad \theta^*=\frac{-b-\sqrt{b^2-4ac}}{2a}$$

Note that we also have, putting some stuff together from above,

$$ax^2+bx+c=a\Big(x^2+\frac{b}{a}+\frac{c}{a}\Big)=a(x-\theta)(x-\theta^*)=a(x^2-(\theta+\theta^*)x+\theta\theta^*)$$

Comparing the coefficients of the two sides we have

$$\begin{aligned}
\text{The sum of the roots} &= \theta+\theta^*=-\frac{b}{a} \\
\text{The product of the roots} &= \theta\theta^* = \frac{c}{a}
\end{aligned}$$

This gives you a fairly easy way to check your work when solving a quadratic equation.

[16] This is not the standard derivation of the quadratic formula but it has the advantage of being correct.

[17] Recall an integral domain is a commuatative ring where cancellation works.

Now we want to create a field for $\sqrt{2}$ to live in. We will build this field on $\mathbb{Q}$, the rational numbers, and we know $\sqrt{2}$ is *not* in $\mathbb{Q}$ so we need to add something to $\mathbb{Q}$ to give $\sqrt{2}$ a home. $\mathbb{Q}$ is called *the ground field*. We are going to add to $\mathbb{Q}$ a new element θ which at first acts like the variable x but for which any occurrence of θ^2 will be replaced by 2. For example $\theta^3 = \theta^2\theta = 2\theta$ and

$$(2+\theta)(5-\theta) = 10 + 3\theta - \theta^2 = 10 + 3\theta - 2 = 8 + 3\theta$$

It is easy to see that everything in this system may be written as

$$r + s\theta \quad \text{where} \quad r \text{ and } s \text{ are in } \mathbb{Q}$$

and more examples of things in this system are

$$2 + 3\theta, \quad 2 = 2 + 0\theta, \quad \frac{13}{11} - \frac{31}{17}\theta, \quad 0 + \frac{17}{13}\theta = \frac{17}{13}\theta$$

This system is denoted by $\mathbb{Q}[\theta]$. It is almost obvious you can add, subtract and multiply elements of $\mathbb{Q}$ and you will get more elements of $\mathbb{Q}$ and that the commutative ring laws are all true. But is $\mathbb{Q}[\theta]$ a *field*? That is the big question. To answer it we must consider the notion of conjugates. Every element of $\mathbb{Q}[\theta]$ which actually has theta in it satisfies a unique (up to multiplication by a constant) quadratic equation $a_2x^2 + a_1x + a_0 = 0$ called its characteristic equation which we will write $ax^2 + bx + c = 0$. If we look at the formula for solving quadratic equations we see that they come in pairs with a positive and negative square root. So let's assume that the other root of the characteristic equation for $r + s\theta$ is $r - s\theta$. Then the two roots add to $2r$ and multiply to $(r+s\theta)(r-s\theta) = r^2 - s^2\theta^2 = r^2 - 2s^2$. Looking back at the relationship between a, b, c and the roots we see (taking $a = 1$) that the characteristic equation must be

$$x^2 - 2rx + (r^2 - 2s^2) = 0$$

Ah, but does it *work* since we found it with an assumption. Substituting $x = r \pm s\theta$ into the equation we get

$$\begin{aligned} r^2 \pm 2rs\theta + s^2\theta^2 - 2r^2 \mp 2rs\theta + r^2 - 2s^2 &\stackrel{?}{=} 0 \\ r^2 + s^2\theta^2 - 2r^2 + r^2 - 2s^2 &\stackrel{?}{=} 0 \\ s^2\theta^2 - 2s^2 &\stackrel{?}{=} 0 \\ s^2 \cdot 2 - 2s^2 &= 0 \end{aligned}$$

So our assumption was *right* and moreover we found the other root $r - s\theta$ of the characteristic equation. Now we define

Def The conjugates of an algebraic number are the roots of its characteristic equation.

We have just shown that the conjugates (including itself) of $r + s\theta$ are $r + s\theta$, $r - s\theta$. In general conjugates of algebraic numbers are not so easy to find, nor is the characteristic equation easy to find, but for our purposes, with quadratic

algebraic numbers, you just flip a sign. In this situation it is permissible to use the terminology *the conjugate* of a quadratic algebraic number, but remember this isn't sensible for a cubic algebraic number which has a cubic characteristic equation and so two conjugates besides itself.

Now notice that the product of the conjugates is a_0, the constant term in the characteristic equation. This gives us a way of proving that $\mathbb{Q}[\theta]$ is indeed a field. Suppose we are faced with

$$\frac{a+b\theta}{c+d\theta} \quad \text{where} \quad a,b,c,d \text{ are in } \mathbb{Q}$$

where $a+b\theta \neq 0$ We multiply numerator and denominator by the conjugate of the denominator to get

$$\frac{a+b\theta}{c+d\theta}\cdot\frac{c-d\theta}{c-d\theta} = \frac{(a+b\theta)(c-d\theta)}{c^2-d^2\theta^2} = \frac{(a+b\theta)(c-d\theta)}{c^2-2d^2}$$

and the last expression is in $\mathbb{Q}[\theta]$ *provided* the denominator is not 0. But if $c^2-2d^2=0$ then $2=c^2/d^2=(c/d)^2$ so 2 is the square of a rational number, which we know is not true. Hence the denominator is *not* 0 and the division is possible, which means

$$(a+b\theta)^{-1} = \frac{1}{a+b\theta} = \frac{a-b\theta}{a^2-2b^2}$$

and every non-zero element of $\mathbb{Q}[\theta]$ has an inverse and we have a field.

Now in this field $\theta^2=2$ so in this field 2 has a square root θ. Notice we can do any bit of arithmetic involving $\sqrt{2}$ in this field, but we can't find a decimal approximation for $\sqrt{2}$ in this context. That requires different equipment. Note also that $\sqrt{2}$ is just a *symbol* for the square root of 2, just like I have been using θ. In fact, I could have written $\sqrt{2}$ in every place in the development where I used θ, but it would have looked a little more crowded.

This method was developed by several mathematicians in the 1840-1875 time frame but it was Kronecker how saw the importance here. The algebraic quantity $\sqrt{2}$ can be put in a field that allows all algebraic manipulations without any use of the decimal expansion. Because decimal expansions are not part of finite mathematics, this gives us a way of dealing with all algebraic numbers without recourse to infinitary methods like decimals. Kronecker felt very strongly that infinitary methods had no place in *real* mathematics, which is reminiscent of the Greek attitude. Note that most irrational numbers are *not* algebraic, for example π because π has no characteristic equation, so the whole process would fail for it. (I would love to prove that π is not algebraic but sadly it takes a small book to do this so you'll have to take my word for it. This fact is closely connected with the old Greek problem of squaring the circle, that is, constructing a square with the same area as a given circle using ruler and compass. It can't be done, but it took over 2000 years to *prove* it couldn't be done. This was finally done by Ferdinand von Lindemann[18] in 1882.) Irrational numbers that

[18]German, 1852-1939. Lindemann was impressed that he had solved this very old problem

are not algebraic are called transcendental numbers and there is a vast theory for them.

Now let's ask the question, suppose instead of $\sqrt{2}$ we were interested in $\sqrt{-1}$. What would be different. Instead of $\theta^2 = 2$ we would have $\theta^2 = -1$, which means every time θ^2 showed up in a calculation we would replace it by -1. Thus

$$\theta^2 = -1 \quad \theta^3 = \theta^2\theta = -\theta \quad \theta^4 = \theta^3\theta = (-\theta)\theta = -\theta^2 = -(-1) = 1$$

after which the powers of θ cycle through $\theta, -1, -\theta, 1$ endlessly. Notice the fourth power of any power of θ is always 1. Since these are different and $x^4 - 1 = 0$ we have found all the fourth roots of 1. The n^{th} roots of 1 are a problem of great interest in algebra.

What else would be different. Absolutely *nothing*. From the algebraic point of view $\sqrt{2}$ and $\sqrt{-1}$ are exactly the same. For each we can construct a field, and we could even make a bigger field that contains them both. They live in the same world. Of course we would need to use two different symbols, for example we could use θ for $\sqrt{2}$ and i for $\sqrt{-1}$. The fact that there is a decimal for $\sqrt{2}$ but no decimal for $\sqrt{-1}$ is completely invisible in the algebraic world and thus causes no philosophical problems. I'll present a couple of little calculations so you can see how it works, as we did for the $\sqrt{2}$ field.

$$\begin{aligned}
(2+i)(3-i) &= 6 + i - i^2 = 6 + i - (-1) = 7 + i \\
(3+2i)(3-2i) &= 9 - 4i^2 = 9 + 4 = 13 \qquad \text{and you thought 13 couldn't be factored!} \\
\frac{7+i}{2+i} &= \frac{7+i}{2+i} \cdot \frac{2-i}{2-i} = \frac{14 - 5i - i^2}{4 - i^2} = \frac{15 - 5i}{5} = 3 - i
\end{aligned}$$

Now we will briefly look at another example where $\theta^2 = -3$. We build the field $\mathbb{Q}[\theta]$ in which $\theta^2 = -3$ in the usual way. Here you might see something you haven't seen before, maybe more than one thing. Let's think about the square root of 1. Actually, since $x^2 - 1 = 0$ is quadratic, it might have two roots, and indeed it does: 1 and -1. Now what does fairness suggest would be the number of *cube* roots of 1. You may have been told there is only one, which is 1. This is a vicious lie. Fairness suggests there ought to be 3 cube roots of 1. Let's check by solving the equation $x^3 - 1 = 0$. I happen to know how to factor $x^3 - 1$ due to many years in the business. Here we go.

$$\begin{aligned}
x^3 - 1 &= 0 \\
(x-1)(x^2 + x + 1) &= 0 \\
(x-1)(x - \frac{-1+\sqrt{-3}}{2})(x - \frac{-1-\sqrt{-3}}{2}) &= 0
\end{aligned}$$

and then decided to take on Fermat's problem, which is much much harder and he made no progress. However, he supervised the doctral degrees of many great mathematicians, among whom were David Hilbert, Hermann Minkowski, Oskar Perron and physicist Arnold Sommerfeld.

Don't believe me? Multiply it out. (I got the factors using the quadratic formula.) Let us abbreviate $(-1+\sqrt{-3})/2$ by ω as is very standard. Then we have (remembering that $\sqrt{-3}\sqrt{-3} = -3$):

$$\begin{aligned}
\omega &= \frac{-1+\sqrt{-3}}{2} \\
\omega^2 &= \frac{-1+\sqrt{-3}}{2}\frac{-1+\sqrt{-3}}{2} = \frac{1-2\sqrt{-3}+(-3)}{4} = \frac{-2-2\sqrt{-3}}{4} = \frac{-1-\sqrt{-3}}{2} \\
\omega^3 &= \frac{-1+\sqrt{-3}}{2}\frac{-1-\sqrt{-3}}{2} = = \frac{1-(-3)}{4} = 1
\end{aligned}$$

So the three cube roots of 1 are ω, ω^2, and $\omega^3 = 1$. And you just saw me check this for ω. For ω^2 notice $(\omega^2)^3 = \omega^6 = (\omega^3)^2 = 1^2 = 1$.

Now let's have another mind blower. ω is a root of $x^2 + x + 1 = 0$. Notice that the coefficient of x^2 is 1. So ω is not only an algebraic number, it is an algebraic *integer*. It lives in $\mathbb{Q}[\sqrt{-3}]$ or what is the same thing $\mathbb{Q}[\omega]$. So a point of some interest here is that when presented in radical form, it is not always easy to identify a quantity as an integer. We know how to do this when the equation is quadratic, but even for cubic equations it is a much harder problem. But I fear we cannot take the time to go into this in detail.

There is one further thing I would like to mention. I said in the section on modular rings and fields that there is a unique finite field for every p^n where p is prime. The $p = 3$ field has three elements $\{0, 1, 2\}$ and addition and multiplication tables are

	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

	1	2
1	1	2
2	2	1

Now I claim that there is a (unique) field of size $3^2 = 9$ which I can build from this using a quadratic extension $\theta = \sqrt{2}$ since 2 is not a square in the 3-field. The elements of the 9-field are $\{0, 1, 2, \theta, 1+\theta, 2+\theta, 2\theta, 1+2\theta, 2+2\theta\}$ and you could construct the addition and multiplication tables yourself remembering that $\theta^2 = 2$. Since this is a finite field, there is an element whose powers generate the field, and $\tau = 1+\theta$ will work. This vastly simplifies making a multiplication table if the elements are listed in order τ^i.

$$\tau^0 = 1 \quad \tau^1 = 1+\theta \quad \tau^2 = 1+2\theta+2 = 2\theta \quad \tau^3 = 2\theta+1$$

$$\tau^4 = (\tau^2)^2 = 2 \quad \tau^5 = 2+2\theta \quad \tau^6 = \theta \quad \tau^7 = 2+\theta \quad \tau^8 = 1$$

You might want to check the above calculations and make the multiplication table just to make sure you are following along correctly.

1.6 Mathematical Induction

First, a warning. The word "induction" used in "mathematical induction" has nothing to do with the use of the word "induction" in any other context. It is

a bad name for a general method of mathematical proof and it is not a form of induction where induction means what it means in other contexts, namely coming at a general law by observing many instances. To take a famous example, when one notes the sun coming up every day, one creates a general law that the sun comes up every day and one is not surprised when it comes up tomorrow. This has *nothing* to do with mathematical induction.

It is tricky to explain mathematical induction without an example but the example is not too clear without the explanation. Either way it's tricky. I'm going to go with the example first, explanation afterwards. We want a formula for the sum of the first n numbers. Mathematical induction won't find you the formula, but if the formula is handed to you on a stone tablet mathematical induction will allow you to prove the formula even if you have no insight into the problem.

Appendix The material up to now might lead you to think of certain additional questions which are slightly tricky. This appendix is for those who would like to dig a little deeper. The philosophical questions have now all been answered so this is more technical material and I include it for the more than usually curious.

Theorem For any positive integer n

$$1+2+3+4+\cdots+(n-2)+(n-1)+n=\frac{n(n+1)}{2}$$

Proof by Mathematical Induction There are two steps in a proof by mathematical induction.
Step 1: Show that your theorem is true when $n=1$. (This is the *base step.*)
Step 2: Assume your theorem is true for j and show it is true for $j+1$
This is called the *induction step.*

Here step 1 consists of checking the theorem works for $n=1$. To do this we write

$$1=\frac{1(1+1)}{2}$$

That is actually true, so step 1 is done. Step 1 is usually easy.

Now for step 2. For step 2 we get to assume that the theorem is true for j. So we consider that

$$1+2+3+4+\cdots+(j-2)+(j-1)+j=\frac{j(j+1)}{2}$$

{Note *This note is not actually part of the proof. I put it in because it helps to know what result we are trying to get. We substitute* $(j+1)$ *for* j *in our theorem and we get*

$$1+2+3+4+\cdots+(j-2)+(j-1)+j+(j+1)=\frac{(j+1)(j+1+1)}{2}$$

This represents what we must get by manipulating the formula for j*. If we get to this we are done.* }

We take our assumption

$$1+2+3+4+\cdots+(j-2)+(j-1)+j=\frac{j(j+1)}{2}$$

and do something to it to try and get the formula above. A natural thing to do is to add $(j+1)$ to both sides to get

$$1+2+3+4+\cdots+(j-2)+(j-1)+j+(j+1)=\frac{j(j+1)}{2}+(j+1)$$

Now the left side is what we want and we have to whip the right side to make it look like what we want. The trick is to go for a common denominator.

$$\begin{aligned}1+2+3+4+\cdots+(j-2)+(j-1)+j+(j+1) &= \frac{j^2+j}{2}+\frac{2j+2}{2}\\ &= \frac{j^2+j+2j+2}{2}\\ &= \frac{j^2+3j+2}{2}\\ &= \frac{(j+1)(j+2)}{2}\end{aligned}$$

and this is where we were trying to get. The proof is now complete.

Now I will try to explain why the proof proves the theorem. We know the theorem is true for $n=1$. The second part of the proof implies that if the theorem is true when $n=1$ then it is true when $n=2$. Combining the two, we know that the theorem is true for $n=2$. Now we use the second part again which implies that if the theorem is true for $n=2$ then it is true for $n=3$. Since we already know it is true for $n=2$ we now know it is true for $n=3$. Repeating this argument over and over we see that the theorem is true for any positive integer.

Here is a second argument to prove the validity of Mathematical Induction. Suppose that there is a positive integer for which the theorem is false. Then there must be *least* positive integer m for which it is false. We know $m\neq 1$ because of part 1 of the proof. Thus the theorem is true for $m-1$ (since m is the *least* integer for which it is false) so we have true for $m-1$, false for m. But this contradicts part 2 of the proof, which shows true for $m-1$ implies true for m. Hence the supposition, that there is a positive integer for which the theorem is false, has lead to a contradiction. Thus the supposition is false and the theorem is true for all positive integers.

The two justifications for mathematical induction look rather different but actually the difference is cosmetic. However, often a cosmetic restatement of something will be much clearer to some people than the alternative, so while they are mathematically basically the same, their psychological impact may be very different.

It is standard to mention the analogy of Mathematical Induction with a row of dominoes. You knock over the first one (step 1) and then each one knocks over the next one (step 2).

Sometimes it is inconvenient or unnatural to begin the induction with $n = 1$. Often it is more natural to start with $n = 0$ and occasionally one starts with $n = 3$ or some other number because of special circumstances. In the last case we would then have proved the theorem for all positive integers $n \geq 3$.

Students are often put off by the fact that in the induction step we assume the theorem is true for j. This feels like assuming what you want to prove, and then proving it. But actually the induction step proves

$$\text{True for } j \Longrightarrow \text{True for } j+1$$

and to prove an implication like this you assume the first thing and prove the second. The difference is a little subtle and you might want to meditate on it a bit.

We note that I have used n in the statement of the theorem and j in step 2. This is good but not so common. Most mathematicians use n in step 2 also. This is slightly confusing but you don't have to do it yourself. When you do enough problems it will start to seem natural.

One of the advantages of mathematical induction is that it is mostly pretty automatic. You start off with a known situation, you change j to $j+1$, and try to get to the new situation by algebra and logic. The situation *can* be difficult but mostly it is routine once you get used to it. This has the slight disadvantage that when reading advanced matheamtics books and a mathematical induction looms in sight, the author will often say something like "an easy (mathematical) induction shows that ....." As you can imagine, this habit could be abused, and sometime is.

Mathematical Induction is often the quickest way to prove a result about positive integers and many mathematicians will jump right to it if they sense it will work. However, there are two disadvantages to keep in mind. First, Mathematical Induction does not find formulas, it only proves them. Finding the result or formula is the creative part; the induction proof is mostly fairly automatic. Second, a proof by Mathematical Induction does not usually give you much insight into *why* a result is true; it just proves it. This is undoubtedly useful, but personally I value the insight a proof by other methods can give you, so Mathematical Induction is not my first choice for proofs. In some disciplines, for example linear algebra, it is very common to be able to take a proof and change it to an algorithm for computing something, but this is almost never possible for a Mathematical Induction proof.

We will now give a couple more examples.

Theorem For any non-negative integer n, $3^{2n} - 1$ is divisible by 8.
Proof step 1. Let $n = 0$ Then $3^{2n} - 1 = 3^0 - 1 = 0$ which is divisible by 8.
step 2. Assume $3^{2j} - 1$ is divisible by 8. Consider

$$\begin{aligned} 3^{2(j+1)} - 1 - (3^{2j} - 1) &= 3^{(2j+2)} - 1 - 3^{2j} + 1 \\ &= 3^{2j} \cdot 3^2 - 3^{2j} \\ &= 3^{2j}(3^2 - 1) \;=\; 3^{2j} \cdot 8 \end{aligned}$$

which is certainly divisible by 8. We have

$$3^{2(j+1)} - 1 = [3^{2(j+1)} - 1 - (3^{2j} - 1)] + [3^{2j} - 1]$$

Thus $3^{2(j+1)} - 1$ is divisible by 8 since it is the sum of two terms each divisible by 8, and the theorem is proved.

As an early example, Francesco Maurolico in his Arithmeticorum libri duo (1575), has the following interesting theorem.

Theorem The sum of the first n odd integers is n^2
Proof 1. The theorem is true when $n = 1$, where it reduces to $1 = 1$
2. Assume it is true for j so that we have

$$1 + 3 + 5 + \ldots + (2j - 1) = j^2$$

The next term in the sum would be $2j + 1$ so we add it to both sides and get

$$\begin{aligned} 1 + 3 + 5 + \ldots + (2j - 1) + (2j + 1) &= j^2 + (2j + 1) \\ &= (j + 1)^2 \end{aligned}$$

which completes the induction proof.

This may be the shortest possible induction proof.

The following is an example of how induction can go wrong, but this is interesting because it goes wrong in a surprising way. My claim that *in any set of horses, all the horses have three legs.* (You may be suspicious of this result, especially if you ride horses.) Here's step 2, the induction step. We assume that in any set of j horses, all the horses have three legs. Now take a set of $j+1$ horses. Number the horses $1, 2, \ldots, j, j + 1$. Remove horse $j + 1$. Horses $1, 2, \ldots, j$ form a set of j horses and thus, by the induction assumption, all have 3 legs. Put horse $j + 1$ back into the herd and remove horse 1. Horses $2, 3, \ldots, j + 1$ are a set of j horses and so all, including horse $j + 1$ have three legs by the induction assumption. We have shown that each of the horses $1, 2, \ldots, j, j + 1$ have 3 legs. Thus step 2 is done. Amazingly, however, the proof is flawed but the flaw is not in the part I did. Find the flaw.

Appendix

I want to give a second proof of our first example

$$1 + 2 + 3 + 4 + \cdots + (n - 2) + (n - 1) + n = \frac{n(n + 1)}{2}$$

This example is historically important because it set C.F. Gauss, the world's greatest mathematician (arguable of course) on the path to fame and fortune. At the age of 8 he was in school and the teacher wished to take a brief rest, so he assigned the class the problem of adding up the first 100 numbers, thinking it would keep the little boys and girls busy for a half hour or so. About 30 seconds later Gauss tossed his slate on the desk with the answer 5050, which you can

get quickly from the formula. Here is how Gauss is thought to have done it.

$$\begin{aligned} \text{Let } S &= 1+2+3+\cdots+98+99+100 \\ \text{then } S &= 100+99+98+\cdots+3+2+1 \\ \text{Adding } 2S &= 101+101+101+\cdots+101+101+101 \;=\; 100\cdot 101 \\ S &= \frac{100\cdot 101}{2} \;=\; 5050 \end{aligned}$$

You can see that by replacing 100 by n this derives the general formula. The teacher realized that he had a student here that was a cut above the usual, and the reputation followed Gauss through school and eventually got him a scholarship to Göttingen University paid for by the Duke of Braunschweig (Brunswick in English).

Notice that this proof is both more direct and gives greater insight into the reason the formula is true than our induction proof. This is usually the case; a direct proof will often be easier and often provide greater insight than a proof by induction. On the other hand, examining the proof by induction, notice it requires nothing but algebra. This proof required a certain degree of cleverness. This is the often true for induction proofs in contrast to other proofs.

A second example of an alternative proof is the following which gives a diagrammatic proof of $1+3+5+\ldots+(2n-1)=n^2$. Here is the diagram when $n=6$; $\quad 1+3+5+7+9+11=6^2=36$:

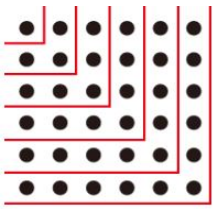


Diagram proof of $1+3+5+\ldots+(2n-1)=n^2$

This technique (diagram and dots) may not look very sophisticated but it can actually be used to give easy proofs of things not so easy to prove in other ways and so deserves respect. It goes back to the Pythagorean brotherhood.

Once in a while it takes great insight to see that a proof by induction is possible at all. A famous example is the proof by Ernst Zermelo that positive integers greater than 1 factor uniquely into prime integers. Zermelo saw that it was possible to give an induction proof of this, which nobody had ever suspected before. The proof uses a variant of the usual mathematical induction which uses the following two steps:
step 1. Show true for n=1 (same as before)
step 2. Assume theorem is true for *all* $k \le j$ and prove true for $j+1$.
The discussion above about why ordinary induction works is just as valid for this form of induction.

1.7 Historical Notes

"Die ganzen Zahlen hat der liebe Gott gemacht, alles anderes ist Menschenwerk" ("The Dear God made the integers, all else is the work of man.") This is Kronecker's most famous quote, as transmitted by Weber in 1893 and it sums up Kronecker's most abiding passion, as we will discuss below.

Leopold Kronecker was born on 7 Dec 1823 in Liegnitz in the Kingdom of Prussia. His parents were from a well to do Jewish family and his early education, along with his younger brother Hugo who became as physiologist, was by private tutor. His math teacher at the Gymnasium, (roughly, high school,) was Ernst Kummer who later had a distinguished career at the University of Berlin, and Kummer recognized his talent and encouraged his mathematics. He went to the University of Berlin where he graduated with a thesis on Algebraic number theory under Dirichlet in 1845. He returned home and took over the business interests of his mother's uncle and married his cousin Fanny Prausnitzer (heir to the business) in 1848. The couple eventually produced six children. He spent about 10 years as a businessman and was quite successful. During this time he continued his mathematical work in his spare time and published a number of papers, the most important of which was a memoir on the solution of algebraic equations using the methods of Evariste Galois, whose work was not well understood before this paper. (This was a very important contribution because up to this time Galoi's work languished in obscurity, but after Kronecker showed how to use it it took flight and became a commonly used powerful tool.) He corresponded with Kummer during this time, who had become a professor at Berlin. Eventually, having secured the family fortune, he retired to a comfortable life in Berlin. On the basis of his publications he was elected to the Berlin Academy in 1861 and this carried with it the right to give lectures at the University of Berlin (no pay) which he decided to do, beginning in 1862. He continued to produce outstanding research and in 1883, when Kummer retired, he finally got his first academic job as professor at the University of Berlin. He was very friendly with Karl Weierstrass though the friendship was later strained by philosophical differences. He supervised the theses of many students, including Kurt Hensel who later invented p-adic numbers which had a great future[19].

Over a period of years Kronecker formed a philosophy of mathematics that we would today call finitist. Kronecker could stomach infinite sets as a manner of speaking or description, but he could not bring himself to regard infinite sets as actual mathematical objects, like 2. The problem here is that without infinite sets you cannot construct the real numbers, as we do in the next chapter. You can construct $\sqrt{2}$, as we saw in this chapter, but you cannot construct π in the way you can construct $\sqrt{2}$ because π is transcendental, which means it is *not* the root of an algebraic equation with integer coefficients. There is a famous story about this. Walking on the campus of Berlin U. one day Kronecker encountered an excited Weierstrass. His old friend asked him "Have you heard? Lindemann

[19]The story that he used his own money to endow a chair at the University of Berlin and became the first person to sit in it appears, sadly, not to be true.

has proved that π is transcendental." "Very interesting," replied Kronecker, "Then π does not exist." The serious problem here is that if you can't construct the real line, you can't do ordinary Calculus, and thus much mathematics and all physics is relegated to the trash heap.

This finitistic attitude is of course completely inconsistent with Georg Cantor's theory of sets (see the chapter on Infinite Numbers.) Cantor was a professor at the University of Halle, in the mathematical backwoods, and accused Kronecker of keeping him from his just deserts, which he thought of as a professorship at the University of Berlin and of keeping his theory of sets from getting the recognition it deserved. This was a plausible story, and is often recounted even today, but no one has ever been able to turn up any evidence the Kronecker put serious effort into opposing the theory of sets or in keeping Cantor in the backwoods. He seems to have been only dimly conscious of Cantor's work which would have had little interest for him. And in fact Cantor's theory of sets gained mathematical recognition about as fast as any mathematical theory ever has; the younger mathematicians all jumped on the set bandwagon as fast as they could and many of the older mathematicians were also quite glad that someone had put together a theory which verified their fuzzy intuitions and produced interesting new ways to look at many things.

It is not clear how much influence Kronecker's finitistic program had on the later development of Intuistionistic Mathematics by L.E.J. Brouwer and his disciples. They were able to reconstruct a substitute for Calculus that was much more difficult and much less powerful but sort of did the job. The vast majority of mathematicians ignored this approach and preferred to jog along as usual. However, the situation changed.

When the digital computer was invented in the 1930s, it was found that 1) the mathematics it could do was exactly the mathematics the Intuitionists had created and 2) the algorithms that were necessary to make the machines do stuff had already been developed for the Intuitionist program where it was called Recursive Mathematics. This was quite a big leg up at the beginning of the computer age. And Kronecker, who was a rather short man, casts a much longer shadow standing atop the digital computer.

So nowadays we have two streams of mathematics; the standard one which represents the classical stream as whipped into shape by Cantor and the finitist stream which descends from Kronecker, (the first important mathematician to see the problem,) the Intuitionists and the computer scientists, who work on those things which can be actually computed and the theory of *what* can be computed.

There is a physical side to this story too. Once string theory failed to perform up to specs, the next candidate for basic physics was Covariant Loop Quantum Gravity. In CLQG, space is not taken to be infinitely divisible like the real line; it is taken to be made up of minimal size pieces of about 10^{-36} cm size. Perhaps *spaceons* might be a good name. The idea is old (we suspect Democritos thought of it) but has only recently become popular, and it can be developed with methods from Quantum Field Theory. This means that after the heroic effort of mathematicians to find a proper mathematical basis for the real line, it

turns out that open subsets of $\mathbb{R}^3$ may not be a suitable model for physical space after all. Moreover, the new theory of spaceons is rather complicated. On the good side, efforts to make spaceon theory compatible with General Relativity are going well and it might not be necessary to go the renormalization route in Quantum Field Theory to avoid infinities which show up in many calculations, and which seem to have their origin in the infinite divisibility of space. See [Rovelli], [Rovelli & Vidotto].

1.8 Problems for Chapter 1

Section 1.1

1. Using the internet to get the information, make a table of the numbers from 1 to 10 in the languages English, German, Hungarian, Hindi. Which one appears to not be related to the others. You can follow up this topic by looking up Indo-European Languages.

2. Give examples of numbers of the types natural, rational, irrational. Is 1.5 a rational number? Is 1.33333333333333..... (the string of dots means "it goes on like this forever"). The question revolves around whether these decimals are the decimals of fractions. This is not a trivial matter. It took over 2000 years to prove the π is not a fraction. (In particular, $\pi \neq \frac{22}{7}$ although $\frac{22}{7}$ is a good enough approximation to π for many practical purposes.)

3. Take some random even numbers, for example 30, 98, 126, 212 and find two prime numbers that add up to each. The internet has lists of prime numbers if you like such things. We will eventually prove the prime numbers go on forever.

Section 1.2

All letters in the problems will be *positive*. You may think of the them as positive integers if you like.

1. Simplify the following with exponents: bbb, bbbbb, bbbbbbbb.

2. Simplify the following with exponents: aba, abba, baba, baacaacb.

3. Simplify the following with fractional exponents: $\sqrt{b}$, $\sqrt[3]{b}$, $\sqrt[n]{b}$.

4. Simplify the following with fractional exponents: $\sqrt{b^4}, \sqrt[2]{b^3}$, $\sqrt[3]{b^{12}}$, $\sqrt[n]{b^m}$.

5. Use fractional exponents to simplify $\sqrt{b}\sqrt[3]{b}$. Give answer both as fractional exponent and in radical form. You have to add some fractions to solve this.

6. What are 1^0 and 0^1? Is $0^n = 0$ for any non-negative integer? Careful; trick question. How about for any positive integer. Is $n^0 = 1$ for any non-negative integer? How about for any positive integer?

Section 1.3

1. Which law is used in the following equations:

Section 1.4

1. We are going to look at the field with the 13 elements $\{0, 1, 2, 3, \ldots, 11, 12\}$. The first thing to do is to make addition and multiplication tables for this field (often called $\mathbb{F}_1 3$.)

2. Using the tables make tables showing the additive and multiplictive inverses of the elements of the field. (0 has no multiplicative inverse.)

3. Using the additive inverse solve the equations $x + 5 = 2$ (add the additive inverse of 5 to both sides.) Also solve $x + 8 = 1$ and $x + 9 = 7$. Remember to use your tables for adding.

4. Same as problem 3. but with multiplicative inverses, solve $3x = 7$, $5x = 9$ and $3x = 6$.

5. Using techniques from problems 3. and 4. solve $4x + 7 = 3$ and $7x + 5 = 8$. It should now be clear to you that all problems of the type $ax + b = c$ with $a \neq 0$ are solvable and the solution is unique.

6. Look down the diagonal of the multiplication table and you will find the squares (mod 13). There are six squares and $\sqrt{a}$ has a value when a is one of these squares. In fact it has two values (except when $a = 0$); for example $\sqrt{4} = \pm 2 = 2, 11$. Find the square roots of all 6 squares.

7. The equation $ax^2 + bx + c = 0$ is a quadratic equation. Typically a quadratic equation will have two solutions, or two *roots* as the professionals say. In section 1.5 we show there is a formula for solving such equations in a field when $a \neq 0$ and $1 + 1 \neq 0$ (as happens in $\mathbb{F}_2$.) The formula is

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = (2a)^{-1} \cdot (-b \pm \sqrt{b^2 - 4ac})$$

where the first formula is traditional but the second is handier in our circumstances. For any quadratic equation, it can be solved exactly when $b^2 - 4ac$ is a perfect square. In this problem is must be 0 or one of the six squares you found in problem 6. If $b^2 - 4ac = 0$ the formula gives only one root but professionals say it has *two equal roots*. That way *every* quadratic equation either has no or two roots. One of the following three quadratic equations has no root; the other two have two roots. Find the roots of the two equations that are solvable.

$$x^2 + 2x + 2 = 0 \qquad x^2 + 9x + 2 = 0 \qquad x^2 + 3x + 12 = 0$$

8. We like to *factor* quadratics which means to write them as a product of two linear terms. This is easy if you have the roots. Call the roots r and s. Then

$$ax^2 + bx + c = a(x - r)(x - s)$$

For example, the roots of $x^2 + 3x + 8 = 0$ are 3 and 7. So

$$x^2 + 3x + 8 = 1(x - 3)(x - 7) = (x + 10)(x + 6)$$

where in the last step I replaced -3 and -7 by their values in $\{0, 1, \ldots, , 11, 12\}$. Multiply out the two factors $(x + 10)$ and $(x + 6)$ using the distributive

law twice and get $x^2 + 3x + 8$. Now factor the two solvable equations from problem 7 and check your factors are correct by multiplying them out. Does that make you more comfortable with saying some quadratic equations have two *equal* roots?

It's worth noting that one can prove that in a field the factorization of polynomials is unique; everybody will get the same factors no matter how they do the factoring.

9. Everything went so well up to now because we were in the *Field* $\mathbb{F}_{13}$. However, the ring (mod 6) is *not* a field; 2 has no inverse. It is a ring. In rings one cannot depend on all the happy stuff we just saw for (mod 13). You have tables for (mod 6) in the text, and we are going to consider the quadratic equation $x^2 + 3x + 2 = 0$ Solve this equation by substituting each of $\{0, 1, 2, 3, 4, 5\}$ for x in the equation. How many roots did you expect? How many roots did you find? Pick any 2 with $r < s$ and use them to factor the equation. You can find three different factorizations this way, so factorization is far from unique. This shows that algebra in a ring is much much more difficult than algebra in a field.

Section 1.5

1. It is obvious that $1+\sqrt{2}$ is an algebraic number. Find it's minimal equation. Hint: Let $x = 1 + \sqrt{2}$. Then write $x - 1 = \sqrt{2}$ and now square both sides. Adjust so that one side is 0 and you have the minimal equation. You can check your answer by solving using the Quadratic formula, and getting $1+\sqrt{2}$. Note you also get the conjugate $1-\sqrt{2}$ which has the same minimal equation. This is an artifact of squaring both sides of the equation above, which may (usually will) introduce extra roots. In application these are called extraneous roots, but they have every right to be there.

2. Find the minimal equation for $\sqrt{2} + \sqrt{3}$ Hint: Let $x = \sqrt{2} + \sqrt{3}$. Then $x - \sqrt{2} = \sqrt{3}$ and now square both sides. In contrast to 1., there will be a $\sqrt{2}$ still on the left side. Move the number to the left and the term with $\sqrt{2}$ to the right and square again. The radicals are now all gone and you can find the minimal equation by moving everything to the left. It looks like $x^4 + 6x^2 + 5 = 0$ but the numbers are different. You can check your answer by setting $x = \sqrt{2} + \sqrt{3}$ on a calculator and then putting it into your minimal equation and seeing if the calculator gets something close to 0. $\sqrt{2} + \sqrt{3}$ is one of four conjugates that are solutions of the minimal equation. See if you can guess the other three. You can check them on a calculator too. You can also check them by hand using the radical forms, which is more in the spirit of this section.

3. Complex numbers which are not algebraic (that is, not roots of polynomials with integer coefficients) are called transcendental numbers. Although most complex numbers are transcendental, they are not so common in real life. Two of them are π and e. You probably know π which gets you

the area πr^2 of a circle with radius r. We use e in Calculus and it is the base of natural logarithms. The decimal value is approximately

$$e = 2.71828182845904523536028747135266249775724709369995957496696 76\ldots$$

but of course the sequence of digits goes on forever and does not repeat, since e is known to be transcendental and therefore irrational. The first number actually proved to be transcendental (by Hermite[20] was $.1010010001000010000010000001\ldots$ and he was able to refine his methods so that he proved that e was also transcendental. Lindemann then used Hermite's methods to prove that π was transcendental. Polynomial manipulation of transcendentals always gives more transcendentals, but there is one extremely important equation (discoved by Euler) that mixes them in non-algebraic ways:

$$e^{i\pi} = -1$$

where $i = \sqrt{-1}$. This is one of about eight formulas called "Euler's Formula". We will prove Euler's formula later in the book. Use Euler's formula to guess the values of $e^{i\frac{\pi}{2}}$ and $e^{2\pi i}$. Also, memorize the decimal for e up to the second 8, or the fourth 8 for enthusiasts. There is more information in this line in section 6.3.

Section 1.6

1. Use mathematical induction to prove that $2^k > k$.

2. Use mathematical induction to prove that $1+r+r^2+r^3+\ldots+r^n = \frac{r^{n+1}-1}{r-1}$. This is the formula that lies behind all compound interest, as, for example, when you purchase a house. (Hint: look for a common denominator.)

3. Use mathematical induction to prove that $1^2+2^2+3^2+\ldots+n^2 = \frac{n(n+1)(2n+1)}{6}$. This was used to find the area under a parabola just before the invention of Calculus. The formula can fairly easily be generalized to a formula for $1^k + 2^k + \ldots + n^k$. The algebra here is a little tricky but remember to factor our expressions that appear in both terms.

4. Use mathematical induction to prove that $1^2 + 3^2 + 5^2 + \ldots + (2n-1)^2 = \frac{(n-1)n(n+1)}{6}$

5. Use mathematical induction to prove that $1^3+2^3+3^3+\ldots+n^3 = \frac{n^2(n+1)^2}{4} = \left(\frac{n(n+1)}{2}\right)^2$. Notice that this is the square of the expression for $1+2+\ldots+n$. This is not a coincidence but the reason lies fairly deep.

[20] Charles Hermite (pronouced air-meet),1822-1901, French. Very creative mathematician who did excellent work in many areas. Without the Hermite polynomials the hydrogen atom would not work at all.

6. Use mathematical induction to prove that $2^n < n!$ for $n \geq 4$. Some help: First $n! = 1 \cdot 2 \cdot 3 \cdots (n-1) \cdot n$. Second you cannot start the induction with $n = 1$. The first step here uses $n = 4$ and then the induction step proves it for $n = 5$, then $n = 6$, then $n = 7$ and onwards. Most inductions start with $n = 0$ or $n = 1$ but be alert to problems like this where you start off for $n = 4$. Often a formula won't work for the first few numbers and you must start the induction with the first number it works for.

7. This problem is a little different in feel from the previous ones but the result is very important. The set $\{a, b, c\}$ has *subsets* made by making sets out of some, all, or none of the elements of the original set. Thus the subsets of $\{a, b, c\}$ are the empty set $\{\}$ and the others $\{a\}$, $\{b\}$, $\{c\}$, $\{b, c\}$, $\{a, c\}$, $\{a, b\}$, $\{a, b, c\}$. Note the last subset contains all the elements of the original set. Now count the subsets. There are $8 = 2^3$ subsets of a set of size 3. Use mathematical induction to prove that a set of size n has 2^n subsets. Hint: pick some element of the set of size $n + 1$ which we'll consider special. How many subsets do not have this element. How many do have it. Rats; I just solved the problem for you, almost.

Chapter 2

NUMBERS WITH GEOMETRY

2.1 A Little History

In this chapter we will discuss the real and the complex numbers. The real numbers are so called to distinguish them from the complex numbers which in olden times were called imaginary numbers, but this is considered rude and insensitive today.

The real numbers have a very long history; in fact they were invented by the Sumerians at the dawn of history. The Sumerians invented the sexagesimal system to represent what we now call real numbers. The Sumerian civilization was slowly replaced by Semitic immigrants whom we refer to as Babylonians. There is no way to tell in detail how much of classical Babylonian mathematics was actually invented by the Sumerians, and for simplicity we will refer to the contributions of both peoples as Babylonian.

The Babylonians did not make an intellectual distinction between the real numbers and the rational numbers because they used the sexagesimal system to represent both. Thus they threw rational and irrational numbers in the same box and saw only real numbers.

Our knowledge of Babylonian mathematics is limited to what chance provides in the way of clay tablets. Since many more tablets were concerned with practical matters rather than theoretical ideas we have very little evidence about Babylonian involvement with pure mathematics. What we have has been developed for practical, in the Babylonian sense of the word, affairs.

The Babylonians were obsessed with the motions of the stars and planets because they thought that their motions predicted the future, and thus they were the inventors of Astrology. To predict the future, you had to know where the sun and planets would be at some time in the future. If the King was contemplating a war, it was important to know if he would win, and he would then turn to his Astrologers, who would compute the positions of the planets and sun at the time of the proposed war and tell him what would happen. If the predictions came out true, well and good for everybody. If they came out false, the astrologers were sent to the mines and new ones promoted. A simple system with the great disadvantage that that motions of the sun and planets have nothing to do with who wins the war, so astrologers develop glib tongues and complex excuses to explain any discrepancies, and also work hard to accumulate information that will help their predictions. Over time they became quite good at it. Neighboring peoples envied this ability to predict the future and the Babylonian system, including the zodiac, eventually spread over all of Eurasia, and with European colonization to the New World too. The Astrology section of your newspaper is a descendant of this Babylonian science, but the accuracy has not increased over time. One of the tricks of astrologers down the ages has been to couch the predictions in somewhat vague terms, so a lot of different things that happen will seem to confirm the predictions. For example you will receive something useful from a tall person. Chances are that one of the mailman, the UPS man or the Fedex man will be tall, and ditto with the mailwoman etc. The desire to know the future is universal among humans, and most people are happy enough with vague or false predictions because they take

the uncertainty out of life, or so it seems. A relevant point here is confirmation bias. One tends to remember far better the times when the Astrologer was accurate than the times he fails, so his successes help him much more than his failures hurt him. The statistical method of blind or double blind experiments, where neither the subjects nor the experimenter knows who gets the medicine and who gets the sugar pill, are designed to counter confirmation bias and other common human errors.

Now it not so easy to predict the motions of the sun and planets, and few people at any time can do it. First and foremost, one needs a very good computational system to do it. In fact one needs a far better computational system for it than one needs to figure out how much dirt you need in a dam and how much it will cost to get the dirt there. We know how the Babylonian system worked and it is far far better than is needed for dams or banking or taxes or feeding the army; in fact the only thing in the Babylonian world that needed this kind of precision was planetary motion for Astrology. They developed their wonderful system almost surely for just this purpose, which makes Astrology the crazy aunt of mathematics. Astrology is nonsense of course, but this wonderful system of computation (which evolved into the decimal system), was partially developed for her. So treat her kindly, but don't believe a word she says. Science has been immeasurably enriched by the mathematics developed for her.

One of the things I want to get across in this book is that the Babylonian system, called the *sexagesimal system* and the decimal system we use today are only cosmetically different, and thus the Babylonians are really the inventors of the real numbers. They did not see the theory behind their system very clearly, and we will discuss this eventually. The key idea here is Cauchy sequence. (A second approach to the reals, which is also interesting, is the Dedekind cut, and we will discuss this in an appendix for the curious.)

2.2 The Babylonian and the Decimal Systems of Representing Numbers

Once you have the integers the next problem to deal with is fractions. The Babylonians were very obsessed with 60 for three reasons; it is a convenient size, it has many divisors, and it is closely related to 360 which is roughly the number of days in the year. It was convenient for the Babylonians to divide the sky into 360 pieces, one for each day, and 30 for each constellation of the zodiac, and then to make the little corrections necessary because of the extra 5 days in the year. If you look at any particular place in the sky at a certain time of night, the constellations will then move approximately 1 degree each night. This is handy for rough predictions, which you can then refine.

At some point some clever Sumerian, (The Sumerians were the predecessors of the Babylonians,) came up with the idea of representing fractions as pieces of 60. Thus $\frac{1}{2}$ was 30 ($\frac{30}{60} = \frac{1}{2}$) and $\frac{1}{3}$ was 20 and $\frac{1}{5}$ was 12 etc. The first integer denominator where this is not so convenient is 7, and we have to deal with that.

So here are some numbers as the Babylonians would see them and they way we see them:

$$\begin{aligned} 2\frac{1}{2} &= 2,30 = 2+\frac{30}{60} \\ 15\frac{1}{3} &= 15,20 = 15+\frac{20}{60} \\ 23\frac{1}{5} &= 23,12 = 23+\frac{12}{60} \\ 23\frac{3}{5} &= 23,36 = 23+\frac{36}{60} \end{aligned}$$

However, this doesn't work so well for $\frac{13}{25}$ since 25 does not divide 60. But 25 does divide $3600 = 60^2$, $\quad 3600 = 25 \cdot 144$. So

$$\frac{13}{25} = \frac{13}{25}\cdot\frac{144}{144} = \frac{1872}{3600} = \frac{31\cdot 60+12}{3600} = \frac{31\cdot 60}{3600}+\frac{12}{3600} = \frac{31}{60}+\frac{12}{3600}$$

so

$$5\frac{13}{25} = 5,31,12 = 5+\frac{31}{60}+\frac{12}{60^2}$$

Now compare this to the decimal approach

$$5\frac{13}{25} = 5.52 = 5+\frac{5}{10}+\frac{2}{10^2}$$

Note they are essentially the same; one uses 60 the other uses 10. Is that something to get all excited about? I don't think so, although the base 10 system has some computational advantages. Base 10 was introduced into the European world by Simon Stevin[1] (1548-1620). Some 'Arabs had also experimented with decimal fractions but their work does not seem to have been known in Europe.

So far everything has been exact but we know there is no way to make $\frac{1}{7}$ into a decimal fraction, because it repeats. Let's do it first for 60. I will do this crudely although there are algorithms to do it. Explanation follows display.

$$\begin{aligned} \frac{1}{7} &= \frac{8}{60}+\frac{1}{105} \\ &= \frac{8}{60}+\frac{34}{60^2}+\frac{1}{12600} \\ &= \frac{8}{60}+\frac{34}{60^2}+\frac{17}{60^3}+\frac{1}{1512000} \\ &= \frac{8}{60}+\frac{34}{60^2}+\frac{17}{60^3}+\frac{8}{60^4}+\frac{1}{22680000} \\ &= \frac{8}{60}+\frac{34}{60^2}+\frac{17}{60^3}+\frac{8}{60^4}+\frac{34}{60^5}+\frac{1}{2721600000} \end{aligned}$$

[1]Simon Stevin did many other things besides decimals. He worked on hydrology, music theory, bookkeeping, and recommended Dutch as the future language of Science. He also put wind sails on a carriage and sailed around the beaches of the Netherlands, which are conveniently flat.

How do I get these numbers? One seventh of 60 is 8.5714.... So I use the 8 as numerator and 60 as denominator for the first term. I subtract 1/7-8/60 and get 1/105. The denominator must be greater than 60. I then divide 105 into 60^2 and get 34.28 so I need $34/60^2$ for the second term and I subtract these two from 1/7 and get 1/12600. I divide 12600 into 60^3 and get 17.142. The 17 is the numerator for the third term $17/60^3$. You continue in this way as long as you wish. There are better ways but we are not actually going to use this again so this is good enough for now. Thus

$$\frac{1}{7} \approx 8, 34, 17, 8, 34$$

with an error

$$\frac{1}{2721600000} \approx 3.67431 \times 10^{-10}$$

This is good enough even to predict planets. One of the advantages of the sexagesimal system is that it takes far fewer terms to get a good approximation.

If you ask your calculator for the decimal for $\frac{1}{7}$, or do the division by hand, you get

$$\begin{aligned}\frac{1}{7} &= .142857142857142857142857142857142857142857\ldots \\ &= \frac{1}{10}+\frac{4}{10^2}+\frac{2}{10^3}+\frac{8}{10^4}+\frac{5}{10^5}+\frac{7}{10^6}+\frac{1}{10^7}+\frac{4}{10^2}+\frac{2}{10^3}+\frac{8}{10^4}+ \\ &\quad +\frac{8}{10^4}+\frac{5}{10^5}+\frac{7}{10^6}+\frac{1}{10^7}+\frac{4}{10^2}+\ldots\end{aligned}$$

which we write as

$$\frac{1}{7} = .\overline{142857}$$

A distinction between rational numbers and irrational numbers is that the decimals (or sexagesimals) always repeat for rational numbers and never repeat for irrational numbers. If you look back at our sexagesimal calculation you will see repetition already starting to set in. We suspect from what we see that

$$\begin{aligned}1/7 &= 8, 34, 17, 8, 34, 17, 8, 34, 17, 8, 34, 17, 8, 34, 17, 8, 34, 17, 8, 34, 17, \ldots \\ &= \overline{8, 34, 17}\end{aligned}$$

and with a little effort we could prove it. This repeating behavior is not too hard to prove but this is not the place to do it.

Also note a decimal expansion is not a good way to show a given number is irrational because 1) you need all the infinitely many terms of the decimal because you don't know how long, if ever, it will be until repetition sets in. and 2) there may be 250,004 terms before repetition sets in so it would be tiresome to compute.

An important point is that the Babylonians never noticed, or never thought it worth writing down on any clay tablet that has come down to us, that rationals repeat and irrationals don't. Both types of numbers were just sequences of

sexagesimal digits, and they threw them all in the same box. That box is the real numbers.

When the world was young and the author of this book was a boy, angles were still measured using sexagesimal fractions as degrees, minutes and seconds but the advent of the hand calculator freed the world from this form of Babylonian oppression and we now (outside of astronomy) we mostly use degrees and decimals of a degree instead of degrees minutes and seconds. However the dead hand of Babylon still lies across our clocks, and it seems unlikely that we will ever see hours and decimals of an hour instead of hours, minutes and seconds when measuring time. It is interesting that almost nobody thinks this is odd.

2.3 Geometric Representation of the Decimals

We now want to correlate numbers written as decimals with points on the line. It is easy to represent integers on the line. We arbitrarily choose a length to be 1 unit and call that the basic length. Then we choose a point on the line to be 0. For simplicity we consider mostly horizontal but sometimes vertical lines. For now it's a horizontal line. From the point we chose for 0 we mark off a basic length to the right[2] and we are at the point corresponding to 1. We mark off a second basic length and we are at the point corresponding to 2. We repeat this as many time as we like and get points corresponding to positive integers. Returning to 0, we mark off a basic length to the left. This will be point for -1. Do it again to get the point for -2. Etc.

Ok, now look at the line segment from 0 to 1, including the end points. If a segment includes its end points it is called *closed*. If the end points of the segment are not there it is called *open*. You wouldn't think this would be important but it turns out to be very important indeed, so learn it. We represent the closed segment from 0 to 1 by $[0,1]$[3]. We now cut the interval [0,1] into ten equal intervals, which are then [0,.1], [.1,.2], [.2,.3],. . .,[.9,1]. Each interval is of length 1/10. The beginning points of these intervals correspond to the decimals 0, .1, .2, .3 . . . .9. A decimal betweens 0 and 1, for example .142857, must lie in the interval [.1,.2]. We can procede in a similar way to divide [.1,.2] into 10 equal intervals [.10, .11],[.11,.12],. . .[.19,.2], each interval of length 1/100. The point that corresponds to .142857 then must line in the interval [.142,.143]. Continuing this process we find that .142857 lies exactly at the beginning of the interval [.142857,.142858]. Everything is fine.

However, if we have a *repeating* decimal

.142857142857142857142857142857142857142857142857142857142857142857142857 . . .

[2]Going to the right is arbitrary. It is very probable that on some planets they go to the left.

[3]Everyone is fine with that, but for open segments most use (0,1) but some use the suggestive but slightly offputting notation]0,1[

things are trickier. Imagine it this way. We have a series of decimals

$$d_1 = .1,\ d_2 = .14\ \ d_3 = .142\ \ d_4 = .1428\ \ d_5 = .14285\ \ d_6 = .142857\ \ d_7 = .1428571$$
$$d_8 = .14285714\ \ d_9 = .142857142\ \ d_{10} = .1428571428\ \ d_{11} = .14285714285\ldots\ldots$$

We continue to subdivide intervals into smaller and smaller intervals of length $1/10^n$ but the process never ends. Empires wax and wane, nations cleave asunder and coalesce, stars get dimmer, galaxies fade and universes darken and die to be replaced by new big bangs, and still we put the next point d_n just a tiny bit to the right at each stage. And the process never stops. If you think about the subdivision process you will see that the points d_n for $n > 6$ can never exceed .142858. They are all trapped between .142857 and .142858 and morever they move to the right. Note that since all the d_n for $n > 6$ lie between .142857 and .142858, their distance from each other is smaller than $.000001 = 10^{-6}$. There is nothing special about 6 here. If we want the terms of the series to differ less than 10^{-20}, we just look at terms further out than d_{20}.

Can we guess where the the points are heading toward. You may have guessed that they are heading toward the point for 1/7 which we get by dividing [0,1] into seven equal parts and taking the beginning of the second division. This whole process is an example of the mathematical concept of *limit*, and we write it

$$\lim_{n\to\infty} d_n = 1/7$$

(This is read "The limit as en approaches infinity of dee sub en is one seventh"). The reading gives a very wrong impression. It actually means the limit of d_n as n increases beyond all bound is one seventh, since clearly one cannot "approach infinity", but this is the way mathematicians speak and it's important to be able to translate from what they say to what they mean.

I will mention in passing that at this time *limit* is the concept upon which Caclulus is constructed. Some people find the limit concept fairly easy and others find it very difficult. Issac Newton tried to find a logical basis for limit based on the concept of motion. It didn't work. No matter if you find it easy or difficult, a logical basis for limit took a very long time to find. Cauchy (1789-1857) came close and the project was completed by Weierstrass (1815-1897). The basic idea here is *approximation*, but we will leave it to your Calculus professor to continue the story.

Now you might say that was an awful lot of work for very little profit and you would be essentially correct; there is nothing new here except the idea of limit. So let's try to squeeze something more out of the stuff we have worked so hard to get.

Up till this moment we have worked entirely with points on the line that correspond to rational numbers, as must always be the case when we deal with repeating decimals. In fact you might think the line is made of just the rational points, but that would be very seriously wrong. Even though between any two distinct rational points there are infinitely many rational points[4] the rational

[4] see problems

line is riddled with holes. We now demonstrate this.

Using your basic length, the line segment [0,1], draw a line segment of this length horizontally and at the right end draw another vertically. Then connect the left end of the first to the top end of the second to get an isosceles right triangle. For right triangles with sides a, b, c with c the long side we have the Pythagorean[5] theorem which says

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

and we then get, if we use l for the length of the diagonal:

$$\begin{aligned} l^2 &= 1^2 + 1^2 \\ l^2 &= 2 \\ l &= \sqrt{2} \end{aligned}$$

Now we take this diagonal, place the left hand end at 0 on the line and the right hand end, let us call it P, will then be at $\sqrt{2}$ which we have proved is *not* a rational number. Now by the method of subdividing we can find the decimal for $\sqrt{2}$ one digit at a time. We can readily see it lies between 1 and 2. We divide [1,2] into ten subdivisions and we find P is in the 5^{th} subdivision [1.4,1.5], so the first decimal digit is .4 and $\sqrt{2} \approx 1.4$. We next subdivide [1.4,1.5] into ten subdivisions and find P is in the second subdivision [1.41,1.42] so $\sqrt{2} \approx 1.41$. We continue the process till we get as many decimal digits as we like, for example

$$\begin{aligned} \sqrt{2} &\approx 1.414213562373 \quad \text{or} \\ \sqrt{2} &\approx 1.4142135623730950488 \quad \text{or} \\ \sqrt{2} &\approx 1.41421356237309504880168872421 \end{aligned}$$

By the nature of the process each of these will be a little less than the actual value of $\sqrt{2}$ so each of their squares will be a little less than 2, but even the first is accurate to about 40 decimal places and so most computer algebra programs will give a square of exactly 2 since they cannot handle so many digits. It is important to realize that computers deal entirely with rational numbers when doing work with decimals, so they cannot be exact. The words used to describe the small errors that even the largest computers must make are *round off error*. This is the result of dealing with essentially infinite matters with machines that have a finite number of storage slots. It is not avoidable and you must be careful about trusting machines. Machines are more trustworthy than politicians or fortune tellers but it is a matter of degree, not a difference in principle. Of course the machines are doing the best they can to tell the truth; it's just not possible.

[5] Pythagóras of Samos, circa 570-485 B.C., founder of the Pythagorean brotherhood and originator of the ideas that all mathematics can be derived using logic from a small number of initial assumptions(axioms) and that nature could be understood through mathematics. Note the accent goes on the o in Greek. The way the name is pronounced in English sounds like we are yelling at him in Greek.

I emphasize once again that the decimal for $\sqrt{2}$ will neither stop nor will repeat, since in either case that would make $\sqrt{2}$ a rational number, which it is not. Thus the decimal goes on forever. This is the standard thing for points on the line; the points corresponding to rational numbers are rare; the ones corresponding to infinite decimals like happens with P are the usual outcomes. With more advanced techniques one can prove that P corresponding to rational numbers can only be found by actually seeking them; if you randomly slice the line with a knife your chances of hitting a rational point are exactly 0. No chance whatever. This is easily proved using an area of analysis (mathematics with limits) called measure theory. This is a little counterintuitive, but ones intuition is not entirely reliable once you start dealing with limits. It is this which causes Philosophers to have a deep suspicion of limits, which they have had ever since Newton and Leibniz started to use them in the late 1600s.

Things are actually rather worse than they seem. There are nowhere near enough English sentences to describe all the points on the line, or what amounts to the same thing, all the infinitely long decimals. And from here one can wander into the land of contradiction, for example the smallest positive decimal that can't be described in English, which I just described. Do not worry about this stuff. It is only of importance to philosophers and the occasional philosophical mathematician, but the rest of us get along quite nicely without bothering our heads about it. In a later section of the book, on infinite numbers, we will prove the part about the English sentences and the decimals.

A historical note is perhaps in order here. Before Calculus with its infinite processes began to infiltrate mathematics, philosophers were in love with mathematics, which is to say the mathematics inherited from the Greeks, especially Euclid. Greek mathematics had forms of Calculus, but they were well disguised and you had to dig really deep to find them. But Euclid's material, quite logically arranged and intuitively clear, was a philosopher's dream. It illustrated the power of logical thinking and was often pointed to as an example of "absolute truth".

In the early 1600s however, two things were happening which were to undermine this partnership between philosophy and mathematics. First, mathematicians had traveled the road that Euclid and the Greeks had pioneered about as far as they could. In order to move further, new methods were necessary, specifically Calculus. Second, long ago Pythagoras had said that Nature could be understood through mathematics, but from 500 BCE to 1600 (although everyone believed it was true because Pythagoras and Plato had said it was,) results along this line had been quite disappointing. There were some successes but even the flight of a cannonball was not really mathematically analyzable. Again, what was needed was Calculus. Once Newton and Leibniz got the ball rolling, a whole new world of applications opened up, both in mathematics itself and in science. It really looked like Pythagoras' dream of understanding the world through mathematics had finally been realized. But the logical certainty that characterized Greek mathematics was lost and the philosophers missed it badly and indeed some attacked the new methods as unreliable, unintuitive, unlogical and probably wicked. Although from the mathematical point of view

eventually something approaching the old logically valid development was restored, the philosophers largely never bought in to the new approaches to the logical underpinning of mathematics and the old relationship of philosophy and mathematics was never restored. It didn't help when later mathematical discoveries knocked the legs out from under absolute truth. We will touch on some of this later.

2.4 Construction of the Real Numbers

For this section it is convenient to define the absolute value $|x|$ of a number. The definition is

$$|x| = \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0 \end{cases}$$

So essentially $|x|$ is the numberical part of x made positive or 0. It is convenient for measuring distances; for example if x_1 and x_2 are two points on the real line, the distance between them is $|x_2 - x_1|$ regardless of which comes first on the line, or whether x_1 and x_2 are positive or negative. Naturally $|x_1 - x_2|$ would work just as well.

This section goes a little deeper than most of our work and can be skipped if you feel overwhelmed since what we do here is not used in the rest of the book. On the other hand it gives an answer to the question "Just what is a real number?" We have given a crude answer to this: a real number is a decimal, usually infinitely long. We can give a better answer, but the answer takes some thought. To motivate the answer, let's recall again how the system works. Let's switch to a different irrational number. Our former irrational was $\sqrt{2}$ which is an algebraic number. This time let's use π which is not an algebraic number. The theorem that π is not an algebraic number took a long time to prove. It is closely connected to one of the three great unsolved problems of Greek mathematics, the problem of squaring the circle[6] which means construct with ruler and compass a square with the same area as a given circle. It is much harder than the other two and was not settled until 1882 when Ferdinand von Lindemann proved that π was not algebraic. So it took about 2200 years to prove that it is impossible to square the circle with ruler and compass.

It is harder than usual to compute decimal approximations of π but it has now become a test for new computers and we have computed 22.4 trillion digits. Part of the reason for such excess enthusiasm is the hope that some sort of regularity will be found, which would be very important. This was what motivated the Chudnovski brothers to build a supercomputer in their New York apartment from materials they bought at local electronics stores. Cables ran from room to room making life a little less convenient for their wives. Alas neither the Chudnovskis nor anyone else has found any regularity in the digits; they appear to be randomly distributed.

[6]The other two are trisecting an angle and doubling the cube, and you are supposed to use only unmarked ruler and compass (no scratching the the ruler with the point of the compass). None of these problems can be solved with ruler and compass construction.

We will work with fewer digits; 30 instead of 22.4 trillion, not counting the whole number 3 at the front. This gives

$$\pi = 3.14159265358979323846264338328 0\ldots$$

Think of this as the first 30 in an infinitely long string of digits[7]. I got this from the computer algebra program Mathematica which did not compute it; it has memorized a few thousand digits of π to satisfy people who keep asking for them. You can ask Wikipedia if you want a LOT of digits. But 30 will do fine to make my point. The infinitely long decimal is not really proper mathematics notation. For that we need to look at the sequence.

$$\begin{aligned} d_0 &= 3 \quad d_1 = 3.1 \quad d_2 = 3.14 \quad d_3 = 3.141 \quad d_4 = 3.1415 \\ d_5 &= 3.14159 \quad d_6 = 3.141592 \quad d_7 = 3.1415926 \\ d_8 &= 3.14159265 \quad d_9 = 3.141592653 \quad d_{10} = 3.1415926535 \end{aligned}$$

This is called an infinite sequence and is a standard mathematical object once you get beyond elementary mathematics. As it clearly visible, this sequence is *converging* to π by which we mean that however small a small positive number $\epsilon > 0$ you give me I will find an integer M so large that $|\pi - d_n| < \epsilon$ whenever $n > M$. In words, for $n > M$ the difference between π and d_n is smaller than ϵ. This is what

$$\lim_{n\to\infty} d_n = \pi$$

actually means. For example if you give me $\epsilon = .0000273 = 2.73 \times 10^{-5}$, I will take $M = 5$. Then

$$\begin{aligned} |\pi - d_5| &= \pi - 3.14159 = .0000026535897\ldots \\ &= 2.65\ldots \times 10^{-6} < 10^{-5} < 2.73 \times 10^{-5} = \epsilon. \end{aligned}$$

and clearly $|\pi - d_n|$ for $n > M = 5$ will be even smaller and so less than ϵ as desired. This kind of gives you the flavor of the subject. Most people find this kind of thinking cumbersome and difficult, but surprisingly it becomes rather easy after a few months of practice. Sadly, there are a few people who never get it.

However, for what we want to do we have to look at the problem slightly differently. The problem with what we did is that we needed π's decimal expansion to show the convergence. This is bad. So we come from a slightly different direction. We want to look at how close the d_n are *to one another*. Specifically, we want to show that given any small $\epsilon > 0$ we can find an M so that if m and n are greater than M then $|d_n - d_m| < \epsilon$. What this means in imprecise words is that the terms of the sequence get close together.

With the same ϵ as before I again take $M = 5$ and lets look at

$$|d_9 - d_6| = 9 = 3.141592653 - 3.141592 = .00000653 = 6.53 \times 10^{-6} < 10^{-5} < \epsilon$$

[7] If you are religious you might think of the infinitely many digits printed on an infinitely long strip of paper whose far end is taped to the handle of God's screen door.

If you play around with this example a bit you'll be able to come up with the M that works whenever someone gives you a small $\epsilon > 0$.

Now you wonder what I am getting at here so I explain again. In the second example we used only the *terms* of the sequence d_n and we did *not* use the *limit* of the sequence. This suggests the following important definition.

Definition A sequence $a_n, \quad n - 1, 2, 3, \ldots$ of numbers is a *Cauchy*[8] *Sequence* if and only if for every $\epsilon > 0$ we can find an M so that

$$\text{if} \quad m, n > M \quad \text{then} \quad |a_m - a_n| < \epsilon$$

This simply means that as m and n get large the terms of the sequence get close together.

The big advantage here is a theorem that says that a sequence converges if and only if it is Cauchy. Thus we can use this to prove a sequence converges *even if we don't know what the limit of the sequence is*, which is the case much of the time. Also it saves us from trying to find out what the limit is when there *is* no limit. However, the situation is actually more complex than this paragraph suggests, as we will see.

We have one more thing to clear up before we come to the coup, which is the definition of a real number. A *zero sequence* is a Cauchy sequence that approaches 0. For example

$$\begin{aligned} z_0 &= 1 \quad z_1 = .1 \quad z_2 = .01 \quad z_3 = .001 \quad z_4 = .0001 \quad z_5 = .00001 \\ z_6 &= .000001 \quad z_7 = .0000001 \quad z_8 = .00000001 \quad z_9 = .000000001 \ldots \end{aligned}$$

To be formal and to get you used to this mode of speaking, a Cauchy sequence is a zero sequence if and only if

for any $\epsilon > 0$ there exists an M so that

$$\text{if } n > M \quad \text{then} \quad |a_n| < \epsilon$$

There is an easy theorem that says if b_n is a Cauchy sequence and a_n is a zero sequence then $b_n + a_n$ is also a Cauchy sequence and

$$\lim_{n\to\infty}(b_n + a_n) = \lim_{n\to\infty} b_n + \lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n + 0 = \lim_{n\to\infty} b_n$$

so that adding a zero sequence to a Cauchy sequence does not affect the Cauchisity nor does it change the limit. Now we introduce the concept of equivalence for Cauchy Sequences. Two Cauchy sequences are called *equivalent* if their difference is a zero sequence. In symbols

The Cauchy sequences b_n and c_n are equivalent (written $b_n \sim c_n$) if and only if the Cauchy sequence $(b_n - c_n)$ is a zero sequence.

[8] Augustine-Louis Cauchy, 1789-1857. Pronounce Cauchy like Coffee but with sh substituted for ff and accent on the last syllable.

Now comes the mentally tricky part. Equivalence of Cauchy sequences is an example of an equivalence relation. These are endemic within higher mathematics; you can't get along without them for even a couple of pages over vast areas of mathematics. Equivalence relations must satisfy the following three requirements (this is in the abstract, the above equivalence for Cauchy Sequences is just an example)

$a \sim a$					reflexivity
$a \sim b$	implies	$b \sim a$			symmetry
$a \sim b$	and	$b \sim c$	implies	$a \sim c$	transitivity

Whenever we have an equivalence relation we have *equivalence classes*. If a is any element then the equivalence class of a is the set of all elements that are equivalent to a, which is denoted by $[a]$ are by $\overline{a}$. We will use $[a]$. To have an example relevant to our work, consider the two Cauchy sequences

$$\begin{aligned} p_0 &= 3 \quad p_1 = 3.1 \quad p_2 = 3.14 \quad p_3 = 3.141 \quad p_4 = 3.1415 \quad p_5 = 3.14159 \\ p_6 &= 3.141592 \quad p_7 = 3.1415926 \quad p_8 = 3.14159265 \quad p_9 = 3.141592653\ldots \end{aligned}$$

and

$$\begin{aligned} q_0 &= 4 \quad q_1 = 3.2 \quad q_2 = 3.15 \quad q_3 = 3.142 \quad q_4 = 3.1416 \quad q_5 = 3.14160 \\ q_6 &= 3.141593 \quad q_7 = 3.1415927 \quad q_8 = 3.14159266 \quad q_9 = 3.141592654\ldots \end{aligned}$$

Let us show that p_n and q_n are equivalent. We look at

$$q_0 - p_0 = 1 \quad q_1 - p_1 = .1 \quad q_2 - p_2 = .01 \quad q_3 - p_3 = .001 \quad q_4 - p_4 = .0001$$

and we see that $q_n - p_n$ is a zero sequence, by coincidence exactly the zero sequence we used as an examples of zero sequences. Thus we see that the two Cauchy sequences p_n and q_n are equivalent; $p_n \sim q_n$ which will also tell us that p_n is in the equivalence class of q_n and vice versa. Note also that $[p_n] = [q_n]$, the equivalence classes are equal. For if r_n is in $[p_n]$ then $r_n \sim p_n$ and we know $p_n \sim q_n$ so transitivity gives us that $r_n \sim q_n$ and so r_n is in $[q_n]$. Reversing the argument we find that if r_n is in $[q_n]$ then r_n is in $[p_n]$, so $[p_n]$ and $[q_n]$ have the same members and thus are equal. This is a general fact; in general two equivalence classes are either equal or have no members in common (in which case we say they are *disjoint*).

We now have all the equipment we need to give the definition of a real number. Let's recall what we have been using:

$$\pi = 3.14159265358979323846264338328 0\ldots$$

This has two disadvantages. First, the infinite decimal is sort of *informal* mathematics and also it relies too heavily on the number 10 to be aesthetically pleasing. For the Babylonians

$$\begin{aligned} \pi &= 3, 8, 29, 44, 0, 47, \ldots \\ &= 3 + 8/60 + 29/60^2 + 44/60^3 + 0/60^4 + 47/60^5 \ldots \end{aligned}$$

and their way and our way of seeing π are only cosmetically different. (I cannot resist pointing out that the term $0/60^4$ says that *pi* comes within $1/60^4 \approx 8 \times 10^{-8} = .00000001$ of being the rational number 84823/27000. Think about it.)

So how are we to get around these unpleasant aspects. Cantor's solution with some help from others was to define π as the equivalence class of the Cauchy sequence

$$\begin{aligned} p_0 &= 3 \quad p_1 = 3.1 \quad p_2 = 3.14 \quad p_3 = 3.141 \quad p_4 = 3.1415 \quad p_5 = 3.14159 \\ p_6 &= 3.141592 \quad p_7 = 3.1415926 \quad p_8 = 3.14159265 \quad p_9 = 3.141592653\ldots \end{aligned}$$

Notice, and this is very important, that the numbers in the sequence are *rational* numbers. Thus $\pi = [p_n]$. There is nothing sacred about *this* Cauchy sequence. Indeed π is equal to the equivalence class of any Cauchy sequence that is equivalent to p_n, and we could easily manufacture one from the Babylonian example

$$q_0 = 3 \quad q_1 = 3 + 8/60 \quad q_2 = 3 + 8/60 + 29/60^2 \quad q_3 = 3 + 8/60 + 29/60^2 + 44/60^3$$

The difference of $p_n - q_n$ is a zero sequence so $[p_n] = [q_n]$ and for practical purposes we can use either sequence. So after this example I can now give the definition

Def A real number is an equivalence class of Cauchy sequences of rational numbers.

To go from the Cauchy sequence to a point on the line, mark the rational numbers in the Cauchy sequence on the line, and they will bunch up at a point, which is the point on the line corresponding to the real number.

To go the other way, from a point P on the line to the Cauchy sequence, select a point corresponding to a rational number within 1 unit of P, that rational number is p_0. Next do the same but within 1/10 to get p_1. Next within 1/100 to get p_2. Continuing in this way we will get a Cauchy sequence for the real number, from which we make the equivalence class of this Cauchy sequence, and this equivalence class is *the* real number which corresponds to the point P on the line.

There is much more to do. The real number corresponding to the rational number 1/7 is the equivalence class of the Cauchy sequence all of whose terms are 1/7. This puts the rational numbers inside the real numbers. Next we must define addition and multiplication of real numbers, none of which is difficult and all of which is tedius. I do this for addition. If r and s are real numbers and r has Cauchy sequence r_n and s has Cauchy sequence s_n we define $r + s$ as the real number with Cauchy sequence $r_n + s_n$. It is easy to show it is Cauchy. The problem is, the Cauchy sequence for r is not unique; there are many many Cauchy sequences for r. It is *critical* that the outcome of addition does *not* depend on which Cauchy sequences we use. To be specific, and to clearly indicate what is to be proved, let r_n and $\tilde{r}_n$ be two Cauchy sequences for r and let s_n and $\tilde{s}_n$ be two Cauchy sequences for s. Then from the definition of addition $r_n + s_n$ and $\tilde{r}_n + \tilde{s}_n$ are both Cauchy sequences for $r + s$. For addition

to make sense, these two Cauchy sequences must be equivalent, so that the equivalence class of the two is the same. Then addition is *well defined*. So we must prove $(r_n + s_n) \sim (\tilde{r}_n + \tilde{s}_n)$. I'm not going to actually do this because it's a little heavy for this book; I just wanted you to understand how we would procede if we wanted to continue the construction of the real numbers. The same thing must be done with multiplication and inverse. Inverse has tricky aspects, which revolve around zero sequences all of which represent the real number 0.

You have seen here an example of the need for proving something involving equivalence classes is well defined. This is remarkably hard for many students so be sure to pay close attention when you get to "well defined" in class. It's also a favorite exam question.

In the end, we would be able to show that the real numbers $\mathbb{R}$ form a field and that would complete the construction. We would then prove several theorems from which we could efficiently develop the theory of limits and then Calculus. As an example we would prove that every Cauchy Sequence of *real* (not just rational) numbers converges. Another example is that any set of numbers which are all less than some number M has a least upper bound. This is the least number that that is above all the numbers in the set, and its existence then allows us to logically develop limits and Calculus. You understand none of this is easy. Some people like it and some don't. It will all be very different from your previous mathematics. However, there are very few basic ideas here so that once you get on top of a couple of examples it becomes relatively easy to go on. So don't get discouraged. You would normally hit this stuff as a college junior or senior.

Once you have done the stuff in the last paragraph, the definition of a real number as an equivalence class of Cauchy sequences of rational numbers fades into the background and the theorems referred to above take over for further development. This is an example of modulization of mathematics. We need a logical foundation for Calculus and what we have discussed about the real numbers provides it. Once that is done, convenient theorems can be proved which make life easier and the original module with the Cauchy sequence definition fades into the background, although Cauchy sequences themselves continue on as the theory develops further. This area of Mathematics is called *real analysis*, which is fundamental for large areas of Mathematics and can be thought of as a further development of Calculus although it precedes Calculus in a logical development, as we have indicated. Most schools develop Calculus using a sort of intuitive concept of limit and then later go back in fill in the basics. This is logically undesirable, but it prevents many prospective mathematics students from becoming discouraged by the heavy theory and becoming statisticians or music majors.

2.5 Appendix. Dedekind Cuts

There is a second approach to real numbers which is simpler than Cauchy sequences but is not as popular because it has less general application. Cauchy sequences occur all through mathematics but Dedekind cuts can be used in only a few places and most mathematicians never find a use for them again.

For psychological reasons it may be best to visualize the rational line as vertical for the purposes of Dedekind cuts though horizontal will work well too if you like. A Dedekind cut is a division of the rational line into two sets A and B with the following properties.

- Every rational number is in A or in B
- No rational number is in both A and in B
- Every rational number in A is less than any rational number in B
- B has no least element.

Let's look at a couple of examples.
Example 1. A contains all negative rational numbers and all positive rational numbers whose square is less than 2.
B contains all positive rational numbers whose square is greater than 2.
This is clearly a Dedekind Cut and A and B give all of the rational line because $\sqrt{2}$ is irrational and thus not on the rational line. This is the cut that corresponds to $\sqrt{2}$.

Example 2. A contains all negative rational numbers and all positive rational numbers which are less than the circumference of a circle whose radius is $\frac{1}{2}$.
B contains all positive rational numbers which are greater than the circumference of a circle whose radius is $\frac{1}{2}$.
This is the Dedekind cut that corresponds to π.

Example 3. A contains all negative rational numbers and all positive rational numbers which are less than or equal to 5.
B contains all positive rational numbers which are greater than 5.
This is the Dedekind cut that corresponds to 5.

Thus we see that the rational numbers fit into the Dedekind cut scene as in the previous example. With this material we are ready to define real number.

Def A real number is a Dedekind Cut.

Now we have to do the same thing that we did after we defined real number with Cauchy sequences. We must define addition, multiplication, inverse and then start proving the basic theorems of real analysis, for example we must prove evey Cauchy sequence of real numbers converges. Moreover, we should probably prove that the two approaches, Equivalence class of Cauchy Sequences of Rationals and Dedekind Cuts, produce the same structure, namely the real numbers $\mathbb{R}$. However, I think I have done my duty towards Dedekind cuts and will stop here.

2.6 Complex Numbers

From the algebraic point of view, as you saw in a previous section, there is no difference whatever between $i = \sqrt{-1}$ and $\sqrt{2}$. For geometric representation of numbers, however, there is quite a difference. It is remarkable how long it took to come up with this representation, since from our point of view it seems completely natural.

Historically, complex numbers are old; there is shaky evidence that even the Babylonians worked with them a bit, but this is not certain. Certainly the Medieval 'Arabs worked with them, for example al-Khwarizmi (died 850 in Baghdad). Strangely from our point of view, it was not the equation $x^2 + 1 = 0$ or $x^2 + 9 = 0$ that forced the consideration of complex numbers. It was cubic equations that forced complex numbers into a sort of shadow existence, because to solve the algebraic solution of $x^3 - 7x + 6 = 0$, which has the three real roots 1,2,-3 one has to take the cube root of a complex number. This is certainly offensive; real equation, real roots, nasty tour of complex numbers to get the solution. But without the complex numbers you can't solve cubic equations. The 'Arabs were almost certainly familiar with this problem but our first documentary evidence is from Girolamo Cardano (1501-1576) who first published the formula for the solution of the cubic in his book the Ars Magna (1545). This book contains Cardano's own investigations and those of Niccoló Fontana (1499-1557) who is generally called Tartaglia which is Italian for stutterer[9]. (Italians of this time period were less politically correct than we are.) Tartaglia gave Cardano his version of the solution under a pledge of secrecy (he said later) but the secrecy was compromised when Cardano published it (with attribution to Tartaglia) in the Ars Magna. There was considerable acrimony over this and it seems to be the first priority fight of which we have knowledge. Complex numbers thus made it into the general knowledge of mathematicians, but were not accorded the status of legitimate numbers and were referred to by ethnic slurs like *imaginary*. They were regarded as a kind of computational trickery, like a magician's slight of hand.

As far as I can tell, the general prejudice against our friends the complex numbers was due to a fixed notion that numbers must be greater than 0 (positive numbers) or 0 or less than 0 (negative numbers). This is the same as the notion that numbers must come on a line. Since complex numbers could not be put into this strait jacket they were not allowed to be numbers from earliest time until about 1800. Descartes gave them the offensive name imaginary numbers and Euler, who used them extensively for many purposes, refused to recognize them as numbers. Around 1800 Caspar Wessel, a Danish surveyor and Jean-Robert Argand a French amateur mathematician tipped to the fact that complex numbers could be represented by points in the plane. At roughly the same time Carl Friederich Gauss had the same idea. Never an enthusiastic publicist, Gauss did not work very hard to spread the word, but word spread anyway. Since no one could deny the respectability of the Euclidean 2-dimensional plane, the pla-

[9]He got the stutter when a French soldier took a slice at him as a child during the siege of Brescia, and damaged his palate

nar representation went a long way to helping the complex numbers to also gain respectability. Eventually, previous knowledge about complex numbers was correlated with the new representation, the enthusiasm of mathematicians grew, and the planar representation fairly quickly became standard. The representation in the plane is still occasionally called the *Argand diagram* for complex numbers. It is also occasionally called the *Gaussian plane*.

From the point of view of the history of mathematics the real question is why this didn't happen in 1500 or 1700 instead of 1800. There were mathematicians active that could have easily put the pieces together, and some came very close, but no one took the final step and publicized it. It is one of the oddest turns in the history of mathematics.

Another question of interest is why the prejudice against complex numbers persists. They get very little time in the elementary mathematics curriculum and because they are introduced so late there is some resistance among the students to them. The planar representation would make it possible to introduce them in seventh or eighth grade and this would greatly facilitate their later use when they become essential for many topics in applied mathematics. There is no need to faithfully continue the prejudice of our forefathers.

From the formal point of view, complex numbers are built from a basis 1 and i by using real coefficients to get $a \cdot 1 + b \cdot i$. When performing algebraic operations with complex numbers the i works just like x *except* that i^2 must be replaced by -1. Thus

$$\begin{aligned}(2-7i)+(4+3i) &= 6-4i\\ (2-7i)-(4-8i) &= 2-4i-4+8i \;=\; -2+4i\\ (2-7i)\times(-3+2i) &= -6+4i+21i-14i^2 \;=\; -6+25i-14\cdot(-1) \;=\; 8+25i\end{aligned}$$

(The parentheses on the left side of the first two are usually not written and the multiplication sign is also usually omitted.)

In general we can write, for $z = a + bi$ and $w = c + di$

$$\begin{aligned}zw \;&=\; (a+bi)(c+di) &&= ac+(ad+bc)i+bdi^2 \;=\; (ac-bd)+(ad+bc)i\\ wz \;&=\; (c+di)(a+bi) &&= ca+(cb+da)i+dbi^2 \;=\; (ca-db)+(cb+da)i\end{aligned}$$

and inspection shows that $zw = wz$; multiplication of complex numbers is commutative. This is important but you were expecting it so it's easy to remember.

Division is a little more complicated and it is convenient to now introduce the conjugate of a complex number.

Def The conjugate of $a + bi$ is $a - bi$.

The notation for conjugate is a bar[10] so

$$\text{if} \quad z \;=\; a+bi \quad \text{then} \quad \overline{z} \;=\; \overline{a+bi} \;=\; a-bi$$

The origin of the term conjugation is related to the fact that the complex number $z = a + bi$ is a root of the equation $z^2 - 2az + a^2 + b^2 = 0$. (Substitute $a + bi$ for

[10] In physics they often use z^* rather than $\overline{z}$. This is probably better, being more consistent with other notation, but there is little hope of changing it in mathematics.

z in the equation if you don't believe me, and it would be good practice). The *other* root of the equation is then $\overline{z} = a - bi$.

It is of great importance that

$$z\overline{z} = (a + bi) \cdot (a - bi) = a^2 - abi + abi - b^2 i^2 = a^2 + b^2 \geq 0$$

That is, for $z \neq 0$, $z\overline{z}$ is a positive real number.

We will also need the equation

$$\overline{wz} = \overline{z}\,\overline{w}$$

This is easily proved from the formula above for wz. It is not strictly necessary to shift the order since complex multiplication is commutative, However, in your further studies you will be dealing with non-commutative situations and there similar equations always require the shift in order. So it's good to start off with the formula that will generalize properly to quaternions and matrices.

Recall

$$\text{if} \quad z \neq 0 \quad \text{then} \quad z\overline{z} > 0$$

Now we can do inverse of z. The trick is to multiply the numerator and denominator by the conjugate of the denominator. If $z = a + bi \neq 0$ then

$$\frac{1}{z} = \frac{1}{z}\frac{\overline{z}}{\overline{z}} = \frac{\overline{z}}{z\overline{z}} = \frac{a}{z\overline{z}} - \frac{b}{z\overline{z}}i$$

and what makes this work is that $z\overline{z}$ is a positive real number so the coefficients of the inverse are real numbers and $1/z$ is a complex number written in standard form.

You now know everything you need to verify that the complex numbers $\mathbb{C}$ are a commutative ring and even more, a field. Thus all your high school algebra, when you get there, will work for complex numbers but remember there is no *order* to complex numbers. One complex number is not bigger or smaller than another unless they both happen to be real numbers.

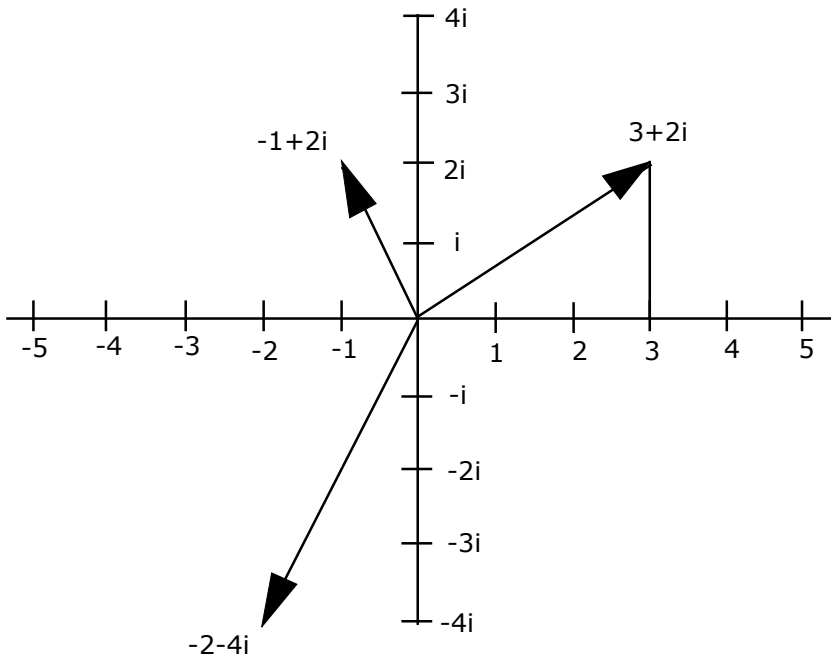
We may as well do division while we are at it. Let $z = a + bi$ and $w = c + di$. Then

$$\frac{w}{z} = w \cdot \frac{1}{z} = w \cdot \frac{\overline{z}}{z\overline{z}} = \frac{w\overline{z}}{z\overline{z}}$$

We will mention in passing that the integers of $\mathbb{C}$ are the $a + bi$ with a and b in $\mathbb{Z}$, that is a and b are ordinary integers. The integers of $\mathbb{C}$ are usually denoted by $\mathbb{Z}[i]$. The arithmetic of these complex integers $\mathbb{Z}[i]$ was first systematically investigated by Carl Friedrich Gauss starting around 1800 and continuing for the rest of his life. The notion of prime number depends on what ring you are working in. To illustrate 5 is a prime in $\mathbb{Z}$ but in $\mathbb{Z}[i]$ it is not a prime: 5=(2+i)(2-i). Also $2 = (1 + i)(1 - i) = -i(1 + i)^2$. The units of the ring $\mathbb{Z}[i]$ are $1, i, -1, -i$ so up to units 2 is a square in $\mathbb{Z}[i]$. As in $\mathbb{Z}$, factorization into primes is unique but because of the units this is not obvious at first glance. For example $2 + i$ and $(-i)(2 + i) = 1 - 2i$ are considered the same prime because

they differ by multiplication by a unit, just as 7 and -7 are considered the same prime in $\mathbb{Z}$.

So now we want to look at the planar representation (or Argand diagram) of complex numbers. Here is the basic picture with a few complex numbers in it.



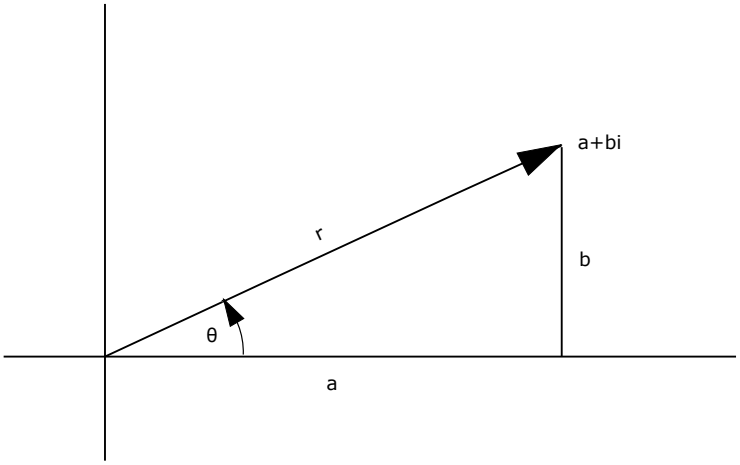
Complex numbers in the plane

The general picture for the geometric representation is below. Here we see a generic complex number $a+bi$ with the a or real coordinate along the horizontal axis and the b or complex coordinate along the vertical axis (but drawn parallel to the vertical axis for didactic purposes[11].

Here we see one of the real advantages of the geometric representation. It turns out that r and θ are extremely important in the theory of complex numbers (as we will see) but how would one every think of θ without the complex representation in the plane.

We begin with r. The Pythagorean theorem tells us that $r^2 = a^2 + b^2$. It turns out that $a^2 + b^2$ is a quantity of great interest also. So we define for $z = a + bi$ the Norm $N(z)$ and modulus r of z by

[11] The complex coordinate is often called the "imaginary coordinate" but this name is now undesirable as it denigrates and insults the complex numbers.



Polar form of complex numbers

$$N(z) = a^2 + b^2 \qquad r = |z| = \sqrt{a^2 + b^2} = \sqrt{N(z)}$$

The angle θ is called the argument of the complex number. Full treatment of the argument requires trigonometry, which we will try to keep to a minimum. The definition of the sin and cos functions for θ are

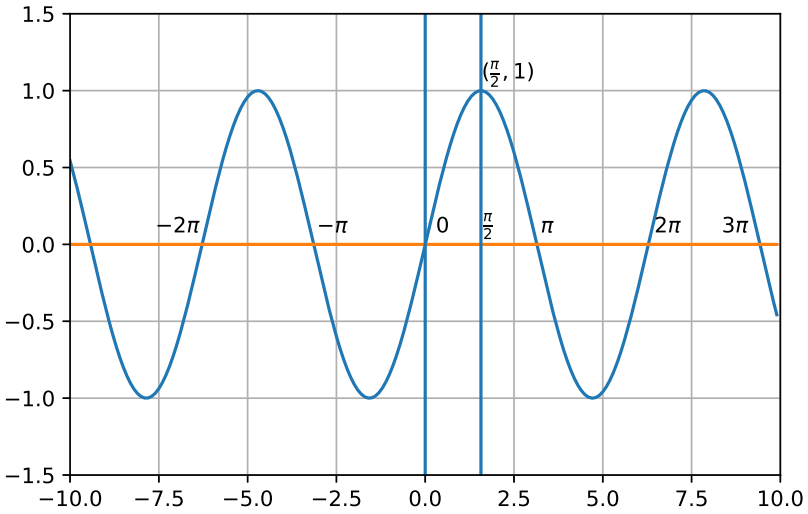
$$\begin{aligned} \cos(\theta) &= \frac{a}{r} \\ \sin(\theta) &= \frac{b}{r} \end{aligned}$$

Thus if you know r and θ you can find a and b by

$$\begin{aligned} a &= r\cos(\theta) \\ b &= r\sin(\theta) \\ a + bi &= r\cos(\theta) + ir\sin(\theta) \;=\; r(\cos(\theta) + i\sin(\theta)) \end{aligned}$$

The last expression is called the *Polar form of a complex number*. If you have a scientific calculator it can give you the values of $\cos(\theta)$ and $\sin(\theta)$ and if it is good scientific calculator it may well have a rectangular-polar conversion button that will go from x, y to r, θ and from r, θ to x, y.

At this point we will take a brief look at the graph of the sin function. We have been measuring the angles in Babylonian degrees, but for graphs this is useless and we measure the x axis in *radians*. Radians are explained more completely in the appendix, but for what we need here we only need to know that 180 degrees is π radians. Thus, by proportion, $90^\circ = \pi/2$ radians and $360^\circ = 2\pi$ radians.



Graph of the function sin(x)

The essential feature of this graph is that it is periodic. This is the reason the sin function is so important; it allows us to control things like springs, strings and airplane wings [12], in fact anything that vibrates. Other things that are important is that the period of the vibration is 2π and the function has outputs between -1 and 1 for real inputs. We see that $\sin(\pi/2) = 1$ and $\sin(\pi) = 0$. The cosine has the same shape for the graph but the graph is shifted $\pi/2$ to the left, so $\cos(0) = 1$ and $\cos(\pi/2) = 0$.

The periodic behavior is captured in the equation

$$\sin(x + 2\pi) = x \qquad \text{for all } x$$

Note that there is no *real* number r so that $\sin(r) = 2$ because the graph never gets that high when the input is real. Some teachers will tell you that $\sin(x) = 2$ has no solution. This is just like $x^2 = -1$ has no solution; what they mean is these equations have no *real* solutions. They do have complex number solutions, and in the problems you will solve $\sin(x) = 2$. Some calculators can handle this complex input to the sine function and some cannot, and some have to have their mode reset to handle it.

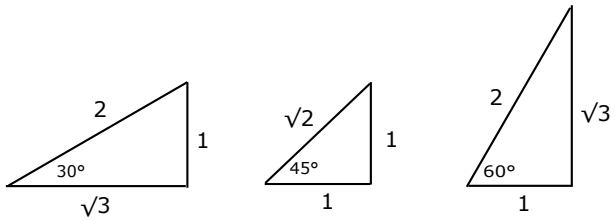
A notation sometimes used in engineering is

$$a + bi = r\angle\theta = r(\cos(\theta) + i\sin(\theta))$$

so we could think of $1\angle\theta = \angle\theta = \cos(\theta) + i\sin(\theta)$

[12] Examples due to Peter Lax.

Now we need some examples. There are three triangles that are useful for this and they are shown in the picture below. The $30°, 90°, 60°$ triangle[13] has sides $\sqrt{3}, 1, 2$, the $45°, 90°, 45°$ triangle has sides $1, 1, \sqrt{2}$ and the $60°, 90°, 30°$ triangle has sides $1, \sqrt{3}, 2$.



Favorite triangles

From the picture we can immediately write down $\cos(30°) = \sqrt{3}/2$ and $\sin(30°) = 1/2$ and thus $\sqrt{3} + i = 2\angle 30°$. Using all three triangles we have

$$\sqrt{3} + i = 2\angle 30° \qquad 1 + i = \sqrt{2}\angle 45° \qquad 1 + i\sqrt{3} = 2\angle 60° \qquad i = 1\angle 90°$$

Addition of complex numbers written in the polar form $r\angle\theta$ is not possible in any reasonable way. Addition must be done in rectangular $a + bi$ form. But multiplication in polar form is *easier* than in rectangular form. Suppose

$$a + bi = r\angle\theta \quad \text{and} \quad c + di = s\angle\phi$$

We look up in a mathematical handbook the addition formulas for cos and sin which are[14]

$$\begin{aligned} \cos(\theta + \phi) &= \cos(\theta)\cos(\phi) - \sin(\theta)\sin(\phi) \\ \sin(\theta + \phi) &= \sin(\theta)\cos(\phi) + \cos(\theta)\sin(\phi) \end{aligned}$$

It is unusual for functions to have addition formulas, so you should remember these exist for later use. We indicate a proof of them in the appendix. Here we will just use them in the following calculation.

$$\begin{aligned} r\angle\theta \cdot s\angle\phi &= r(\cos(\theta) + i\sin(\theta)) \cdot s(\cos(\phi) + i\sin(\phi)) \\ &= rs(\cos(\theta)\cos(\phi) - \sin(\theta)\sin(\phi)) + i(\sin(\theta)\cos(\phi) + \cos(\theta)\sin(\phi)) \\ &= rs(\cos(\theta + \phi) + i\sin(\theta + \phi)) \end{aligned}$$

so when you multiply complex numbers in polar form you *multiply* the moduli r and s and you *add* the arguments θ and ϕ. Note that if $r = s = 1$ this comes to $\angle\theta \cdot \angle\phi = \angle(\theta + \phi)$. If we repeat this over and over we have de Moivre's[15]

[13] To get the numbers chop an equilateral triangle of side 2 in half.

[14] ϕ is now generally pronounced fee not fy in mathematics.

[15] French 1667-1754. A Huguenot(French Protestant), he was forced by religious prosecution to leave France. The French edict of 1685 revoking toleration of Protestants scattered Huguenot mathematicians abroad and raised the general level of mathematics in several countries. De Moivre spent the the second half of his life in England and became a friend of Halley and Newton. However, due to English prejudice against the French he could never get a University job.

theorem

$$(\angle\theta)^n = \angle(n\theta) \qquad (\cos(\theta) + i\sin(\theta))^n = (\cos(n\theta) + i\sin(n\theta))$$

Let's do an example of this. From the previous example we see that

$$\frac{1+i}{\sqrt{2}} = 1\angle 45^\circ$$

If we square this we have

$$\left(\frac{1+i}{\sqrt{2}}\right)^2 = (\angle 45^\circ)^2 = \angle 90^\circ = i$$

so without the slightest effort we have found the solution of $x^2 = i$, that is we have found the square root of i. Naturally the other one is just the negative.

The square roots of i are

$$\sqrt{i} = \frac{1+i}{\sqrt{2}}, \quad -\frac{1+i}{\sqrt{2}} = \frac{-1-i}{\sqrt{2}}$$

Our method for finding the second square root was crude; let's see if we can find a better method. Recall that adding 360° to an angle does not change the position of the angle. So let's take for the position of i not 90° but $90^\circ + 360^\circ = 450^\circ$. If we cut this angle in half to get the square root we have $\angle(450/2)^\circ = \angle 225^\circ$. So the second square root of i should be

$$\angle 225^\circ = \cos(225^\circ) + i\sin(225^\circ) = -\frac{1}{\sqrt{2}} - i\frac{1}{\sqrt{2}}$$

which is what we got before, as you can see by drawing a triangle for 225° and noting it is the same shape as for 45° but is in the third quadrant so x and y are negative.

The method we used for the second square root, taking the negative, can be generalized. If we have one n^{th} of a number we can find all the others by multiplying by the n^{th} roots of 1. In the square root case the square roots of 1 are 1 and -1, which explains taking the negative. For cube roots multiply one cube root of a number by the cube roots of 1, (see below)

$$1 \qquad \omega = \frac{-1+i\sqrt{3}}{2} \qquad \omega^2 = \frac{-1-i\sqrt{3}}{2}$$

to get all three cube roots of a number.

For example the cube roots of 8 are

$$2 \qquad 2\omega = -1+i\sqrt{3} \qquad 2\omega^2 = -1-i\sqrt{3}$$

If you remember someone telling you that 8 has only *one* cube root they were talking about real numbers only. You might run them down and show them the

other two. It would be good for you to cube 2ω by hand and see that you really do get 8.

Looking at the formula $(r\angle\theta)^n = r^n\angle n\theta$ we see

$$(\sqrt[n]{r}\angle\frac{\theta}{n})^n = r\angle\theta$$

and thus we have a formula for one of the n^{th} roots of a complex number written in polar form. To get all the roots there are two choices. One can replace θ in the formula by each of

$$\theta + 0\cdot 360^\circ \quad \theta + 1\cdot 360^\circ \quad \theta + 2\cdot 360^\circ \quad \ldots \theta + (n-2)\cdot 360^\circ \quad \theta + (n-1)\cdot 360^\circ$$

or one can multiply the one value you have by each of the n^{th} roots of 1.

A couple of examples. The cube roots of one are

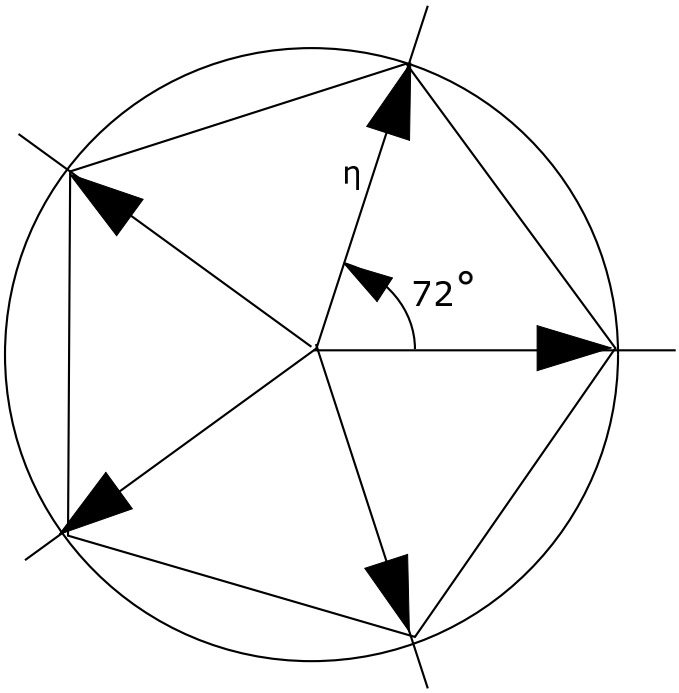
$$\begin{aligned} \angle 0 &= \cos(0) + i\sin(0) = 1 \\ \omega = \angle 120^\circ &= \cos(120^\circ) + i\sin(120^\circ) = -\frac{1}{2} + i\frac{\sqrt{3}}{2} \\ \omega^2 = \angle 240^\circ &= \cos(240^\circ) + i\sin(240^\circ) = -\frac{1}{2} - i\frac{\sqrt{3}}{2} \end{aligned}$$

The fourth roots of 1 are of course $\angle 0^\circ = 1$, $\angle 90^\circ = i$, $\angle 180^\circ = -1$, $\angle 270^\circ = -i$. The fifth roots of 1 are much more complicated to do completely, although the decimal values are no problem. We have $360^\circ/5 = 72^\circ$ so the first fifth root is

$$\eta = \angle 72^\circ = \cos(72^\circ) + i\sin(72^\circ) = 0.309017 + 0.951057i$$

(be sure to set your calculator in degree mode; if you are getting -.9672505 you are in radian mode; change the mode to degree) and the others are

$$\begin{aligned} \eta^2 &= \angle 2\cdot 72 = \angle 144^\circ = -0.809017 + 0.587785i \\ \eta^3 &= \angle 3\cdot 72 = \angle 216^\circ = -0.809017 - 0.587785i \\ \eta^4 &= \angle 4\cdot 72 = \angle 288^\circ = \ \ 0.309017 - 0.951057i \\ \eta^5 &= \angle 5\cdot 72 = \angle 360^\circ = 1 \end{aligned}$$



The fifth roots of one

Notice how I have drawn a pentagon between connecting the noses of the fifth roots. Gauss noticed this when he was 18 and then solved a problem that had been open for about 2000 years. To describe this I am now going to solve the fifth root of 1 problem algebraically. You might want to skim till the answer.

We start with $z^5 - 1 = 0$ and factor easily, since we know $z - 1$ is a factor, to get $(z-1)(z^4 + z^3 + z^2 + z + 1) = 0$ If $z^4 + z^3 + z^2 + z + 1$ factors it must factor into $(z^2 + az + 1)(z^2 + bz + a)$. I multiply these out and get

$$z^4 + (a+b)z^3 + (2+ab)z^2 + (a+b)z + 1 = z^4 + z^3 + z^2 + z + 1$$

which gives, by comparing coefficients, $a + b = 1$ and $2 + ab = 1$. I eliminate b by substituting $b = 1 - a$ into $ab = -1$ to get $a^2 - a - 1 = 0$ and get

$$a = \frac{1+\sqrt{5}}{2} \qquad b = 1 - a = \frac{1-\sqrt{5}}{2}$$

using the quadratic formula. This gives

$$\left(z^2 + \frac{1+\sqrt{5}}{2}z + 1\right)\left(z^2 + \frac{1-\sqrt{5}}{2}z + 1\right) = z^4 + z^3 + z^2 + z + 1 = 0$$

as the equation for the other four fifth roots. I want the equation for η which has positive real part. The second solution will be $\overline{\eta} = \eta^4$ since the roots must multiply to one and be conjugate. The two roots add to a positive number

which must equal the negative of the coefficient of z. This means η is a root of $z^2 + \frac{1-\sqrt{5}}{2}z + 1 = 0$. We use the quadratic formula again on this equation and after some trivial simplification we get the two solutions.

$$\begin{aligned} z &= \frac{1}{4}\Big(-1+\sqrt{5}+i\sqrt{2(5+\sqrt{5})}\Big) \quad \text{or} \\ z &= \frac{1}{4}\Big(-1+\sqrt{5}-i\sqrt{2(5+\sqrt{5})}\Big) \end{aligned}$$

η is the one with the positive complex part so the first one. If we calculate the decimal for η we get

$$\eta = \frac{1}{4}\Big(-1+\sqrt{5}+i\sqrt{2(5+\sqrt{5})}\Big) = 0.309017 + 0.951057i$$

which was the value we got from trigonometry.

Now for the big news. Our fifth root η consists of rational numbers and square roots of numbers built from rational numbers and square roots. It is easy to show that, having chosen a unit length line, the only lines that can be constructed from that line with ruler and (unmarked) straightedge are lines of length made from rational numbers and square roots. Thus the formula for η shows that we can construct the pentagon with ruler and compass. This was known to the ancients who did it in a different way.

The open problem since antiquity was what other p-gons with prime p could be constructed with ruler and compass. This is the same problem as what other p-roots could be expressed using nothing but rational operations and square roots, like we did for $p = 5$. The 18 year old Gauss proved that the constructible p-gons where precisely the Fermat Primes, that is those of the form

$$p = 2^{2^n} + 1$$

Fermat conjectured that all numbers of this form were prime, but this is not at all the case. In fact the only known Fermat primes are

n	0	1	2	3	4
$2^{2^n}+1$	3	5	17	257	65537

We think that all subsequent numbers in the series are composite, not prime, and this has been checked by computer. This is one of those things that each new and better computer is set to checking up to the limit of the number of digits it can handle. No other primes have ever turned up. Fairly quickly after the 18 year old Gauss published this result, a construction of the 17-gon was found. In fact, you can sort of see from our construction of the 5-gon how one might attack the problem. The fact that a 17-gon *could* be constructed was a complete surprise; no one had ever suspected this was possible. Eventually the 257-gon was constructed. At least one person invested a large part of his life attempting unsuccessfully to construct the 65537-gon. As far as I am aware no one has yet succeeded. Here is the ultimate homework problem.

The takeaway here is that there are deep and surprising connections between the complex numbers and geometry.

Another connection with geometry is that complex numbers admit of dual aspect; one aspect is as vectors in the plane, a second aspect is as rotators. We have seen that

$$1\angle\theta \cdot r\angle\phi = r\angle(\theta + \phi)$$

which can be interpreted to mean that $1\angle\theta = \cos(\theta + i\sin(\theta))$ is a complex number that *rotates* every vector in the plane by the angle θ[16]. For example $\angle 45^\circ = \frac{\sqrt{2}}{2}(1+i)$ rotates every vector by 45° counterclockwise while $\angle 90^\circ = i$ rotates every vector by 90° counterclockwise. Note that these rotator vectors are all of unit length. While occasionally useful, rotations in the plane are pretty easy to control, so this is not a matter of great importance. However, rotations in 3-space are another matter entirely; they are very complicated to control! We take this up in the next chapter.

Another matter of great importance is illustrated by the process by which we got here: natural numbers, integers, rational numbers, real numbers, complex numbers. In almost all cases we were pushed to the next level by the insolvability of an equation. For example $x + 2 = 0$ requires negative integers, and $x^2 = 2$ requires real numbers. Finally we had $x^2 = -1$ which forced us to invent, or discover, i. Most people when they get this far wonder about $\sqrt{i}$ but as we saw above this is just $(1+i)/\sqrt{2}$ so no problem there. In fact the process of solving equations will never take us any further, because of the

Fundamental Theorem of Algebra (Gauss) If a polynomial $p(z)$ has complex coefficients then the equation $p(z) = 0$ has a root which is a complex number.

This guarantees that an equation like

$$(2+\sqrt{3}i)z^3 + (\pi - 3\sqrt{17}i)z^2 + \exp(2\sqrt[3]{2})z + 107 = 0$$

will have a solution z which is a complex number. There exist methods for finding numerical solutions for z which nowadays are implemented on computers.

It is easy to show, using the fundamental theorem of algebra, that if the polynomial $p(z)$ has order n (which implies that the coefficient of z^n is not 0) then $p(z)$ has exactly n roots, which may not be distinct. This really means that there are n complex numbers $r_1, r_2, \ldots, r_n$, possibly not all distinct, so that

$$p(x) = a_n(x - r_1)(x - r_2)\cdots(x - r_n)$$

Thus the process of adding numbers to an existing system to get a new system so that we can solve an equation ends with the complex numbers. This suggests that the complex numbers are a natural stopping point in the development of numbers, and should be considered the basic numbers of mathematics. Mathematicians divide up into three sets depending on whether they accept the proceeding sentence. One set considers the complex numbers as basic and the real numbers as inhabitants of a particular line that runs through the complex

[16] In old times this was called a *versor*.

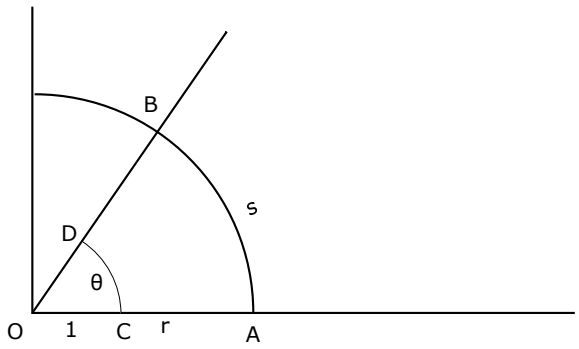
plane (the real axis). Another set considers the real numbers as the basic mathematical numbers and the complex numbers as an extension for special purposes. The third set (small) considers the algebraic numbers as the only philosophically justified numbers and all others as tools for physics and engineering. Kronecker would have fallen into this third set. To illustrate the point, there is story where the two friends, Weierstrass and Kronecker, met on the campus of the University of Berlin. Weierstrass said "Kronecker, have you heard?, Lindemann has proved that π is not an algebraic number." Kronecker replied "Very interesting. Then π does not exist."

Although we cannot go into it here, there are tight connections between complex numbers and Physics. Although complex numbers are occasionally used in classical physics for convenience, they are not essential there. The situation is quite different in Quantum physics, where complex numbers are built into the very fabric of the theory. Real numbers simply do not suffice for a quantum description of anything. Your TV, your cell phone, your stereo system, your computer and increasingly also things like your refrigerator, are quantum devices and their mathematics is based on the complex numbers. Be grateful. Without complex numbers instead of playing your ipod you would be hand cranking a phonograph with a big horn to play vinyl records.

2.7 Appendix: Radian measure and Euler's formula

This section will be a little unsatisfying for you because we must make use of results which we cannot prove. Nevertheless the results are sufficiently interesting that it's worth departing from our usual standards.

First we will introduce radian measure. The degree measure we usually use for measuring angles is left over from Babylonian astronomy. It works pretty well for most purposes but there are some things for which it is unsuitable, as we will see.



Radian Measure of θ is s/r

The radian measure of an angle θ is defined to be the length cut off by the angle θ on the unit circle which is the length of the arc CD of the unit circle in the picture. The arc AB of length s cut off by the angle θ on the circle of radius r has by proportion the value $s = r\theta$, and thus the radian measure of θ is s/r. It is also useful to know that the area of the sector OAB is $\frac{1}{2}r^2\theta$ using the radian measure of θ.

We can find this length easily by proportion: the length around a circle is $2\pi r$ so if $r = 1$ the length around the circle is 2π and this is the radian measure of a complete circle, and the Babylonian angle around the circle is 360° so we have the equation for the angle measurement $2\pi = 360^\circ$. Thus $\frac{2\pi}{360^\circ} = 1$ and we can easily change from degree measure to radian measure as follows:

$$
\begin{array}{lcccll}
0^\circ & = & 0^\circ \cdot \frac{2\pi}{360^\circ} & = & 0 & \text{radians} \\
30^\circ & = & 30^\circ \cdot \frac{2\pi}{360^\circ} & = & \frac{\pi}{6} & \text{radians} \\
45^\circ & = & 45^\circ \cdot \frac{2\pi}{360^\circ} & = & \frac{\pi}{4} & \text{radians} \\
60^\circ & = & 60^\circ \cdot \frac{2\pi}{360^\circ} & = & \frac{\pi}{3} & \text{radians} \\
90^\circ & = & 90^\circ \cdot \frac{2\pi}{360^\circ} & = & \frac{\pi}{2} & \text{radians} \\
120^\circ & = & 120^\circ \cdot \frac{2\pi}{360^\circ} & = & \frac{2\pi}{3} & \text{radians} \\
135^\circ & = & 135^\circ \cdot \frac{2\pi}{360^\circ} & = & \frac{3\pi}{4} & \text{radians} \\
150^\circ & = & 150^\circ \cdot \frac{2\pi}{360^\circ} & = & \frac{5\pi}{6} & \text{radians} \\
180^\circ & = & 180^\circ \cdot \frac{2\pi}{360^\circ} & = & \pi & \text{radians} \\
210^\circ & = & 210^\circ \cdot \frac{2\pi}{360^\circ} & = & \frac{7\pi}{6} & \text{radians} \\
225^\circ & = & 225^\circ \cdot \frac{2\pi}{360^\circ} & = & \frac{5\pi}{4} & \text{radians} \\
240^\circ & = & 240^\circ \cdot \frac{2\pi}{360^\circ} & = & \frac{4\pi}{3} & \text{radians} \\
270^\circ & = & 270^\circ \cdot \frac{2\pi}{360^\circ} & = & \frac{3\pi}{2} & \text{radians}
\end{array}
$$

It is customary for mathematicians to use $\pi/6$ rather than the decimal value 0.52359877559829887307710723054658 as it is easier to understand and using $\pi = 180^\circ$ we can convert from $\pi/6$ to $180^\circ/6 = 30^\circ$ very easily.

Background for the material that follows.

The following material requires a little Calculus for it's logical development. We will simply use the formula taken from reference works. Calculus is divided into three parts, Differential Calculus, Integral Calculus, and Infinite Series. Infinite series is rendered difficult by the question of whether an infinite series converges, which means the finite sums approach a finite value rather than going off to infinity or oscillating around different values. When you take the relevant course, usually a part of Calc II or Calc III, you will learn a variety of tests that prove a series converges. In the following we will meet series which converge due to one of the simplest tests, the ratio test. Moreover, they will converge for any complex number z. So the situation is maximally simple.

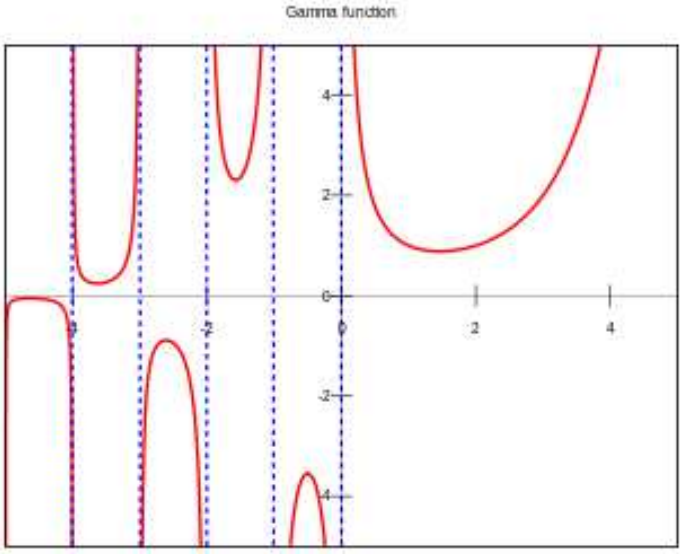
The series may be derived in a number of ways but the usual is to use Taylor's series which gives a general method for expanding functions in series. We will not pursue this but just present the series as they can be found in any Mathematics Handbook.

Infinite series often require the use of factorials, which are simple and worth learning because they play a big part in probability and statistics. Here are

some factorials:

$$\begin{aligned} 1! &= 1 \qquad 2! = 1\cdot 2 = 2 \qquad 3! = 1\cdot 2\cdot 3 = 6 \qquad 4! = 1\cdot 2\cdot 3\cdot 4 = 24 \\ 5! &= 1\cdot 2\cdot 3\cdot 4\cdot 5 = 120 \qquad 6! = 1\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6 = 720 \qquad 7! = 7\cdot 6! = 5040 \\ 8! &= 40320 \qquad 9! = 362880 \qquad 10! = 3628800 \qquad 11! = 39916800 \end{aligned}$$

It is very interesting although irrelevant to our purposes that there is a function Γ which gives the values of $n!$ by $n! = \Gamma(n+1)$. Here is the graph of the Γ function:



Graph of the Gamma Function

Note that the size of $n!$ grows very quickly. This is most important because it makes the terms in the infinite series become small quickly and this helps the series to converge. As an application we note that $n!$ is the number of ways to rearrange a sequence of n distinct letters, so for example a,b,c,...,j,k can be rearranged in 11! = 39,916,800 ways. You might want to test this with fewer letters[17].

In any mathematical handbook you will find the following formulas. It is *critical* that in substituting into these series that z be measured in *radians*.

$$\begin{aligned} \cos(z) &= 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \frac{z^6}{6!} + \frac{z^8}{8!} - \frac{z^{10}}{10!} + \dots \\ \sin(z) &= z - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} + \frac{z^9}{9!} = \frac{z^{11}}{11!} + \dots \end{aligned}$$

[17] We note for the philosophically inclined that 7! =5040 is the number of citizens in Plato's ideal state since the citizenry can be divided into teams of size n for $n = 1, 2, 3, \dots 10$

So for example to find $\sin(30^\circ)$ we must use $z = \frac{\pi}{6} \approx 0.523599$. Then we have, with the approximations getting better as we go down,

$$\begin{aligned}
\sin\frac{\pi}{6} &\approx \frac{\pi}{6} = .523599 \\
\sin\frac{\pi}{6} &\approx \frac{\pi}{6} - \frac{(\frac{\pi}{6})^3}{3!} = 0.499674 \\
\sin\frac{\pi}{6} &\approx \frac{\pi}{6} - \frac{(\frac{\pi}{6})^3}{3!} + \frac{(\frac{\pi}{6})^5}{5!} = 0.500002 \\
\sin\frac{\pi}{6} &\approx \frac{\pi}{6} - \frac{(\frac{\pi}{6})^3}{3!} + \frac{(\frac{\pi}{6})^5}{5!} - \frac{(\frac{\pi}{6})^7}{7!} = 0.5
\end{aligned}$$

The last calculation with 4 terms appears to be exact but that is because the calculator ran out of digits. Any finite number of terms can never be exact, but you notice it can be exact enough for all practical purposes. Did you ever wonder how tables of trigonometric functions were calculated or how your calculator does it? This is one of the ways it could be done.

Incidentally the formulas above will work for any z in $\mathbb{C}$, not just for real ones. If your calculator or computer algebra program can do complex numbers you might want to try $\sin(1.5708 - 1.31696i)$ for a surprise. If you want to do this as I did above, using the series, you must use more terms than I used above; you need to go out to $z^{11}/11!$ to get a good approximation of the answer. In general, the bigger the input (in terms of r) the more terms you need.

For all science and engineering there is one more function that is critical, called the exponential function. It is based on the number

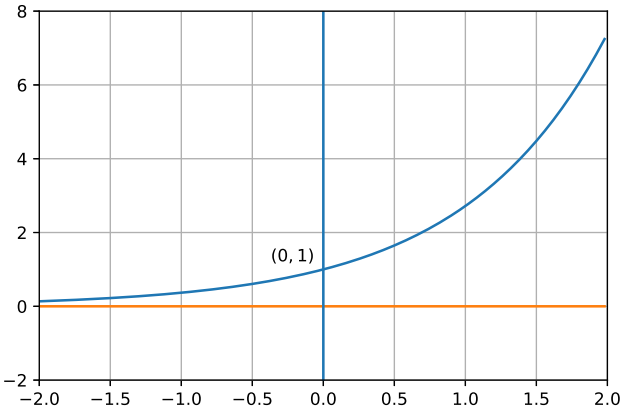
$$e = 2.7182818284590452354\ldots$$

which is not an algebraic number (like π) and the decimals go on forever and do not repeat. You might wonder why anyone would choose such a number but the answer requires differential equations. Essentially it is because $w = e^z$ is a solution of one of the simplest differential equations $\frac{dw}{dz} = w$ and from there great numbers of other properties and applications follows. The number e is precisely the right number to make it work. For real numbers $y = e^x$ the differential equation says the the slope of the tangent line is equal to the y coordinate, but this is not much help. When people say something is *growing exponentially* they mean its graph is like e^{kx} with some constant k

It is often typographically inconvenient that the variable z is the exponent in e^z so we have another notation for e^z to make typesetting easier. It is

$$\exp(z) = e^z$$

We will use both to get you used to it.



Graph of the Exponential Function

We note that there is no symmetry to the graph and that it decreases very quickly as x becomes large and negative and increases very rapidly as x becomes large and positive. For example, $\exp(-5) = .00673\ldots$ and $\exp(5) = 148.41315\ldots$. Another fact of great importance is *the exponential function $exp(z)$ never outputs* 0 for any input $z \in \mathbb{C}$.

Looking carefully at the graph we see that for any number $y > 0$ we can find an x for which $e^x = y$. To do this graphically, find the point y on the vertical axis and draw a horizontal line through y. Find where this horizontal line hits the graph, drop down vertically to the horizontal axis and you have found x. There is a function, ln on your calculator which performs this duty. The definition is

Def For real numbers x and $y > 0$

$$x = \ln y \text{ if and only if } e^x = y$$

These can also be written as

$$\ln(e^x) = \ln(\exp(x)) = x \qquad e^{\ln y} = \exp(\ln y) = y \text{ for } y > 0$$

We say that e^x maps the $\mathbb{R}$ onto the positive reals $\mathbb{R}^+$ and it does this in a one to one manner; for each x there is only one $y > 0$ and for each $y > 0$ there is only one x.

Now an example. Suppose you wish to solve $e^x = 3$. Then from the definition we must have $x = \ln 3$. Pull out your calculator and find $\ln 3$ getting $1.098612288\ldots$. Next (using the exp or e^x button) find $e^{1.098612288}$ and you get 3. If you didn't get exactly 3 it was because you didn't use the whole decimal for $\ln 3$ but stopped it at the 9th place. If you want an exact example solve $e^x = 1$ by finding $\ln 1 = 0$ exactly. We say that exp and ln are *inverse functions* of one another.

The function ln is called the (natural) logarithm function but few younger people nowadays say natural. (ln stands for *l*ogarithmus *n*aturalis, latin for natural logarithm.) For all *mathematical* purposes it is the *only* logarithm function. However you will notice a button labeled log on your calculator. *NEVER PUSH THIS BUTTON* as it will get you the wrong answer in math classes. This function log has the same relation to 10^x and ln has to e^x. These two functions, 10^x and $\log x$ were popular in my youth (except we had no calculators and could not push buttons; we had to use tables). The functions 10^x and $\log x$ are now used by four kinds of people. 1) Chemists calculating pH 2) Audiologists calculating decibels (a measure of loudness) 3) Astronomers calculating stellar magnitude and 4) Geologists calculating the strength of an earthquake on the Richter scale. If you are not one of these people enjoying one of these tasks, DO NOT PUSH THE log BUTTON.

Now we ask what is the infinite series for e^z. It is (get out the math handbook again)

$$\exp(z) = e^z = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \frac{z^4}{4!} + \frac{z^5}{5!} + \frac{z^6}{6!} + \frac{z^7}{7!} + \frac{z^8}{8!} + \frac{z^9}{9!} + \dots$$

This formula works for all complex numbers but the bigger the r the more terms you have to take. Let's try $z = \frac{\pi}{6}i \approx 0.523599i$ and calculate a few terms.

$$\begin{aligned}
e^{.523599i} &\approx 1 + z = 1. + 0.523599i \\
e^{.523599i} &\approx 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} = 0.862922 + 0.499674i \\
e^{.523599i} &\approx 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \frac{z^4}{4!} + \frac{z^5}{5!} = 0.866054 + 0.500002i \\
e^{.523599i} &\approx 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \frac{z^4}{4!} + \frac{z^5}{5!} + \frac{z^6}{6!} + \frac{z^7}{7!} = 0.866025 + 0.5i
\end{aligned}$$

Now those numbers look familiar. We recall that $\cos(30^\circ) = \cos\frac{\pi}{6} = \frac{\sqrt{3}}{2} \approx .866025$ and $\sin(30^\circ) = \sin\frac{\pi}{6} = \frac{1}{2} = .5$ So what we have shown is that

$$\exp\left(i\frac{\pi}{6}\right) = e^{i\frac{\pi}{6}} = \cos\left(\frac{\pi}{6}\right) + i\sin\left(\frac{\pi}{6}\right)$$

This is an example of Euler's formula[18]. In general

$$\exp(i\theta) = e^{i\theta} = \cos(\theta) + i\sin(\theta)$$

It is interesting that such a famous and important formula is quite easy to prove. Recall that

$$i^{4k+j} = i^{4k}i^j = (i^4)^k i^j = 1^k i^j = i^j$$

[18]There are many different formulas called Euler's formula. This is probably the most important.

so that i^n cycles through the four values $i, -1 = i^2, -i = i^3, 1 = i^4$. Thus

$$\begin{aligned} e^{i\theta} &= 1 + i\theta + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \frac{(i\theta)^4}{4!} + \frac{(i\theta)^5}{5!} + \frac{(i\theta)^6}{6!} + \frac{(i\theta)^7}{7!} + \frac{(i\theta)^8}{8!} + \dots \\ &= 1 + i\theta - \frac{\theta^2}{2!} - i\frac{\theta^3}{3!} + \frac{\theta^4}{4!} + i\frac{\theta^5}{5!} - \frac{\theta^6}{6!} - i\frac{\theta^7}{7!} + \frac{\theta^8}{8!} + \dots \\ &= (1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \dots) + i(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \frac{\theta^7}{7!} + \dots) \\ &= \cos(\theta) + i\sin(\theta) \end{aligned}$$

I want to emphasize that in the real domain the exponential function e^x and the trigonometric functions $\sin(\theta)$ and $\cos(\theta)$ have very little to do with each other. However, as Euler's formula shows, in the complex domain they are intimately related. This relationship suggests that the periodicity of sine and cosine might have some effect on the exponential function. In fact

$$\exp(z + 2\pi i) = e^{z+2\pi i} = e^z e^{2\pi i} = e^z(\cos 2\pi + i \sin 2\pi) = e^z \cdot 1 = e^z$$

Thus the exponential function has a *complex* period which of course we can't see in the real graph above. Hence there are always infinitely many complex solutions z to the equation $e^z = w$ for any complex $w \neq 0$. In fact, if $w = r\angle\theta = r(\cos\theta + i\sin\theta)$ then

$$\text{The solution of } e^z = w \text{ is } z = \ln r + i\theta + 2n\pi i$$

where $\ln r$ is the natural logarithm of r and n is any integer. We will look at this more carefully in the problems. We can write this more efficiently by extending the function ln to complex inputs so we then have

$$\text{For } z = r\angle\theta \text{ we have } \ln z = \ln r + i\theta + 2n\pi i$$

For fun let's do a couple of examples.

$$\begin{aligned} \ln 1 = \ln(1\angle 0) &= \ln 1 + 0i + 2n\pi i = 2n\pi i \\ \ln i = \ln(1\angle\pi/2) &= \ln 1 + \frac{\pi}{2}i + 2n\pi i = \frac{\pi}{2}i + 2n\pi i \\ \ln(-1) = \ln(1\angle\pi) &= \ln 1 + \pi i + 2n\pi i = \pi i + 2n\pi i \\ \ln(-2) = \ln(2\angle\pi) &= \ln 2 + \pi i + 2n\pi i = .693147\ldots + \pi i + 2n\pi i \\ \ln((1+i) = \ln(\sqrt{2}\angle\pi/4) &= \ln\sqrt{2} + \frac{\pi}{4}i + 2n\pi i = \frac{1}{2}\ln 2 + \frac{\pi}{4}i + 2n\pi i \end{aligned}$$

So much for the notion that negative numbers have no logarithms. The do not have *real* logarithms but they certainly have logarithms which are complex numbers.

We will derive one more set of formulas that emphasizes what one can do with Euler's formula. First however we derive and important trigonometric formula. Note first that the series for $\cos(z)$ has only terms with even powers

of z, and thus $\cos(z)$ is an even function which means that $\cos(-z) = \cos(z)$. Similarly the series for $\sin(z)$ has only terms with odd powers of z and thus $\sin(z)$ is an odd function which means that $\sin(-z) = -\sin(z)$. This is just as true for complex inputs, as we have proved, as for real inputs and these are very important facts. Using these we show immediately that

$$\exp(-iz) = e^{-iz} = \cos(-z) + i\sin(-z) = \cos(z) - i\sin(z)$$

and from this we have

$$\begin{aligned}
\exp(iz) = e^{iz} &= \cos(z) + i\sin(z) \\
\exp(-iz) = e^{-iz} &= \cos(z) - i\sin(z) \\
e^{iz} + e^{-iz} &= 2\cos(z) \qquad \text{adding the first two equations} \\
e^{iz} - e^{-iz} &= 2i\sin(z) \qquad \text{subtracting the first two equations}
\end{aligned}$$

from which we get the two important formulas

$$\begin{aligned}
\cos(z) &= \frac{e^{iz} + e^{-iz}}{2} \\
\sin(z) &= \frac{e^{iz} - e^{-iz}}{2i}
\end{aligned}$$

Also notice that[19]

$$\begin{aligned}
1 &= e^0 = e^{z+(-z)} = e^z \cdot e^{-z} = (\cos(z) + i\sin(z))(\cos(z) - i\sin(z)) \\
&= \cos^2(z) - (i^2)\sin^2(z) = \cos^2(z) + \sin^2(z)
\end{aligned}$$

which is one of the most important trig formulas. It is also a disguised form of the Pythagorean theorem. There are hundreds of trig formulas which can be derived this way. See if you can derive these important addition formulas in a similar way by looking at e^{u+v} in two ways [20].

$$\begin{aligned}
\cos(u+v) &= \cos(u)\cos(v) - \sin(u)\sin(v) \\
\sin(u+v) &= \sin(u)\cos(v) + \cos(u)\sin(v)
\end{aligned}$$

I hope you have enjoyed this small selection of the wonderful world of complex function theory. There is a junior-senior level college course called *Complex Function Theory* or *Complex Analysis* which does much more along this line. Sadly the most wonderful things I cannot really present at this stage of your education, so don't miss this course if you have the chance. There are hundreds of books on Complex Analysis and while you are taking the course, in addition to your own book, have a look at the really nice two volume set Theory of Functions vols I, II by Konrad Knopp. After finishing the course you can then read the beautiful but more difficult book Complex Analysis by Lars Alfors.

There are many scientific applications of the theory of complex variables and there are many fine books on this. One of my favorites is Polya and Latta's book Complex Variables. Annoyingly the wonderful book by Shabat and Fuks is still available only in Russian.

[19] Out of ancient habit $(\sin(z))^2$ is written $\sin^2(z)$

[20] Addition formulas like these for functions are very rare and precious

2.8 Problems for Chapter 2

Section 2.1

1. Let's do a little astrology. If you are a Leo you were born between 23 July and 22 Aug. Why is that? Because 2000 years ago at that period in the year the sun was in the constellation Leo. (You couldn't see Leo because the sun overwhelmed the starlight but they had good maps of the stars so they knew.) Now if you have ever played with tops you know that as they slow down they start to wobble. This is called *procession* and is very important in physics, even quantum physics. The Earth behaves in many ways like a top and also has procession, called the *procession of equinoxes*. This causes the constellations to slowly move around the sky for a given date in the year. The period of procession is 25,920 years, so after 25,920 years the constellations are back in the position they started. The system was started up approximately 4500 years ago, but the birthday days were firmed up around the beginning of our era. Thus the dates determining your sign were determined about 2000 years ago. Thus the constellations have shifted position by

$$\frac{2000}{25,920} \cdot 360^\circ \text{ degrees approximately}$$

a) How many degrees is this? b) how many degrees in the sky does each constellation take up? (There are 12 constellations, one for each month.) c. So approximately how many constellations off the original are we now? d) Here is the constellation list:

Aries, Taurus, Gemini, Cancer, Leo, Virgo, Libra,
Scorpio, Sagittarius, Capricorn, Aquarius, and Pisces.

Aries goes from 15 April to 15 May and the system marches through the other constellations in order with similar dates. The other thing you need to know is the direction of wobble of the Earth, which is such that the if your Babylonian sign is somewhere in the list, the sun when you were born was actually in a constellation to the left. How far to the left? You already answered this in question "c". e) So in what constellation was the sun when you were born? f) Does it seem likely given what you now know that your Babylonian astrological sign accurately predicts anything about you?

Section 2.2

1. Find the decimal representation of the Babylonian sexagesimal fraction 2,20. Notice this is a little ambiguous. It could also be the representation of $\frac{7}{180}$. Why? The Babylonians told them apart by context, but it was always a bit of a problem for them.

2. Find the fraction representation of the Babylonian sexagesimal fraction 2,20,14. What is the decimal representation of this number. Remember the use of the overbar for repetition!

3. Find the sexagesimal and decimal representation of 5/7. Use the overbar for repetition.

4. Enthusiasts only. a) Using a calculator fine the sexagesimal representation of π. Use the decimal representation 3.1415926535. Answer: 3, 8, 29, 44, 0, 47, 25, 53, 7, 24,... but you probably won't get what I list as the answer because the sexagesimal representation is more accurate than the decimal I gave you for pi, which is rounded off in the last digit. b) If you are computer literate write a program to change a fraction or decimal into a sexagesimal expansion. Python or Mathematica will do it relatively easily.

5. We are now going to find the fraction correspoinding to a repeating decimal. This problem will be concerned with a repeating decimal in (0,1], that is $0 <$ decimal ≤ 1. An example should make this clear. Let us take for example .132513251325.....=.$\overline{1325}$. The first thing to know is the multiplier. Since there are 4 digits in the repeated part we use 1 followed by 4 0's, which is 10000. We then have:

$$\begin{aligned} x &= .132513251325132513251325..... \\ 10000x &= 1325.132513251325132513251325....\text{(first line times 10000)} \\ 10000x - x &= 1325.132513251325132513251325..... - .132513251325132513251325..... \\ 9999x &= 1325 \\ x &= \frac{1325}{9999} \end{aligned}$$

It is not so easy to tell if this fraction reduces or not. We take up this question in Chapter 6.
a) use this procedure to find the fraction for .333333333333...
b) same for .4545454545.... Notice you can reduce the fraction here.
c) same for .5151515151.... Reduce the fraction
d) same for .58258258258.... Reduce the fraction.

6. Now for 23.1734545454545.... We already know what x =.45454545.... is from b) above. Hence we have 23.1734545454545.... = 23 + 173/1000 + x/1000 since x/1000 = .00045454545.....
a) Find the fraction for 27.13265151515151.....
b) Check all the answers for 5. and 6. using a hand calculator or computer. Many computers now have little calculators built into them somewhere.

Section 2.3

1. Draw a horizontal line on a sheet of paper. a) Choose a zero point and a unit length and mark off 1,2,3,4,5 and then -1,-2,-3,-4,-5. Label the points with the numbers (and 0). b) Locate, draw the points (tiny filled in circles)

and label the points for the numbers. 1/2, -3/2, 4/3, two and two fifths, -13/4.

2. An important function in number theory is the floor function. The definition is, for a real number x,

The floor of $x = \lfloor x \rfloor$ = The largest integer less than or equal to x

a) Determine the floors: $\lfloor 1 \rfloor$, $\lfloor 2 \rfloor$, $\lfloor 0 \rfloor$, $\lfloor -1 \rfloor$, $\lfloor -1 \rfloor$, $\lfloor -2 \rfloor$
b) Careful; there's a twist. Determine the floors (refer to the line in problem 1) $\lfloor 1/2 \rfloor$, $\lfloor 4/3 \rfloor$, $\lfloor 12/5 \rfloor$, $\lfloor 2.25 \rfloor$, $\lfloor -1/2 \rfloor$, $\lfloor -13/4 \rfloor$, $\lfloor -2.3333 \rfloor$, $\lfloor -3.6 \rfloor$,
c) Draw a graph of $y = \lfloor x \rfloor$. Put a solid dot on the graph above each integer indicating the correct value.
d) Describe in words how to find $\lfloor x \rfloor$ using words.

3. a)Give 1000 decimals that are rational numbers and that all lie in the closed interval $[3.77, 3.78]$.
b) Now using the fact that $\sqrt{2}$ is irrational and the answer to a) give describe 1000 numbers which are irrational and lie in the same interval. Be careful not to overrun the right end.
c) Keeping in mind that there is nothing very special about 2, give a second 1000 irrational numbers whose decimals lie in the same interval.

Section 2.6

1. First we just do a bit of practice with complex number operations. Just keep in mind that $i^2 = -1$. There is really nothing else to remember for most complex arithmetic
a) Add $1 - 7i$ and $2 + 3i$
b) Subtract $4 - 2i$ from $5 + 7i$
c) Multiply $3 + 5i$ times $4 - 2i$
d) Multiply $2 - 5i$ times $2 + 5i$. What number have you factored?
e) Multiply $2 - 5i$ times $6 + 15i$ Why is this a real number?
f) Divide $22 + 20i$ by $3 - 2i$. Remember that $\frac{w}{z} = \frac{w\overline{z}}{z\overline{z}}$ this problem has been cooked to have a nice answer.
f) Divide $3 - 7i$) by $(2 + 3i)$. This is the more typical situation.
g) Find $(2 + 3i)^2$ and $(2 + 3i)^3$.
h) Put $z = 2 + 3i$ into the expression $z^2 - 4z + 13$. What do you get? What does this mean?
i) Find the two roots z_1 and z_2 of $z^2 - 4z + 13$. Now find $z_1 + z_2$ and $z_1 z_2$.
j) Find $(z - z_1)(z - z_2)$ with z_1, z_2 from i). Multiply it out. Notice you have factored $z^2 - 4z + 13$. Odds are great someone will someday tell you this is impossible. As usual they mean "impossible with *real* numbers."
Some Answers (not in order) $22 + 14i$, $z^2 - 4z + 13$ 29, 87, $1 + 9i$, 0, $3 - 4i$, $2 + 8i$, $(-1/13)(15 + 23i)$, $-5 + 12i$

2. Prove $\overline{wz} = \overline{zw}$ by computing both sides and seeing they come out the same.

3. If z lies on the unit circle ($|z| = r = 1$ which is equivalent to $z\overline{z} = 1$) what is $1/z$?

4. Select the correct answer: a) $2+i > 1-2i$, b) $2+i < 1-2i$, c) $2+i = 1-2i$ d) The question is stupid because complex number do not have an order on them.

5. Let $z = 4+i$. Graph z and $\overline{z}$ in the complex plane. What are the relationships of the r and the θ for z and for $\overline{z}$?

6. Looking at the figure labled "Complex numbers in the plane"
a) Add $-2-4i$ and $3+2i$.
b) Without stretching it or turning it, move the vector $-2-4i$ so it's tail is placed on the nose of $3+2i$. At what point does it's nose end up? What is the complex number that goes from the origin to the point you just found? This is called graphical addition of vectors or equivalently complex numbers.

7. We are going to find some roots of 1. For problems a, b, c, you want the *smallest* θ. All other roots will be the powers $(1\angle\theta)^n$ for that θ. a) What angle would an 4th root of 1 have? You need $1\angle\theta)^4 1 = 1\angle 360°$.
b) What angle would a 6th root of 1 have?
c) What angle would a 12th root of 1 have?
d) Using $1\angle\theta = 1(\cos\theta + i\sin\theta)$ compute the four 4th roots of 1.
e) Using the favorite triangles compute the six 6th roots of 1. Give answers in radical form
f) Same but the twelve 12th roots of 1. Answers in radical form.
g) Use the answers for f) to draw a regular duodecagon (12 sided figure).
h) If you wanted to draw a regular octagon what roots of 1 would you find?

8. There are many applications of complex numbers to number theory and we are going to look at a simple one. Some numbers are the sum of two squares of integers, like $29 = 5^2 + 2^2$, and some, like 7 are *not* then sum of two squares. Prove that the if two numbers like 13 and 29, are the sum of two squares then their product 377 is also the sum of two squares. The trick is to reinterpret being the sum of two squares as being the norm of a complex number $13 = 3^2 + 2^2 = N(3+2i)$ and then use the fact that $N(wz) = N(w)N(z)$. So now use this to find the two squares that add up to 377. Notice that 13 is also $N(3-2i)$. Use this to find *another* two squares that add up to 377.
If you factor a positive integer n into prime numbers

$$n = 2^{\alpha} p_1^{\beta_1} \cdots p_r^{\beta_r} q_1^{\gamma_1} \cdots q_s^{\beta_s}$$

where p_i is a prime of the form $p_i = 4k+1$ and q_j is a prime of the form $q_j = 4k+3$ then n is a sum of 2 squares if and only if all of $\beta_1 \ldots \beta_s$ are even. The *only if* part is easy but you can see that the *if* part can be

proved if we can prove that any prime of the form $4k+1$ is the sum of 2 squares and this is a bit more difficult. There is even a formula for each of the two squares. Note that the value of α, the exponent of 2 doesn't matter. With a great deal more work we can even figure out *in how many ways* n is a sum of two squares.

Section 2.7 Appendix
For the problems in this section it is necessary to have a calculator which can deal with the exponential function and the sine and cosine functions in radians. It is handy but not critical that this be a physical device. Most computers now already have, or have access to, an onboard calculator and you can use this to do the problems. A computer algebra package like Mathematica, Maple or Matlab will also work.

1. a) Without looking at the table in the book convert $45°, 60°, 90°, 120°, 150°, 180°, 240°, 170°, 330°$ to radians, (expressed as fractions of π, for example $\frac{2\pi}{3}$). You can do this by hand by multiplying the given angle by $\frac{\pi}{180°}$ and simplifying the fraction. Check your answers by looking at the table. This is busywork but you will find the experience helpful.
b) Similarly, change the following angles in radians to angles in degrees: $\frac{\pi}{4}, \frac{3\pi}{4}, \frac{5\pi}{6}, \frac{7\pi}{6}, \frac{7\pi}{3}$. It is desirable to be able to do these popular angles in your head, but it's not critically important.

2. Using the calculation of $\sin(\frac{\pi}{6})$ by use of an infinite series as a model calculate $\cos(\frac{\pi}{6})$ in the same way. In ancient times whole tables of trigonometric functions were calculated in this way, normally by students.

3. Calculate an approximation to the value of e by using the infinite series for e^x and letting $x = 1$. Use the first 10 terms of the series so your answer will be fairly accurate.

4. Use a graphing calculator to make a graph of $\ln x$. Compare to graph of e^x. Can you see the connection? Graph e^x, $\ln x$, and the line $y = x$ all on the same graph. Now do you see the connection? This works for any pair of functions inverse to one another.

5. Using the rules for exponents and the rule $y = e^x$ if and only if $\ln y = x$, prove that $\ln(yz) = \ln y + \ln z$, $\ln y^k = k \cdot \ln y$ $\ln(1/y) = -\ln y$, $\ln(y/z) = \ln y - \ln z$.

6. Using $x = \log y$ if and only if $10^x = y$ prove that $\log y = \ln y / \ln 10 = (\ln y)/2.30258509....$ This shows that the difference between $\ln y$ and $\log y$ is totally trivial and why mathematicians do not bother with log. Another point is the Calculus formulas

$$\text{The derivative of } \ln x \quad \text{is} \quad \frac{1}{x}$$

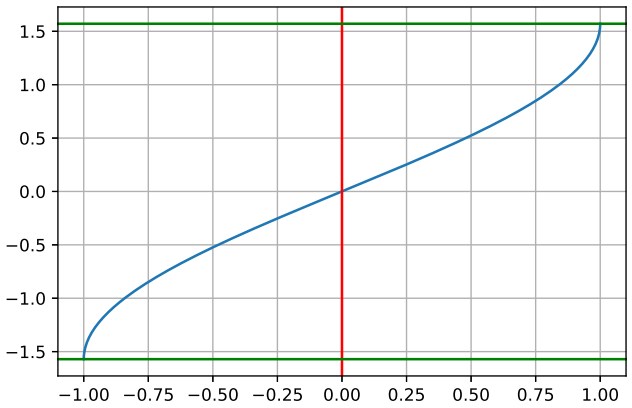
$$\text{The derivative of } \log x \quad \text{is} \quad \frac{1}{2.30258509\, x}$$

Which formula do you prefer?

7. Find the natural logarithm of 2. If you got .301029995... you pushed the wrong button. Try again. (This is a constant danger and you might put a bit of masking tape over the log button to avoid the problem.)

8. Use the fact that $\ln(x)$ and e^x are inverse functions to solve the following problems.
a) $e^x = 5$
b) $\ln x = 3$
c) $2e^{3x} = 8$
d) $4\ln(5x) = 2$
e) $7^x = 5$
f) $7^{x^2} = 4$

9. Two functions are inverse functions when $f(x) = y$ if and only if $x = g(y)$. *However* there are tricky aspects. For example, with $\exp(x)$ and $\ln(y)$ it only works for real numbers when $y > 0$. If there is periodicity involved it is much more complicated. The inverse function of $\sin(x)$ is written $\sin^{-1}(x)$ or $\arcsin(x)$. The graph of arcsin looks like



Graph of the arcsin Function

Note that the only inputs with real outputs are $-1 \le x \le 1$ and all the outputs are in the interval $-\pi/2 \le y \le \pi/2$. If you input a real number greater than 1 the output is complex and does not show on the graph.
a) Find the exact value of $\arcsin\frac{1}{2}$ by thought. Perhaps draw a triangle.
b) Find the approximate value of $\arcsin\frac{1}{2}$ using a calculator in radian mode. Multiply your answer by 6. Look familiar.
c) Find the approximate value of $\arcsin\frac{1}{2}$ using a calculator in degree mode.

Note: The calculator has three modes; degree, radian and grad. Unless you are building a highway or railway line do not use grad.
d) Use a calculator to find sin(1.4323). Be sure you are in radian mode!! Answer is .990425
e) Now find the arcsin of the answer to d). Is this a surprise?
f) Use a calculator to find sin(2.51327). Answer is .587789
g) Now find the arcsin of the answer to f). The correct answer is .628023. However, in recent times some calculators have been taught to remember that the .587789 came from the sine of 2.51327, and so it spits back the 2.51327. It is trying to be helpful but this could seriously screw things up and so is not a desirable feature of a calculator.
h) Convert the angles 2.51327 radians and .628023 radians to degrees and then draw them on an x, y rectangular graph. Try to see why they have the same sine.

10. Use the infinite series for Cosine given in the text to find $\cos(30^\circ)$. Remember to first change the 30° to radians. Four terms should be enough for a good approximation.

11. Prove that $z = \ln r + i\theta + 2n\pi i$ are all solutions to $e^z = w = r(\cos\theta + i\sin\theta)$ by substituting the expression for z into e^z and using the exponent laws.

11. Using the formulas for $\cos(z)$ and $\sin(z)$ in terms of e^{iz} and e^{-iz}, show that $2\cos(z)\sin(z) = \sin(2z)$.

12. We are going to solve $\sin(z) = 2$.
a) Using the formula for $\sin(z)$ in terms of e^{iz} and e^{-iz}, get to $e^{iz} - e^{-iz} = 4i$.
b) Multiply by something which is never 0 and rearrange to get $e^{2iz} - 4ie^{iz} - 1 = 0$.
c) Make the substitution $u = e^{iz}$ to get the quadratic equation $u^2 - 4iu - 1 = 0$
d) Solve the quadratic equation for u.
e) Replace u by e^{iz} in the previous equation for u to get $e^{iz} = (2 \pm \sqrt{3})i$.
f) Choose the plus sign and solve for z to get $z = \frac{1}{i}\ln[(2+\sqrt{3})i]$
g) Simplify to get $z = \frac{\pi}{2} - i\ln(2+\sqrt{3}) = 1.57080 - 1.31696i$.

Chapter 3

QUATERNIONS

3.1 Introduction

We saw how by adding i to the real numbers we came up with a much richer structure, the complex numbers, and these numbers were intimately associated with the plane. We might ask, since we live in three dimensional space, might there be a larger system that would do for our space $\mathbb{R}^3$ what the complex numbers do for the plane $\mathbb{R}^2$. The answer is, sort of.

This question exercised the imagination of the great Irish mathematical physicist William Roman Hamilton. He pondered it for several years, and then one day as he and his family were walking near Droichead Broome (Broom bridge, Broom spelled many ways in English) the answer came to him. Whipping out his penknife he carved the basic equations into the granite bridge. They of course quickly faded but the moment is commemorated by a stone plaque (later vandalized by those with an unCeltic lack of enthusiasm for mathematics), which reads

> Here as he walked by
> on the 16th of October 1843
> Sir William Rowan Hamilton
> in a flash of genius discovered
> the fundamental formula for
> quaternion multiplication
> $i^2 = j^2 = k^2 = ijk = -1$
> & cut it on a stone of this bridge.

This was not exactly what Hamilton had set out to find; he wanted a 3-dimensional algebra to describe 3-space. However, such an algebra does not exist. He had to be satisfied with a four dimensional algebra though in time he came to appreciate that for dealing with rotations things *had* to be that way.

Besides having an extra dimension, quaternions had a second surprising property; they were not commutative; $ij = -ji$ as we will see in the next section. It was this property which had held up the discovery for many years, but Hamilton eventually saw he must give up commutativity if he were to make any progress. Aside from this strange property, quaternions are in fact pretty well behaved (they are associative) and because they are rather attractive many mathematicians, after some reluctance, were willing to accept them in spite of the non-commutativity, which at the time was quite shocking but is no longer regarded with horror. Nowadays non-commutativity is seen as pretty normal for algebraic systems.

Hamilton, was a superb mathematical physicist. In 1788 Lagrange (1736-1813, Italian [1]) had invented a new method of doing classical mechanics based

[1] Lagrange's great grandfather was French; his other ancestors were Italian, although he did have a tendency, when convenient, to think of himself as French. He worked for 20 years in Berlin and then in 1787, at the age of 51, moved to Paris. While in Berlin he wrote the magnificent Mechanique Analytique, which provided a new way, via a function called the Lagrangian and using the Calculus of Variations, of doing physics. He worked on the committee which came up with the metic system.

on a function now called the Lagrangian. Going on from there, Hamilton invented what we now call the Hamiltonian function, which is the total energy of a system expressed in terms of position and momentum. Lagrangians and Hamiltonians replaced the clumsy methods descended from Newton and created a whole now way of doing classical mechanics. But, far more important, Hamiltonian methods proved to be a way to transition from classical mechanics to quantum mechanics. There are not many ways to cross this chasm.

However, once he, so to speak, crossed the quaternion bridge, he spent the rest of his life developing their theory, writing two two-volume sets of large books about them[2] and inspiring a number of mathematicians to carry on his quaternion work.

Opinions differ violently on this midlife career change of Hamilton. Some feel that he abandoned his true calling in mathematical physics and wasted the rest of his life on the relatively unimportant quaternions. However, even if quaternions never fulfill their early promise themselves, they did something of epochal significance. They broke the hold of the commutative law of multiplication on the minds of mathematicians, just as the invention of non-Euclidean geometry had unchained the mathematical world from the shackles of Euclidean Geometry and showed the path to far larger domains for mathematicians to investigate. Bolyai and Lobachevski freed geometry and Hamilton (and Grassmann) freed algebra.

In the processes of living, order matters! Put on your shoes then your socks differs fundamentally from put on your socks then your shoes. If we want to build a mathematics that can mirror the world, we must have some sort of non-commutativity somewhere that can handle this. Quaternions were the beginning of this, and since they are so charming in their way, after them non-commutativity of multiplication became, relatively quickly, accepted, and then expected. It turns out that relatively few systems are commutative, although they are often highly important, like $\mathbb{R}$ or $\mathbb{C}$ or the commutative rings used in Algebraic Geometry. When the symbols represent transformations of some sort, like rotations in space, commutativity is not to be expected. Order matters.

As for the Quaternions themselves, opinions differ on them too. They have had ups and downs over the years. Originally they were used for a sort of vector algebra. James Clerk Maxwell's first edition of his book on Electromagnetic theory used them this way, but his second edition returned to the use of coordinates which was definitely as step backwards. Other forms of vector algebra, using the scalar and vector product, took over the things that quaternions could do in electromagnetics and other elementary forms of physics. Vector algebra is a crude tool, but it suffices for much of physics. Quaternions have been used for relativity. By chance (?) they fit special relativity well. I believe they don't work as well for general relativity.

[2] At a certain point in time the English foreign office wanted to curry favor with Eamon de Valera, ex-mathematician and Premier of Eire (Ireland). They sought a suitable present to convey their esteem and they chose one of the two two-volume books by Hamilton on Quaternions. De Valera (a quaternion enthusiast) carried the set with him wherever he went for the rest of his life.

In algebra they have a important role as building blocks of many other systems. They also play the role of constants when they appear as entries in matrices. Frobenius proved that the quaternions and complex numbers are the only division rings that contain $\mathbb{R}$, and this makes them stand out from among the many other algebras.

Moreover, there is still a possibility that in time Quaternions will be found to play a more fundamental role in mathematics and physics than they do at present. For example they are important in the theory of Dirac Operators, which lies on the borderland between mathematics and physics.

You may well ask whether Quaternions are *Numbers*. This is not a question with a precise answer. What we want to know is do Quaternions act enough like numbers to qualify for membership in the club. They are not commutative, and for that reason many current members would look askance at them. But, except for commutativity, they do in fact act very much like numbers, enough in fact for them to have a "number theory". The great mathematician Adolf Hurwitz wrote a small book on this subject. They also help out in the number theory of ordinary integers, as you can see in a subsequent section in this book. The author first met them in 1961 in a course in *Number* Theory taught by Michael Artin. So, without committing ourselves completely to the notion that quaternions are numbers, we consider they are numberlike enough to sneak into this book.

The Author and W. R. Hamilton in the library
of Queen's college Baile Átha Cliath (Dublin)

3.2 Algebraic Definition and Properties of Quaternions

In this section we will build the quaternions algebraically. This was not how Hamilton approached it. In the next section we will look at things geometrically which will be more Hamilton's line of country.

The primary complex number is i and it satisfies the basic equation $i^2 = -1$. If we want a bigger system, we will have to add another unit, call it j. Fairness requires $j^2 = -1$, since i and j ought to somehow have symmetric roles[3]. If we have $ij = ji$ we will have a commutative system, which would not control rotations in 3-space since they are not commutative. (Also, it wouldn't turn out to be a very nice system but we won't dwell on that.) So we decide to make the system "almost" commutative by replacing $ij = ji$ by $ij = -ji$. We also require, which is standard in the circumstances, that real numbers commute with i and j, so that, for example, $j2 = 2j$. Thus we can summarize the basic equations of quaternions as

$$i^2 = j^2 = -1 \quad \text{and} \quad ji = -ij$$

From these, and the standard associative and distributive laws, all else follows. To recover the standard presentation (which is somewhat redundant) we add the definition

Def k=ij

and we must then derive k^2 and ijk.

$$ijk = ijij = ij(-ji) = ij(-1)ji = (-1)ijji = -i(-1)i = ii = -1$$

and

$$k^2 = ijk = -1 \quad \text{from the previous calculation}$$

We have now recovered Hamilton's definition from the plaque. There is a little more to do. We already know that $ij = k$ and $ji = -ij = -k$. We then have

$$-ki = jii = j(-1) = -j \quad \text{which implies} \quad ki = j$$

and

$$ik = iij = -1j = -j$$

All we need now are the k and j equations.

$$kj = ijj = i(-1) = -i$$

and

$$jk = j(-ji) = -jji = -(-1)i = i \quad \text{using} \quad k = ij = -ji$$

Now we have the complete multiplication table

[3]This immediately closes off the possibility the the Quaternions could form a field, since in a field $X^2 + 1 = 0$ can have only two solutions, but here we have at least four solutions $\pm i$ and $\pm j$. This shows the importance of commutativity in the theory of polynomial equations.

	i	j	k
i	-1	k	-j
j	-k	-1	i
k	j	-i	-1

A way to remember this is to write

$$ijkijk$$

and then notice that if you want jk you look for it in the row, going left to right, If you find it, the product is the following letter i. If you *don't* find it, you can find it by going in the opposite direction right to left and the product is the letter to the left but with a minus sign: $kj = -i$. You can also draw the letters i, j, k on a circle and it'll work much the same.

Def A quaternion is an expression of the form

$$A = a_0 \cdot 1 + a_1 \cdot i + a_2 \cdot j + a_3 \cdot k$$

where the a_i are real numbers. (The dots are usually omitted as is the 1 in the first term.)

Thus the following are quaternions:

$$2 + 3i - j - \frac{1}{3}k, \quad 2 + i - j, \quad j, \quad \pi + 3k, \quad \cos(\theta) + \sin(\theta)(\frac{2}{3}i + \frac{2}{3}j - \frac{1}{3}k)$$

The set of all quaternions has the code letter $\mathbb{H}$ which is probably in honor of Hamilton.

We now compute the formula for the multiplication of two quaternions. Let, with the more relaxed notation

$$A = a_0 + a_1 i + a_2 j + a_3 k \qquad B = b_0 + b_1 i + b_2 j + b_3 k$$

Then, (using the multiplication table and being fanatically careful about order,)

$$\begin{aligned} AB &= (a_0 + a_1 i + a_2 j + a_3 k)(b_0 + b_1 i + b_2 j + b_3 k) \\ &= a_0(b_0 + b_1 i + b_2 j + b_3 k) + a_1 i(b_0 + b_1 i + b_2 j + b_3 k) \\ &\quad + a_2 j(b_0 + b_1 i + b_2 j + b_3 k) + a_3 k(b_0 + b_1 i + b_2 j + b_3 k) \\ &= a_0 b_0 + a_0 b_1 i + a_0 b_2 j + a_0 b_3 k + a_1 b_0 i - a_1 b_1 + a_1 b_2 k - a_1 b_3 j \\ &\quad + a_2 b_0 j - a_2 b_1 k - a_2 b_2 + a_2 b_3 i + a_3 b_0 k + a_3 b_1 j - a_3 b_2 i - a_3 b_3 \\ &= (a_0 b_0 - a_1 b_1 - a_2 b_2 - a_3 b_3) + (a_0 b_1 + a_1 b_0 + a_2 b_3 - a_3 b_2) i \\ &\quad + (a_0 b_2 + a_2 b_0 + a_3 b_1 - a_1 b_3) j + (a_0 b_3 + a_3 b_0 + a_1 b_2 - a_2 b_1) k \end{aligned}$$

Now in analogy with complex numbers we define the conjugate of a quaternion.

Def if $A = a_0 + a_1 i + a_2 j + a_3 k$ then the conjugate of A is

$$\overline{A} = a_0 - a_1 i - a_2 j - a_3 k$$

Note the following important fact. If we know $A = \overline{A}$ then we know $a_i = -a_i$ for $i = 1, 2, 3$ which means $a_i = 0$ for $i = 1, 2, 3$ which means that $A = a_0$ Such a quaternion is called a *real quaternion* and does not differ from the real number a_0 except spiritually. (The fanatic will want to check that multiplying by the real number and the real quaternion give the same output.)

Next we must derive the critical equation linking multiplication and conjugation. This is annoying and you can just skim it.

$$\begin{aligned}
AB &= (a_0b_0 - a_1b_1 - a_2b_2 - a_3b_3) + (a_0b_1 + a_1b_0 + a_2b_3 - a_3b_2)i \\
&\quad +(a_0b_2 + a_2b_0 + a_3b_1 - a_1b_3)j + (a_0b_3 + a_3b_0 + a_1b_2 - a_2b_1)k \\
\overline{AB} &= (a_0b_0 - a_1b_1 - a_2b_2 - a_3b_3) - (a_0b_1 + a_1b_0 + a_2b_3 - a_3b_2)i \\
&\quad -(a_0b_2 + a_2b_0 + a_3b_1 - a_1b_3)j - (a_0b_3 + a_3b_0 + a_1b_2 - a_2b_1)k \\
\overline{B} &= b_0 - b_1i - b_2j - b_3k \qquad \overline{A} = a_0 - a_1i - a_2j - a_3k \\
\overline{B}\,\overline{A} &= (b_0a_0 - b_1a_1 - b_2a_2 - b_3a_3) + (-b_0a_1 - b_1a_0 + b_2a_3 - b_3a_2)i \\
&\quad +(-b_0a_2 - b_2a_0 + b_3a_1 - b_1a_3)j + (-b_0a_3 - b_3a_0 + b_1a_2 - b_2a_1)k
\end{aligned}$$

Comparing the equations for $\overline{AB}$ and $\overline{B}\,\overline{A}$ we see they are equal. Thus we have proved the important equation

$$\overline{AB} = \overline{B}\,\overline{A}$$

I want to add a bit of context here. An operation T is an involution if doing it twice to any input gives the input back again, so $TTu = u$ Conjugation is an involution. Almost always involutions applied to products require a shift in order, that is they satisfy $T(uv) = T(v)T(u)$. Since conjugation is an involution, we expect $\overline{AB} = \overline{B}\ \overline{A}$ so then it was just a matter of doing the annoying calculation to *prove* the expected result.

Now we can get the harvest from this bit of work. Recall the norm of a complex number: $N(a+bi) = (a+bi)(\overline{a+bi}) = a^2 + b^2$. We have the same sort of equation here.

Def The *norm* of a quaternion A is $N(A) = A\overline{A}$.

We note that

$$\overline{N(A)} = \overline{A\overline{A}} = \overline{\overline{A}}\,\overline{A} = A\overline{A} = N(A)$$

so $N(A)$ is a real quaternion and just a glance at the above equations and setting $B = \overline{A}$ shows that

$$N(A) = a_0^2 + a_1^2 + a_2^2 + a_3^2$$

which is important. Note that, trivially, $N(\overline{A}) = N(A)$. Also note in the calculation where we used $\overline{AB} = \overline{B}\,\overline{A}$ and how essential it was to the calculation. This is again true in the following calculation:

$$N(AB) = (AB)\overline{AB} = AB\overline{B}\,\overline{A} = A(B\overline{B})\overline{A} = A(N(B))\overline{A}$$

Now, noting that $N(B)$ is a real quaternion, that is essentially a real number, it commutes with $\overline{A}$ to give us

$$N(AB) = A\overline{A}(N(B)) = N(A)N(B)$$

which is an important equation. As an example of just one use, suppose we note that

$$7 = 2^2 + 1^2 + 1^2 + 1^2 \qquad \text{and} \qquad 11 = 3^2 + 1^2 + 1^2 + 0^2$$

We would like to find four squares that add up to $7 \cdot 11 = 77$. It is easy. We note $7 = N(2 + i + j + k)$ and $11 = N(3 + i + j)$. Then

$$77 = 7{\cdot}11 = N(2{+}i{+}j{+}k)N(3{+}i{+}j) = N(2{+}i{+}j{+}k)(3{+}i{+}j)) = N(4{+}4i{+}6j{+}3k)$$

so

$$77 = 4^2 + 4^2 + 6^2 + 3^2$$

You can get other sets of 4 number that add to 77 by changing a sign in one of the quaternions, like

$$77 = 7{\cdot}11 = N(2{-}i{+}j{+}k)N(3{+}i{+}j) = N((2{-}i{+}j{+}k)(3{+}i{+}j)) = N(6{-}2i{+}6j{+}1k)$$

so

$$77 = 6^2 + 2^2 + 6^2 + 1^2$$

What we have shown is that if two numbers are each the sum of 4 squares so is their product. So to prove any number is the sum of 4 squares it is enough to prove that any prime is the sum of 4 squares, and quaternions are helpful in that proof also, but to do it we would need much more equipment.

A most important property of Quaternions is that non-zero quaternions have inverses. This is easy. We calculate

$$A \cdot \frac{\overline{A}}{N(A)} = \frac{A\overline{A}}{N(A)} = \frac{N(A)}{N(A)} = 1$$

Also, since $\overline{A}A = N(A)$ this works on the other side:

$$\frac{\overline{A}}{N(A)} \cdot A = 1$$

so we have an element that is an inverse on both sides and we can set

$$A^{-1} = \frac{\overline{A}}{N(A)}$$

Note how analogous this is to complex numbers. It also means that the set of quaternions $\mathbb{H}$ satisfies M3 in the laws for a Ring. But the quaternions do not satisfy M4, which is commutativity. Thus they are not quite a Field, but close.

Def A ring which satisfies M3 is called a *Division Ring*.

The way I have set this up a field is a commutative division ring. If a division ring is *not* commutative it is often called a skew field, but we will not use this term.

Although a division ring is quite close to being a field when looked at in terms of axioms, it is not so close when you consider properties. For example the quadratic equation $X^2 + 1 = 0$ can have infinitely many solutions in a division ring but only two in a field.

Clearly the quaternions are a division ring and probably they are the easiest non-commutative example of the concept.

Algebras

This section is not essential for the rest of the book but it is sort of interesting and may sort out some confusion.

Some rings have an additional property in that they have *scalars* that function like the real numbers do with vectors. These scalars must come from a field and commute with every element of the ring. I will give a precise definition.

Def Let R be a ring. The center $Z(R)$ is the set of all elements of the ring which commute with every element of the ring.

For example, the identity 1 of the ring is always an element of the center. Clearly $Z(R) = R$ is the same as the R being commutative (satisfies M4). Now we can define an *algebra*.

Def Let R be a ring and F be a field. R is an *algebra over* F if and only if F is a subset of $Z(R)$.

Thus the elements α of F commute with all elements r of R so $\alpha r = r\alpha$. We have actually been working with this concept for awhile. $\mathbb{R}$ is an algebra over itself. $\mathbb{C}$ is an algebra over $\mathbb{R}$ and also an algebra over $\mathbb{C}$. $\mathbb{H}$, the set of quaternions, is an algebra over $\mathbb{R}$ but *it is* **not** an algebra over $\mathbb{C}$ because the element i in $\mathbb{C}$ does not commute with the element j in $\mathbb{H}$ since $ij = -ji$.

With this equipment we can now state a very important theorem:

Theorem(Frobenius[4]) There are only three finite dimensional division algebras over $\mathbb{R}$ and they are $\mathbb{R}$, $\mathbb{C}$ and $\mathbb{H}$.

Stated in words, the only finite dimensional division algebras over the real numbers are the real numbers, the complex numbers and the quaternions.

Finite Dimensional here means the basic elements of the algebra, like 1,i for $\mathbb{C}$ over $\mathbb{R}$, or 1, i,j,k for the Quaternions over $\mathbb{R}$, are finite in number. This suggests that we have now exhausted the numberlike objects which can be created by adding new elements to $\mathbb{R}$.

However, I feel I should mention the Octonions, which have eight basis elements analogous to the four 1,i,j,k for Quaternions. These are usually called

$$e_0 = 1,\ e_1,\ e_2,\ e_3,\ e_4,\ e_5,\ e_6,\ e_7$$

and have the multiplication laws

$$e_i e_j = \begin{cases} e_j & \text{if} \quad i = 0 \\ e_i & \text{if} \quad j = 0 \\ -\delta_{ij} e_0 + \epsilon_{ijk} e_k \end{cases}$$

where δ_{ij} is the Kronecker delta (remember this; you'll run into it a lot)

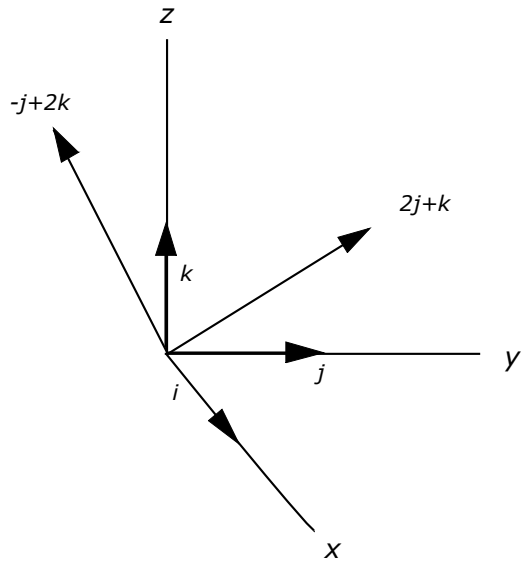
$$\delta_{ij} = \begin{cases} 0 & \text{if} \quad i \neq j \\ 1 & \text{if} \quad i = j \end{cases}$$

[4]Ferdinand Otto Frobenius, 1849-1917, German Mathematician, who made important contributions in many areas of mathematics. For example he created the modern theory of group characters and group representations, and the always unpopular method of Frobenius for singular differential equations.

and ϵ_{ijk} is completely antisymmetric (switches signs when any of ijk are swapped) and is equal to 1 for $ijk = 123,\ 145,\ 176,\ 246,\ 257,\ 347,\ 365$ and is 0 if any two of ijk are the same. From this one can build a multiplication table all very much like the quaternions. However, Frobenius' theorem above tells us that the octonions cannot be a division algebra, so something must be wrong. What is wrong is that the octonions are not associative. For example $(e_i e_j)e_k = -e_i(e_j e_k) \neq e_i(e_j e_k)$. In my opinion this lack of associativity puts octonions outside the world of numbers. However, this is only *my* opinion . Octonions (symbol $\mathbb{O}$) have many numberlike properties and an enthusiastic if not too numerous following among mathematicians. So if someday *they* become accepted as "numbers," then the Sedenions $\mathbb{S}$ (16 base elements) will be clamoring at the gates.

3.3 Relations of Quaternions to Geometry

This section will be much closer to Hamilton's line of thought.



The basis vectors i, j, k

In the figure above you see the x, y, z axes (only the positive parts.) Hamilton thought of i, j, k as being unit vectors along the x, y, z axes. As with complex numbers, quaternions partake of dual import; they are first, vectors and second rotation operators. The situation is not entirely straightforward, but we will start where things are simplest. Hamilton envisaged i as the operator that rotates by 90° in the plane *perpendicular* to i or to the axis of i, the x axis. The rotation operator is on the *left,* so $ij = k$ means if you rotate j by 90° in the plane perpendicular to i you will get k, which is clear from the picture. Next, $ik = -j$ means that rotating k by 90° in the plane perpendicular to i you will

get $-j$. The basic entries in the multiplication table thus express simple facts about the xyz or ijk tripod of axes. You should look at the picture and imagine rotating 90° around the y axis using left multiplication by j, so that $jk = i$ and $ji = -k$. Don't read on until you see this, in both senses of see. Now do it with k. You might find it helpful to point the thumb of your *right* hand along the z axis and your fingers will curl from i to j and from j to $-i$. You see here the power of non-commutative multiplication. How on earth could we describe this with a commutative system?

To avoid making the calculations more messy than they need to be, and also to indicate some connections between quaternions and vector algebra, we will take a slight excursion through vector algebra.

3.3.1 Pure Vector Algebra

By pure vector algebra we mean the algebra of vectors used in Calculus and applied mathematics which does not mention quaternions. We give a quick introduction to this discipline in this section and then we will indicate the connections with quaternions in a subsequent section. We do make some historical remarks about the connections with quaternions in this section.

I have used the notation $\vec{\mathbf{v}}$ for vectors which is over elaborate. It is more standard to simply use $\mathbf{v}$ but when you work by yourself on paper it is not easy to do boldface, and so a standard substitute for boldface in handwritten work is to put the arrow on top. I have put the arrow on top of the boldface in the hope it will get you used to both conventions simultaneously.

As I said above, we will indicate vectors in this section by $\vec{\mathbf{v}}$ and write them in terms of $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$ where the hat instead of the arrow means unit vector. (This notation is common in physics but scarce in mathematics.) Vectors will be written in terms of $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$ as follows

$$\vec{\mathbf{v}} = v_1\hat{\mathbf{i}} + v_2\hat{\mathbf{j}} + v_3\hat{\mathbf{k}}$$

and similarly for $\vec{\mathbf{w}}, \vec{\mathbf{u}}$, etc.

The Scalar or Dot Product

The Scalar or Dot product inputs two vectors and outputs a scaler, that is a real number. If

$$\vec{\mathbf{u}} = u_1\hat{\mathbf{i}} + u_2\hat{\mathbf{j}} + u_3\hat{\mathbf{k}} \qquad \text{and} \qquad \vec{\mathbf{v}} = v_1\hat{\mathbf{i}} + v_2\hat{\mathbf{j}} + v_3\hat{\mathbf{k}}$$

then

$$\vec{\mathbf{u}} \cdot \vec{\mathbf{v}} = u_1v_1 + u_2v_2 + u_3v_3$$

The length of a vector is

$$|\vec{\mathbf{u}}| = \sqrt{\vec{\mathbf{u}} \cdot \vec{\mathbf{u}}} = \sqrt{u_1^2 + u_2^2 + u_3^2}$$

Note $|\vec{\mathbf{u}}| > 0$ for any vector $\vec{\mathbf{u}} \neq 0$. One can now use the law of cosines to prove that for two non-zero vectors

$$\vec{\mathbf{u}} \cdot \vec{\mathbf{v}} = |\vec{\mathbf{u}}| \; |\vec{\mathbf{v}}| \; \cos(\theta)$$

where θ is the angle in the range $0 \leq \theta \leq 180^\circ$ between the vectors. (We will often use θ for the angle between two vectors. Which two vectors should be clear from context.) We immediately have for the angle θ between $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$

$$\cos(\theta) = \frac{\vec{\mathbf{u}} \cdot \vec{\mathbf{v}}}{|\vec{\mathbf{u}}| \; |\vec{\mathbf{v}}|}$$

It is then clear that for two non-zero vectors

$$\vec{\mathbf{u}} \cdot \vec{\mathbf{v}} = 0 \quad \text{if and only if} \quad \vec{\mathbf{u}} \text{ is perpendicular to } \vec{\mathbf{v}}$$

This might be the most important use of the dot product.

Note that the dot product is commutative:

$$\vec{\mathbf{u}} \cdot \vec{\mathbf{v}} = \vec{\mathbf{v}} \cdot \vec{\mathbf{u}}$$

It is worth mentioning that if the scalars are *complex numbers* then the definition needs some conjugation in it. For quantum mechanics the definition is

$$\vec{\mathbf{u}} \cdot \vec{\mathbf{v}} = \overline{u_1} v_1 + \overline{u_2} v_2 + \overline{u_3} v_3$$

In mathematics it is more customary (unfortunately) to conjugate the second vector. The physics tradition is better but we will not be using either.

The Vector or Cross Product

The cross product started life in the following way. Let us multiply $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$ as *quaternions*. We then have

$$\begin{aligned} \vec{\mathbf{u}}\vec{\mathbf{v}} &= (u_1\hat{\mathbf{i}} + u_2\hat{\mathbf{j}} + u_3\hat{\mathbf{k}})(v_1\hat{\mathbf{i}} + v_2\hat{\mathbf{j}} + v_3\hat{\mathbf{k}}) \\ &= -u_1v_1 - u_2v_2 - u_3v_3 + (u_2v_3 - u_2v_2)\hat{\mathbf{i}} + (u_3v_1 - u_1v_3)\hat{\mathbf{j}} + (u_1v_2 - u_2v_1)\hat{\mathbf{k}} \end{aligned}$$

The vector part of this expression was denoted by $V(\vec{\mathbf{u}}\vec{\mathbf{v}})$ by Hamilton and his disciples. Josiah Willard Gibbs(1839-1903), who was mildly hostile to quaternions and leaned more in the direction of Grassmann, defined it independently of quaternions as

$$\vec{\mathbf{u}} \times \vec{\mathbf{v}} = (u_2v_3 - u_2v_2)\hat{\mathbf{i}} + (u_3v_1 - u_1v_3)\hat{\mathbf{j}} + (u_1v_2 - u_2v_1)\hat{\mathbf{k}}$$

and an influential textbook based on Gibb's lectures at Yale was written by E. B. Wilson. This showed that much of what quaternions did could be done without quaternions in what seemed a simpler way using the dot and cross product.

Thus we have three products for vectors connected by

$$\vec{\mathbf{u}}\vec{\mathbf{v}} = -\vec{\mathbf{u}} \cdot \vec{\mathbf{v}} + \vec{\mathbf{u}} \times \vec{\mathbf{v}}$$

It is sometimes handy to write the cross product in determinant form. The later chapter on Matrices has a short section on determinants which you might want to look at now. In determinant form the cross product can be written

$$\vec{\mathbf{u}} \times \vec{\mathbf{v}} = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = \begin{vmatrix} u_1 & v_1 & \hat{\mathbf{i}} \\ u_2 & v_2 & \hat{\mathbf{j}} \\ u_3 & v_3 & \hat{\mathbf{k}} \end{vmatrix}$$

The first determinant is the standard definition but the second one is preferable because it generalizes to any dimension, where of course instead of two vectors you must input $n-1$ vectors[5]. However, we will stick with $n = 3$ dimensions.

Properties of the (three dimensional) cross product

1 $\vec{\mathbf{u}} \times \vec{\mathbf{v}}$ is perpendicular to $\vec{\mathbf{u}}$ amd $\vec{\mathbf{v}}$

2 $|\vec{\mathbf{u}} \times \vec{\mathbf{v}}| = |\vec{\mathbf{u}}|\ |\vec{\mathbf{v}}|\ \sin(\theta)$

3 The trio of vectors $\vec{\mathbf{u}}$, $\vec{\mathbf{v}}$, $\vec{\mathbf{u}} \times \vec{\mathbf{v}}$ forms a right handed system like $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$.

4 $\vec{\mathbf{u}} \times \vec{\mathbf{v}} = -\vec{\mathbf{v}} \times \vec{\mathbf{u}}$

Item 2 shows us the useful fact that $\vec{\mathbf{u}} \times \vec{\mathbf{u}} = 0$ since $\theta = 0$.
Item 3 means that if you make the fingers of your *right* hand curl from $\vec{\mathbf{u}}$ to $\vec{\mathbf{v}}$ then your thumb will point in the direction of $\vec{\mathbf{u}} \times \vec{\mathbf{v}}$. This is called the *right hand rule*.
Item4 says that the cross product is almost commutative but gets a minus sign when the order is switched. You can see this from item 3, since when you switch the order of the inputs your thumb points the other way, but the length (item 2) remains the same. This property $\vec{\mathbf{u}} \times \vec{\mathbf{v}} = -\vec{\mathbf{v}} \times \vec{\mathbf{u}}$ is called anticommutativity and is fairly common in mathematics.

A defect of the cross product is that it is not associative. However $\mathbb{R}^3$ with the cross product is an example of a class of non-associative algebras called Lie Algebras[6]. We will not go into this further as it is a bit advanced. I only mention it because Lie Algebras are fairly standard objects so the cross product is not a mathematical outlier; it just isn't the sort of system we study in this book.

A closely associated product, called the scalar triple product or box product, is

$$[\vec{\mathbf{u}}\vec{\mathbf{v}}\vec{\mathbf{w}}] = \vec{\mathbf{u}} \cdot \vec{\mathbf{v}} \times \vec{\mathbf{w}} = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$$

[5]The oft repeated statement that the cross product will not generalize to other dimensions than 3 is true only if you insist on inputting just two vectors. If you input $n-1$ vectors a perfectly serviceable cross product (defined by the second determinant) exists for n dimensions.

[6]Sophus Lie, 1842-1899, was a Norwegian Mathematician, who, despite getting a late start in his mathematical career, made fundamental contributions with his Lie Groups and Lie Algebras to mathematics. This theory has proved to be quite central to modern mathematics but is a little difficult to get into. Lie Algebras were independently invented by Wilhelm Killing, German, 1847-1923, about 1880 and he made important contributions to the subject including a number of conjectures which turned out to be true.

It has the interesting property of being invariant under cyclic permutation and changing sign for non-cyclic permutation.

$$\begin{aligned}[\vec{\mathbf{u}}\vec{\mathbf{v}}\vec{\mathbf{w}}] &= [\vec{\mathbf{v}}\vec{\mathbf{w}}\vec{\mathbf{u}}] = [\vec{\mathbf{w}}\vec{\mathbf{u}}\vec{\mathbf{v}}] = \\ = -[\vec{\mathbf{v}}\vec{\mathbf{u}}\vec{\mathbf{w}}] &= -[\vec{\mathbf{w}}\vec{\mathbf{v}}\vec{\mathbf{u}}] = -[\vec{\mathbf{u}}\vec{\mathbf{w}}\vec{\mathbf{v}}]\end{aligned}$$

Knowing this can be very handy in applications. Also note that somewhere in there is

$$\vec{\mathbf{u}} \cdot \vec{\mathbf{v}} \times \vec{\mathbf{w}} = \vec{\mathbf{u}} \times \vec{\mathbf{v}} \cdot \vec{\mathbf{w}}$$

since $\vec{\mathbf{u}} \times \vec{\mathbf{v}} \cdot \vec{\mathbf{w}} = \vec{\mathbf{w}} \cdot \vec{\mathbf{u}} \times \vec{\mathbf{v}}$. Note that the parentheses in $\vec{\mathbf{u}} \cdot \vec{\mathbf{v}} \times \vec{\mathbf{w}}$ must be $\vec{\mathbf{u}} \cdot (\vec{\mathbf{v}} \times \vec{\mathbf{w}})$ since $(\vec{\mathbf{u}} \cdot \vec{\mathbf{v}}) \times \vec{\mathbf{w}}$ would make no sense; it has a scalar as input to a vector product whereas vector products input only two vectors, not a scalar and a vector.

An important identity (called the vector triple product law) which is rather hard to prove is

$$\vec{\mathbf{u}} \times (\vec{\mathbf{v}} \times \vec{\mathbf{w}}) = (\vec{\mathbf{u}} \cdot \vec{\mathbf{w}})\vec{\mathbf{v}} - (\vec{\mathbf{u}} \cdot \vec{\mathbf{v}})\vec{\mathbf{w}}$$

This can be proved by brute force in about half a page but such a proof is unsatisfying. We will prove this later using the associativity of quaternions. The above formula is essentially the same as Grassmann's formula

$$(\vec{\mathbf{t}} \times \vec{\mathbf{u}}) \cdot (\vec{\mathbf{v}} \times \vec{\mathbf{w}}) = \begin{vmatrix} \vec{\mathbf{t}} \cdot \vec{\mathbf{v}} & \vec{\mathbf{t}} \cdot \vec{\mathbf{w}} \\ \vec{\mathbf{u}} \cdot \vec{\mathbf{v}} & \vec{\mathbf{u}} \cdot \vec{\mathbf{w}} \end{vmatrix}$$

since from one you can easily prove the other.

Lie Algebras (mentioned above) have a fundamental identity which is the Lie Algebras substitute for the associative law. In our notation it is

$$\vec{\mathbf{u}} \times (\vec{\mathbf{v}} \times \vec{\mathbf{w}}) + \vec{\mathbf{v}} \times (\vec{\mathbf{w}} \times \vec{\mathbf{u}}) + \vec{\mathbf{w}} \times (\vec{\mathbf{u}} \times \vec{\mathbf{v}}) = 0$$

(Note the cyclic order!) It is very easy to prove from the vector triple product law.

3.3.2 Connections Between the Various Products

We have already seen the connection

$$\vec{\mathbf{u}}\vec{\mathbf{v}} = -\vec{\mathbf{u}} \cdot \vec{\mathbf{v}} + \vec{\mathbf{u}} \times \vec{\mathbf{v}}$$

We can now easily generalize this. Let the two quaternions U and V be written as

$$U = u_0 + \vec{\mathbf{u}} \qquad V = v_0 + \vec{\mathbf{v}}$$

where, as before

$$\vec{\mathbf{u}} = u_1\hat{\mathbf{i}} + u_2\hat{\mathbf{j}} + u_3\hat{\mathbf{k}}$$

and similarly for $\vec{\mathbf{v}}$. Then using associative and distributive laws we have

$$\begin{aligned} UV &= (u_0 + \vec{\mathbf{u}})(v_0 + \vec{\mathbf{v}}) \\ &= u_0v_0 + u_0\vec{\mathbf{v}} + v_0\vec{\mathbf{u}} + \vec{\mathbf{u}}\vec{\mathbf{v}} \\ &= u_0v_0 - \vec{\mathbf{u}} \cdot \vec{\mathbf{v}} + u_0\vec{\mathbf{v}} + v_0\vec{\mathbf{u}} + \vec{\mathbf{u}} \times \vec{\mathbf{v}} \end{aligned}$$

where the first two terms are scalar part and the last three the vector part of UV.

We introduce the concept of the size $||U||$ of a quaternion

$$||U|| = \sqrt{N(U)} = \sqrt{u_0^2 + u_1^2 + u_2^2 + u_3^2} = \sqrt{u_0^2 + |\vec{\mathbf{u}}|^2}$$

For a vector quaternion ($u_0 = 0$) the size and the length are the same. We will have a lot of uses for quaternions of size 1 which we call versors. These are the quaternions useful in rotation.

Def A versor is quaternion of unit size.

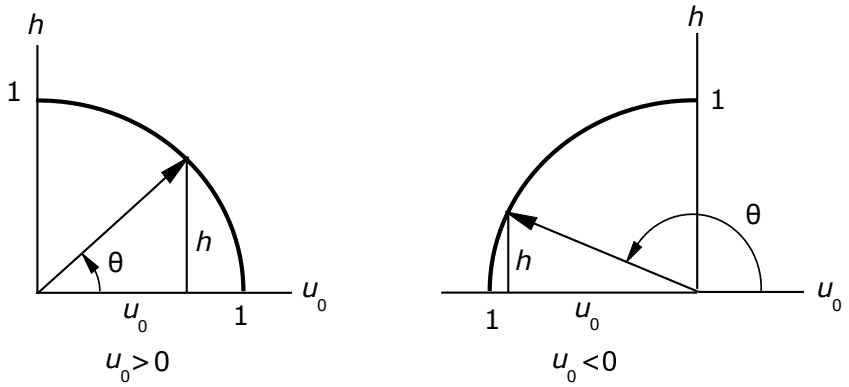
Versors can be written in a special way. Let the versor U be written as $U = u_0 + \vec{\mathbf{u}}$. Let $h = |\vec{\mathbf{u}}|$. Then $\vec{\mathbf{u}}/h$ is a unit vector which we denote, as usual, by $\hat{\mathbf{u}}$. Thus $\vec{\mathbf{u}} = h\hat{\mathbf{u}}$ and

$$U = u_0 + h\hat{\mathbf{u}}$$

and

$$1 = ||U|| = \sqrt{u_0^2 + |h\hat{\mathbf{u}}|^2} = \sqrt{u_0^2 + h^2}$$

We have $h \geq 0$ by its nature as a length, but u_0 can be positive or negative. The previous equation allows us to draw a triangle in the coordinate plane with u_0 on the horizontal axis and h perpendicular to the horizontal axis and the hypotenuse of 1.



Conversion process for Versors

If we take θ to be as shown, we have $u_0 = \cos(\theta)$ and $h = \sin(\theta)$. Note that by construction $0 \leq \theta \leq 180°$. Thus

$$U = u_0 + h\hat{\mathbf{u}} = \cos(\theta) + \sin(\theta)\hat{\mathbf{u}} \qquad 0 \leq \theta \leq 180°$$

and this is the form we like for a versor. We want a special notation for such a versor which will be

$$T_\theta = \cos(\theta) + \sin(\theta)\hat{\mathbf{u}}$$

where hopefully $\hat{\mathbf{u}}$ is clear from the context. If absolutely necessary we can also write

$$T_{\theta,\hat{\mathbf{u}}} = \cos(\theta) + \sin(\theta)\hat{\mathbf{u}}$$

Now we want to determine what T_θ does to a vector perpendicular to $\hat{\mathbf{u}}$. So we take $\vec{\mathbf{v}}$ perpendicular to $\hat{\mathbf{u}}$ which means that $\vec{\mathbf{v}} \cdot \hat{\mathbf{u}} = 0$ We then do the multiplication

$$\begin{aligned} T_\theta\vec{\mathbf{v}} &= (\cos(\theta) + \sin(\theta)\hat{\mathbf{u}})\vec{\mathbf{v}} \\ &= 0 - \sin(\theta)\hat{\mathbf{u}} \cdot \vec{\mathbf{v}} + \cos(\theta)\vec{\mathbf{v}} + \sin(\theta)\hat{\mathbf{u}} \times \vec{\mathbf{v}} \\ &= \cos(\theta)\vec{\mathbf{v}} + \sin(\theta)\hat{\mathbf{u}} \times \vec{\mathbf{v}} \end{aligned}$$

Since $\hat{\mathbf{u}}$ and $\vec{\mathbf{v}}$ are perpendicular, the length of $\hat{\mathbf{u}} \times \vec{\mathbf{v}}$ is $|\hat{\mathbf{u}}|\,|\vec{\mathbf{v}}|\sin(90°) = |\vec{\mathbf{v}}|$. Thus in the plane Π perpendicular to $\hat{\mathbf{u}}$ the two vectors $\vec{\mathbf{v}}$ and $\hat{\mathbf{u}} \times \vec{\mathbf{v}}$ have the same length and are perpendicular to one another and positively oriented (fingers from $\vec{\mathbf{v}}$ to $\hat{\mathbf{u}} \times \vec{\mathbf{v}}$ implies thumb goes up $\hat{\mathbf{u}}$) and thus they form a sort of positively oriented coordinate basis for Π. Note also that, since $\vec{\mathbf{v}}$ and $\hat{\mathbf{u}} \times \vec{\mathbf{v}}$ are perpendicular and $|\hat{\mathbf{u}} \times \vec{\mathbf{v}}| = |\vec{\mathbf{v}}|$ the Pythagorean theorem gives us

$$\begin{aligned} |T_\theta\vec{\mathbf{v}}|^2 &= |\cos(\theta)\vec{\mathbf{v}}|^2 + |\sin(\theta)\hat{\mathbf{u}} \times \vec{\mathbf{v}}|^2 \\ &= \sin^2(\theta)|\vec{\mathbf{v}}|^2 + \cos^2(\theta)|\vec{\mathbf{v}}|^2 \\ &= |\vec{\mathbf{v}}|^2 \\ |T_\theta\vec{\mathbf{v}}| &= |\vec{\mathbf{v}}| \end{aligned}$$

The cosine of the angle ϕ between $T_\theta\vec{\mathbf{v}}$ and $\vec{\mathbf{v}}$ is given by

$$\begin{aligned} \cos(\phi) &= \frac{T_\theta\vec{\mathbf{v}} \cdot \vec{\mathbf{v}}}{|T_\theta\vec{\mathbf{v}}|\,|\vec{\mathbf{v}}|} \\ &= \frac{(\cos(\theta)\vec{\mathbf{v}} + \sin(\theta)\hat{\mathbf{u}} \times \vec{\mathbf{v}}) \cdot \vec{\mathbf{v}}}{|T_\theta\vec{\mathbf{v}}|\,|\vec{\mathbf{v}}|} \\ &= \frac{\cos(\theta)\vec{\mathbf{v}} \cdot \vec{\mathbf{v}} + 0}{|\vec{\mathbf{v}}|\,|\vec{\mathbf{v}}|} \\ &= \cos(\theta) \end{aligned}$$

So the angle between $\vec{\mathbf{v}}$ and $T_\theta\vec{\mathbf{v}}$ is θ and since they are the same length, $T_\theta\vec{\mathbf{v}}$ is the rotation of $\vec{\mathbf{v}}$ through the angle θ.

We would like to know what effect T_θ would have on the *right* rather than the left. We would have

$$\begin{aligned} \vec{\mathbf{v}}T_\theta &= \vec{\mathbf{v}}(\cos(\theta) + \sin(\theta)\hat{\mathbf{u}}) \\ &= 0 - \sin(\theta)\hat{\mathbf{u}} \cdot \vec{\mathbf{v}} + \cos(\theta)\vec{\mathbf{v}} + \sin(\theta)\vec{\mathbf{v}} \times \hat{\mathbf{u}} \\ &= \cos(\theta)\vec{\mathbf{v}} + \sin(\theta)\vec{\mathbf{v}} \times \hat{\mathbf{u}} \\ &= \cos(\theta)\vec{\mathbf{v}} - \sin(\theta)\hat{\mathbf{u}} \times \vec{\mathbf{v}} \\ &= \cos(-\theta)\vec{\mathbf{v}} + \sin(-\theta)\hat{\mathbf{u}} \times \vec{\mathbf{v}} \end{aligned}$$

using $\cos(-\theta) = \cos(\theta)$ and $\sin(-\theta) = -\sin(\theta)$. Although this does not conform to our rule $0 \leq \theta \leq 180^\circ$ it seems clear that that right multiplication by T_θ rotates $\vec{\mathbf{v}}$ by $-\theta$. This same effect could be accomplished by multiplying by T_θ^{-1} on the left, so we have

$$T_\theta^{-1}\vec{\mathbf{v}} = \vec{\mathbf{v}}T_\theta$$

We showed a while back that the quaternion T_θ has size 1 which means that $N(T_\theta) = 1$. Recall that the multiplicative inverse of a quaternion is given by $A^{-1} = \frac{\overline{A}}{N(A)}$. Thus

$$T_\theta^{-1} = \frac{\overline{T_\theta}}{N(T_\theta)} = \overline{T_\theta} = \cos(\theta) - \sin(\theta)\hat{\mathbf{u}} = \cos(\theta) + \sin(\theta)(-\hat{\mathbf{u}})$$

where the last equation shows that T_θ^{-1} can be written as a proper versor when the unit vector $\hat{\mathbf{u}}$ has been replaced by it's negative. Try to visualize this; same angle θ but $\hat{\mathbf{u}}$ in the opposite direction.

The equation above is not quite the one we want. We will fix it.

$$\begin{aligned} T_\theta^{-1}\vec{\mathbf{v}} &= \vec{\mathbf{v}}T_\theta \\ T_\theta^{-1}\vec{\mathbf{v}}T_\theta^{-1} &= \vec{\mathbf{v}} \\ \vec{\mathbf{v}}T_\theta^{-1} &= T_\theta\vec{\mathbf{v}} \\ T_\theta\vec{\mathbf{v}}T_\theta^{-1} &= T_\theta T_\theta\vec{\mathbf{v}} \;=\; T_{2\theta}\vec{\mathbf{v}} \end{aligned}$$

If we substitute $\theta/2$ for θ we have the important equation

$$T_{\theta/2}\vec{\mathbf{v}}T_{\theta/2}^{-1} = T_\theta\vec{\mathbf{v}}$$

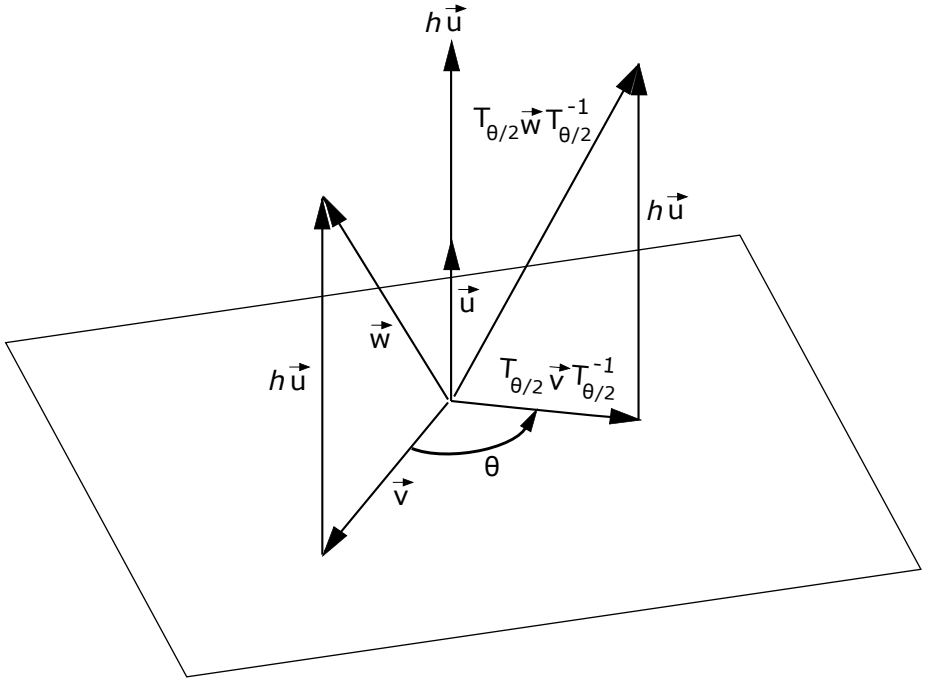
where, you recall, $0 \leq \theta \leq 180^\circ$ and $\vec{\mathbf{v}}$ is perpendicular to $\hat{\mathbf{u}}$. Now $T_{\theta/2}\vec{\mathbf{v}}T_{\theta/2}^{-1}$ certainly looks like a less convenient formula for the rotation of $\vec{\mathbf{v}}$ through an angle of θ around $\hat{\mathbf{u}}$ but in fact it is better, as we shall see. To appreciate this we need to apply this formula to $\hat{\mathbf{u}}$ itself. This is very easy. First note

$$\hat{\mathbf{u}}\hat{\mathbf{u}} = -\hat{\mathbf{u}}\cdot\hat{\mathbf{u}} + \hat{\mathbf{u}}\times\hat{\mathbf{u}} = -1 + 0 = -1$$

so that $\hat{\mathbf{u}}^2 = -1$ just like $i^2 = -1$. All unit vectors square to -1, not just i, j and k. Fairness would have predicted this. Next

$$\begin{aligned} T_{\theta/2}\hat{\mathbf{u}}\,T_{\theta/2}^{-1} &= T_{\theta/2}\hat{\mathbf{u}}(\cos(\theta) - \sin(\theta)\hat{\mathbf{u}}) \\ &= T_{\theta/2}(\cos(\theta)\hat{\mathbf{u}} - \sin(\theta)\hat{\mathbf{u}}^2) \\ &= (\cos(\theta) + \sin(\theta)\hat{\mathbf{u}})(\sin(\theta) + \cos(\theta)\hat{\mathbf{u}}) \\ &= \cos(\theta)\sin(\theta) + \cos^2(\theta)\hat{\mathbf{u}}) + \sin^2(\theta)\hat{\mathbf{u}}) + \sin(\theta)\cos(\theta)u^2 \\ &= \cos(\theta)\sin(\theta) + \hat{\mathbf{u}} - \sin(\theta)\cos(\theta) \\ &= \hat{\mathbf{u}} \end{aligned}$$

Thus $\hat{\mathbf{u}}$ is invariant under this form of the rotation operator.



Rotating with a Versor

Now comes the coup. An arbitrary vector $\vec{\mathbf{w}}$ can be resolved into a sum of a vector $h\hat{\mathbf{u}}$ which is a multiple of $\hat{\mathbf{u}}$ and a vector $\vec{\mathbf{v}}$ in Π perpendicular to $\hat{\mathbf{u}}$. This is done as follows. We set $h = \vec{\mathbf{w}} \cdot \hat{\mathbf{u}}$ and $\vec{\mathbf{v}} = \vec{\mathbf{w}} - h\hat{\mathbf{u}}$. It is obvious that $h\hat{\mathbf{u}}$ and $\vec{\mathbf{v}}$ add to $\hat{\mathbf{w}}$, so it remains to show that $\vec{\mathbf{v}}$ is perpendicular to $\hat{\mathbf{u}}$. This is easy.

$$\begin{aligned} \vec{\mathbf{v}} \cdot \hat{\mathbf{u}} &= (\vec{\mathbf{w}} - h\hat{\mathbf{u}}) \cdot \hat{\mathbf{u}} \\ &= \vec{\mathbf{w}} \cdot \hat{\mathbf{u}} - h\hat{\mathbf{u}} \cdot \hat{\mathbf{u}} \\ &= h - h \ = \ 0 \end{aligned}$$

Now we rotate w

$$\begin{aligned} T_{\theta/2}\vec{\mathbf{w}}\, T_{\theta/2}^{-1} &= T_{\theta/2}(\vec{\mathbf{v}} + h\hat{\mathbf{u}})\, T_{\theta/2}^{-1} \\ &= T_{\theta/2}\vec{\mathbf{v}}\, T_{\theta/2}^{-1} + hT_{\theta/2}\hat{\mathbf{u}}\, T_{\theta/2}^{-1} \\ &= T_{\theta/2}\vec{\mathbf{v}}\, T_{\theta/2}^{-1} + h\hat{\mathbf{u}} \end{aligned}$$

and this is $\vec{\mathbf{w}}$ rotated around $\hat{\mathbf{u}}$ by θ. We see that the reason we must use the more complicated form $T_{\theta/2}\vec{\mathbf{w}}\, T_{\theta/2}^{-1}$ is so we will get $h\hat{\mathbf{u}}$ back at the end in order to reconstitute the rotated $\vec{\mathbf{w}}$ from the rotated $\vec{\mathbf{v}}$. So we see from the picture that

$$T_{\theta/2}\vec{\mathbf{w}}\, T_{\theta/2}^{-1}$$

gives us $\vec{\mathbf{w}}$ rotated by the angle θ around $\hat{\mathbf{u}}$. This is the technology that allows us to show pictures of a rotating earth on a computer screen.

3.4 Vector Identities Via Quaternions

An underappreciated facet of quaternions is that they can be used to derive many vector identities, some of which are difficult to derive by pure vector methods. We will derive a number of these identities, previously mentioned, here.

We first derive Jacobi's identity

$$u \times (v \times w) + v \times (w \times u) + w \times (u \times v) = 0$$

We derive this from the associative law for quaternions $\vec{\mathbf{u}}(\vec{\mathbf{v}}\vec{\mathbf{v}}) = (\vec{\mathbf{u}}\vec{\mathbf{v}})\vec{\mathbf{w}}$. (It is interesting that the associative law for quaternions lets us derive the substitute for the associative law in the (non-associative Lie) algebra $\mathbb{R}$ with cross product. There is a bit of a mystery here.) First we decode $\vec{\mathbf{u}}(\vec{\mathbf{v}}\vec{\mathbf{w}})$ and $(\vec{\mathbf{u}}\vec{\mathbf{v}})\vec{\mathbf{w}}$.

$$\begin{aligned}
\vec{\mathbf{u}}(\vec{\mathbf{v}}\vec{\mathbf{w}}) &= \vec{\mathbf{u}}(-\vec{\mathbf{v}} \cdot \vec{\mathbf{w}} + \vec{\mathbf{v}} \times \vec{\mathbf{w}}) \\
&= -(\vec{\mathbf{v}} \cdot \vec{\mathbf{w}})\vec{\mathbf{u}} - \vec{\mathbf{u}} \cdot \vec{\mathbf{v}} \times \vec{\mathbf{w}} + \vec{\mathbf{u}} \times (\vec{\mathbf{v}} \times \vec{\mathbf{w}}) \\
(\vec{\mathbf{u}}\vec{\mathbf{v}})\vec{\mathbf{w}} &= (-\vec{\mathbf{u}} \cdot \vec{\mathbf{v}} + \vec{\mathbf{u}} \times \vec{\mathbf{v}})\vec{\mathbf{w}} \\
&= -(\vec{\mathbf{u}} \cdot \vec{\mathbf{v}})\vec{\mathbf{w}} - \vec{\mathbf{u}} \times \vec{\mathbf{v}} \cdot \vec{\mathbf{w}} + (\vec{\mathbf{u}} \times \vec{\mathbf{v}}) \times \vec{\mathbf{w}}
\end{aligned}$$

Since $\vec{\mathbf{u}}(\vec{\mathbf{v}}\vec{\mathbf{w}}) = (\vec{\mathbf{u}}\vec{\mathbf{v}})\vec{\mathbf{w}}$, when we subtract these two equations we have 0. Also we remember that we showed earlier that $\vec{\mathbf{u}} \cdot \vec{\mathbf{v}} \times \vec{\mathbf{w}} = \vec{\mathbf{u}} \times \vec{\mathbf{v}} \cdot \vec{\mathbf{w}}$.

$$\begin{aligned}
0 &= \vec{\mathbf{u}}(\vec{\mathbf{v}}\vec{\mathbf{w}}) - (\vec{\mathbf{u}}\vec{\mathbf{v}})\vec{\mathbf{w}} \\
&= -(\vec{\mathbf{v}} \cdot \vec{\mathbf{w}})\vec{\mathbf{u}} - \vec{\mathbf{u}} \cdot \vec{\mathbf{v}} \times \vec{\mathbf{w}} + \vec{\mathbf{u}} \times (\vec{\mathbf{v}} \times \vec{\mathbf{w}}) + (\vec{\mathbf{u}} \cdot \vec{\mathbf{v}})\vec{\mathbf{w}} + \vec{\mathbf{u}} \times \vec{\mathbf{v}} \cdot \vec{\mathbf{w}} - (\vec{\mathbf{u}} \times \vec{\mathbf{v}}) \times \vec{\mathbf{w}} \\
&= -(\vec{\mathbf{v}} \cdot \vec{\mathbf{w}})\vec{\mathbf{u}} + (\vec{\mathbf{u}} \cdot \vec{\mathbf{v}})\vec{\mathbf{w}} + \vec{\mathbf{u}} \times (\vec{\mathbf{v}} \times \vec{\mathbf{w}}) - (\vec{\mathbf{u}} \times \vec{\mathbf{v}}) \times \vec{\mathbf{w}}
\end{aligned}$$

so we have

$$\vec{\mathbf{u}} \times (\vec{\mathbf{v}} \times \vec{\mathbf{w}}) - (\vec{\mathbf{u}} \times \vec{\mathbf{v}}) \times \vec{\mathbf{w}} = (\vec{\mathbf{v}} \cdot \vec{\mathbf{w}})\vec{\mathbf{u}} - (\vec{\mathbf{u}} \cdot \vec{\mathbf{v}})\vec{\mathbf{w}}$$

Now we do a trick which has many applications. There is nothing special about the the vectors $\vec{\mathbf{u}}, \vec{\mathbf{v}}, \vec{\mathbf{w}}$ so we can change the vectors in the equation according to the cyclic permutation $u \to v,\ v \to w,\ w \to u$ and get three similar equations which we will then add up.

$$\begin{aligned}
\vec{\mathbf{u}} \times (\vec{\mathbf{v}} \times \vec{\mathbf{w}}) - (\vec{\mathbf{u}} \times \vec{\mathbf{v}}) \times \vec{\mathbf{w}} &= (\vec{\mathbf{v}} \cdot \vec{\mathbf{w}})\vec{\mathbf{u}} - (\vec{\mathbf{u}} \cdot \vec{\mathbf{v}})\vec{\mathbf{w}} \\
\vec{\mathbf{v}} \times (\vec{\mathbf{w}} \times \vec{\mathbf{u}}) - (\vec{\mathbf{v}} \times \vec{\mathbf{w}}) \times \vec{\mathbf{u}} &= (\vec{\mathbf{w}} \cdot \vec{\mathbf{u}})\vec{\mathbf{v}} - (\vec{\mathbf{v}} \cdot \vec{\mathbf{w}})\vec{\mathbf{u}} \\
\vec{\mathbf{w}} \times (\vec{\mathbf{u}} \times \vec{\mathbf{v}}) - (\vec{\mathbf{w}} \times \vec{\mathbf{u}}) \times \vec{\mathbf{v}} &= (\vec{\mathbf{u}} \cdot \vec{\mathbf{v}})\vec{\mathbf{w}} - (\vec{\mathbf{w}} \cdot \vec{\mathbf{u}})\vec{\mathbf{v}}
\end{aligned}$$

When we add the three equations the terms on the right side all cancel out, and we have, using $\vec{\mathbf{r}} \times \vec{\mathbf{s}} = -\vec{\mathbf{s}} \times \vec{\mathbf{r}}$ etc.

$$\begin{aligned}
\vec{\mathbf{u}} \times (\vec{\mathbf{v}} \times \vec{\mathbf{w}}) &+ \vec{\mathbf{w}} \times (\vec{\mathbf{u}} \times \vec{\mathbf{v}}) \\
+\vec{\mathbf{v}} \times (\vec{\mathbf{w}} \times \vec{\mathbf{u}}) &+ \vec{\mathbf{u}} \times (\vec{\mathbf{v}} \times \vec{\mathbf{w}}) \\
+\vec{\mathbf{w}} \times (\vec{\mathbf{u}} \times \vec{\mathbf{v}}) &+ \vec{\mathbf{v}} \times (\vec{\mathbf{w}} \times \vec{\mathbf{u}}) = 0
\end{aligned}$$

Since each term in the sum occurs twice, we divide by 2 and get the very important Jacobi identity

$$\vec{\mathbf{u}} \times (\vec{\mathbf{v}} \times \vec{\mathbf{w}}) + \vec{\mathbf{v}} \times (\vec{\mathbf{w}} \times \vec{\mathbf{u}}) + \vec{\mathbf{w}} \times (\vec{\mathbf{u}} \times \vec{\mathbf{v}}) = 0$$

Then we have, using the third equation above in the next to the last step,

$$\begin{aligned} \vec{\mathbf{u}} \times (\vec{\mathbf{v}} \times \vec{\mathbf{w}}) &= -\vec{\mathbf{v}} \times (\vec{\mathbf{w}} \times \vec{\mathbf{u}}) - \vec{\mathbf{w}} \times (\vec{\mathbf{u}} \times \vec{\mathbf{v}}) \\ &= (\vec{\mathbf{w}} \times \vec{\mathbf{u}}) \times \vec{\mathbf{v}} - \vec{\mathbf{w}} \times (\vec{\mathbf{u}} \times \vec{\mathbf{v}}) \\ &= -(\vec{\mathbf{u}} \cdot \vec{\mathbf{v}})\vec{\mathbf{w}} + (\vec{\mathbf{w}} \cdot \vec{\mathbf{u}})\vec{\mathbf{v}} \\ &= (\vec{\mathbf{u}} \cdot \vec{\mathbf{w}})\vec{\mathbf{v}} - (\vec{\mathbf{u}} \cdot \vec{\mathbf{v}})\vec{\mathbf{w}} \end{aligned}$$

which is the vector triple product law. The trick of adding up the same equation rewritten by permuting its variables is used a number of times in algebra and geometry. It takes a lot of writing but all the steps are easy and the results often important.

One other thing that can be done using a similar technique is finding the component of a vector perpendicular to a given vector. Let $\vec{\mathbf{u}}$ be the given vector and $\vec{\mathbf{v}}$ be any vector. We wish to resolve $\vec{\mathbf{v}}$ as the sum of two vectors $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_2$ so that $\vec{\mathbf{v}} = \vec{\mathbf{v}}_1 + \vec{\mathbf{v}}_2$, $\vec{\mathbf{v}}_1$ is parallel to $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}_2$ is perpendicular to $\vec{\mathbf{u}}$.

We see immediately that $\vec{\mathbf{v}}_1 = h\vec{\mathbf{u}}$. We have

$$\begin{aligned} \vec{\mathbf{v}} \cdot \vec{\mathbf{u}} &= (\vec{\mathbf{v}}_1 + \vec{\mathbf{v}}_2) \cdot \vec{\mathbf{u}} \\ &= h\vec{\mathbf{u}} \cdot \vec{\mathbf{u}} + 0 \\ &= h|\vec{\mathbf{u}}|^2 \\ h &= \frac{\vec{\mathbf{v}} \cdot \vec{\mathbf{u}}}{|\vec{\mathbf{u}}|^2} \end{aligned}$$

Thus

$$\begin{aligned} \vec{\mathbf{v}} &= h\vec{\mathbf{u}} + \vec{\mathbf{v}}_2 = h\vec{\mathbf{u}} + (\vec{\mathbf{v}} - h\vec{\mathbf{u}}) \\ \vec{\mathbf{v}}_2 &= \vec{\mathbf{v}} - h\vec{\mathbf{u}} \end{aligned}$$

and we see

$$\vec{\mathbf{v}}_2 \cdot \cdot \vec{\mathbf{u}} = (\vec{\mathbf{v}} - h\vec{\mathbf{u}}) \cdot \vec{\mathbf{u}} = \vec{\mathbf{v}} \cdot \vec{\mathbf{u}} - h|\vec{\mathbf{u}}|^2 = \vec{\mathbf{v}} \cdot \vec{\mathbf{u}} - \vec{\mathbf{v}} \cdot \vec{\mathbf{u}} = 0$$

as we wanted. Now we note that by associativity, $(\vec{\mathbf{v}}\vec{\mathbf{u}})\vec{\mathbf{u}} = \vec{\mathbf{v}}(\vec{\mathbf{u}}\vec{\mathbf{u}})$. Decoding this we have

$$\begin{aligned} (-\vec{\mathbf{v}} \cdot \vec{\mathbf{u}} + \vec{\mathbf{v}} \times \vec{\mathbf{u}})\vec{\mathbf{u}} &= \vec{\mathbf{v}}(-\vec{\mathbf{u}} \cdot \vec{\mathbf{u}} + \vec{\mathbf{u}} \times \vec{\mathbf{u}}) \\ -(\vec{\mathbf{v}} \cdot \vec{\mathbf{u}})\vec{\mathbf{u}} - \vec{\mathbf{v}} \times \vec{\mathbf{u}} \cdot \vec{\mathbf{u}} + (\vec{\mathbf{v}} \times \vec{\mathbf{u}}) \times \vec{\mathbf{u}} &= -|\vec{\mathbf{u}}|^2\vec{\mathbf{v}} \\ |\vec{\mathbf{u}}|^2\vec{\mathbf{v}} - (\vec{\mathbf{v}} \cdot \vec{\mathbf{u}})\vec{\mathbf{u}} &= -(\vec{\mathbf{v}} \times \vec{\mathbf{u}}) \times \vec{\mathbf{u}} = \vec{\mathbf{u}} \times (\vec{\mathbf{v}} \times \vec{\mathbf{u}}) \\ \vec{\mathbf{v}}_2 = \vec{\mathbf{v}} - h\vec{\mathbf{u}} = \vec{\mathbf{v}} - \frac{\vec{\mathbf{v}} \cdot \vec{\mathbf{u}}}{|\vec{\mathbf{u}}|^2}\vec{\mathbf{u}} &= \frac{\vec{\mathbf{u}} \times (\vec{\mathbf{v}} \times \vec{\mathbf{u}})}{|\vec{\mathbf{u}}|^2} \end{aligned}$$

so the formula for the perpendicular component is

$$v_2 = \frac{\vec{\mathbf{u}} \times (\vec{\mathbf{v}} \times \vec{\mathbf{u}})}{|\vec{\mathbf{u}}|^2} = \frac{(\vec{\mathbf{u}} \times \vec{\mathbf{v}}) \times \vec{\mathbf{u}}}{|\vec{\mathbf{u}}|^2}$$

and as we see in the picture $\vec{\mathbf{w}}$ has been rotated by θ around the vector $\vec{\mathbf{u}}$.

3.5 Problems for Chapter 3

Section 3.1

1. What sort of problems was Hamilton trying to solve with the invention of quaternions? What surprised Hamilton about the solution?

2. What property of quaternions was surprising and upsetting for people in Hamilton's time but now is considered perfectly reasonable and commonplace?

3. It is extremely important that quaternion multiplicaton is associative. It is easy to prove that if all combinations of three basis vectors are associative then the whole system is associative. Prove that the following small selection of the 27 non-trivial possibilities satisfy the associative law by calculating each side and showing they are the same.

$$\begin{aligned} (ij)k &= i(jk) \\ (jk)j &= j(kj) \\ (ki)i &= k(ii) \\ (ki)j &= k(ij) \end{aligned}$$

Obviously this is a crude methodology and in a later problem we will show the associativity in a much more satisfying (and easy) way.

4. The *center* of an algebra consists of the elements a which commute with *every* element of the algebra. What is the center of the the quaternion algebra?

Section 3.2

1. Put i at 12 o:clock, j at 4 o:clock, and k at 8 o:clock. Draw arcs between the letters and put arrowheads on the end of the arcs pointing in a clockwise direction. Now figure out how this diagram tells you the fundamental multiplication relations between the basis vectors i, j and k.

2. Multiply the quaternions
a) $2 + 3i + 4j - k$ and $3 - 2i - j - 3k$. Check your work by finding the norms and seing if the norm of the product is the product of the norms.
b) $2 + 3i - j - 2k$ and $2 - 3i + j + 2k$
c) find the inverse of the quaternion $2 + i - 2j + k$.

Chapter 4

INFINITE NUMBERS

This chapter on infinities is divided into three parts, Part A and Part B and Part C. Each concerns infinite numbers but they are different kinds of infinite numbers and the three parts are totally independent of one another. Part A concerns the infinite numbers used in counting, and Part B concerns the infinite, and infinitely small, numbers that were used in the early development of Calculus. Thus part B will have just the tinest bit of Calculus, and a certain amount of history, and much more. Part C, which you will probably come in contact with first in your mathematical education, could be called symbolic infinity, and can be regarded as a manner of speaking. It is by far the simplest, and the least interesting but is a source of much confusion. You won't be as confused as your classmates since you will have read this book.

PART A

4.1 Introduction

Between 1874 and 1884 Georg Cantor (1845-1918) created set theory. Before Cantor interest in set theory was low and the material elementary but Cantor changed all that with his introduction of the idea that infinite sets could have different sizes. This was radically new. Cantor first defined what it meant for two infinite sets to have the same number of elements and then was able to show that there are more real numbers than there are integers. We will discuss all this in some detail, because though considered advanced by some, it is actually all fairly simple at the beginning. Many of Cantor's papers are very readable and have been collected in [Cantor].

Cantor's work generated both great enthusiasm (Hilbert, Hurwitz, Hadamard and many others) and deep antagonism (Kronecker, Brouwer, Poincare, Weyl). By the time Cantor was publishing his papers on set theory Kronecker had already begun a program to remove infinite sets from mathematics, so Cantor's and Kronecker's programs came into direct conflict. As Kronecker was important, Cantor's career suffered setbacks, though perhaps not as many or as severe as Cantor claimed. The early attacks hindered the acceptance of set theory by many mathematicians, but again not nearly as much as Cantor and his partisans claimed and in fact Cantor's set theory was accepted by the majority of mathematicians by 1900. Hilbert's 1900 quote "From the paradise created for us by Cantor no one shall drive us forth,[1]" describes the general feeling. A fair minded person is in fact astounded by how quickly Cantor's ideas became standard mathematics. Most originators have to wait far longer. (Cantor suffered from some sort of mental disease (bipolar?), which colored his view of the matter. Also Cantor was overly affected by the negative opinions of some philosophers but these had little effect on the situation in Mathematics.) Ironically, set theory eventually came to be seen as the standard basic discipline from which higher forms of mathematics are derived.

[1] As is standard with Paradises there is a snake. We will discuss the snake at a later time.

The resistance to Cantor's set theory eventually gelled into a variety of related movements, Intuitionism (Brouwer, Heyting), finitism, computability etc. The problem with all the alternatives to Cantor's form of set theory is that they make the development of ordinary garden variety Calculus either impossible or very much harder. Ordinary folks felt that Calculus was already hard enough, and philosophical objections of doubtful value did not justify flogging students through a more difficult approach. Thus while finitists and other resisters continue in a small way, the great majority of mathematicians are happy enough doing their work using what have become the standard methods despite various annoying features in the foundations. We will mention some of the annoying features at appropriate times.

Although they represent only a small minority of mathematicians, those who form the resistance to the set theory descended from Cantor are often quite respected mathematicians and can make quite a lot of noise. For example Errett Bishop (1928-1983, American)published a book FOUNDATIONS OF CONSTRUCTIVE ANALYSIS in 1967 which stirred up some interest in what I have called finitist methods. However, one always falls back on the problem of making a lot of standard mathematics harder, and although there are certainly payoffs (explicit computability) the price of taking this road is high.

I remark there there is some similarity here to the situation in physics, where the interpretation of what Quantum Theory really means is far from clear, but it is still possible to calculate what one needs to make lasers, understand how quarks make up protons and neutrons, design faster and more efficient computers and manufacture ever more destructive bombs. In both cases vagueness at the foundations does not seem to have any effect on progress, although the occasional physicist has the occasional sleepless night, and one hears the howls of philosophers in the distance.

In fairness I must mention that when the digital computer was invented it was not necessary to develop a lot of new math to help it. All the mathematics computers would ever need had been developed already by the followers of Kronecker and Brouwer in their efforts to reconstruct mathematics along finitistic lines. Their mathematics was exactly what computers could do and they had shown how computers should do it. Though Kronecker was a very short man, standing atop the digital computer he today casts a remarkably long shadow.

Cantor was deeply interested in the philosophical impact of his work, though many philosophers were deeply suspicious. His interest in the philosophical aspects of his ideas and their reception among the philosophers may have contributed to his sense that his work was underappreciated, which was very far from true. Worrying about what the philosophers thought may not have helped his mental problems[2].

One final point I would like to emphasize is that Cantor was not led to his investigations in set theory by vague or philosophical speculation. His original

[2]The interaction between mathematics and philosophy has been bumpy and has not made either discipline happier. Much the same is true of physics; philosophical worries almost surely contributed to Ludwig Boltzmann's suicide (1906). Boltzman was a pioneer in the molecular theory of matter.

interest was the convergence of the Fourier series of a function to the function. (The Fourier series of a function is an infinite series which when summed up gives the value of the function, hopefully. This was of enormous importance in both mathematics and mathematical physics and has many practical uses.) Cantor was seduced into set theory by questions about the set of points at which convergence *failed*. From there he went on to arbitrary sets of real numbers and their properties, and from there to sets in general. This is the normal sort of development in mathematics; from specific problems to development of new tools and theories. It is very rare for a new theory to be plucked whole from thin air, although many people unfairly view mathematics that way.

4.2 The Algebra of Sets

Although Leibniz played around a little with sets in the late 1600s, the persons who first set up our current system were George Boole and Gerhardt Schröder. The concept of set is intuitively relatively clear although a sophisticated treatment requires care. Taking the naive view, a set has elements. We will use upper case letters for sets and lower case letters for elements to the extent this is possible. We will use braces { and } to make containers for sets, so the set A whose elements are 1,3, and 5 will be written

$$A = \{1, 3, 5\}$$

Sets can contain absolutely any objects (including other sets although although this requires care and can be troublesome). Thus we can have

$$B = \{a, b, c, d, e\} \text{ or } C = \{\text{Hillary Clinton}, \text{Melania Trump}, 125, f\}$$

The elements of a set need not be related but they usually will be in what we do. For typographical convenience most of our sets will have numbers and letters as elements.

This might be an appropriate moment to mention one of the dangers of set theory. Although the idea of sets seems intuitively clear, it easily leads to contradictions if care is not exercised, a lot of care. The early days of set theory, say 1880 to 1907, produced a number of contradictions. It was annoying that though one had a feeling that the contradictions had things in common no single prohibition seemed to solve the problem. One of the ideas that can be helpful is to forbid a set from containg itself, so you don't want $A = \{2, 5, A\}$. However, this is not as easy as it looks since one could have $A = \{2, 5, B\}$ and $B = \{3, 7, A\}$. There is a way to stop all this sort of pathology from happening called the Axiom of Foundation (Fundierungsaxiom) which is part of the standard Zermelo Axioms for Zermelo set theory.

Now you might suspect from the above that there are *varieties* of set theory. The one commonly used in mathematics is the Zermelo set theory based on Zermelo's axioms. There are several other varieties, one of which traces back to Bertrand Russel and Alfred North Whitehead and there are a couple of

set theorys invented by Williard van Orman Quine in the Russel/Whitehead traditon. In some of these theories $A = \{2, 5, A\}$ would be allowed. All the theories go to great lengths to avoid contradictions, but Kurt Gödel showed that it is impossible to prove that a strong system is contradiction free. (Strong here means strong enough to develop arithmetic.) So most mathematicians coast along with Zermelo's system and have faith that it is consistent.

We now return to the development of set thory and you need to have faith that everything I do can be justified in Zermelo's version of set theory.

Def Two sets are identical if they have the same elements.

To illustrate this point, note that $\{a, b, c, d\} = \{a, a, b, c, c, d\}$ since both sets have elements a, b, c, d. Repeated elements should just be thrown out of the set as this will not affect the membership.

We introduce a new symbol to indicate that an element is in a set. Often the letter ϵ is used but there is a special symbol with no other use and we will use that. Thus if $A = \{a, b, 3, 9\}$ we will indicate that 3 is in the set A by $3 \in A$ and will indicate that c is *not* in the set by the symbol $c \notin A$. This symbol in one form or another is very pervasive in modern mathematics since mathematicians tend to talk in terms of sets and their elements.

Also, there is a notation which veries in detail for building sets from descriptions. For example, the set of prime positive integers could be described by $P = \{x \in \mathbb{Z} \mid x \text{ is prime and } x > 0\}$. The part before the vertical line has a variable (here x) and an optional big set to indicate the general range of the variable, here $\mathbb{Z}$. After the vertical line there is a description involving the variable x. Elements of $\mathbb{Z}$ for example 7, are substituted for x and the resulting statement is either true or false. If "7 is prime and $7 > 0$" is true then 7 in in the set P or, in symbols, $7 \in P$. If you put 6 in for x you get "6 is prime and $6 > 0$" and the statement is false and 6 is not in P, or in symbols $6 \notin P$. We can then write out the elements in a list, sort of, and we have

$$P = \{x \in \mathbb{Z} | x \text{ is prime and } x > 0\} = \{2, 3, 5, 7, 11, 13, 17, \ldots\}$$

The three dots in the set mean that if you are smart enough or experienced enough you can see what the rest of the list will look like but since it is infinitely long I am not going to list all of them. For example in the set

$$F = \{0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, \ldots\}$$

it may not be obvious to you what the next elements are, but, if you are smart enough, you recognize that each element is the sum of the previous two elements and thus you can carry on as long as you like, and if you are experienced enough you will recognize this as the Fibonacci sequence, which has lots of interesting properties, and is defined by $F_0 = 0$, $F_1 = 1$ and $F_{n+1} = F_{n-1} + F_n$. If you are a super genius you see that the formula is

$$F_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right]$$

but nobody is *that* smart; you have to already know it from somewhere.

I mention that $\{x \in \mathbb{Z} \mid x \text{ is prime and } x > 0\}$ is one way of making the set of prime numbers but there are many many ways to do this using different descriptions. The set descriptor way of making a set is far from unique. The list, however, is unique except for order. You can move the elements around as you like. $\{a, b, c, d\} = \{b, c, a, d\}$.

Even this early in the development of set theory problems can arise. It may not be possible to evaluate the condition due to lack of knowledge, which may be temporary or permanent. The first is illustrated by the example

$$\{x \mid (x = 1 \text{ if there is an even number not the sum of two primes}) \text{ otherwise } x = 0\}$$

This set is either $\{0\}$ or it is $\{1\}$ but just now we can't tell which it is because we don't know if every even number is the sum of two primes.

The second is illustrated by

$$\{x \mid (x = 1 \text{ if Julius Caesar had epilepsy}) \text{ otherwise } x = 0\}$$

It is probable that the set is $\{1\}$ but we cannot ever be sure because it is buried too far in the past to test.

This problem has led some mathematicians (and philosophers) to insist that the condition be something definite and decidable. But it is hard to define exactly what that means, so the cure may be worse than the disease. We will avoid the problem by not using sets with conditions we cannot decide the truth of.

Any mathematical theory consists of objects, relations and operations. Relations are between objects, like $3 < 5$. Operations combine a number of objects to make new objects, like 2+3=5. We now introduce these for sets

Given two sets we can find their *union*. If $A = \{1, 2, 3, 4, 5\}$ and $B = \{4, 5, 6, 7, 9\}$ then the union $A \cup B$ is the elements that are in one or the other set. Here $A \cup B = \{1, 2, 3, 4, 5, 6, 7, 9\}$. Another example: recall that $\mathbb{N}^+ = \{1, 2, 3, \ldots\}$ so

$$\mathbb{Z} = \mathbb{N}^+ \cup \{x \mid -x \in \mathbb{N}^+\} \cup \{0\}$$

The second set is of course the negative integers. Make sure you see why. Notice that we can use the new set notation to define union.

Def $A \cup B = \{x \mid x \in A \text{ or } x \in B\}$

Notice that in mathematics when we use "or" it is true even if both alternatives are true. This differs a bit from ordinary language where we don't tend to use or if both alternatives are true. In mathematics "I will eat beans or I will eat carrots" is true if you load up your plate with both and eat them all.

The next operation is called intersection and the symbol is $\cap$.

Def $A \cap B = \{x \mid x \in A \text{ and } x \in B\}$

Thus the intersection of A and B is the elements that are in each of the sets.

Example: If $A = \{a, b, c, d, e\}$ and $B = \{a, c, e, g, i, k\}$ then $A \cap B = \{a, c, e\}$.

A very important special case is when there are *no* elements that are in both sets; the intersection is the empty set. There is a special notation for the empty set; it is $\emptyset$. Thus $\emptyset = \{\ \}$. We will have great use of the empty set when we define the natural numbers in terms of set theory. Also the idea of empty intersection leads to an important new concept.

Def Two sets A and B are *disjoint* if and only if $A \cap B = \emptyset$.

Since sets are determined by their elements, there is only *one* empty set. The set of integers which are both even and odd is equal to the set of married bachelors or the set of humans who had no mothers. Once again, there are difficulties; consider $\{x \mid x \text{ is a horse and } x \text{ can fly}\}$. This is usually thought of as the empty set, but what should we do with Pegasus, Bellerophon's trusty steed. This leads to the question of what exists. The set of things a person thinks exist are called the person's *ontology*. We are never going to all agree on a common ontology, but for specific purposes we may decide to provisionally limit ourselves to a common ontology called the *Universe of Discourse*, and we may or may not decide to include Pegasus. For purposes of this book poor Pegasus is *out*. You can see that we are here treading close to the fence separating mathematics from philosophy, and it is time to tiptoe quietly away.

We pause a moment to consider $\cup$ and $\cap$. From the point of view of teaching newcomers to the subject, to have the symbols resemble one another so closely is very poor pedagogy. However, mathematicians love symmetry, and it turns out that any true equation with $\cup$ and $\cap$ remains true if you turn them both over. For example

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

are both true and you may recognize them as distributive laws. So the beautiful symmetry shown here trumps the difficulty students have when the symbols look pretty much the same but have very different meanings. However, I will give you a mnemonic to remember which is which. In $A \cup B$ you should think of dumping the elements of A and B together into a large pot, represented by the $\cup$. On the other hand, in $A \cap B$ if you try to dump both A and B onto the $\cap$ most will roll off the $\cap$; the only ones which balance on top of the $\cap$ are the ones that are in both sets. OK, that should keep it straight, I hope.

There are two other set operations which we include for completeness; the first is the the set difference.

Def The (set) difference of two sets A and B is $A - B = \{x \mid x \in a \text{ and } x \notin B\}$

Thus to get $A - B$ we remove from A any element that is in B. See if you can see why the equation $A - B = A - (A \cap B)$ is true.

The second set operation is useless except for one really surprising thing. The opearation is

Def The (set) symmetric difference is $A \triangle B = (A - B) \cup (B - A)$

Is this a useful operation? Not very much. But there is one really interesting thing about it. Suppose we give ourselves a Universal set $\mathbf{U}$ which contains all the things under discussion (see below for more on this.) Now suppose we make some notational changes. We set $A+B = A \triangle B$ and $A \cdot B = A \cap B$. We let $0 = \emptyset$ and $1 = \mathbf{U}$. Then we get a kind of a miracle; The set $\mathbf{U}$ with the operations $+$ and $\cdot$ and 0 and 1 as defined is a ***RING*** as we defined in Chapter 1. I'd love to give you this as a problem but the distributive laws are a bit of a pain to prove. The thing to recognize here is that this is a ring without any relationship to numbers, so the algebraic concept ring is not dependent on the elements of the ring being numerical in any way.

Now for set relations. There is only one at the moment but it is very important. We say A is a subset of B if and only if any element of A is also in B. The symbol is $\subseteq$. Thus

Def $A \subseteq B$ if and only if (for any x, if $x \in A$ then $x \in B$)

For example, the set of primes is a subset of the set of integers. The set of squares is a subset of the set of rectangles. The set of dogs is a subset of the set of mammals. The set of politicians is a subset of the set of (fill in as you like).

Note that if $A = B$ then $A \subseteq B$ and $B \subseteq A$. Note the converse is also true: if $A \subseteq B$ and $B \subseteq A$ then $A = B$. This is a very common way of proving two sets are equal; show each is a subset of the other.

There is a symbol for

A is a subset of B but A is not equal to B

which is $A \subset B$ but this is not a very useful concept and we won't use it[3].

It is very important to keep the difference between element and subset clear. It is useful to have the concept of singleton set here. A singleton set is a set with only one element, for example $\{a\}$ is a set whose only element is a, and naturally $a \in \{a\}$. Here is one way to think of it. The element a is a thing. The set $\{a\}$ is a box with one thing, a, in it. $\{a, b, c\}$ is a box with three things in it. See the difference. Naked thing vs. thing in a box. (If you ever have to teach this stuff this is one of the things that you'll have difficulty getting across; humans have difficulty distinguishing element and subset.)

One more time. If $a \in A$ then $\{a\} \subseteq A$. If a is an element of A then $\{a\}$ is a subset of A. And backwards; if $\{a\}$ is a subset of A then a is an element of A.

There is one other complex of ideas which we will not be using but is always included in treatments of set theory. Read on if you are curious; otherwise skip to the next section.

For many purposes it is convenient to deliniate the objects under discussion and create a universal set $\mathbf{U}$ that contains them all. If one is working with numbers then one might use $\mathbf{U} = \mathbb{R}$ or $\mathbf{U} = \mathbb{C}$ or if working with plane geometry one might take $\mathbf{U}$ to be the set of points, lines, geometric figures, etc in the plane, or one might restrict $\mathbf{U}$ to be just the points of $\mathbb{R}^2$. So let us say we have selected

[3]Some mathematicians use $\subset$ in the way I use $\subseteq$. You should always check which convention the book you are reading is using.

a universal set **U** for our purposes. (It is not really practical to say that **U** is everything there is because we then run up against the Pegasus problem again.)

Having selected **U**, we can then introduce the idea of complement[4].

Def The *compliment* of a set A, denoted A', is $\mathbf{U} - A$.

Thus A' is the set of things under discussion that are *not* in A. In a mammology class it would be reasonable on some days to take **U** to be the set of mammals. Then if P is the set of pandas, P' is the set of mammals which are not pandas. If D is the set of dogs then $D \subseteq P'$. If B is the set of bears, then $P \cup B$ is the set of pandas and bears together, or the set of bearlike beasts. Then $(P \cup B)'$ is the set of mammals that are neither pandas nor bears, which is $P' \cap B'$. This is an example of De Morgan's laws which are

Theorem De Morgan's Laws

1. $(A \cup B)' = A' \cap B'$

2. $(A \cap B)' = A' \cup B'$

These are a set theory form of the logical laws

not(P and Q) is equivalent to (not P or not Q)
not(P or Q) is equivalent to (not P and not Q)

Both the set and logical forms can be useful in real life. Lawyers use them in legal documents. Politicians misusing them have brought down whole empires.

4.3 Equivalence for Sets

Now we want to learn to count sets. The basic idea here is the *equivalence* of sets. This is not a great choice of word for the idea but we are stuck with it. Basically, two sets are equivalent if and only if they have the same number of elements. But this assumes that we can count. We want to define equivalence without using numbers.

Remember the (possibly mythical) primitive tribe that could only count to three? Oog has many sheep and he wishes to give some of them to his two sons Tlog and Fnog. He wants to keep the peace so he wants to give the same "number" of sheep to each but there are no numbers big enough (more than 3) to use. Tlog arrives to get his sheep, but Fnog is busy raiding a tribe on the other side of the mountain. They turn to Swinthilla, the youngest of the family, who is more than 3 years old, hoping she can provide a solution. She can. Swinthilla advises them to collect a fair number of rocks. Then as each sheep is given to Tlog, a rock is put in a special pile. When Oog feels he has given Tlog enough, Tlog returns to his family followed by his sheep which are kept together by his sheep dog, Spot. Swinthilla returns to her research, which

[4]We are using a prime to denote compliment A', but many other symbols are also used for the same purpose, like $\overline{A}$ or $\mathbf{C}A$. $\overline{A}$ is a particularly bad choice because it conflicts with the notation for closure in topology.

concerns the invention of a number beyond 3. There is little support in her family for her endeavors. As we have seen, suggestions for new numbers are never met with enthusiasm but Swinthilla persists, like a good mathematician. Her only worry is, will the new number be weaponized.

Eventually Fnog arrives to collect his sheep, and rocks are taken off the pile as the sheep are handed over to Fnog, who wished he knew what was coming so he could have dumped a couple of extra rocks on the special pile. When the pile is empty, the proper "number" of sheep have been given to Fnog and he returs home with his sheep kept together by his sheep dog Nospot.

The mathematical object being used here is a one-to-one correspondence or bijection. (Bijection is the modern term for it.) In both cases to each rock corresponds a sheep, and to each sheep a rock. A second bijection between Tlog's sheep and Fnog's sheep can be constructed making correspond to each of Tlog's sheep the sheep of Fnog that corresponds to the same rock. Thus a bijection between Fnog's sheep and Tlog's sheep can be established and so they both have the same number of sheep. Tlog's set of sheep is said to be *equivalent* to Fnog's set of sheep.

Def A *Bijection* between two sets A and B is a correspondence between the elements of A and the elements of B such that to each element of A there corresponds a unique element of B and to each element of B there corresponds a unique element of A.

Def Two sets A and B are *equivalent* if and only if there exists a bijection from A to B

Example: The set $A = \{a, b, c\}$ and the set $B = \{0, 1, 2\}$ are equivalent as shown by the following table which illustrates the bijection.

$$\begin{array}{ccc} b & a & c \\ | & | & | \\ 0 & 1 & 2 \end{array}$$

Note that order doesn't matter at all.

There is a synonym for equivalent; Two sets A and B have the same number of elements if and only if they are equivalent if and only if there is a bijection between them. We don't really need all this names but they are customary. The phrase "have the same number of elements" should be thought of as a fixed item and should be used with only tiny variation, like A has the same number of elements as B. A and B are in bijection, while correct, is probably a little less clear than the others.

4.4 Definition of the Natural Numbers

Before we begin it might be useful to have a little philosophical discussion. Imagine a garden party of philosophers, everyone chatting happily. You take an apple, carve on it "What Exists" and throw it into the party. Within minutes wives are pulling their husbands away from fist fights.

The technical term for what exists is *ontology*. A person's ontology is the set of what he believes exists. These questions are deep. Consider Julious Caesar. He certainly doesn't exist in the ordinary sense of existing *now*, but he certainly exists in space-time. The city of Troy didn't exist in 1850; it was a myth. But now it exists, as ruins, because Heinrich Schliemann dug it up. The kings of the Chinese Shàng dynasty used to not exist; mythical again, but now they exist as historical entities, like Julius. Nobody is quite sure about the kings of the previous Xià dynasty. And what about fictional characters like Captain Ahab and his pet whale. Another favorite is Dumbo, the elephant that can fly with his big ears. Or what about 2. Does 2 exit? Where? If 2 exists, what about i? What about the characters in historical novels, which are based on definitely once existing people, like Julius. So existence is a slippery concept, and this has implications for mathematics.

The philosopher Bertrand Russel and the mathematician Alfred North Whitehead wrote a book about the foundations of mathematics, called Principia Mathematica. (The title was consciously modeled on the title of Newton's physics book.) Russel and Whitehead were trying to base the foundations of mathematics on the theory of sets. This was not the first such attempt; Frege had written a 3 volume work with the same aim. Russel torpedoed it with a postcard. We discuss this later. Here we are interested in his definition of 2. Russel and Whitehead defined 2 as the set of all pairs. Thus

$$2 = \{z \mid z = \{x, y\}\}$$

Thus $\{MotherTeresa, AlbertEinstein\} \in 2$. Also $\{a, b\} \in 2$, $\{\frac{1}{2}, \frac{9}{10}\} \in 2$. But a mathematician might well object that pairs consist of two things that exist, and what exists is slippery. Is $\{MotherTeresa, Dumbo\}$ a pair and is it in 2? And there are other treacherous hazards. Is $\{1, 2\}$ a pair? Looks innocent until you look closer, and there is the thing we are trying to define sitting inside itself.

Whitehead and Russel were trying to build mathematics with a set theory that applies to ordinary life. But the existence or non-existence of Dumbo is not really a mathematical question; should it be part of the mathematical object 2. And the last sentence of the previous paragraph leads into very treacherous waters indeed. Need we bathe in them.

It is at this point that Mathematics and the Philosophy of Mathematics part company. The philosophy of Mathematics is done by philosophers who have great sympathy with Whitehead and Russel's approach. They still are inclined to follow the Whitehead/Russel path. Mathematicians, however, soon tired of the complications of this approach and sought a simpler way. In 1908 the German mathematician Ernst Zermelo, 1871-1953, analyzed exactly what part of set theory was needed in order to build the integers and Calculus, and set up a system of axioms to do this. This is the mathematical way to solve problems; set up axiom systems. We will discuss the axiom system later. For now we only need to know it justifies the empty set $\emptyset$ and sets built from already existing sets by stuffing them together into sets. His definition of the integers was perfectly good but not quite as good as an improvement developed by the

Hungarian mathematician John von Neumann 1903-1957. Here is the definition of the non-negative integers (more precisely the cardinal numbers) according to von Neumann:

$$\begin{aligned}
0 &= \emptyset = \{\,\} \\
1 &= \{0\} = \{\{\,\}\} \\
2 &= \{0,1\} = \{\{\,\},\{\{\,\}\}\} \\
3 &= \{0,1,2\} \\
4 &= \{0,1,2,3\} \\
5 &= \{0,1,2,3,4\} \\
6 &= \{0,1,2,3,4,5\} \\
\vdots &\quad \vdots \\
\aleph_0 &= \{0,1,2,3,4,5,\ldots\}
\end{aligned}$$

The symbol at the end is the first letter of the Hebrew alphabet[5] with a subscript 0. There are $\aleph$s with higher subscripts but we will have almost nothing to do with them. $\aleph_0$ is the cardinal number of things that come in infinitely long lists, as we will see. To be clear, $\aleph_0$ is the set of non negative integers. Notice that von Neumann's definition makes $\aleph_0$ look like the other numbers, but bigger.

We are now in a position to count the elements of some sets. The cardinal number of a set is the cardinal number it can be put in one to one correspondence with. For example $A = \{a, b, r, s\}$ has four elements because of the following one to one correspondence:

$$\begin{array}{ccccccc}
A & = & a & b & r & s \\
 & & | & | & | & | \\
4 & = & 0 & 1 & 2 & 3
\end{array}$$

I hope you begin to see the cleverness of this definition, which is characteristic of von Neumann. At this point we need to introduce a notation for the cardinal number of a set. We use the unimaginative notation $n(S)$ for the cardinal number of the set S. Thus with $A = \{a, b, r, s\}$ we have $n(A) = 4$ because A is in one to one correspondence with the cardinal number 4. Alternative notations are

$$n(A) = |A| = \overline{\overline{A}}$$

The last is Cantor's notation but it is typographically inconvenient (although I prefer it). We will stick with $n(A)$.

Going on, we wish to find the cardinal number of the set E of non-negative even integers. It is $\aleph_0$, as the following shows.

$$\begin{array}{ccccccccc}
E & = & 0 & 2 & 4 & 6 & \ldots \\
 & & | & | & | & | & \ldots \\
\aleph_0 & = & 0 & 1 & 2 & 3 & \ldots
\end{array}$$

[5] pronounced alef with a like the a in rah as in the cheer, or Ra the Egyptian God. With the subscript it is pronounced alef-null.

We can even describe the bijection algebraically: an x in $\aleph_0$ corresponds to $2x$ in E. Note that we can write our result as $n(E) = \aleph_0$ which might be read as: The cardinal number of the set of non-negative even integers is aleph null.

Now you would probably like to see a bigger infinity, like perhaps the number of fractions. Surely there must be more fractions than there are integers. But in fact this is not true. There are $\aleph_0$ fractions. I'll show this for the positive fractions.

$$\begin{array}{ccccccccccc}
1/1 & \rightarrow & 1/2 & & 1/3 & \rightarrow & 1/4 & \rightarrow & 1/5 & \rightarrow & \\
 & \swarrow & & \nearrow & & \swarrow & & \nearrow & & \swarrow & \\
2/1 & & 2/2 & & 2/3 & & 2/4 & & 2/5 & & \\
\downarrow & \nearrow & & \swarrow & & \nearrow & & \swarrow & & \nearrow & \\
3/1 & & 3/2 & & 3/3 & & 3/4 & & 3/5 & & \\
 & \swarrow & & \nearrow & & \swarrow & & \nearrow & & \swarrow & \\
4/1 & & 4/2 & & 4/3 & & 4/4 & & 4/5 & & \\
\downarrow & \nearrow & & \swarrow & & \nearrow & & \swarrow & & \nearrow & \\
5/1 & & 5/2 & & 5/3 & & 5/4 & & 5/5 & &
\end{array}$$

From the array we can now write out all the fractions in a list in two steps. First step follow the arrows:

$$1/1, 1/2, 2/1, 3/1, 2/2, 1/3, 1/4, 2/3, 3/2, 4/1, 5/1, 4/2, /3/3, 2/4, 1/5, 1/6, 2/5, 3/4, 4/3, 5/2, 6/1, \ldots$$

The second step is to remove the repetitions, like $1/1, 2/2, 3/3, \ldots$ and just leave the first one getting

$$1/1, 1/2, 2/1, 3/1, 1/3, 1/4, 2/3, 3/2, 4/1, 5/1, 1/5, 1/6, 2/5, 3/4, 4/3, 5/2, 6/1, \ldots$$

This list we then put in one to one correspondence with $\aleph_0$, proving there are exactly $\aleph_0$ positive fractions. Although it is almost unbelievable, there is actually a formula where you plug in an non-negative integer and it outputs the corresponding fraction, but it is too complicated to present here.

If you want the negative fractions and 0 too, you use the list

$$0/1, 1/1, -1/1, 1/2, -1/2, 2/1, -2/1, 3/1, -3/1, 1/3, -1/3, 1/4, -1/4, 2/3, -2/3, \ldots$$

showing there are $\aleph_0$ fractions.

It is also possible to show that the set of algebraic numbers has cardinal $\aleph_0$ but the demonstration would take a lot of work so maybe I'll just indicate how it works. There are ways to list all the algebraic equations with integer coefficients, and each such equation has only finitely many solutions, so that is the basic idea. The details are not particularly interesting.

For our next trick I must explain how to tell if one cardinal number is bigger than another. This requires a couple of definitions. There are some little things to prove here which I will omit. And *remember* that a subset of a set A includes the possibility that the subset is A itself.

Def The cardinal number $\mathfrak{a}$ is less than or equal to the cardinal $\mathfrak{b}$ (notation $\mathfrak{a} \leq \mathfrak{b}$) if and only if $\mathfrak{a}$ is in one to one correspondence with a subset of $\mathfrak{b}$.

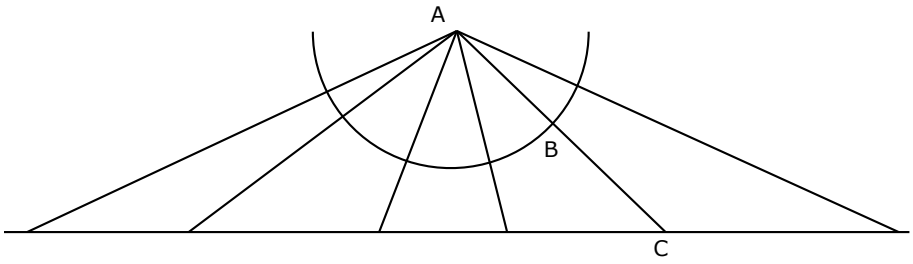
Thus $3 \leq \aleph_0$. To get $\mathfrak{a} < \mathfrak{b}$ we have to show that $\mathfrak{a}$ is in one to one correspondence with a subset of $\mathfrak{b}$ but $\mathfrak{b}$ is *not* in one to one correspondence with a subset of $\mathfrak{a}$. Thus

Def $\mathfrak{a} < \mathfrak{b}$ if and only if $\mathfrak{a} \leq \mathfrak{b}$ and $\mathfrak{b} \not\leq \mathfrak{a}$.

To make this consistent with normal usage (and what we would expect) we need: if $\mathfrak{a} \leq \mathfrak{b}$ and $\mathfrak{b} \leq \mathfrak{a}$ then $\mathfrak{a} = \mathfrak{b}$. This is true, but surprisingly hard to prove. Cantor himself was unable to prove it but it was proved by Ernst Schröder 1841-1902, but with an error, corrected by Felix Bernstein 1878-1956, the son of one of Cantor's friends. The currently popular proof is attributed to the Hungarian mathematician Julius König 1849-1913 but is too complicated to reproduce here.

I also note that we need not compare the actual cardinal numbers; it suffices to compare sets that are equivalent to the cardinal numbers should this be more convenient.

By now you may be thinking that all infinite sets have cardinal number $\aleph_0$ but then this whole chapter would be pointless, so at this point I feel I owe you an infinite set with cardinal larger than $\aleph_0$. This result will follow from our systematic discussion later but the proof is so famous and the method so ingenious that we will present the more elementary form here. This is called Cantor's diagonal procedure. The set in question is $\mathbb{R}$. Suppose that $\mathbb{R}$ has cardinal $\aleph_0$. Then it would be possible to produce a one to one correspondence between the reals and $\aleph_0$. I will show with the following example that this is not possible. It doesn't matter what list you give me; I will find with Cantor's diagonal procedure a real number not in the list. I will do this for a list of decimals between 0 and 1 not including the end points. This is good enough because the set U of decimals between 0 and one is in one to one correspondence with $\mathbb{R}$, as shown by the following diagram where the semicircle is $U = \{x \in \mathbb{R} \mid 0 < x < 1\}$ bent into the form of a semicircle.



Equivalence of $\mathbb{R}$ and U

In the diagram the point B on U corresponds to the point C on the real line $\mathbb{R}$, and one sees the one to one correspondence immediately. This can also be done with the function $\tan(\frac{\pi x}{2})$ which takes U to $\mathbb{R}$ and is a bijection as can be seen from the graph of the tangent function.

Note in passing that we have shown there are exactly as many points between 0 and 1 as there are on the entire real line.

If the reals could be put in a list, by means of the one to one correspondence above, then the reals from 0 to 1 could be put in a list. I now show this is impossible by giving a number not in the list. For convenience ignore the first line in the list. This is a *particular* list but you can see the method would work for any such list.

Suppose the list is

0	.11111111111111111111111111111111111111111111...
1	.12121212121212121212121212121212121212121212...
2	.13567836856724785096735689573954739678673 8...
3	.08902367589278723469839064392740943859 6127...
4	.90573902750948560729084098670249076702340 7...
5	.78858346950894357983569036793847539046838 7...
6	.22412389388435897039850346840697230595086 3...
7	.99999999989749286498273985692847948126433...
8	.31245668365892390004702759082340974085937 0...
9	.77777999999990000009999777555555555555444...
⋮	⋮

Starting at row labled 1, I take the first digit 1, from the row labled 2 I take the second digit 3, from the row labeled 3 I take the third digit 9, and forming a decimal from these digits I get

$$.139783989\ldots$$

Now I modify the above decimal by the following (traditional) method. If the digit is 1 through 8 I add one to it. If the digit is 9 I replace it by 0. I then get the decimal

$$Z = .240894090\ldots$$

I can show this decimal is not in the list.
SZ cannot be the decimal in the line labeled 1 because the first digit is wrong
Z cannot be the decimal in the line labeled 2 because the second digit is wrong
Z cannot be the decimal in the line labeled 3 because the third digit is wrong
Z cannot be the decimal in the line labeled 4 because the fourth digit is wrong
...
Z cannot be the decimal in the line labeled 1008 because the 1008^{th} digit is wrong
...

Thus I have shown that Z cannot be a decimal in the list and hence the list is incomplete and the correspondence between U and $\aleph_0$ fails to be one to one. (If you are worried about it being the decimal labeled 0 you can change the first

digit of Z to be different from the decimal labeled 0 and the decimal labeled 1 and still have a decimal not in the list.)

Since $\aleph_0$ is equivalent to a subset of $\mathbb{R}$ but $\mathbb{R}$ is not equivalent a subset of $\aleph_0$ (else it would be finite or equivalent to $\aleph_0$) we have that $\aleph_0 <$ cardinal number of the reals $n(\mathbb{R})$. The convention going back to Cantor is that this number is denoted by the lower case German letter $\mathfrak{c}$ (for continuum, another word for the reals or the real line) and we have shown

$$\aleph_0 < \mathfrak{c}$$

This was the huge surprise that Cantor brought to the world, with essentially this proof. Up until then everyone thought infinite was just infinite, but we have shown there are different sizes of infinite. Cantor's proof was attacked on various grounds but none of the attacks were successful and the proof now stands as a triumph of mathematical reasoning and the method, called Cantor's diagonal method, has become an important tool in mathematics. We will present some of the results of this method in a subsequent section. Prepare to be surprised some more.

Let us now quickly take stock of the supply of cardinal numbers we have discovered (or created or invented). There are the finite cardinals $0, 1, 2, 3, \ldots$ and then comes $\aleph_0$ and then $\mathfrak{c}$ where I have listed them in increasing order. These are the cardinal numbers most useful in mathematics.

Are there infinite numbers beyond $\mathfrak{c}$? Indeed there are, and the sequence never ends. We will prove this. We can also ask if there are any infinite numbers between $\aleph_0$ and $\mathfrak{c}$. Alas we cannot prove this in Zermelo set theory. This is the snake in paradise which I mentioned earlier. The system is not completely determined by the axioms, and in an annoying way. We just have to live with this, or add the axiom $\aleph_1 = \mathfrak{c}$. There are very fine philosophical arguments supporting or denying this. But no proofs are possible.

I want to do one more thing at this point. I want to prove that there are just as many points in a square as there are on one side of it. This proof is also due to Cantor.

Before we begin, let's remember that $.9999999999999999\ldots = 1$. Recall how we proved this:

$$\begin{aligned} x &= .9999999999999999\ldots \\ 10x &= 9.9999999999999999\ldots \\ 9x &= 9 \qquad \text{subtracting first from second} \\ x &= 1 \end{aligned}$$

The square we will work with is the square in the plane with corners, going counterclockwise (0,0),(1,0),(1,1).(0,1), often called the unit square. We are going to include the edges, so the set is

$$\{(x, y) \mid 0 \leq x \leq 1 \text{ and } 0 \leq y \leq 1\}$$

Thus the values of x and y lie between

$$.000000000000000000\ldots \text{ and } .999999999999999999999999\ldots$$

Now we set up a one to one correspondence (bijection) between the square and the closed unit interval $\{z \mid 0 \leq z \leq 1\}$. The digits of x and y are interleaved to get z as in this example:

$$(.1415926535\ldots, .7182818284\ldots) \longrightarrow .17411852982168523854\ldots$$

Thus the first, third, fifth, seventh, ninth etc digits of z come from x and the second, fourth, sixth, eighth, tenth, digits of z come from y, all digits being kept in order. The process is reversible; if we have a z in the closed unit interval we can reconstruct the x and y it came from in the obvious way:

$$.123456789123456789 \longrightarrow (.135792468\ldots, .246813579\ldots)$$

and we see that the second procedure is just the inverse of the first procedure. The first, second, third, etc. digits of x come from the first, third, fifth etc. digits of z and similarly for y. Thus we have set up a one to one correspondence between the decimal representations of the closed unit square and and the decimal representation of the closed unit interval, which shows that they have the same number of points. (The use of closed objects which include their edges is a merely technical matter which simplifies the proof slightly but is inessential.)

This is probably contrary to what your intuition is telling you. Intuition is a rather poor guide to infinite arithmetic, but surprisingly after a while you do get used to it and your intuitive predictions become more reliable. It was results like this, which contradicted many peoples' intuition, that hindered the initial acceptance of Cantor's theory of sets. However, with time the usefulness of the concepts and the lack of any alternative theory overwhelmed the objections and more and more mathematicians came to accept it, so that nowadays it feels almost "natural".

4.5 Cardinal Arithmetic

We will now present the definitions of addition, multiplication and exponentiation for cardinal numbers. This is relatively easy. I will prove a couple of easy cases.

We begin with addition.

Def Suppose $\mathfrak{a} = n(A)$ and $\mathfrak{b} = n(B)$ and $A \cap B = \emptyset$. Then $\mathfrak{a} + \mathfrak{b} = n(A \cup B)$.

For example, if we wish to add 3 and 5 we choose *disjoint* sets A and B with $3 = n(A)$ and $5 = n(B)$ I can take $A = \{a, b, c\}$ and $B = \{1, 2, 4, 5, 6\}$. Then

$$3 + 5 = n(A \cup B) = n(\{a, b, c, 1, 2, 4, 5, 6\}) = 8$$

since I can put $A \cup B$ in one to one correspondence with $8 = \{0, 1, 2, 3, 4, 5, 6, 7\}$. For $3 + \aleph_0$ I select $A = \{a, b, c\}$ and $B = \aleph_0$ and then form the one to one correspondence of $A \cup B$ with $\aleph_0$ as follows

$$\begin{array}{cccccccccc} a & b & c & 0 & 1 & 2 & 3 & 4 & 5 & \dots \\ | & | & | & | & | & | & | & | & | & \dots \\ 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & \dots \end{array}$$

Hence

$$3 + \aleph_0 = n(A \cup B) = \aleph_0$$

It is just as easy to do $\aleph_0 + \aleph_0$. Here I take $A = \{1, 2, 3, 4, 5, 6, \ldots\}$ and $B = \{-1, -2, -3, -4, -5, -6, \ldots\}$ and then find the cardinal of $A \cup B$ as follows

$$\begin{array}{cccccccccccc} 1 & -1 & 2 & -2 & 3 & -3 & 4 & -4 & 5 & -5 & 6 & \dots \\ | & | & | & | & | & | & | & | & | & | & | & \dots \\ 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & \dots \end{array}$$

and thus we see that

$$\aleph_0 + \aleph_0 = n(A \cup B) = \aleph_0 \quad \Longrightarrow \quad 2\aleph_0 = \aleph_0$$

The general rule for addition of transfinite cardinals is that the sum equals the bigger of the two, as in $3 + \aleph_0 = \aleph_0$ but we have done enough addition.

Next comes multiplication. To do this I must introduce a new actor on the stage, the *Cartesian Product*. This is very easy. If A and B are two sets then their Cartesian product $A \times B$ is the set of all pairs (x, y) where $x \in A$ and $y \in B$, or to put it more efficiently

$$A \times B = \{(x, y) \mid x \in A \text{ and } y \in B\}$$

You may well be able to guess what comes next.

Def Suppose $\mathfrak{a} = n(A)$ and $\mathfrak{b} = n(B)$. Then $\mathfrak{a} \cdot \mathfrak{b} = n(A \times B)$.

For example let $A = \{a, b\}$ and $B = \{1, 2, 3\}$. Then $2 = n(A)$, $3 = n(B)$ and

$$A \times B = \{(a, 1), (a, 3), (a, 3), (b, 1), (b, 2), (b, 3)\}$$

So we see $2 \cdot 3 = n(A \times B) = 6$.

Now let us do something very similar to get $\aleph_0 \cdot \aleph_0$. We take a set A with cardinality $\aleph_0$ and then

$$\aleph_0 \cdot \aleph_0 = n(A \times A)$$

For A we take the set of positive integers $\{1, 2, 3, 4, 5, \ldots\}$ and we form $A \times A$ which we arrange below.

$$\begin{array}{ccccccccc} (1,1) & \rightarrow & (1,2) & & (1,3) & \rightarrow & (1,4) & \rightarrow & (1,5) & \rightarrow \\ & \swarrow & & \nearrow & & \swarrow & & \nearrow & & \swarrow \\ (2,1) & & (2,2) & & (2,3) & & (2,4) & & (2,5) & \\ \downarrow & \nearrow & & \swarrow & & \nearrow & & \swarrow & & \nearrow \\ (3,1) & & (3,2) & & (3,3) & & (3,4) & & (3,5) & \\ & \swarrow & & \nearrow & & \swarrow & & \nearrow & & \swarrow \\ (4,1) & & (4,2) & & (4,3) & & (4,4) & & (4,5) & \\ \downarrow & \nearrow & & \swarrow & & \nearrow & & \swarrow & & \nearrow \\ (5,1) & & (5,2) & & (5,3) & & (5,4) & & (5,5) & \end{array}$$

Clearly the pairs can now be put in a list so $n(A \times A) = \aleph_0$,

$$\begin{array}{lccccccccc} A \times A & = \{ & (1,1) & (1,2) & (2,1) & (3,1) & (2,2) & (1,3) & (1,4) & (2,3) & \ldots \} \\ & & | & | & | & | & | & | & | & | & \\ \aleph_0 & = \{ & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & \ldots \} \end{array}$$

Thus we have

$$\aleph_0 \cdot \aleph_0 = n(A \times A) = \aleph_0$$

We can do $\mathfrak{c} \cdot \mathfrak{c}$ very easily since we have already done the work. Let us for convenience name the closed unit interval U, so $U = \{x \mid 0 \leq x \leq 1\}$. We know that $n(U) = \mathfrak{c}$. We also know that $n(U \times U) = n(U)$, since we proved this earlier. Thus we have

$$\mathfrak{c} \cdot \mathfrak{c} = n(U \times U) = n(U) = \mathfrak{c}$$

which is boring. Also

$$\mathfrak{c}^3 = \mathfrak{c}^2 \cdot \mathfrak{c} = \mathfrak{c} \cdot \mathfrak{c} = \mathfrak{c}$$

and repeition would prove that $\mathfrak{c}^n = \mathfrak{c}$ for any positive integer n.

Our next project is exponentiation, which is far the most interesting of the three operations. However, before we can take it on we must first deal with the idea of *function*. The function is one of the most important mathematical ideas and this is a convenient moment for you to learn about it. However, you already know most of it and it is only a matter of firming up your intuitive knowledge.

A function from set A to set B is a machine that takes inputs from A and ejects outputs in B. For example the square function from $\mathbb{R}$ takes real number inputs and outputs their squares, in $\mathbb{R}$. A function usually has a name, often f or F and the full notation is

$$f : \mathbb{R} \to \mathbb{R} \qquad\qquad f(x) = x^2$$

There are many variants in notation for functions each having their special use. For example

$$x \mapsto x^2$$

is another notation for the function f.

If f maps elements in A to elements in B the notation is

$$f : A \to B$$

A is called the *domain* of the function (notation $\mathrm{Dom}(f)$) and B is called the *range space* or *codomain*. The actual set of outputs is called the *range* of f and is a subset of B. It need not be, and often is not, all of B. The function $f(x) = x^2$ has codomain $\mathbb{R}$ (as given above) but has range $\{x \in \mathbb{R} \mid x \geq 0\} \subseteq \mathbb{R}$. However, unsymmetrically, every element of A must be an allowable input[6]. Note that a user could change the codomain to match the range if this didn't cause any

[6]There are occasional exceptions, especially in Functional Analysis.

problems, like messing up the uniformity of a presentation. It is often not such a good idea.

Note the very important property of functions that each input has exactly one output; no more, no less. Thus the often seen $g(x) = \pm\sqrt{x}$ is *not* a function since $g(4) = \pm 2$ is two outputs, not one. Such things used to be called *multi-valued* functions, but as Hermann Grassmann pointed out about 1843, this is very bad because it messes up function composition. The tendency now is to avoid the term. Note also that if the function has formula $x \mapsto \sqrt{x}$ then it is *not* a function from $\mathbb{R}$ to $\mathbb{R}$ because negative inputs have no outputs *in* $\mathbb{R}$ and it also is not a function from $\mathbb{C}$ to $\mathbb{C}$ unless you have some way of picking which of the two outputs you want, a perilous business. The takeaway is that domain and range of functions are slightly tricky concepts and often give beginning students difficulties.

One way of both clarifying the function concept and bringing it into set theory is to regard a function as a set of ordered pairs $\langle a, b\rangle$. It took a while to figure out that ordered pairs could be defined within set theory, but Kazimierz Kuratowski and Norbert Wiener showed that they can be defined as follows

Def $\langle a, b\rangle = \{\{a\}, \{a, b\}\}$

This is representative of a class of definitions which are never used once they have served a very limited purpose. Here is the purpose:

Theorem $\langle a, b\rangle = \langle c, d\rangle$ if and only if $a = c$ and $b = d$.

This is the whole story about ordered pairs and once you have proved this theorem the definition is never used again. It's purpose is twofold; to pull ordered pairs firmly into set theory and to be used in your proof of this theorem.

Now we show how a function is a set of ordered pairs. Let

$$\begin{aligned} h : \{1,2,3,4\} &\rightarrow \{a,b,c,d,e,f\} \\ h &= \{\langle 1,c\rangle, \langle 2,d\rangle \langle 3,d\rangle \langle 4,f\rangle\} \\ \text{Then} \quad h(1) = c, \quad h(2) &= d, \quad h(3) = d, \quad h(4) = f \end{aligned}$$

As you can see, there is exactly *one* pair for every element of $\mathrm{Dom}(h) = \{1,2,3,4\}$), but not all the elements of the codomain $\{a,b,c,d,e,f\}$ are outputs. The range of h, $\mathrm{Ran}(h) = \{c,d,f\}$, is the set of outputs.

When a set contains only pairs it is called a *pairset*. The set for a function is a pairset but not all pairsets are are associated to functions. Because a pairset for a function has exactly one pair for each element of the domain, we can tell if a pairset is a function by the following test:
A pairset h is a function if and only if

$$\text{If } \langle a, r\rangle \in h \text{ and } \langle a, s\rangle \in h \text{ then } r = s$$

There is a slight complication in identifying a function h with its pairset. We can recover the domain of the function from the first elements of the pairs in the pairset and the range from the second elements, but the pairset does not in general determine the codomain. In practice this seldom causes problems

because functions are almost always introduced in the form $f : A \to B$ which identifies the codomain as B.

It is often more convenient for psychological reasons to present the pairs in a tabular form, especially if one is dealing with more than one function. This is done as follows; the function h shown above can be presented in tabular form as

1	2	3	4
c	d	d	f

It should be obvious how to go from tabular form to pairset and vice versa.

Another way of specifying a functions is by giving a formula. This is the way you are probably already familiar with. Consider the function $S : \mathbb{R} \to \mathbb{R}$ which is the square function. The formula for the function is

$$S(x) = x^2 \qquad \text{or} \qquad x \mapsto x^2$$

Note the arrow in the second form has a vertical line at the left end. The pairset for S is

$$\{\langle x, y\rangle \in \mathbb{R}\times\mathbb{R} \mid y = x^2\} = \{\langle 0,0\rangle, \langle 2,4\rangle, \langle \pi, \pi^2\rangle, \langle 1.2, 1.44\rangle, \langle -7, 49\rangle, \langle -2, 4\rangle, \langle -\frac{2}{3}, \frac{4}{9}\rangle, \ldots\}$$

There are many more things to know about functions and there is an appendix to this chapter that goes into a little more detail but we already have all we need for our purposes. We need one more bit of notation.

Notation The set of all functions from the set A to the set B is denoted by B^A.

We will now look at the set 2^A for $A = \{a, b, c\}$. We will write the functions out in tabular form. Recall $2 = \{0, 1\}$

	a	b	c	
f_1	0	0	0	$\{\ \}$
f_2	0	0	1	$\{c\}$
f_3	0	1	0	$\{b\}$
f_4	0	1	1	$\{b, c\}$
f_5	1	0	0	$\{a\}$
f_6	1	0	1	$\{a, c\}$
f_7	1	1	0	$\{a, b\}$
f_8	1	1	1	$\{a, b, c\}$

There are some very important things to see in this table. First, we see that their are 8 functions in 2^A. This is consistent with the following definition for exponentiation of cardinal numbers. Recall that we use German letters for cardinal numbers whether finite or infinite, (but that $\mathfrak{c}$ has the fixed meaning cardinal of the continuum $\mathbb{R}$).

Def Let $\mathfrak{a} = n(A)$ and $\mathfrak{b} = n(B)$. Then $\mathfrak{b}^\mathfrak{a} = n(B^A)$. That is, $\mathfrak{b}^\mathfrak{a}$ is the cardinal number of the set of functions from A to B.

The above table gives us all the possible functions from A to $2 = \{0, 1\}$. Since $3 = n(A) = n(\{a, b, c\})$ and $2 = n(\{0, 1\})$ we have

$$2^3 = n(2^A) = n(\{f_1, f_2, f_3, f_4, f_5, f_6, f_7, f_8, \}) = 8$$

Thus we know how to exponentiate cardinal numbers. However, there is more to see in the table. Note the subsets of $A = \{a, b, c\}$ written to the right of each row of the table. The subsets consist of those elements of A for which the function f_i has output 1. The functions $f_i,\ i = 1, \ldots, 8$ are called characteristic functions. The general definition is

Def Let A be a set. The characteristic function f_A is the function whose values are

$$f_A(x) = \begin{cases} 0 & \text{if } x \notin A \\ 1 & \text{if } x \in A \end{cases}$$

So we see that $f_4(x) = 1$ for $x = b, c$ and thus f_4 is the characteristic function for the subset $\{b, c\}$.

This is true in general; if A is any set then the functions in 2^A, that is the functions from A to $2 = \{0, 1\}$, are all characteristic functions of the subsets of A. The characteristic functions are in one to one correspondence to the subsets of A. Hence a set A whose cardinal number is $\mathfrak{a}$ will have $2^{\mathfrak{a}}$ subsets. As we saw above, $A = \{a, b, c\}$ has $2^3 = 8$ subsets. There is a standard notation for the set of subsets of a set.

Def Let A be a set. The set of subsets of A is denoted by $\mathfrak{P}(A)$ and is called the powerset of A.

So, for example, for $A = \{a, b, c\}$ the powerset of A is

$$\mathfrak{P}(A) = \{\{\ \}, \{a\}, \{b\}, \{c\}, \{b, c\}, \{a, c\}, \{a, b\}, \{a, b.c\}\}$$

What makes all this *really* interesting is the following Theorem.

Theorem The cardinal number of a set is always smaller than the cardinal number of it's powerset:

$$n(A) < n(\mathfrak{P}(A))$$

Proof The correspondence $x \mapsto \{x\}$ shows that A is in one to one correspondence with a subset of $\mathfrak{P}(A)$ so $n(A) \leq n(\mathfrak{P}(A))$. We need to show that there is no one to one correspondence between A and $\mathfrak{P}(A)$ which proves $n(A) \neq n(\mathfrak{P}(A))$. So let us assume that there *is* a one to one correspondence between A and $\mathfrak{P}(A)$ and let us denote the subset corresponding to $x \in A$ by $F(x) \subseteq A$. Of course $F(x) \in \mathfrak{P}(A)$. Now we use an abstract form of Cantor's diagonal procedure. (It is not obvious that this is related to Cantor's diagonal procedure.) We form a subset $B \subseteq A$ according to the following rule:

$$x \in B \quad \text{if and only if} \quad x \notin F(x)$$

In words, B is the subset of A whose elements x are *not* elements of their correlated subsets $F(x)$.

Now B is a subset and thus is correlated to some element $b \in A$, that is $B = F(b)$. So the questions is, is b in B. Well if $b \in B$ then $b \in F(b)$ so it is an element of its correlated subset which is a contradiction to the definition of B. So let's try $b \notin B = F(B)$. Well then b satisfies the condition for being in B (not in it's correlated subset) and so $b \in B$. Another contradiction. So $b \in B$ and $b \notin B$ both lead to contradictions. Thus the correspondence F between A and $\mathfrak{P}(A)$ cannot exist since I found a set B that does not correspond to any $b \in A$. Thus the one to one correspondence between A and $\mathfrak{P}(A)$ cannot exist and $n(A) \neq n(\mathfrak{P}(A))$ Hence $n(A) < n(\mathfrak{P}(A))$.

Note this is the first time that an arithmetic operation has resulted in something with a larger cardinal number than the cardinals of the inputs to the operation. This is the key to ever larger cardinals. Note there cannot be a largest cardinal $\mathfrak{m}$ because $2^{\mathfrak{m}}$ is larger than $\mathfrak{m}$ by the theorem.

The two cardinal numbers of greatest interest in ordinary mathematics are $\aleph_0$ and $\mathfrak{c} = n(\mathbb{R})$. We will now show that these numbers are connected. First we show

Theorem $\mathfrak{c} = 10^{\aleph_0}$.

Proof: Recall that $10 = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ and also recall that $\aleph_0 = n(\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, \ldots\})$ so that $10^{\aleph_0}$ is the cardinal number of the set of functions from $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, \ldots\}$ to $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$. Here are three typical such functions:

	1	2	3	4	5	6	7	8	9	10	11	12	13	...
f_1	7	8	9	4	7	9	6	0	6	8	3	0	4	...
f_2	9	8	6	7	4	9	6	7	2	3	9	4	3	...
f_3	3	9	0	5	8	3	0	6	6	3	0	5	8	...

Now we set up a one to one correspondence between all these functions and the decimals between 0 and 1 as follows:

$$\begin{aligned} f_1 &\longleftrightarrow .7894796068304\ldots \\ f_2 &\longleftrightarrow .9867496723943\ldots \\ f_2 &\longleftrightarrow .3905830663058\ldots \end{aligned}$$

Thus we have a one to one correspondence between the elements (functions) of $10^{\aleph_0}$ and the decimals of $\{x \mid 0 \leq x \leq 1\}$ and thus we see that we have $n(10^{\aleph_0}) = n(\{x \mid 0 \leq x \leq 1\}) = n(\mathbb{R}) = \mathfrak{c}$ as required. (Recall that $n(10^{\aleph_0}) = 10^{\aleph_0}$ by definition of $10^{\aleph_0}$.)

Now notice that we have

$$2 \leq 10 \leq \aleph_0 \leq 2^{\aleph_0}$$

because we always have $\mathfrak{a} < 2^{\mathfrak{a}}$ but I use $\leq$ instead of $<$ because $<$ would not survive the next step but $\leq$ does. So we now exponentiate to the $\aleph_0$ power all of the last inequalities:

$$2^{\aleph_0} \leq 10^{\aleph_0} \leq {\aleph_0}^{\aleph_0} \leq (2^{\aleph_0})^{\aleph_0} = 2^{(\aleph_0 \cdot \aleph_0)} = 2^{\aleph_0}$$

Since the beginning and end of this chain of inequalities are the same, all the inequalities must be equalities and we have the interesting

$$2^{\aleph_0} = 10^{\aleph_0} = {\aleph_0}^{\aleph_0} = \mathfrak{c}$$

where the last equality comes from the already proved $10^{\aleph_0} = \mathfrak{c}$. We also note the interesting

$$\mathfrak{c}^{\aleph_0} = (2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0 \cdot \aleph_0} = 2^{\aleph_0} = \mathfrak{c}$$

This is the cardinal number of the set of real valued *continuous* functions but it would take us to far afield to explain why. The cardinal of the set of real valued functions (no continuity assumed) is of course

$$\mathfrak{c}^{\mathfrak{c}} = (2^{\aleph_0})^{\mathfrak{c}} = 2^{\aleph_0 \cdot \mathfrak{c}} = 2^{\mathfrak{c}}$$

using the principle (which we didn't prove) that a product of cardinals is the larger of the two. Note that the cardinal of the set of real valued functions on $\mathbb{R}$ is equal to the cardinal of the set $\mathfrak{P}(\mathbb{R})$ of subsets of $\mathbb{R}$ though the one to one correspondence of these sets is by no means obvious.

This concludes our introduction to the world of transfinite numbers. A more complete version of our results can be found in the very old book SET THEORY by Felix Hausdorff. There is little more to be had in English on the subject. A translation of the 3rd edition is available from Amazon. The first edition has a much more detailed treatment but is available only in German. It is also available from Amazon with the title Grunzüge der Mengenlehre.

4.6 Appendix on Set Philosophy

Although set theory is considered by most mathematicians as an acceptable foundation for mathematics it is not without its problems We cannot go deeply into these matters which would require an entire book. But we can have a quick look in a relatively few pages.

4.6.1 Too liberal a policy

The first to attempt to actually explicitly build mathematics on set theory was Gottlob Frege (1848-1925) who did this is three volumes. He based his theory on the principle of extentionality, which simply says that for every property (like red) there is a set of things for which this property is true, in this case the set R of red things. Similarly for the property of being a cat, so that C, the set of cats, is the set of things that come out true when substituted for x in the sentence

$$x \text{ is a cat}$$

So we can say

$$\text{For all } x \;\; x \in C \longleftrightarrow x \text{ is a cat}$$

Thus Spot the dog is not in C because "Spot is a cat" is not true. But Puff the cat is in C because "Puff is a cat" is true.

It would seem that the principle of extentionality is indeed harmless and a suitable basis for a 3 volume work but this is not true. It is not harmless; it is fatally flawed, as Bertrand Russel pointed out. For let the property be x is not a member of itself. In symbols $x \notin x$. This is certainly the usual situation. Let F be the set of things which satisfy this property:

$$\text{For all x } \; x \in F \longleftrightarrow x \notin x$$

Now the x that we choose to test is F itself. We have, substituting F in for x

$$F \in F \longleftrightarrow F \notin F$$

which is as contradictory as contradictions can be. Hence the F cannot exist, since we cannot put it in for x without generating a contradiction.

Russel and Alfred North Whitehead thought they had figured out a way to stop the contradiction, called type theory. The 0 type was the elements, which were *not* sets. Then type 1 was sets whose elements where type 0, like $\{a, b, c\}$. Type 2 was sets whose elements were type 1, like $\{\{a, b, c\}, \{b, d\}, \{a, c\}, \{a, b, c, z\}\}$. Type 3 were sets whose elements were type 2, like $\{\{\{a, b, c\}, \{b, d\}\}, \{\{a, b\}, \{a, b, c, z\}\}\}$. You can then see how the pattern continues. This will keep out sets x for which $x \in x$ because they won't fit into the type structure.

It turned our, however, that to build mathematics quite a bit of tinkering with this simple system was necessary, but they persevered and produced a three volume set called Principia Mathematica. To show the problems this sort of approach generates, the American philosopher Willard von Orman Quine found a contradiction in the 1930s. This contradiction was not fatal and a paste-in was prepared for the current edition of Principia to fix it, and later editions contained a more elegant fix, so the system continued and continues to this day as the favored approach to the foundations of mathematics by Philosophers of Mathematics.

The approach followed with this system can be characterized as "let in as much as you can without generating a contradiction." There are many such systems now but they generate little interest among mathematicians. Philosophers, in contrast, tend to prefer such systems.

4.6.2 The Axiomatic Approach

Mathematicians greeted the publication of Whitehead and Russel's Principia Mathematica with initial interest. Most were happy that mathematics now had a solid foundation but most were disappointed that the solution was so complicated and unintuitive.

Ernst Zermelo (1871-1953, German) published (1908) another solution to the set theory problem with a different mind set. Zermelo's strategy was to avoid the contradictions by letting into his set theory only the objects necessary for building mathematics from set theory. Pegasus, Julius Caesar, all the hydrogen

atoms, Queen Victoria, etc where all OUT. The empty set was IN. The axioms as originally published (more or less) are

I Axiom of Extension. Two sets are equal if they have the same elements

II Axiom of Elementary Sets: a) there is an empty set $\emptyset = \{\ \}$, b) for any object a there is a set $\{a\}$ whose only element is a, c) if a and b are any two objects then there is a set $\{a, b\}$ whose only two elements are a and b. (Here any set is an object).

III Axiom of Separation. Given a set and some property there is a set whose members are exactly the elements of the given set for which the property is true. (Example: if the set is the set of foxes and the propery is being gray then the axiom guarantees the set of gray foxes exists)

IV Axiom of the Power Set: Given a set there is a set whose elements are exactly the subsets of the given set.

V Axiom of Union. Given a set S whose elements are sets $A_1, A_2, \ldots$ there is a set $\bigcup S$ whose elements are exactly the elements of the sets $A_1, A_2, \ldots$. Example: if $S = \{\{a, b\}, \{c, d\}, \{a, c, f\}\}$ then $\bigcup S = \{a, b, c, d, f\}$.

VI Axiom of Choice. Given a set S whose elements are disjoint sets there is a set which has as elements exactly one element from each of the sets in S. Example: $S = \{\{a, b\}, \{c, d, f\}, \{r, s, t, u, v\}, \{1, 5\}\}$ then the set this axiom says exists might be $\{a, f, t, 1\}$ or it might be $\{b, f, r, 5\}$ or lots of other possibilities, but at least one of these exists.

VII Axiom of Infinity. Given a set S the *successor* of S is $\sigma(S) = S \cup \{S\}$. The axiom says that there is a set I satisfying the properties a) $\emptyset \in I$, b) if $a \in I$ then $\sigma(a) \in I$. (With the von-Neumann definition of the integers this gets the non negative integers N into I and thus guarantees I is infinite which is the point of the axiom; we wouldn't want all our sets to be finite sets.)

These axioms will probably do the job but to give ourselves a little more safety we want an axiom that keeps $x \in x$ or $x \in y \in x$ or similar unpleasant objects from sneaking in. Thus we would like to add the

VIII Axiom of Regularity. Every non-empty S set there is an element $a \in S$ for which $a \cap S = \emptyset$.

Although it is not obvious, the Axiom of regularity prevents $x \in x$ and also keeps away infinite descending chains of elementhood: you cannot have

$$\ldots \in a_6 \in a_5 \in a_4 \in a_3 \in a_2 \in a_1 \in S$$

(all of these are sets).

And there is one other thing we can do. We can add

IX Axiom of the Continuum. $2^{\aleph_0} = \aleph_1$

We have not discussed $\aleph_1$ but Cantor showed how to create a sequence of alephs which include all the infinite numbers. Axiom IX cannot be proved from the other axioms but assuming it results in a maximally simple set theory. This choice is called *Cantorian set theory.* It is possible to set $2^{\aleph_0}$ to be other choices,

but it would require much explanation to do justice to this question. Such set theories are called *Non-Cantorian set theories*. No consensus has emerged on what to do, but the default is the Cantorian set theory.

And that's it girls and boys, upon that rock we can erect mathmatics, from Arithmetic to the cohomology of profinite groups or Abelian varieties. Historically, nobody was really excited in 1908 because they had been shellshocked by all the contradictions in Cantor's too liberal set theory. Hermann Weyl bet someone that within 20 years a contradiction would appear in this theory too. He lost. After 40 years and no contradiction mathematicians felt the system might really hold up. By 1980 we were pretty sure the problem of foundations had been solved for good. The centennial passed in 2008 without the worldwide celebration the event deserved. Nobody now expects a contradiction to appear and these axioms still suffice to build all the mathematics built in the last century, with the exception that certain studies of large cardinal numbers require some addition axioms. These are of little relevance to ordinary mathematics and I will not go into the subject here.

This is not to say everyone is happy with the situation. There are technically upsetting things, like the existence of strange systems which are not mathematics in which all the axioms are true (Löwenheim-Skolem theorem), or some counterintuitive theorems (Banach-Tarski theorem; You can take a sphere the size of a pea, cut it into finitely many pieces, and reassemble it into a sphere the size of the sun with no empty space in the reassembly. This one traces right back to the Axiom of Choice.) Then there are odd questions of meaning. There are $\aleph_0$ sentences in the English languge and $2^{\aleph_0}$ real numbers. Since $2^{\aleph_0} > \aleph_0$, most real numbers cannot be described in English. The sensitive find this upsetting. Also upsetting is the least positive real number that cannot be described in English, which I just described in English. The trick is that there is no such number; for every positive real number not describable in English there is a smaller positive real number not describable in English.

Let me elaborate slightly on one of these points. The ancient Greeks wanted an axiom system for plane geometry which would have plane geometry as its only model. A situation like this were a set of statements has only *one* model is called a categorical set of statements. Using first order logic such sets are rare and their models trivial. Thus the Greek dream for plane geometry is not possible. We actually will this later in this chapter with the real numbers. We will take *all* the true statements about the real numbers and then add infinitesimals. As we will see, no set of statements about the real numbers will ban the infinitesimals. That is, the set of true statements about the real numbers is not categorical. There is more than one model; the ordinary reals and the reals with infinitesimals. Like most of the results of research into formal logic this is a little sad. But things are as they are, and we must try to be happy anyway.

One might ask if it would be helpful to use a stronger form of logic than first-order logic. For example we could quantify over predicates. But this then allows much more vicious beasts into the arena and makes it much less predictable what

effects will show up in the long run. The general feeling is that it isn't worth the trouble and there is little interest, but here and there you can find a lonely enthusiast for this approach.

David Hilbert (1862-1943) conceived the brilliant idea of making mathematics into a formal structure, which could then be studied in the same way physics is studied, using mathematics. But Hilbert's brilliant plan was to use only finistic mathematics to study mathematics, and to prove that no contradiction could arise in mathematics with the proof using only the safe and universally accepted finitistic mathematics. Alas, it was not to be. Kurt Gödel (1906-1978 born in the Austro-Hungarian Empire) proved that it is not only impossible to do this but you can't prove ordinary mathematics is contradiction free even if you allow ourself the resources of ordinary mathematics to do the proof (which would not be very comforting anyway). This has all been described as God gave us a consistent mathematics but Satan keeps us from *proving* the consistency and thus keeps us uncomfortable.

PART B

4.7 Some History

Although there were some forms of Calculus used in ancient times, especially by Archimedes, these forms did not explicitly use infinitely small numbers and so do not concern us. The person who begins our story is Bonaventura Cavalieri, who began the use of infinitely small numbers, but whose methods were less than reliable. Reliability was achieved by John Wallis, Isaac Barrow and James Gregory and also by Pierre de Fermat. Isaac Barrow actually wrote a book, Geometrical Investigations, but in 1669 he resigned his mathematics professorship in order to pursue his theological interests, and he did not want to be bothered with seeing his book through the publication process. He arranged for Isaac Newton to get his mathematics professorship and also arranged for Newton to see his book through the press, and at this point Issac Barrow leaves mathematics forever. Newton always claimed that he had been the first to find Calculus, and he certainly found a lot of it in Isaac Barrow's book, including an obscure form of the Fundamental Theorem of Calculus. This key theorem connects two of the three pillars of Calculus, Differential and Integral Calculus[7].

Meanwhile the other main actor in the drama, Gottfried Wilhelm Leibniz[8] was journeying back and forth across Europe on various errands for his employer, the Elector of Hannover. On missions to England he visited mathematicians, as he did everywhere, and inquired if there were any new developments. It is by no means certain, but it is possible, that he picked up hints of the Barrow/Newton developments. He worked on Calculus for several years and eventually in 1684 published his research in the first issue of Acta Eruditorum (he was the editor) and this was the first really public appearance of Calculus. The several years

[7]The third pillar is infinite series.
[8]Leibniz, like Schulz, has no t before the z.

work he had put in on the subject resulted in a fantastically well adapted notation that worked so well we still use it today. Every generation improves on it and the improvements are forgotten by the subsequent generation but Leibniz's notation soldiers on, cleverly concealing difficulties and making it possible for students of lesser gifts to still gain the use of this marvelous tool.

Newton usually concealed the infinitely small numbers at the base of Calculus, preferring to use arguments similar to the classical Greeks, but nobody else was interested in this obfuscation of what was really going on and infinitely small numbers (called infinitesimals from here on out) were commonly used. But there were very tricky aspects.

Some mathematicians and almost all philosophers were seriously annoyed at the loss of the classical beauty of Greek mathematics, and set up howls, but most mathematicians were very excited by the wonderful new things that could be done with the infinitesimal methods, not only in mathematics but all over science. So the shaky foundations were mostly ignored.

However, the son of a friend of a noted Philosopher, Bishop Berkeley, became so enamored with Newton's Gravitation that he lost his Christian faith and on his deathbed refused final Christian rites. Bishop Berkeley was so annoyed by this that he sought for some flaw in Newton's system of the world, and he found it. It was the infinitesimals. He called them, in a famous phrase, the ghosts of departed quantities, and denied they could exist. This was convincing for the philosophers, but the mathematicians were having far too good a time with their new toys to let mere philosophers spoil it all, so they largely ignored the objections. It was more than a hundred years before the problems in the foundations of Calculus caused enough rot that the edifice began to totter, and then work began on fixing it. You have already seen some of the work necessary to fix the problems; remember the Cauchy sequences used to define the reals.

Beginning around 1820 a movement in mathematics began to replace infinitesimals by the concept of limit. The history of this is quite convoluted and the definition of limit can be traced as far back as Leibniz. We will not follow this trail, however, as we do not want to follow the history of Calculus but instead the further history of infinitesimals.

4.8 Early Calculus an Infinitesimals

In this section we will have a brief look at the early use of infinitesimals and will more or less follow Fermat. The problem which Fermat was attacking was finding the tangent line to a curve, and the curve we will take is the parabola $y = x^2$ and we want the tangent line at the point $(1, 1)$, that is the point $x = 1$, $y = 1$. The method is to use a secant line, which is a line connecting two points on a curve. For example we might connect the point $(1, 1)$ and the point $(2, 4)$ on the curve which is the graph of $y = x^2$. The *slope* of a line through two points (x_1, y_1) and (x_2, y_2) is given by the formula

$$m = \frac{y_2 - y_1}{x_2 - x_1}$$

where m is the usual letter for slope. (The subscripts on x and y are just labels; x_1 means the x for the first point and y_2 means the y for the second point. For the points $(1,1)$ and $(2,4)$ we have

$$m = \frac{4-1}{2-1} = 3$$

When you know a point (x_1, y_1) on a line and its slope m the equation of the line is given by

$$y = m(x - x_1) + y_1$$

so in our example the equation of the secant line is

$$y = 3(x-1) + 1 = 3x - 2$$

Plug $(1,1)$ and $(2,4)$ into this equation and you will see it is satisfied; the line goes through both points.

Now to find the equation of the tangent line we just need it's slope, and we can approximate that by replacing the second point $(2,4)$ by a much closer point, say $(1.1, 1.21)$. Then the slope is

$$m = \frac{1.21-1}{1.1-1} = \frac{.21}{.1} = 2.1$$

We could even make a table using x's closer and closer to 1. Thus

x	2	1.1	1.01	1.001
y	4	2.21	1.0201	1.002002
m	3	2.1	2.01	2.001

At this point those with a mathematical turn of mind might suspect that the slope off the tangent line is 2, and that would be correct. You would have used here an intuitive form of the limit concept.

For reasons no terribly clear, this approach to the problem was not comfortable in the 1600s and they preferred to you the method which I now describe. We take an infinitely small number which I will designate by dx although Fermat himself used E. (I am using Leibniz's notation which was commonly done in Europe after 1684) The dx is an infinitesimal thought of a associated with the variable x. Then the two points on parabola $y = x^2$ that we are interested in are $(1,1)$ and $(1+dx, 1+dy)$, where dy is another infinitesimal which we now calculate.=:

$$\begin{aligned} 1 + dy &= (1+dx)^2 = 1 + 2dx + dx^2 \\ dy &= 2dx + dx^2 \end{aligned}$$

We then have

$$m = \frac{1+dy-1}{1+dx-1} = \frac{dy}{dx} = \frac{2dx + dx^2}{dx} = 2 + dx$$

Since dx is an infinitesimal and $m = 2 + dx$, it is reasonable to toss the infinitely small dx overboard and say the slope of the tangent line is $m = 2$. If this

strikes *you* as very suspicious, imagine what the philosophers thought of it. Nevertheless it was perfectly acceptable mathematics for quite a long time, although worry about it grew with the years.

Eventually two approaches developed. One involved concealing the cheating by the use of arcane vocabulary. The second was to find alternate methods. For example Lagrange advocated a method involving infinite series. These approaches merely papered over the difficulties, which were causing more and more trouble. By 1810 Calculus had grown into a large and powerful set of tools, now called analysis, in which various methods interacted, and while the best mathematicians had the taste and discernment to avoid the sinkholes forming it was difficult for lesser talents to avoid contradictions and mistakes due to the difficulty of manipulating infinitesimals when more than one of them was present. Cauchy led the way in finding new methods to get reliable results, and others soon followed, and the infinitesimals faded away in pure mathematics though in physics and chemistry and even in mathematical physics infinitesimal methods are used to this day, often in deriving basic equations, and they still work tolerably well.

4.9 The Fate of Infinitesimals

We have had a brief look at how infinitesimals were used in their heyday and how they were criticized by philosophers, who often claimed that no such beasts were possible. So one might reasonably ask the question: Were the philosophers that claimed infinitesimals were impossible correct? Is it impossible to build Calculus in logical fashion using infinitesimals. The answer to this is that YES, IT IS POSSIBLE. The demonstration that it is indeed possible is relatively new (last century) and uses entirely new tools.

We have not been very forthcoming about what an infinitesimal *is*. Let us clarify this. For our purposes and infinitesimal is a "number" with the following properties:

Def An infinitesimal is a quantity that is greater than 0 and smaller than any positive real number.

There was early on some hope that the infinitesimals might lead to contradictions, but no one was every able to show such a thing, and there is hardly any evidence that anyone was trying, since people don't like to publicize their failure. Doubtless Bishop Berkeley, who loathed infinitesimals, had a try at it, but had to be satisfied with calling them bad names ("ghosts of departed quantities").

Do infinitesimals exist? This, oddly, is a choice. If you don't like them there is nothing that obligates you to deal with them. No mathematics forces you to accept infinitesimals. But suppose you like the little fellas, and want to believe in them. What do you have to know in order to admit infinitesimals into your mathematical universe. ***You have to know that they will not lead to contradictions.*** But how can you know this in advance? The remainder of this chapter will discuss this problem, and show that it is solvable. Arguably

this was one of the first serious contributions of formal logic to mathematics itself, as opposed to proving things about mathematics.

4.10 Formal Logic

Leibniz was fond of Chinese Characters and fond of algebra. With this background it occurred to him that perhaps it would be possible to set up a symbol system to make sure that one did not make mistakes in logic. Logic could perhaps be reduced to algebraic computation, and perhaps one could even derive truths algebraically from previously known truths by computation. As with many of his great ideas he never got around to following up on this brilliant idea, but eventually Ernst Schröder(1841-1902) and George Boole(1815-1864) began the project and many others contributed to it. Two main competing symbol system were developed with numberless variants and some mixing of the two. One system was developed to completion by Bertrand Russel and Alfred North Whitehead and used in their Principia Mathematica, and this is the system preferred by philosophers. A second system was developed in Germany and France and this is the system generally used by mathematicians. Be aware that there are several minor variants of this. Naturally I will use the mathematical system. I will give a couple of examples so you know what I am talking about.

We symbolize variables that range over our universe of discourse by x, y, z. We symbolize constants by $a, b, c, \ldots, 0, 1, 2, \ldots$ although we will not use constants much in our examples, since everything we say is true whether or not the constants are part of the language. We also use the symbol $=$ which retains always its normal meaning and is considered part of the logical language. We symbolize statements (Rover is a dog, Fluffy is a cat) by letters p, q, r and we symbolize objects and predicates like this

Dx **or** $D(x)$ x is a dog

Mx **or** $M(x)$ x is a mammal

Sx **or** $S(x)$ x is a snake

and here are the logical symbols

p, q, r statesments without variables (Rover is a dog)

$\neg$ not

$\vee$ or

$\wedge$ and

$\rightarrow$ if ... then ...

$\leftrightarrow$ ... if and only if ...

$\forall$ for all

$\exists$ there exists

Thus with the values give above for p, q etc. we have

$p \wedge q$ Rover is a dog *and* Spot is a cat

$p \vee q$ Rover is a dog *or* Spot is a cat

$p \to q$ *if* Rover is a dog *then* Spot is a cat

$p \to \neg q$ *if* Rover is a dog *then* Spot is *not* a cat

$\neg(p \to q)$ *it is not true that if* Rover is a dog *then* Spot is a cat

$(\forall x)(Dx \to Mx)$ for all x, if x is a dog then x is a mammal (All dogs are mammals)

$(\exists x)(Sx \wedge Mx)$ there is an x so that x is a snake and x is a mammal (Some snakes are mammals)

Note that not all of these are true. However, all matters of fact can be diagrammed with the above equipment. (This excludes, wanting, wishing, believing, and other wishy washy stuff. The system was invented for dealing with mathematics and science where statements are true or false. The system isn't really meant to deal with $(\exists x)(God(x)) \wedge (\forall x)(God(x) \wedge God(y) \to x = y))$ which says there is exactly one God because this is a matter of belief, not fact. Of course people disagree about what statements are facts. The system is meant to be used by groups of people who agree on what the facts are and wish to derive consequences from these facts, although it's a little different in mathematics.) These difficulties are irrelevant for our purposes.

I also must mention that the system presented is not capable of handling all of mathematics; there are theorems for which it is inadequate. We can avoid much annoying technical trivia by simply saying that this system is called *first order logic*. There is a second order logic which allows you to quantify over predicates. The most important example of this is, using $M(x)$ to symbolize 'x is a set',

$$\text{Frege's Axiom} \qquad (\forall F)(\exists x)(M(x) \wedge (\forall y)(y \in x \leftrightarrow F(y)))$$

which says for that for every predicate $F(x)$ there is a set x so that something y is in the set x if and only if the predicate $F(y)$ is true. For example, if $F(y)$ means 'y is a dog' then x is the set of dogs and something is in x if and only if it is a dog. Though this seems clear and unobjectionable it is false. There are predicates for which there is no such set. We will examine this in the problems. This is the reason we need an axiom system, for example the system of Zermelo, to build mathematics from set theory, although Bertrand

Russel and Alfred North Whitehead labored mightily to go down Frege's path, seemingly successfully if not beautifully.

Second order logic is much more complicated than first order logic and almost nothing that we say for first order logic will be true for second order logic, nor for our purposes have we any need for second order logic. Second order logic is convered in [Hermes].

4.11 Syntactics and Semantics

There are two different ways that a schema (a sentence diagram) can be valid, which means always true. We will take for our example

$$(\forall x)(P(x) \wedge Q(x) \rightarrow P(x))$$

This says that if $P(x)$ and $Q(x)$ are both true then $P(x)$ is true and moreover this is true for every x in the universe, or domain of discourse. Not very exciting but valid. Now the syntactic way of dealing with this is to give an algebraic proof. I give it, but only as a sample, not something you are going to understand. This is what Leibniz dreamed of doing.

$$\begin{aligned} P(x) \wedge Q(x) \\ P(x) \\ P(x) \wedge Q(x) \rightarrow P(x) \\ (\forall x)(P(x) \wedge Q(x) \rightarrow P(x)) \end{aligned}$$

This is a proof of $(\forall x)(P(x) \wedge Q(x) \rightarrow P(x))$ by means of a system called *natural deduction*. There are many systems, all complex, and to learn to use any one requires study of a whole book on symbolic logic. There are quite a number of such books, and as long as they have *some* system of derivation like the above they will all work. My favorite is [Quine]. There are some books on symbolic logic that do not go this deep though.

The second method is the semantic method. This consists of choosing some Universe for the variables to live in, *which must be nonempty,* and then giving some interpretation of the predicates. This could be true or false, but if the schema comes out true *for every interpretation of the predicate in every nonempty universe,* then the schema is valid, which you recall means always true.

For example I interpret $P(x)$ to be 'x is a dog' and $Q(x)$ to be 'x is a mammal' and then the schema comes out true

For every x, if x is a dog and x is a mammal then x is a dog.

Note that if an inhabitant of the planet Tralfamadore interprets the symbols so it comes out

For every x, if x is a grzeb and x is a dnorg then x is a grzeb.

it is still true. It's the shape of the sentence that makes it always true, not the actual things and predicates under consideration. So we see that this schema

is valid in all universes with all possible interpretations, and this is a *symantic* test for validity.

Now here is the wonderful thing about first order logic. These two methods of testing for validity of a schema turn out to make the same schemata[9] valid. If it's valid by the syntactic test it's valid by the semantic test and vice versa. *This is not true for more complex systems of logic, for example second order logic.* For proofs of these things you can consult [Hermes] but this is a complicated subject and this is a hard book for beginners.

4.12 Models and Consistency

Suppose we have a set of sentence schemata. There are two possibilities for such a set; it may be possible to derive a contradiction from the set, for example $p \wedge \neg p$, (which cannot be true under any circumstances) in which case the set is called inconsistent, or it may not. If one cannot derive a contradiction from the set of schemata, the set is called consistent.

From the syntactic point of view, a set Σ of schemata is consistent if it is not possible to derive a contradiction from the elements of the set.

From the semantic point of view a set Σ of schemata is consistent if and only if an interpretation of the symbols is possible in which all the schemata come out true, and this interpretation is called a model of the set of schemata.

Once again, the two tests give the same result; a set Σ of first order schemata is consistent by the syntactic test if and only if it is consistent by the semantic test.

Note that an inconsistent set cannot have a model, so having a model is a test for consistency. For example, consider the set

$$(\forall x)(D(x) \rightarrow M(x))$$
$$\neg(\exists x)(M(x) \wedge S(x))$$
$$(\forall x)(S(x) \rightarrow \neg M(x))$$

I now prove that this set is consistent by exhibiting a model; the universe will be our very own, $D(x)$ will be 'x is a dog', $M(x)$ will be 'x is a Mammal' and $S(x)$ will be 'x is a snake'. then the three sentences, when interpreted with this interpretation, mean

All dogs are mammals
There does not exist anything which is a mammal and a snake
All snakes are not mammals

Since all of these sentences are true in our universe, we have come up with an interpretation in which all the sentences of the set are true. Hence our universe is a *model* of the set, which shows the set is consistent. Or, going the other way it illustrates that a consistent set has a model.

[9]Schemata is the Greek plural of Schema. The Latin versions are singular schemum, plural schema which makes things confusing.

Now the fact that the two tests for consistency of a set Σ of first order schemata give the same results has an interesting consequence which is of great importance for our purposes. set Σ of first order schemata is inconsistent if and only if from it a contradiction may be reached by the algebraic proof process. Now it is characteristic of a proof that it has only finitely many lines; otherwise how could we check it? So if Σ is inconsistent we could extract from Σ a finite subset $\Sigma 1 \subseteq \Sigma$ from which a contradiction could be derived. Turning this around, Σ is consistent if and only if ever finite subset of Σ is consistent. Remember, we derived this result using the syntactic test of consistency.

Going over to the semantic test, we can rewrite this as the following, since a set of first order schemata is consistent if and only if it has a model. So we have the theorem

Theorem A set Σ of first order schemata is consistent if and only if every finite subset of Σ is consistent.

or to rewrite it slightly

Theorem A set Σ of first order schemata has a model if and only if every finite subset of Σ has a model.

For finite sets Σ of first order schemata this is of no value but for infinite sets Σ we can extract interesting things. The whole point of the journey here was to see the following example.

We set up a formal system for the real numbers $\mathbb{R}$ by taking the basic operations $+, \cdot, <$ and rewriting them

$$\begin{array}{rcl} P(x,y,z) & \text{means} & x+y=z \\ M(x,y,z) & \text{means} & x \cdot y = z \\ L(x,y) & \text{means} & x < y \\ N(n) & \text{means} & n \text{ is an integer} \end{array}$$

With this vocabulary (and including the constants 0,1) we can write all the (first order) statements about $\mathbb{R}$. For example

$$\begin{array}{rcl} (\forall x)(P(x,0,x)) & \text{means for all x} & x+0=x \\ (\forall x)(M(x,1,x)) & \text{means for all x} & x \cdot 1 = x \\ (\forall x)(\forall y)(\forall (z)(M(x,y,z) \leftrightarrow M(y,x,z) & \text{means for all x,y} & x \cdot y = y \cdot x \\ (\forall x)(\exists y)(\neg(x=0) \rightarrow M(x,y)=1) & \text{means for all x} & x \text{ has an inverse} \\ (\forall z_1)(\forall z_2)(M(x,z_1,1) \wedge M(x,z_2,1) \rightarrow z_1 = z_2) & \text{means} & \text{the inverse is unique} \end{array}$$

We introduce the notation $1/x$ for the unique inverse of x (provided $x \neq 0$).

Now suppose we let Σ_0 be the set of first order statements in the above vocabulary that are true in $\mathbb{R}$. Next we let Σ_{11} be the set $\{L(0,x), L(x,1)\}$ which just says $0 < x < 1$. For integers n we let Σ_{1n} be the set $\{L(0,x), L(x,1/n)\}$ which just says $0 < x < 1/n$.

Next we let

$$\Sigma_n = \Sigma_0 \cup \Sigma_{11} \cup \Sigma_{12} \cup \Sigma_{13} \cup \Sigma_{14} \cup \Sigma_{15} \cup \ldots \cup \Sigma_{1n}$$

A model for Σ_n must satisfy all the sentences the real numbers satisfy and have an x which satisfies $0 < x < 1,\ 0 < x < 1/2,\ 0 < x < 1/3,\ 0 < x < 1/4,$ $0 < x < 1/5, \ldots, 0 < x < 1/n$. I have a model that will work. I take $\mathbb{R}$ as the set and for x I take $x = 1/(n+1)$. Then all the sentences of Σ_n are true in my interpretation.

Now do you see where I am going with this? I let

$$\Sigma_\infty = \Sigma_0 \cup \Sigma_{11} \cup \Sigma_{12} \cup \Sigma_{13} \cup \Sigma_{14} \cup \Sigma_{15} \cup \ldots$$

so Σ_∞ has a schema for $0 < x < 1/n$ *for every* n. Now let us take a finite subset Σ of Σ_∞, which will have some of the schemata of Σ_0 and some schemata that interpret as $0 < x < 1/n$. Since there are only finite many sentence schemata in Σ there must be one of the form $0 < x < 1/n$ with largest n, and we will denote by m this largest n. For example the largest n might be 783 so $m = 783$. Then we seek a model for Σ. I have it. We take the real numbers $\mathbb{R}$ as our set and we interpret all the symbols of Σ as usual in the real numbers which leaves x to interpret. I interpret x as $\frac{1}{784}$ in our example or as $\frac{1}{m+1}$ in general. Then all the sentences of Σ are true in our interpretation and it is then a model of Σ.

We have shown that every finite subset of Σ_∞ has a model. But we know that if every finite subset has a model, then Σ_∞ itself has a model. What do we know about this model. We know that every true statement about $\mathbb{R}$ is true in this model and also that there is an x which satisfies $0 < x < \frac{1}{n}$ for every integer n and thus the number interpreted for x is an infinitesimal. We have proved that infinitesimals cannot cause an inconsistency because we have exhibited a model that has them.

This may have seemed a lot of work simply to prove that infinitesimals are consistent with the real numbers, but the result is philosophically important. It says that if you wish to develop Calculus using infinitesimals Bishop Berkeley cannot stop you. Has this been done? Well it was certainly done at the beginning of Calculus; it was how the subject was originally developed. However recently there have been modern Calculus books written using the infinitesimal technique, the first being [Keisler] and a very recent one being [Dawson].

Now you may ask, if it can be done this way why doesn't everybody do it this way? This brings us to a sad fact about Calculus. Look at the following (without understanding much of course).

$$\sum_{i=1}^{\infty} \int_0^\infty f_i(x)\,dx = \int_0^\infty \sum_{i=1}^{\infty} f_i(x)\,dx$$

This is an equation we hope is true, but it isn't always. It concerns the combination of two processes that each require either limits or infinitesimals to calculate. They are

$$\sum_{i=1}^{\infty} \qquad \text{an infinite sum}$$

$$\int_0^\infty \ldots\,dx \qquad \text{an integral}$$

However, when the two are *combined*, it becomes important to know if the order of combination matters, (sometimes yes, sometimes no) and the game gets significantly harder. It was exactly this that around 1810 caused people to start to consider more seriously the foundations of Calculus. It took about 50 years to come up with conditions that allowed one to interchange the two limit processes (order doesn't matter). The investigations started with Cauchy around 1820 and were pretty well completed for elementary applications by Weierstrass in the 1850s. Questions of this sort make up a good part of the subject now called *real analysis* which is just the higher end of Calculus, although it no longer looks much like Calculus in some of its areas. (There is also an enchanting complex analog, called *complex analysis.*)

Real analysis can be done either using limits or using infinitesimals, but either way it is difficult. A case can be made that when only one limit process is on offer, the infinitesimals are somewhat easier to deal with and the subject may be easier to learn. However, when you have to deal with more than one limit process then new concepts must be introduced to handle the more complicated situation, and the current consensus is that the limit approach is then the easier road. In fact, it may only be a case of the limit approach being the more familiar road, but it is true that the infinitesimal road requires leaning some fairly subtle techniques of mathematical logic, and these may be less geometrically immediate than the limit techniques. But time will tell. In a hundred years infinitesimals may be everywhere and limits rather rare. The takeaway is that both techniques work, and mathematics is enriched by having both.

As regards infinite numbers, these are simply the reciprocals of infinitesimals. We know that

$$\text{if } a < b \text{ then } \frac{1}{a} > \frac{1}{b}$$

(note flip in the inequality sign). Since if α is an infinitesimal we know $0 < \alpha < 1/n$ for every positive integer n so we must have

$$\frac{1}{\alpha} > n \text{ for every positive integer } n$$

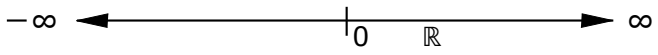
and thus $1/\alpha$ is infinite. It is possible to investigate the new system (essentially $\mathbb{R}$ with infinitesimals and infinite numbers added) in detail but we have gone far enough down this road. I mention that if you think about how they were constructed it is obvious that they form a field.

It is absolutely essential to realize that the infinite numbers we have just constructed have *nothing to do* with the infinite numbers from part A of this chapter, that is, with $\aleph_0$, $\mathfrak{c}$ or $2^{\aleph_0}$. Like chess and football, they are totally different games and knowledge of one is not necessarily of much help in playing the other.

PART C

4.13 Introduction

In this section we deal with the symbolic infinity which one runs into all over mathematics. How real this infinity, symbolized by ∞ is, depends on the mathematician you are speaking with. Some regard it as a mere manner of speaking and some give it a kind of shadow existence. For example, the mathematical symbol $n \to \infty$ can be thought of as simply saying that "n increases beyond all bound." This is the official line; if cornered by a philosopher you fall back on this as the meaning of $n \to \infty$ and say we use $n \to \infty$ because it is a handy way of speaking. This is indeed the real attitude of many mathematicians. However, to be honest, other mathematicians regard ∞ as a not very real point out at the end of the real line $\mathbb{R}$. This is a good time to draw some pictures.



The extended real line

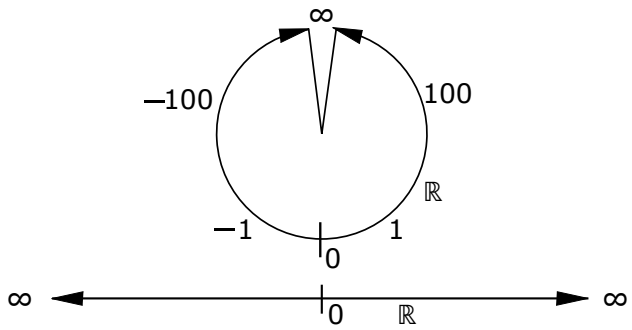
This is the standard picture, with ∞ at the right end of the real line and $-\infty$ at the left end of the real line. This picture comes with an arithmetic. Let a be any real number. It seems amusing to use the computer science term nan (which means *not a number* in computer science) to mean here no definite output, which just means the operation is undefined. Then we have, with $a \in \mathbb{R}$

$$
\begin{aligned}
\infty \pm a &= \infty \\
-\infty \pm a &= -\infty \\
\infty - \infty &= \text{nan} \quad \text{This subtraction is undefined} \\
\text{if } a > 0 \text{ then } a \times \infty &= \infty \\
\text{if } a < 0 \text{ then } a \times \infty &= -\infty \\
\text{if } a > 0 \text{ then } a \times -\infty &= -\infty \\
\text{if } a < 0 \text{ then } a \times -\infty &= \infty \\
\text{if } a = 0 \text{ then } a \times \pm\infty &= \text{nan} \quad \text{This multiplication is undefined} \\
\frac{a}{\pm\infty} &= 0 \\
\frac{\pm\infty}{\pm\infty} &= \text{nan} \quad \text{This division is undefined} \\
\text{if } a > 0 \text{ then } \frac{a}{0} &= \infty \\
\text{if } a < 0 \text{ then } \frac{a}{0} &= -\infty
\end{aligned}
$$

As usual $0/0$ remains undefined. The emotional meaning of the $\pm\infty$ in this picture is that ∞ is *far away to the right* and $-\infty$ is *far away to the left*. Notice with the above arithmetic rules the extended real line is not a field, or a ring, or any kind of decent algebraic structure.

It is sometimes useful to think of ∞ as just far away on the real line at either end, which means that ∞ and $-\infty$ refer to the same (shadow) point and that

the ends of the real line approach one another. This requires a more complicated picture.



The extended real line, another view

The lower picture shows at the bottom the conventional real line with unsigned ∞ at either end. The upper picture shows the real line bent around so that the two ends are going to the same place ∞. (The lines connecting the arrows to the center of the circle have no meaning.) For some kinds of studies this picture is more suitable than the first picture we looked at. It depends on the situation. An arithmetic can be set up analogous to the one for the first form of extended real line, and is simpler than the first one.

This would be the point to mention and embarrassing lapse in mathematical notation; namely the symbol $n \to \infty$ can mean two different things, which you must disentangle from context. It can mean $n \to +\infty$ (n gets big to the right) or it can mean $n \to \pm\infty$ (n gets big without specifying a sign; the absolute value of n gets big). It would be nice if mathematicians adopted the notation $n \to +\infty$ for the first case instead of the ambiguous $n \to \infty$ but there seems no hope that this will happen soon, so be aware of the problem. It is especially important when describing graphs of functions.

A similar picture to this one can be drawn for the complex plane which can be thought of as folded up and the hole at the top filled in with a single point ∞. More carefully, think of the lines through the origin in the complex plane folded up just as the real line above is, and all having the same ∞ at the ends. Can you see that this is essentially a sphere? It is called the Riemann sphere and is critical in complex analysis and also a great deal of fun.

4.14 Problems for Chapter 4

Recall that cardinal numbers which are *not* finite are called transfinite cardinals.

Section 4.1

1. What did Cantor discover that suddenly made set theory interesting?

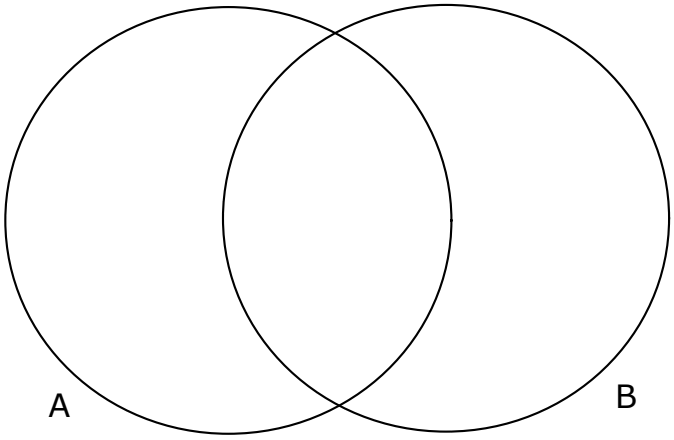
2. What was the original reaction to set theory and how long did it take before set theory became part of ordinary mathematics?

3. What was the objection that opponents of set theory had?

4. What is the great inconvenience if you replace Cantor's set theory by one of the finitistic alternatives?

5. What mathematical problem led Cantor from ordinary mathematics into set theory and it's problems.

Section 4.2



Venn Diagram for Two Sets

1. Copy the Venn diagram[10] and shade $A \cap B$. This would be the area thatis in A *and* in B.

2. Copy the Venn diagram and shade $A \cup B$. This would be the area that is in A *or* in B or in both.

3. Copy the Venn diagram and shade $A - B$. This would be the area that is in A but not in B.

4. Copy the Venn diagram and shade $B - A$. This would be the area that is in B but not in A.

[10]Sometimes called an Euler diagram.

5. Copy the Venn diagram. Using 3. and 4. you can easily shade $A \triangle B = (A - B) \cup (B - A)$.

6. Notice how the diagram makes set containments obvious. For example $A \cap B \subseteq A$ and $A \subseteq A \cup B$.

7. Let $A = a, c, e, g, i, k$ and $B = g, h, i, j, k, l, m, n$. Put the correct letters into the areas of the Venn diagram.

8. Draw a Venn diagram with three circles and lable the circles A, B and C. First shade $A \cap (B \cup C)$. Next (new diagram) shade $A \cap B$ with vertical lines and then $A \cap C$ with horizontal lines. The shadings of the two diagrams should match up. Variants of this technique can be used to prove set identities, like $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ which I hope you recognize as a distributive law.

9. The following is a version of a classical problem that can be found in every Math for Poets book ever written. A Tibetan student has been assigned to sample the eating habits of herdsmen. He and his dog Woof (a massive Tibetan Mastiff) go out to the camp and he takes notes as the herdsmen eat dinner. The choices are yak, mutton and tsampa[11] Unfortunately Woof eats some of the homework but we will try to reconstruct the data from what the student was able to salvage. He had already started working on it and the part that remains contains the following data:

3 of the herdsmen wanted all three items

7 of the herdsmen wanted yak and mutton (including the 3 that wanted all three choices)

21 of the herdsmen wanted yak and tsampa (including the 3 that wanted all three choices)

18 of the herdsmen wanted mutton and tsampa (including the 3 that wanted all three choices)

29 of the herdsmen wanted yak (including those who wanted yak in combination with other choices)

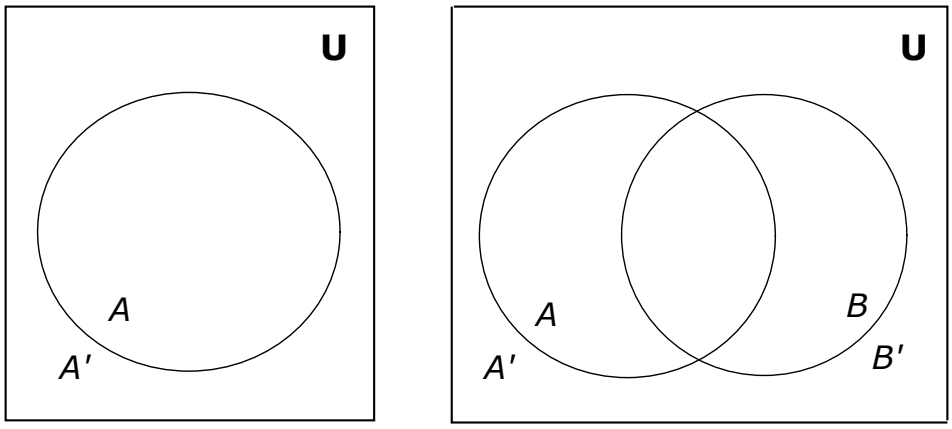
27 of the herdsmen wanted mutton(including those who wanted mutton in combination with other choices)

39 of the herdsmen wanted tsampa(including those who wanted tsampa in combination with other choices)

So now you draw the usual three circles, label them yak, mutton and tsampa, and then from the above information fill in all the regions. You can now answer the following two questions. a) How many herdsmen wanted yak and nothing else? and b) how many herdsmen were in the survey?

[11] A kind of parched barley which is nourishing and lasts a long time.

10 (De Morgan's Laws) We are now going to draw diagrams which include the Universal Set **U**. Here are the appropriate pictures; the left one has the set A inside **U** and the right has the two sets A and B inside **U**. Remember that A' contains all the elements of **U** that are outside of A. Copy the left hand diagram and shade A', the *outside* of the circle.



Venn Diagrams including the Universal set **U**

We will now prove one of De Morgan's laws of set theory (and logic) by shading. Make two copies of the right hand diagram. In the first, keep your eye on $A \cap B$ and shade everything outside of $A \cap B$; this is $(A \cap B)'$. In the second diagram, shade A' with horizontal strokes and B' with vertical strokes. Then $A' \cup B'$ will be all the area shaded, regardless of horizontal or vertical. Notice in the two diagrams idential areas are shaded, proving $(A \cap B)' = A' \cup B'$.

Recall that turning the $\cup$ and $\cap$ upside down in a correct formula will yield a correct formula. Prove this new formula with the methods we just used. When doing intersections it's best to shade the intersecting sets in different directions and the doubly shaded region is the intersection.

Section 4.3

1. Show by putting $A = \{r, s, t, u, m\}$ in one to one correspondence with $B = \{\alpha, \gamma, \mu, \rho, \omega\}$ that the two sets have the same number of elements.

2. Using another word we have shown the two sets are _____. (Fill in word.)

3. What does the following picture

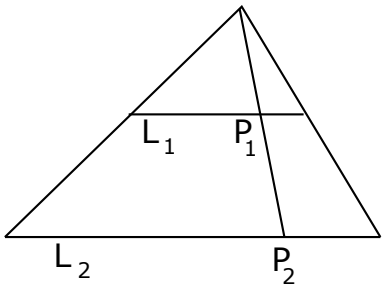


Diagram showing *what* about L_1 and L_2?

show you about the two lines L_1 and L_2? Are there fewer points on the shorter line or the same number of points as on the longer line? Why?

4. Leaving $\mathfrak{c}$ out of consideration, as it requires too much explanation, consider the finite cardinal numbers and $\aleph_0$ and the relation $\leq$ among them.

Section 4.4

1. Give the proper cardinal number for each of the sets a) $\{a, b, a, d, e, a, f, d\}$ (careful; trick question) b)$\{x \in \mathbb{Z} \mid 10 \text{ divides } x\}$ c) The set of real numbers from 2 to 6.

2. Leaving $\mathfrak{c}$ out of consideration, as it requires too much explanation, consider the finite cardinal numbers and $\aleph_0$ and the relation $\leq$ among them. We introduce the notation $A \subsetneqq B$ which means $A \subseteq B$ and $A \neq B$. We defined $<$ in the text but there is also another way to go. Consider 2 and 5 and $2 < 5$. Using the definitions of 2 and 5 as sets could we define $2 < 5$ by one of $2 \in 5$ or $2 \subsetneqq 5$? Which one should we use or does it matter. (This is one of the nice things about von Neumann's definition of the cardinal numbers and is very useful theoretically.) Does the same trick work when you replace 5 by $\aleph_0$? Do you think we might be able to develop the theory so that for any two cardinals $\mathfrak{a}$ and $\mathfrak{b}$

$$\mathfrak{a} < \mathfrak{b} \text{ if and only if } \mathfrak{a} \in \mathfrak{b} \text{ if and only if } \mathfrak{a} \subsetneqq \mathfrak{b}$$

(If we use the slightly ambiguous symbol $\subset$ as we defined it in this book (the same as $\subsetneqq$) this comes out slightly better looking

$$\mathfrak{a} < \mathfrak{b} \text{ if and only if } \mathfrak{a} \in \mathfrak{b} \text{ if and only if } \mathfrak{a} \subset \mathfrak{b}$$

and a little more systematic.)

3. Prove that the set of finite decimals (those which terminate after a finite number of steps) is $\aleph_0$. Do this from your knownledge of the cardinal of the fractions and the rules of $\leq$ for cardinal numbers, so you don't have to set up any one-to-one correspondences.

Section 4.5

1. Using sets as in the text, prove the following: a) $3 + 4 = 7$, b) $3 \cdot 4 = 12$, c) $3^2 = 9$ (Do c. by finding all the functions from a set of size 2 to a set of size 3.)

2. Same as 1 but with $\aleph_0$. a) $3 + \aleph_0 = \aleph_0$ (You need a set of size 3 and a set disjoint from that set of size $\aleph_0$.) b) $\aleph_0 + \aleph_0 = \aleph_0$ (You need two disjoint sets of size $\aleph_0$. Might I suggest the positive and negative integers.) c) Use the Cartesian product of sets to prove $3 \cdot \aleph_0 = \aleph_0$. d) The next problem would naturally be $\aleph_0 \cdot \aleph_0 = \aleph_0$ but lucky for you it's in the text.

3. It takes some proving but not too much to show that the laws of exponents that we learned in section 1.2 continue to hold for transfinite numbers. With these we can prove a great many things, but first I want to give you two important laws. For any cardinal numbers $\mathfrak{a}$, $\mathfrak{d}$ and $\mathfrak{e}$:

$$\text{if } \mathfrak{d} \leq \mathfrak{e} \text{ then } \mathfrak{d}^{\mathfrak{a}} \leq \mathfrak{e}^{\mathfrak{a}}$$

$$\text{if } \mathfrak{d} \leq \mathfrak{e} \text{ then } \mathfrak{a}^{\mathfrak{d}} \leq \mathfrak{a}^{\mathfrak{e}}$$

It is *extremely important* to realize that if $\leq$ is replaced by $<$ the preceding result may be false, as you will soon show. Inequalities with $<$ remain trustworthy only for finite cardinals.

Starting with the sequence of inequalities $2 \leq 10 \leq \aleph_0 \leq 2^{\aleph_0} = \mathfrak{c}$ raise all the terms to the $\aleph_0$ power. Use power laws to simplify the final term, and then notice that the first and last term in the sequence are identical, so that all the terms in the new sequence are equal. Was that not beautiful? Notice that you get $\mathfrak{c}^{\aleph_0} = \mathfrak{c}$ out of this for free.

4. Notice that while $2 < 10$ we have $2^{\aleph_0} \not< 10^{\aleph_0}$ since both are equal to $\mathfrak{c}$ as we showed in the previous problem. This shows that in general you cannot replace $\leq$ by $<$ in the general law quoted above.

Section 4.5

1. We are going to use infinitesimals the way Fermat did to solve an ancient problem. A line that connects two points on a curve is called a secant line. If we have two points (x_1, y_1) and (x_2, y_2) the *slope* m of the secant line through these two points is

$$m = \frac{y_2 - y_1}{x_2 - x_1}$$

Given a point (x_0, y_0) on a line and its slope m the equation of the secant line is

$$y = m(x - x_0) + y_0$$

See if you can see why (first move y_0 to the left side and then think similar triangles. It helps to draw pictures on a graph.) Write the equation of a line through $(1, 1)$ with slope $m = 2$. Simplify.

2 Now we want to find the tangent line to the parabola $y = x^2$ at $(1, 1)$. Let us take an infinitesimal Δx, and find the point $(1 + \Delta x, 1 + \Delta y)$ where of course Δy is another infinitesimal. Since $(1 + \Delta x, 1 + \Delta y)$ is on the curve $y = x^2$, we must have $1 + \Delta y = (1 + \Delta x)^2$. Then

$$\Delta y = (1 + \Delta x)^2 - 1 = 1 + 2\Delta x + (\Delta x)^2 - 1 = 2\Delta x + (\Delta x)^2$$

The slope of the line from $(1, 1)$ to $(1 + \Delta x, 1 + \Delta y)$ is then

$$m = \frac{1 + \Delta y - 1}{1 + \Delta x - 1} = \frac{\Delta y}{\Delta x} = \frac{2\Delta x + (\Delta x)^2}{\Delta x} = 2 + \Delta x$$

Now the two points $(1, 1)$ and $(1 + \Delta x, 1 + \Delta y)$ are infinitely close to one another and so the secant line with $m = 2 + \Delta x$ must be infinitely close to the tangent line. Thus the slope of the tangent line must be $m = 2$. Now copy all this to find the slope of the tangent line at the point $(3, 9)$ on the curve $y = x^2$ Also find the equation of the tangent line. That was probably your first Calculus problem. We now have more efficient methods.

Chapter 5

MATRICES

5.1 Introduction

Matrices are definitely not numbers in any reasonable sense of the term. However, matrices can be used to counterfeit virtually any numerical system, and this is important. It is called a matrix representation of the counterfeited system. In this chapter we will first introduce matrices and their arithmetic, which is that of a ring, and then discuss how to make sets of matrices act like numbers, for example complex numbers. You see this from time to time in mathematics but seldom is it explained how the representations are obtained, and we will do this in detail. It is quite interesting and has many applications throughout science and engineering and as far afield as economics and other social sciences.

We will work almost entirely with 2 by 2 matrices so that we don't get bogged down in onerous calculation. For larger matrices it is really convenient to have a calculator or computer program that will do the arithmetic. From time to time we will up the ante to 3 by 3 and 4 by 4 matrices, very briefly.

Matrices have a long history but much of it is concerned with the subdivision of matrix theory which is determinants. (We will discuss these too.) Determinants and allied areas go very far back in Chinese mathematics as do elimination methods. Cardano has a little to say about this in the Ars Magna (1545) Leibniz, between 1700 and 1710, investigated array methods. Matrices and array methods were more or less synonymous until Arthur Caley (1821-1895) discovered how to multiply matrices. This was the beginning of matrices leading a separate mathematical life apart from mere organization and recording of data. Cayley wrote a large amount about Matrices and many other things, being one of the most prolific mathematical authors. He and J. J. Sylvester 1814-1897 virtually created matrix theory as a separate discipline. Essential additional contributions were made by Ferdinand Georg Frobenius (1849-1917). It now forms a large part of the even larger discipline called linear algebra, which is all-pervasive in modern mathematics from statistics to number theory and is essential for Quantum theory in physics. An interesting feature of linear algebra is that everyone feels he can write a book on the subject are there are a vast number, but there are rather few good books.

I will share a personal story which I think is enlightening. Once at dinner with the science fiction author Jack Vance, a person of vast erudition and skepticism, I was explaining some of the wonders of matrices. He broke in and turned to his son John, an engineer, and asked him if the matrices Schulz was expounding were really all that important and could he think of an application that used them. John was silent for ten seconds, and Jack broke in with "I thought so. Schulz was overselling his stuff as usual." John's reply was "No dad, I was just trying to think of someplace where we *don't* use them.

So I hope in this chapter of the book you will enjoy our little foray into the world of matrices. This is only the foothills of a wonderful range of mountains which I encourage you to explore in your future education.

Historical note: The words matrix (plural matrices) was first used (I think) by J. J. Sylvester who was investigating the subdeterminants of a rectangular array of numbers. Since these determinants came out of the array, which in

some sense is their mother, he used the Latin word matrix (womb) to describe the array.

5.2 Basic Properties and Arithmetic

Matrices are rectangular arrays of "numbers". Usually these are real numbers or integers, but it is not uncommon to find complex numbers (essential for Quantum physics) and nowadays one encounters matrices whose elements are Quaternions, or virtually any fields or division rings. Even greater generalizations are possible but need not concern us. The basic arithmetic rules are the same for all.

The first thing to note is that matrices have a shape. The matrix (we will use upper case letters from the front of the alphabet for matrices)

$$A = \begin{pmatrix} 2 & -3 & 11 \\ -1 & 0 & \frac{4}{3} \end{pmatrix}$$

has *shape* 2×3. This means 2 up and 3 over[1]. Matrices may be added if they have the same shape, and then they are added position by position. For example if

$$B = \begin{pmatrix} -3 & 4 & -11 \\ -5 & 0 & \frac{7}{3} \end{pmatrix}$$

then

$$A + B = \begin{pmatrix} -1 & 1 & 0 \\ -6 & 0 & \frac{11}{3} \end{pmatrix}$$

and

$$A - B = \begin{pmatrix} 5 & -7 & 22 \\ 4 & 0 & -\frac{3}{3} \end{pmatrix} = \begin{pmatrix} 5 & -7 & 22 \\ 4 & 0 & -1 \end{pmatrix}$$

Multiplication of matrices is a little more difficult. First, one must understand matrices can be multiplied if and only if they have *compatible shapes*. The shapes must have the form $m \times n$ and $n \times p$ and the resulting matrix will come out $n \times p$. (The inside terms (which are the same) die and the outside terms give the shape of the product.) Now we give some examples. We begin with the easiest cases. Let

$$A = \begin{pmatrix} -3 & 4 \end{pmatrix} \qquad B = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

$$1 \times 2 \qquad 2 \times 1$$

[1] In the social sciences the order is reversed; this would be a 3×2 matrix; 3 over and 2 down.

Then AB will be a 1×1 matrix formed by multiplying the numbers in corresponding positions (first and second) and adding up the results. Here goes:

$$AB = \begin{pmatrix} (-3) \cdot 2 + 4 \cdot 3 \end{pmatrix} = \begin{pmatrix} 6 \end{pmatrix}$$

Now we up the ante. We will let A be a 2×2 matrix, and each row of the matrix will attack B just as above, ignoring the other row, and then we will have a 2×1 matrix for output.

$$A = \begin{pmatrix} -3 & 4 \\ 2 & -2 \end{pmatrix} \qquad B = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

$$2 \times 2 \qquad 2 \times 1$$

$$AB = \begin{pmatrix} (-3) \cdot 2 + 4 \cdot 3 \\ 2 \cdot 2 + (-2) \cdot 3 \end{pmatrix} = \begin{pmatrix} 6 \\ -2 \end{pmatrix}$$

Note how the first row calculation is exactly the same as the previous example. Now we will let B be a 2×3 matrix and so the product will be a 2×3 matrix.

$$A = \begin{pmatrix} -3 & 4 \\ 2 & -2 \end{pmatrix} \qquad B = \begin{pmatrix} 2 & 5 & -3 \\ 3 & -1 & 0 \end{pmatrix}$$

Then, doing each column separately we have

$$\begin{pmatrix} -3 & 4 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 6 \\ -2 \end{pmatrix} \qquad \begin{pmatrix} -3 & 4 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} 5 \\ -1 \end{pmatrix} = \begin{pmatrix} -19 \\ 12 \end{pmatrix}$$

$$\begin{pmatrix} -3 & 4 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} -3 \\ 0 \end{pmatrix} = \begin{pmatrix} 9 \\ -6 \end{pmatrix}$$

so

$$AB = \begin{pmatrix} -3 & 4 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} 2 & 5 & -3 \\ 3 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 6 & -19 & 9 \\ -2 & 12 & -6 \end{pmatrix}$$

Notice that the entry in the first row second column of AB depends only on the first row of matrix A and the second column of B. Thus from the previous matrix equation we can immediately extract

$$\begin{pmatrix} -3 & 4 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} 2 & -3 \\ 3 & 0 \end{pmatrix} = \begin{pmatrix} 6 & 9 \\ -2 & -6 \end{pmatrix}$$

using the first and third columns of B.

With these operations the $n \times n$ matrices (that is, square matrices of any size n) form a ring, as we discussed long ago. Recall that this means we have the associative and distributive laws, but not the commutative law. Some but not all matrices have multiplicative inverses. The identity for matrix multiplication is the matrix with ones down the principal diagonal and zeros elsewhere.

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

The $n \times n$ matrices are a little better than a ring because you can multiply matrices by numbers from the same source as the elements of the matrices. The numbers commute with the matrices. This kind of algebraic system is called and *Algebra*.

$$3\begin{pmatrix} -3 & 4 \\ 2 & -2 \end{pmatrix} = \begin{pmatrix} -9 & 12 \\ 6 & -6 \end{pmatrix}$$

Note $a(AB) = (aA)B = A(aB)$.

An important fact to keep in mind is that matrices have *divisors of 0*. In a ring, a is a left divisor of 0 and b is a right divisor of 0 if and only if $ab = 0$. For matrices an example is

$$\begin{pmatrix} 4 & 6 \\ 6 & 9 \end{pmatrix}\begin{pmatrix} 3 & -6 \\ -2 & 4 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Before doing inverse matrices we must take a slight detour through the world of determinants. Determinants in various forms are very old, and were probably first used by the Chinese. They have become more useful nowadays since we don't have to compute their values ourselves but can assign that tasks to our computers. However, even with computers there is a limit to the size of the determinants that they can handle, since determinants are so computation intensive.

5.3 Determinants

Determinants are an old, complex and important subject. A determinant is like a Phillip's head screwdriver; when one needs a determinant, nothing else will work. They are closely related to an underappreciated area of mathematics called Grassmann Algebra. Indeed their natural role in mathematics is as the coefficients of elements of the Grassmann Algebra, but we cannot go into this because it is both complicated and a bit tricky and would double the size of the book. So we will limit ourselves to a few definitions and easy theorems which we will need in our work. And of course Cramer's rule.

The determinant function inputs square matrices and outputs numbers, in the sense of whatever numbers are in the matrix array. For two by two matrices the formula is easy to remember and use:

Def The determinant of a 2×2 square matrix is given by

$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

You must clearly distinguish between the following two objects

$$\begin{pmatrix} 2 & 5 \\ -4 & 3 \end{pmatrix} \qquad \text{and} \qquad \begin{vmatrix} 2 & 5 \\ -4 & 3 \end{vmatrix}$$

The first is a matrix, an array of numbers, which does not have a numerical value. The second is the number 26, somewhat disguised.

I must mention that the notation above for a matrix, while easy for ordinary humans, was very difficultfor the typesetters in the days of typesetting by hand. Thus there was an alternate notation for matrices that typesetters found more congenial. It looked like this

$$\begin{bmatrix} 2 & 5 \\ -4 & 3 \end{bmatrix} = \begin{pmatrix} 2 & 5 \\ -4 & 3 \end{pmatrix}$$

With the advent of computerized typesetting there is no longer any need for this altenate notation, expecially since it is harder for humans to handwrite and makes confusion with the determiant more likely. When something like this becomes obsolete the computer term is *deprecated* which means we suggest you do not use it even if it works.

Although most of our later work will use only determinants of 2×2 matrices, some properties of determinants are easier to see with 3×3 determinants. Hence It is worth spending a couple of minutes finding the determinant of a 3×3 square matrix. First you select a row or a column of the determinant. I will arbitrarily select the 3rd row (and boldface it). The terminology here is *expanding off the third row.*

$$\begin{vmatrix} 4 & 3 & 1 \\ 2 & -4 & -6 \\ \mathbf{1} & \mathbf{-3} & \mathbf{-2} \end{vmatrix}$$

Next I will show the minor of each element in the third row. The minor of the first element in the third row is boldfaced. You find it by crossing out the row and the column of the first element.

$$\begin{vmatrix} 4 & \mathbf{3} & \mathbf{1} \\ 2 & \mathbf{-4} & \mathbf{-6} \\ 1 & -3 & -2 \end{vmatrix}$$

Here are the elements and their minors.

$$1, \quad \begin{vmatrix} 3 & 1 \\ -4 & -6 \end{vmatrix} \qquad -3, \quad \begin{vmatrix} 4 & 1 \\ 2 & -6 \end{vmatrix} \qquad -2, \quad \begin{vmatrix} 4 & 3 \\ 2 & -4 \end{vmatrix}$$

Now for the cofactors. Each of the elements of the third row has a *position.* The 1 at the beginning of the row is in the third row first column so its position is (3,1). The other two elements of the row have positions (3,2) and (3,3). To get the cofactors you multiply the minors by $(-1)^{(\text{sum of the position numbers})}$. Thus the cofactors of 1, −3, −2 are

$$(-1)^{3+1} \begin{vmatrix} 3 & 1 \\ -4 & -6 \end{vmatrix} \qquad (-1)^{3+2} \begin{vmatrix} 4 & 1 \\ 2 & -6 \end{vmatrix} \qquad (-1)^{3+3} \begin{vmatrix} 4 & 3 \\ 2 & -4 \end{vmatrix}$$

Next multiply the elements times the cofactors, compute the determinants and add up the three terms to get

$$1 \cdot 1 \cdot (-14) + (-3) \cdot (-1) \cdot (-26) + (-2) \cdot 1 \cdot (-22) = -14 - 78 + 44 = -48$$

This -48 is the value of the determinant.

To compute the value of a 4×4 this way one must compute four 3×3 determinants and this is a nasty task and best left to a machine. Even worse is finding the inverse matrix for a 4×4 which requires 16 3×3 determinants. Thus there has been a minor industry for many decades finding improved ways to do these calculations since for applications one may well have to work with 40×40 matrices and these strain even computing machines and round off error is a real problem.

Back to 3×3 determinants. Suppose you exchange two adjacent rows or two adjacent columns of a determinant $\det(A)$. If you compute the value before and after the switch using one of the columns switched, the signs will be reversed on the cofactors, so switching two adjacent columns in a determinant will change it's sign. Then it turns out they don't have to be adjacent because you can switch any two columns with an odd number of switches of adjacent columns (try it). Thus switching any two columns changes the sign. Now suppose there are two identical columns. When you compute the determinant it doesn't know you switched the columns so the answer is the same. On the other hand, switching the columns changes the sign. So $\det(A)$=-$\det(A)$ and this can only be true if $\det(A)$=0. Hence we have the important principle

If a matrix has two identical rows or two identical columns then it is equal to 0

From the method of evaluating a determinant by expanding off a row or a column it is obvious that

If a determinant has a row or column of 0's it is equal to 0.

If you multiply a determinant by some number you get the same result as if you multiplied any row or any column (just one) by that number. Thus

$$5 \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = \begin{vmatrix} a & b & 5c \\ d & e & 5f \\ g & h & 5i \end{vmatrix}$$

You can see how this must be true because I can expand the second determinant by the third column and then the two sides are obviously the same. Suppose now we have two determinants where two columns are the same and the third is different:

$$\begin{vmatrix} a & b & r \\ d & e & s \\ g & h & t \end{vmatrix} \quad \text{and} \quad \begin{vmatrix} a & b & u \\ d & e & v \\ g & h & w \end{vmatrix}$$

If we expand both of these off the last columns, the two will have the same minors and we can combine the pairs with the same minor, so

$$\begin{vmatrix} a & b & r \\ d & e & s \\ g & h & t \end{vmatrix} + \begin{vmatrix} a & b & u \\ d & e & v \\ g & h & w \end{vmatrix} = \begin{vmatrix} a & b & r+u \\ d & e & s+v \\ g & h & t+w \end{vmatrix}$$

It is easy to see we can introduce multipliers into this to get

$$m\begin{vmatrix} a & b & r \\ d & e & s \\ g & h & t \end{vmatrix} + n\begin{vmatrix} a & b & u \\ d & e & v \\ g & h & w \end{vmatrix} = \begin{vmatrix} a & b & mr+nu \\ d & e & ms+nv \\ g & h & mt+nw \end{vmatrix}$$

With these rules we can write down a very useful additional rule. We know (because two columns are identical) that

$$\begin{vmatrix} c & b & c \\ f & e & f \\ i & h & i \end{vmatrix} = 0$$

Thus

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} + 0 = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} + m\begin{vmatrix} c & b & c \\ f & e & f \\ i & h & i \end{vmatrix}$$

$$= \begin{vmatrix} a+mc & b & c \\ d+mf & e & f \\ g+mi & h & i \end{vmatrix}$$

and we have the important rule that if we add a multiple of a column (or row) of a determinant to a different column (or row) we will not change the value of the determinant. This can save you a lot of work. As an example let us do again the old example

$$\begin{vmatrix} 4 & 3 & 1 \\ 2 & -4 & -6 \\ 1 & -3 & -2 \end{vmatrix}$$

where we know the value is -48. We first add twice the first column to the second to get

$$\begin{vmatrix} 4 & 3+2\cdot 4 & 1 \\ 2 & -4+2\cdot 2 & -6 \\ 1 & -3+2\cdot 1 & -2 \end{vmatrix} = \begin{vmatrix} 4 & 11 & 1 \\ 2 & 0 & -6 \\ 1 & -1 & -2 \end{vmatrix}$$

I now add 11 times the last row to the top row to get

$$\begin{vmatrix} 4+11\cdot 1 & 11+11\cdot(-1) & 1+11\cdot(-2) \\ 2 & 0 & -6 \\ 1 & -1 & -2 \end{vmatrix} = \begin{vmatrix} 15 & 0 & -21 \\ 2 & 0 & -6 \\ 1 & -1 & -2 \end{vmatrix}$$

Now because I have two 0s in a column, the expansion of that column is very simple; I have

$$(-1)^{1+2}\cdot 0\cdot\begin{vmatrix} & & \\ & & \\ & & \end{vmatrix} + (-1)^{2+2}\cdot 0\cdot\begin{vmatrix} & & \\ & & \\ & & \end{vmatrix} + (-1)^{3+2}\cdot(-1)\cdot\begin{vmatrix} 15 & -21 \\ 2 & -6 \end{vmatrix}$$

$$= 0+0+(-1)(-1)\big(15\cdot(-6)-2\cdot(-21)\big) = -90+42 = -48$$

where I have not bothered to fill in the minors for the first two terms since they are multiplied by 0. Thus we needed to compute only *one* 2×2 determinant. This can save time computing the value of a 3×3 determinant but all computations involving determinants are prone to human error so *be careful* or, better, use a machine!

5.3.1 Cramer's Rule

Cramer's rule is a method of solving systems of equations which have a unique solution. It is often not the easiest way to get to the solution but it often has value for theoretical investigations. We will use a 3×3 system to illustrate the rule because it is easier to see what is happening.

Let the system be

$$\begin{aligned} ax + by + cz &= r \\ dx + ey + fz &= s \\ gx + hy + oz &= t \end{aligned}$$

You first form the determinant[2] Δ from the coefficients.

$$\Delta = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}$$

Next you select a variable, say y, to solve for. You form a new determinant

$$\Delta_y = \begin{vmatrix} a & r & c \\ d & s & f \\ g & t & i \end{vmatrix}$$

by replcing the coefficients b, e, h of y by the right hand side of the equations r, s, t. Then if $\Delta \neq 0$ we have

$$y = \frac{\Delta_y}{\Delta}$$

[2]pronounced delta

The proof is truly easy. We have

$$y = \frac{y\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}}{\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}} = \frac{\begin{vmatrix} a & by & c \\ d & ey & f \\ g & hy & i \end{vmatrix}}{\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}} = \frac{\begin{vmatrix} a & ax+by+cz & c \\ d & dx+ey+fz & f \\ g & gx+hy+iz & i \end{vmatrix}}{\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}}$$

$$= \frac{\begin{vmatrix} a & r & c \\ d & s & f \\ g & t & i \end{vmatrix}}{\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}}$$

where in the calculation I used the original equations at the end and the properties of determinants in the first line.

Clearly the size of the equation system doesn't matter. Here is an example for a two by two system, where the method is as good as any other.

$$\begin{aligned} 2x + 7y &= 11 \\ 4x - 5y &= 1 \end{aligned}$$

We have

$$\Delta = \begin{vmatrix} 2 & 7 \\ 4 & -5 \end{vmatrix} \qquad \Delta_x = \begin{vmatrix} 11 & 7 \\ 1 & -5 \end{vmatrix} \qquad \Delta_y = \begin{vmatrix} 2 & 11 \\ 4 & 1 \end{vmatrix}$$

so

$$x = \frac{\Delta_x}{\Delta} = \frac{-62}{-38} = \frac{31}{19} \qquad y = \frac{\Delta_y}{\Delta} = \frac{-42}{-38} = \frac{21}{19}$$

which you can easily check is correct.

If $\Delta = 0$ the system may have a solution or may not. The method for attacking such problems is called Gaussian elimination. The method is simple in concept but difficult to describe in detail, and since we don't actually need it we will omit a description. If you want to pursue this topic the phrase to look up is *reduced row echelon form*. Gaussian elimination is a powerful and useful technique which you will learn in your linear algebra class. There are difficulties in implementing it on computers because of round off error and much work has gone into the programming for it.

5.3.2 Inverse Matrices and Applications

Not all matrices have multiplicative inverses but we will see that if the determinant of a square matrix is non-zero then the matrix does indeed have an inverse

which we can compute with some effort. If a matrix is not square it has no (two sided) inverse.

We begin with an interesting observation. Recall the computation of the determinant

$$\begin{vmatrix} 4 & 3 & 1 \\ 2 & -4 & -6 \\ \mathbf{1} & \mathbf{-3} & \mathbf{-2} \end{vmatrix}$$

where we expanded by cofactors off the last row. The cofactors of the last row are

$$(-1)^{3+1}\begin{vmatrix} 3 & 1 \\ -4 & -6 \end{vmatrix} = -14, \quad (-1)^{3+2}\begin{vmatrix} 4 & 1 \\ 2 & -6 \end{vmatrix} = 26, \quad (-1)^{3+3}\begin{vmatrix} 4 & 3 \\ 2 & -4 \end{vmatrix} = -22$$

As we saw if we multiply these cofactors by the numbers in the last row

$$1(-14) + (-3)(26) + (-2)(-22) = -14 - 78 + 44 = -48$$

we get the determinant -48. However if we multiply the cofactors by the elements in the *second* row we get

$$2(-14) + (-4)(26) + (-6)(-22) = -28 - 104 + 132 = 0$$

This is predictable, since this is an expansion not of the original determinant but of the determinant

$$\begin{vmatrix} 4 & 3 & 1 \\ 2 & -4 & -6 \\ \mathbf{2} & \mathbf{-4} & \mathbf{-6} \end{vmatrix}$$

which will come out 0 because it has two equal rows. Thus the mantra is

Multiply a row (or column) times it's own cofactors; get determinant.
Multiply a row (or column) times another rows cofactors; get 0

This material can be organized in an interesting way if we follow the following recipe to get the *adjoint* $\mathrm{Adj}(A)$ of a square matrix.

1 Replace each element by its minor

2 Put in the signs to get the cofactors

3 Switch the rows to columns

The result is the adjoint of the original square matrix. We now give an example using the recipe on our usual square matrix.

$$\begin{pmatrix} 4 & 3 & 1 \\ 2 & -4 & -6 \\ 1 & -3 & -2 \end{pmatrix} \to \begin{pmatrix} -10 & 2 & -2 \\ -3 & -9 & -15 \\ -14 & -26 & -22 \end{pmatrix} \to \begin{pmatrix} -10 & -2 & -2 \\ 3 & -9 & 15 \\ -14 & 26 & -22 \end{pmatrix}$$

$$\to \begin{pmatrix} -10 & 3 & -14 \\ -2 & -9 & 26 \\ -2 & 15 & -22 \end{pmatrix}$$

Now if we multiply the original matrix times its adjoint we have lined things up so that rows will be multiplying columns of their own or other rows cofactors, and so we naturally get

$$\begin{pmatrix} 4 & 3 & 1 \\ 2 & -4 & -6 \\ 1 & -3 & -2 \end{pmatrix} \begin{pmatrix} -10 & 3 & -14 \\ -2 & -9 & 26 \\ -2 & 15 & -22 \end{pmatrix} = \begin{pmatrix} -48 & 0 & 0 \\ 0 & -48 & 0 \\ 0 & 0 & -48 \end{pmatrix}$$

Thus we have

Theorem If a matrix A is multiplied times its adjoint $\mathrm{Adj}(A)$ (0n either side) the result is a matrix with the determinant of A down the diagonal and 0's elsewhere.

Now if $\det(A) \neq 0$ then we can divide by $\det(A)$ and the previous theorem tells us

Theorem if $\det(A) \neq 0$ then

$$A^{-1} = \frac{1}{\det(A)} \mathrm{Adj}(A)$$

The process is reduces to a much simpler one for 2×2 matrices. If you follow all the steps above for

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

you will get

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

The mantra for the inverse of a 2×2 matrix is to swap the entries on the principal diagonal, change the signs on the other two, and divide by the determinant. This is very easy. Now let us solve a 2 by 2 system of equations using a matrix inverse. The original problem is

$$\begin{aligned} 3x - 7y &= 5 \\ 9x + 4y &= 11 \end{aligned}$$

The first step is to replace the system by a single matrix equation.

$$\begin{pmatrix} 3x - 7y \\ 9x + 4y \end{pmatrix} = \begin{pmatrix} 5 \\ 11 \end{pmatrix}$$

The second step is to decouple the constants from the variables using matrix multiplication

$$\begin{pmatrix} 3 & -7 \\ 9 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 5 \\ 11 \end{pmatrix}$$

What a coincidence that matrix multiplications does the decoupling so well. Now recall how we solve $2x = 6$. We multiply both sides by the inverse $2^{-1} = \frac{1}{2}$

getting $1x = \frac{1}{2} \cdot 6$ and thus $x = 3$. The same trick, multiplying by the inverse, works for our problem too.

$$\frac{1}{75}\begin{pmatrix} 4 & 7 \\ -9 & 3 \end{pmatrix}\begin{pmatrix} 3 & -7 \\ 9 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{75}\begin{pmatrix} 4 & 7 \\ -9 & 3 \end{pmatrix}\begin{pmatrix} 5 \\ 11 \end{pmatrix}$$

$$\frac{1}{75}\begin{pmatrix} 75 & 0 \\ 0 & 75 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{75}\begin{pmatrix} 97 \\ -12 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{75}\begin{pmatrix} 97 \\ -12 \end{pmatrix}$$

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{97}{75} \\ \frac{-12}{75} \end{pmatrix}$$

It works just as well for 3×3 or 40×40 systems provided the determinant of the coefficient matrix is non-zero. Of course there are serious computational difficulties as the coefficient matrices get larger. But theoretically the problem is solved.

I now have to give you a theorem that I really don't want to prove. There is a proof that would work at this level but it depends on some material we haven't covered and you probably don't want me to cover. The 2×2 case is easily done by brute force if you'd like to try your hand at the algebra, but it's a little boring. This theorem relates the determinant of the product of two matrices to the determinants of each of them. Things are as happy as they could be.

Theorem Let A and B be two square matrices of the same size. Then

$$\det(AB) = \det(A)\det(B)$$

This is one of those theorems where if you have the proper tools available it is quite trivial and if you don't it is miserably hard. The tool is, once again, Grassmann Algebra. I know about 3 proofs that don't use Grassmann Algebra but all of them are long and boring and, worse, give you no idea at all why the theorem is true. In linear algebra most of the proofs actually *do* tell you why the theorem is true, but for this one only the Grassmann algebra proof does this. The others just verify it. Someday, if you follow a mathematical career, some teacher will lead you through one of these proofs, but it won't be me.

5.4 Matrix Representation of Numbers

In this section we will show that certain sets of matrices act exactly like number systems. We will explain how to come up with these matrices, without digging too deeply into the theory. Some of what I say you will have to take on faith but all of it will be useful later.

Let us start with an example, ordinary three space $\mathbb{R}^3$. Suppose we let unit vectors point out the x, y and z axes and we will call these $\hat{\mathbf{i}}, \hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$. It is customary to make vectors boldface in math texts but the custom is not always

followed and, of course, cannot be used in handwriting. A second custom then is to put a little arrow on top of the letter to indicate it is a vector $\vec{\mathbf{v}}$. A third custom, more usual in physics than in mathematics, is to replace the arrow with a hat when the vector is a unit vector. We will follow all these customs simultaneously.

Any vector in $\mathbb{R}^3$ can be written as a sum $\vec{\mathbf{v}} = v_1\hat{\mathbf{i}} + v_2\hat{\mathbf{j}} + v_3\hat{\mathbf{k}}$. This is called *expressing the vector in terms of the basis* $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$. There is nothing sacred about these three vectors and other vectors could be used instead but humans seem to prefer these and often refer to them as "the natural basis" and we will follow the custom though not the terminology. When I taught classes in Linear Algebra I called $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, $\hat{\mathbf{k}}$ the *childrens' basis* to indicate that more knowledgeable users might well choose to use other bases.

In the neighborhood of vectors the coefficients v_1, v_2, v_3 are not called numbers but instead are called scalars, from ancient habit. In ordinary use of vectors the scalers commute with the vectors so that $2\vec{\mathbf{v}} = \vec{\mathbf{v}}2$ and it does not matter at all whether the scalars are written on the left or right of the vectors. However, in the representation game there are systems in which the quantities we designate as "scalars" do *not* commute with the vectors and then it matters and the scalars really belong on the right. Forgetting this can lead to sad results and much frustration, since one seems to be doing everything right and still it doesn't work. So for certain things we do we will be writing the coefficients v_i on the right where they really belong.

In this section we want to illustrate the idea of a representation by matrices, and so we will represent vectors by *columns*. This choice was made by Cayley back in the early days of matrices and has to do with writing our functions to the left of their arguments: f is to the left of x in $f(x)$, and $g(f(x))$ means take x, first do f to it, and then do g to the output from f. When things work this way it is natural to have what represents the argument go to the right of the matrix that represents the activity, and so it must be a column for matrix multiplication to work.

Occasionally people try to change the system and naturally this just makes everything worse. The system as it is was set up by many very very smart people and is used by almost everyone, so putting up with its inconveniences is just part of life. AND, changing the system does not get rid of the inconveniences; it just puts them in different places. So do your best to get along with such people but don't loan them money.

When we are setting up representations we will write the scalars on the right side of the vectors. At the critical point where this matters I will point it out (forcefully).

Now the idea of a matrix representation is that computations can be done more conveniently by humans and computers with the matrices than using the actual objects. This is not obvious at first but will become clear later and actually was one of the reason matrices were invented in the first place.

So to begin our journey I will represent to vectors in $\mathbb{R}^3$ by matrices.

$$\vec{\mathbf{v}} = \hat{\mathbf{i}}v_1 + \hat{\mathbf{j}}v_2 + \hat{\mathbf{k}}v_3 \longleftrightarrow \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

For example

$$\vec{\mathbf{v}} = \hat{\mathbf{i}}2 + \hat{\mathbf{j}}(-3) + \hat{\mathbf{k}}5 \longleftrightarrow \begin{pmatrix} 2 \\ -3 \\ 5 \end{pmatrix}$$

and

$$\vec{\mathbf{w}} = \hat{\mathbf{i}}4 + \hat{\mathbf{j}}7 + \hat{\mathbf{k}}(-2) \longleftrightarrow \begin{pmatrix} 4 \\ 7 \\ -2 \end{pmatrix}$$

Then

$$\vec{\mathbf{v}} + \vec{\mathbf{w}} = \hat{\mathbf{i}}6 + \hat{\mathbf{j}}4 + \hat{\mathbf{k}}3 \longleftrightarrow \begin{pmatrix} 6 \\ 4 \\ 3 \end{pmatrix}$$

The matrices on the right act just like the vectors on the left, but they are much easier to use in a computer since all computers have storage arrays that can easily be made to act like matrices. In this case there is not much going on here; just a cosmetically different way of writing the vectors. In the next subsection we will see a more interesting application of the same ideas.

5.4.1 Representation of the complex numbers

What corresponds to the $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$ in the previous subsection are for complex numbers 1 and i because any complex number can be written in term of 1 and i, for example $2 + 3i = 1 \cdot 2 + i \cdot 3$. Also recall that complex numbers are more than vectors with basis $1, i$; they also do things to other complex numbers. Recall that $\frac{1}{\sqrt{2}}(1 + i)$ rotates complex numbers by 45°. Activities are represented in the matrix game by matrices, in this case square matrices, and I now show you how to find the matrices by example. We need to represent the basis elements 1 and i; once we have these we can get any other complex number. Here is the method, which is perfectly general. You represent multiplication by 1 and i and extract the scalars as shown. First multiplication by 1, which is trivial

$$\begin{array}{rcl} 1 \cdot 1 &=& 1 \cdot 1 + i \cdot \mathbf{0} \\ 1 \cdot i &=& 1 \cdot 0 + i \cdot 1 \end{array} \longleftrightarrow \begin{pmatrix} 1 & 0 \\ \mathbf{0} & 1 \end{pmatrix}$$

Note that I boldfaced one of the 0s so you could see that when making the matrix you switch the entries in the rows to columns of the matrix[3]. This was dull, since we just got the identity matrix. But the next one is more interesting.

$$\begin{array}{rcl} i \cdot 1 &=& 1 \cdot 0 + i \cdot 1 \\ i \cdot i &=& 1 \cdot (-1) + i \cdot 0 \end{array} \longleftrightarrow \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

[3]Just DO IT! Also note we did exactly the same thing when we represented vectors with column matrices

Now we have, without going through the process, although we could,

$$2+7i = 1\cdot 2 + i\cdot 7 \longleftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} 2 + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} 7 = \begin{pmatrix} 2 & -7 \\ 7 & 2 \end{pmatrix}$$

The same process will give us

$$a+bi \longleftrightarrow \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

and that is the matrix representation of complex numbers. Now watch this:

$$2+7i \longleftrightarrow \begin{pmatrix} 2 & -7 \\ 7 & 2 \end{pmatrix}$$

$$3-2i \longleftrightarrow \begin{pmatrix} 3 & 2 \\ -2 & 3 \end{pmatrix}$$

$$(2+7i)(3-2i) \longleftrightarrow \begin{pmatrix} 2 & -7 \\ 7 & 2 \end{pmatrix} \begin{pmatrix} 3 & 2 \\ -2 & 3 \end{pmatrix}$$

$$20+17i \longleftrightarrow \begin{pmatrix} 20 & -17 \\ 17 & 20 \end{pmatrix}$$

Where on each side I have done the multiplication.

Now take stock of what we have done here. All complex number arithmetic can be done by the matrix representatives. The inverse of the matrix works like the inverse of the complex number also. And note

$$\det \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = a^2 + b^2$$

which is the norm of the complex number. Real numbers are represented by diagonal matrics

$$14 = 1\cdot 14 \longleftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} 14 = \begin{pmatrix} 14 & 0 \\ 0 & 14 \end{pmatrix}$$

Thus every computation that we do with complex numbers written in the form $a+bi$ can be done using two by two matrices with real number entries.

So, should a philosopher object to computations done in the $a+bi$ form on the grounds that $i^2 = -1$ and *there is no square root of -1* you can say OK, I'll just use my matrices which use only the real numbers you are so in love with, and I'll use $a+bi$ as an abbreviation for

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

Then all the computations with $a+bi$ become abbreviations of matrix calculations. Surely you don't object to matrices with real entries do you Mr. Philosopher?

Another small point before we go on. If you recall multiplication by i turns a complex number 90° counterclockwise. Hence representing the complex number to be turned by a column vector we have the representation of the rotated vector. If the complex number is $2+3i$ we find the rotated complex number from

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} -3 \\ 2 \end{pmatrix}$$

so the rotated complex number is $-3+2i$. More on this later.

5.4.2 Real Representation of the Quaternions

Using the identical methods to those used in the real representation of complex numbers we can use the basis $1, \hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$ to find a real representation of the quaternions. Since these are 4×4 matrices they are not so useful to humans as they were in the complex case, but they work fine for computers. I will do the matrix for $\hat{\mathbf{i}}$ and just tell you the ones for $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$. You can work these out for yourself easily enough.

$$\begin{aligned}
\hat{\mathbf{i}} \cdot 1 &= 1 \cdot 0 + \hat{\mathbf{i}} \cdot 1 + \hat{\mathbf{j}} \cdot 0 + \hat{\mathbf{j}} \cdot 0 \\
\hat{\mathbf{i}} \cdot \hat{\mathbf{i}} &= 1 \cdot (-1) + \hat{\mathbf{i}} \cdot 0 + \hat{\mathbf{j}} \cdot 0 + \hat{\mathbf{j}} \cdot 0 \\
\hat{\mathbf{i}} \cdot \hat{\mathbf{j}} &= 1 \cdot 0 + \hat{\mathbf{i}} \cdot 0 + \hat{\mathbf{j}} \cdot 0 + \hat{\mathbf{j}} \cdot 1 \\
\hat{\mathbf{i}} \cdot \hat{\mathbf{k}} &= 1 \cdot 0 + \hat{\mathbf{i}} \cdot 0 + \hat{\mathbf{j}} \cdot (-1) + \hat{\mathbf{j}} \cdot 0
\end{aligned}$$

Thus the matrix for $\hat{\mathbf{i}}$ is

$$\hat{\mathbf{i}} \longleftrightarrow \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

The matrices for $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$ are

$$\hat{\mathbf{j}} \longleftrightarrow \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \qquad \hat{\mathbf{k}} \longleftrightarrow \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

You can verify these work by checking

$$\hat{\mathbf{i}}^2 = -1 \qquad \hat{\mathbf{j}}^2 = -1 \qquad \hat{\mathbf{i}}\hat{\mathbf{j}} = -\hat{\mathbf{j}}\hat{\mathbf{i}}$$

which are the defining relations for the quaternions. Here -1 means the negative of the identity matrix of course.

5.4.3 Complex Representation of the Quaternions

In this section we find a representation of the quaternions as 2×2 matrices with *complex* number entries. This is a valuable resource but there are tricky aspects which I emphasize.

The basis used to find this representation will be $1, \hat{\mathbf{j}}$ and the "scalers" will be the complex numbers where we write a complex number as $a + b\hat{\mathbf{i}}$. Since $\hat{\mathbf{i}}$ the quaternion and i the complex number both square to -1, we can consider them the same thing. When we write down the matrices I will swap $\hat{\mathbf{i}}$ for i as is customary.

Naturally 1 has the identity 2×2 matrix for it's matrix representation. We won't bother to derive this again.

So here we go deriving the matrix for left multiplication by $\hat{\mathbf{i}}$. Remember the mantra "Action on the left, coefficients on the right" and stuff will work. Behind this is the associative law of multiplication.

$$\begin{aligned} \hat{\mathbf{i}} \cdot 1 &= \hat{\mathbf{i}} \cdot 1 + \hat{\mathbf{j}} \cdot 0 = 1 \cdot \hat{\mathbf{i}} + \hat{\mathbf{j}} \cdot 0 \\ \hat{\mathbf{i}} \cdot \hat{\mathbf{j}} &= 1 \cdot 0 + \hat{\mathbf{i}} \cdot \hat{\mathbf{j}} = 1 \cdot 0 + \hat{\mathbf{j}} \cdot (-\hat{\mathbf{i}}) \end{aligned}$$

where in the second line we had to conpensate for the basic element $\hat{\mathbf{j}}$ moving to the left side by using $\hat{\mathbf{i}} \cdot \hat{\mathbf{j}} = -\hat{\mathbf{j}} \cdot \hat{\mathbf{i}}$. We now read off the matrix corresponding to $\hat{\mathbf{i}}$ (remembering to switch rows to columns) as

$$\hat{\mathbf{i}} \longleftrightarrow \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

We do the same for $\hat{\mathbf{j}}$.

$$\begin{aligned} \hat{\mathbf{j}} \cdot 1 &= 0 \cdot 1 + 1 \cdot \hat{\mathbf{j}} = 1 \cdot 0 + \hat{\mathbf{j}} \cdot 1 \\ \hat{\mathbf{j}} \cdot \hat{\mathbf{j}} &= -1 \cdot 1 + 0 \cdot \hat{\mathbf{j}} = 1 \cdot (-1) + \hat{\mathbf{j}} \cdot 0 \end{aligned}$$

No compensations necessary this time. We read off the matrix as

$$\hat{\mathbf{j}} \longleftrightarrow \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Now for $\hat{\mathbf{k}}$.

$$\begin{aligned} \hat{\mathbf{k}} \cdot 1 &= 0 \cdot 1 + \hat{\mathbf{i}} \cdot \hat{\mathbf{j}} = 1 \cdot 0 + \hat{\mathbf{j}} \cdot (-\hat{\mathbf{i}}) \\ \hat{\mathbf{k}} \cdot \hat{\mathbf{j}} &= -\hat{\mathbf{i}} \cdot 1 + 0 \cdot \hat{\mathbf{j}} = 1 \cdot (-\hat{\mathbf{i}}) + \hat{\mathbf{j}} \cdot 0 \end{aligned}$$

so the matrix is

$$\hat{\mathbf{k}} \longleftrightarrow \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix}$$

Now let's check some things to see if it works.

$$\begin{array}{ccc} \hat{\mathbf{i}} & \hat{\mathbf{j}} & = \quad \hat{\mathbf{k}} \\ \updownarrow & \updownarrow & \quad \updownarrow \\ \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} & \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} & = \quad \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} \end{array}$$

Another check:

$$\begin{array}{ccc} \hat{\mathbf{k}} & \hat{\mathbf{k}} & -1 \\ \updownarrow & \updownarrow & \updownarrow \\ \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} & \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} & = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \end{array}$$

Notice that the scaler -1 is represented by -1 times the identity matrix but the "scaler" i does not have this property, because it doesn't commute with the basis elements like a real scaler does. Nevertheless, our methodology continues to work if you just remember the mantra.

We can also now write down the matrix for any quaternion just by adding corresponding terms.

$$\begin{aligned} a+b\hat{\mathbf{i}}+b\hat{\mathbf{k}}+c\hat{\mathbf{k}} \quad &\longleftrightarrow \quad a\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} + c\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} + d\begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} \\ &= \quad \begin{pmatrix} a+bi & -c-di \\ c-di & a-bi \end{pmatrix} \end{aligned}$$

Let's look at the determinat of this matrix:

$$\det\begin{pmatrix} a+bi & -c-di \\ c-di & a-bi \end{pmatrix} = (a+bi)(a-bi)-(-c-di)(c-di) = a^2+b^2+c^2+d^2$$

which is the norm of the corresponding quaternion. This can be used to give an easy proof of the result $N(AB) = N(A)N(B)$ for A, B quaternions based on the equation $\det(AB) = \det(A)\det(B)$ where A and B are matrices.

5.5 Problems for Chapter 5

Section 5.2

Here are some problems to get you used to manipulating matrices. I've thrown in some odd ones to get you used to what can happen.

1.
$$\begin{pmatrix} 2 & -3 & 1 \\ -1 & 0 & \frac{4}{3} \end{pmatrix} + \begin{pmatrix} -2 & 5 & -2 \\ -5 & 0 & \frac{-7}{3} \end{pmatrix}$$

2.
$$\begin{pmatrix} 2 & -1 \\ -3 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ -3 \end{pmatrix}$$

3.
$$\begin{pmatrix} 2 & -1 \\ -3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ -3 & 1 \end{pmatrix}$$

4.
$$\begin{pmatrix} 2 & -1 \\ -3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 & -1 \\ -3 & 1 & 3 \end{pmatrix}$$

5.
$$\begin{pmatrix} 2 & -1 \\ -4 & 2 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix}$$

Recall that if $ab = 0$ and $a, b \neq 0$ then a is called a *left zero divisor* and b is called a right zero divisor, and we have a pair of them here. Recall there are no zero divisors in a field or in a division ring.

6. To get the transpose $A^\intercal$ of a matrix A you change all the rows to columns. Here we are going to practice this and also find something interesting. Let

$$A = \begin{pmatrix} 2 & -1 \\ -3 & 4 \end{pmatrix} \text{ and } B = \begin{pmatrix} 4 & 1 \\ -1 & 2 \end{pmatrix}$$

Note
$$A^\intercal = \begin{pmatrix} 2 & -3 \\ -1 & 4 \end{pmatrix}$$

Find the following: $B^\intercal$, AB, $(AB)^\intercal$, $A^\intercal B^\intercal$, $B^\intercal A^\intercal$. Make a conjecture about how transpose interweaves with the product of matrices. Clearly $(AB)^\intercal \neq A^\intercal B^\intercal$. But from your calculation what does $(AB)^\intercal$ appear to equal?

7. Using 6., show that if A is a left divisor of 0 then $A^\intercal$ is a right divisor of 0. Do this and the next problem *symbolically* using the law you conjectured in 6. Don't write out any actual matrices.

8. A matrix A is *symmetric* if and only if $A = A^{\intercal}$. Two matrices A and B *commute* if and only if $AB = BA$. Show that if A and B are symmetric and A and B commute then AB is symmetric.

9. Let

$$A = \begin{pmatrix} 2 & -1 \\ -5 & 3 \end{pmatrix} \text{ and } B = \begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix}$$

Find AB and BA. When the product of two square matrices is the identity matrix the two matrices are called inverses of one another. The order of multiplication does not matter.

9. As we mentioned at the beginning of the chapter, the entries of matrices can be complex numbers or quaternions or even more general kinds of numbers. Matrices with quaternions have some peculiar difficulties and have never attained popularity, but matrices with complex numbers are the computational workhorses of quantum theory. The arithmetic works just like with real numbers. Here's an example:

$$A = \begin{pmatrix} 2i & -1-i \\ -2 & 3+i \end{pmatrix} \text{ and } B = \begin{pmatrix} 3-i & i \\ -i & 2+i \end{pmatrix}$$

Find $A + B$, AB and BA.

Section 5.3

1. Find the value of the following:

$$\det \begin{pmatrix} 4 & 3 \\ 5 & -2 \end{pmatrix} \qquad \begin{vmatrix} 3 & 6 \\ 2 & 4 \end{vmatrix} \qquad \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} \qquad \begin{vmatrix} -4 & 6 \\ 2 & 5 \end{vmatrix} \qquad \begin{vmatrix} i & 2+i \\ 2i & -3 \end{vmatrix}$$

2. Find the value of the following determinant which is the same one as was used in the text, but this time expand off the top row.

$$\begin{vmatrix} 4 & 3 & 1 \\ 2 & -4 & -6 \\ 1 & -3 & -2 \end{vmatrix}$$

3. Once again find the value of the same determinant (the one used in the text), but this time expand off the middle column. It works the same way.

$$\begin{vmatrix} 4 & 3 & 1 \\ 2 & -4 & -6 \\ 1 & -3 & -2 \end{vmatrix}$$

Naturally the answers you got for should be the same as in the text.

4. We are going to evaluate the same determinant again but this time we will use a little trickery. We will multiply the first column by 3 and add it to the second column. Then we will multiply the first column by 2 and add it to the third column. These operations do not change the value of the determinant. Then we will evaluate by expanding off the third row, which means you need only find one cofactor.

$$\begin{vmatrix} 4 & 3 & 1 \\ 2 & -4 & -6 \\ 1 & -3 & -2 \end{vmatrix} = \begin{vmatrix} 4 & 15 & 1 \\ 2 & 2 & -6 \\ 1 & 0 & -2 \end{vmatrix} = \begin{vmatrix} 4 & 15 & ?? \\ 2 & 2 & ?? \\ 1 & 0 & ?? \end{vmatrix}$$

$$= 1 \cdot (+1) \cdot \begin{vmatrix} 15 & 9 \\ 2 & ? \end{vmatrix} = -48$$

You now see how much work can be saved by adding multiples of a row to another row or adding multiples of a column to another column.

5. Now we will solve a two by two system by Cramer's rule.

$$\begin{aligned} 2x + 5y &= 8 \\ 3x - 2y &= -7 \end{aligned}$$

Check your answer!

6. There is a kind of $n \times n$ matrix where you put the first n^2 positive integers into it in order. For $n = 3$ you get

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}$$

It is interesting that for $n \geq 3$ this determinant is always 0. To see this, subtract the first row from the second and the second row from the third. Would this work similarly for any size $n \geq 3$? Why does it fail for $n = 2$?

7. Solve the two by two system using Cramer's rule:

$$\begin{aligned} 4x + 3y &= 10 \\ 2x - 4y &= -6 \end{aligned}$$

8. Solve the two by two system by inverting the matrix:

$$\begin{aligned} 4x + 3y &= 10 \\ 2x - 4y &= -6 \end{aligned}$$

9. Solve the three by three system using Cramer's rule:

$$\begin{aligned} 4x + 3y + 1z &= 13 \\ 2x - 4y - 6z &= -24 \\ 1x - 3y - 2z &= -11 \end{aligned}$$

Check your answer.

10. Solve the three by three system in problem 9. by inverting the matrix.

Note: For three by three systems of equations you will note that neither Cramer's rule nor the matrix inverse method is easy. There is another method which has some advantages over these methods called Gaussian Elimination, although it is still not fun. If you are curious you can look Gaussian elimination up on the internet.

Section 5.4

Since we did the representation of complex numbers in the text, there isn't much left for you to do with that. However, we can take a different field and have you come up with a representation for it. It will all work almost the same way except that here and there some numbers will change.

1. We are going to represent the numbers of the field $\mathbb{Q}[\sqrt{2}]$ by matrices. You need to find the matrices representing the basis elements for $\mathbb{Q}[\sqrt{2}]$ as I did in the text for $1, i$. These are of course $1, \sqrt{2}$. 1 works the same as in the text but there are some small changes for $\sqrt{2}$. When you have found the matrix representing $\sqrt{2}$, multiply it by itself and you should get the matrix representing 2.

2. Now that you have the matrices representing 1 and $\sqrt{2}$, you can easily find the matrix repreenting $a + b\sqrt{2} = 1 \cdot a + \sqrt{2} \cdot b$.

3. Find the matrices representing $1 + \sqrt{2}$ and $1 - \sqrt{2}$ and multiply them. Does what you get make sense?

4. We are now going to deal with the cube roots of 1. This has a few curve balls. To find the cube roots, we start with the equation $x^3 - 1 = 0$ and factor the left side into $(x-1)(x^3+x+1) = 0$. One cube root is obviously 1. The second cube root is obtained by solving the quadratic using the quadratic formula in the text. We call this cube root ω, and

$$\omega = \frac{-1 + \sqrt{-3}}{2} = \frac{-1 + i\sqrt{3}}{2}$$

What we want to do in this problem is find a matrix representation of the numbers in $\mathbb{Q}[\omega]$ *never using the above formula for* ω. This is the way a professional would do it; the algebraic way. We know one thing about ω because we got ω by solving the equation $x^2 + x + 1 = 0$ so we know $\omega^2 + \omega + 1 = 0$ which tells us that

$$\omega^2 = -1 - \omega$$

Next we multiply this by ω getting

$$\omega^3 = -\omega - \omega^2 = -\omega - (-1 - \omega) = 1$$

which is no surprise. Now it is time for the matrix representation. The basis we use is, of course, $1, \omega$. We copy the process used in the text getting (not forgetting to switch the rows to columns) and get

$$1 \longleftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Now we do this for ω, that is, we multiply 1 and ω by ω and then re-express this in terms of 1 and ω, which if you are clever you can find above in what I did. Schematically it would look like this

$$\begin{aligned} \omega \cdot 1 &= 1 \cdot a + \omega \cdot b \\ \omega \cdot \omega &= 1 \cdot c + \omega \cdot d \end{aligned}$$

Your job is to figure out what numbers go in where I have a, b, c, d. When you have done that you get the matrix for ω as

$$\omega \longleftrightarrow \begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

Note the numbers in the matrix have the rows and columns swapped from the equations above. Now you have found the matrix for ω.

5. Continuing with the previous problem, first multiply the matrix ω times itself to get the matrix for ω^2. Check this by finding the matrix for

$$-1 - \omega \longleftrightarrow -(\text{matrix for } 1) \; - \; (\text{matrix for } \omega)$$

This should be the same as the matrix for ω^2. If it is, everything is going great. If it is not you have made a mistake. Go find it. (Part of being a mathematician is 1. devising self checks like this and 2. finding the screw up when check fails.)

6. Now that we are sure we have the correct matrices, find the matrix for any member $a + b\omega \in \mathbb{Q}[\omega]$ by multiplying a times the matrix for 1 and b times the matrix for ω to get

$$a + b\omega \longleftrightarrow \begin{pmatrix} & , \\ , & \end{pmatrix}$$

where you fill in the matrix with entries involving a and b. Partial answer:

$$a + b\omega \longleftrightarrow \begin{pmatrix} & , \\ , & a - b \end{pmatrix}$$

7. Now multiply $(3 + 2\omega)$ times $(4 - 3\omega)$ using the following procedure. Convert each of them to a matrix, multiply the two matrices, and then extract the product from the matrix. *You* may not think this is wonderful but a

computer would because computers *love* to multiply matices. Now you can check your work by dividing the product you got by $(4-3\omega)$. The best way to do this is multiplying the matrix for the product by the *inverse* of the matrix for $(4-3\omega)$. You should get the matrix for $(3+2\omega)$. If you don't, find the screwup. Note that you know how solve this problem with the matrices but you *don't* know how to solve it with the $a+b\omega$ kind of numbers. Extract from your calculations somewhere the fact that

$$\frac{1}{4-3\omega} = \frac{7+3\omega}{37}$$

8. We have found algebras of dimension 1,2,4. We might ask if there are any similar algebras of dimension 8. The answer is, yes and no. Yes there is an algebra with many properties in common with the quaternions, called the *Octonions* with code letter $\mathbb{O}$. However, there is a theorem that says the only associative division algebras contining $\mathbb{R}$ are the Reals, Complex numbers and Quaternions. Hence there is something not quite right about the Octonions $\mathbb{O}$. Given that the Octonions have inverses. So two questions: a) what algebraic property so the Octonions not have and b) Is a matrix representation of the Octonions possible?

Chapter 6

SOME NUMBER THEORY

6.1 Introduction

My thought in including this chapter is that a book on numbers would benefit from having a chapter that dug a little deeper into some subject and since number theory is fairly neglected in elementary education and since it it very interesting it seemed a natural choice for a final chapter. First we will do some very easy things by looking at patterns of dots. Then we are going to use some of the material we have already learned to develop a complex of results centering on the Euclidean algorithm, and we are going to introduce methods for efficiently computing various things. Among other things we will be able to find all solutions of $ax+by=c$ where $a,b,c\in\mathbb{Z}$. This sort of equation shows up in puzzles like in how many ways can you use dimes and quarters to get \$2.85?

After these elementary things we will dig a little deeper. Naturally as the results get deeper we will not be able to prove everything, but even so we will be able to understand most things easily enough. I have made a selection of material which I think is very interesting and sometimes rather surprising. Most of this material is proved in the better grade of number theory textbooks and my hope is that some of you will find the material inspiring and wish to pursue it further.

The most basic relation in number theory is divisiblity.

Def Let a and $b\neq 0$ be integers. Then b divides a (in symbols $b|a$) if and only if there exists an integer c so that $a=bc$.

6.2 Dot patterns

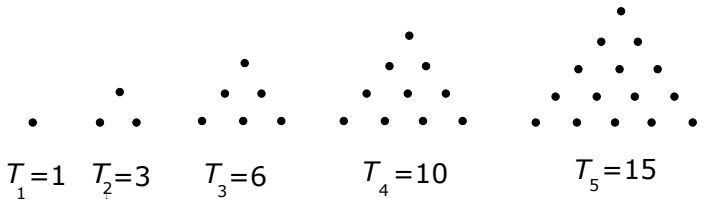
Some of the oldest results in number theory relate to numbers connected with patterns of dots. The were investigated by the Pythagorean Brotherhood, a mathematically oriented secret society founded by Pythagoras who settled in The Greek colony of Crotōn[1] in the instep of the South Italian boot about 525 BCE. Eventually the Brotherhood took over the city, which later led to a democratic revolution and the Brotherhood was scattered over many Greek cities where they spread their mathematical knowledge far and wide. Two of the tenets of the Brotherhood were that Nature could be understood through mathematics and that mathematics should be developed by logical deduction from a small number of initial assumptions (called posutlates or axioms). Plato seeded the first of these into his philosophy and Euclid developed the second. Both became bedrock principles of western culture, although it was 2000 years before the first became clear with the beginnings of modern science. Nobody was surprised though, since everybody had always known (since Plato) that it was true, even if nobody knew how to actually do it.

Originally the Brotherhood attempted to base all knowledge on the whole numbers but this was not too successful since the world turned out to be more

[1] The original spelling of Crotōn was Qrotōn, as this was so early that the spelling conventions had not taken their modern form. They continued to use the Q spelling on their coins long after it was obsolete for everything else.

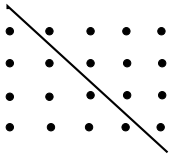
complicated that they thought it was. However, another of their principles, that down deep the world is simple, still guides modern physics. We have much to thank them for, and we owe much to those who guarded and passed on their knowledge through a thousand years of pointless veneration of ancient superstitions. From 400 to 1400 the only new knowledge to appear in Europe was imported from the 'Arabs, who got much of it from India. The new knowledge actually generated in Europe from 400 to 1400 can be balanced on the head of a pin[2].

One of the things that fascinated the Pythagoreans were figurate numbers. Here are the triangular numbers T_n which count triangular patterns of dots.



Triangular Numbers

Naturally we would like a formula for T_n. We will do this as Pythgaoras would have. We draw a rectangle of n by $n+1$ dots (in this case $n=4$) and



Finding formula for T_n

then draw a diagonal line as shown. We see that the rectangle of n by $n+1$ dots is split into two identical T_n's, (in this case $n=4$) so we know that

$$\begin{aligned} T_n + T_n &= n(n+1) \\ 2T_n &= n(n+1) \\ T_n &= \frac{n(n+1)}{2} \end{aligned}$$

Now let's get one other result from these things, a formula which has many uses.

[2]This overstates the case slightly. For example Fibonacci (c.1170-c.1250) actually did do some original work, as well as popularizing the modern forms of digits, which were derived from the 'Arabic forms.

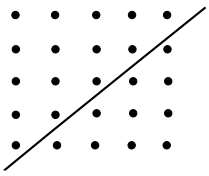
Notice that if we count the dots by rows in the triangles above we see

$$\begin{aligned} T_1 &= 1 \\ T_2 &= 1+2 \\ T_3 &= 1+2+3 \\ T_4 &= 1+2+3+4 \\ \cdots &\quad \cdots \quad \cdots \\ T_n &= 1+2+3+4+\cdots+n \end{aligned}$$

Hence we have proved the useful formula

$$1+2+3+4+\cdots+n = \frac{n(n+1)}{2}$$

We can prove another amusing formula by looking at dot patterns as follows.



Decomposing a Square of Dots

We want to decompose a square of dots (here $n = 5$) into two triagular numbers. As we can see in the illustration, one of the triangles has the same size as the square (here it is 5) and the other has one less (here it is 4). The general case is just the same, and we have shown that

$$n^2 = T_{n-1} + T_n$$

This relates to another theorem of great interest: any natural number n is the sum of three triangular numbers (where $T_0 = 0$ is allowed). In contrast to our theorem, which is simple, this is a very hard theorem to prove. It was first proved by Gauss and recorded in his notebook as $n = \triangle + \triangle + \triangle$.

Many other interesting things can be done with dot patterns and they are still of use in modern studies, but we must move on.

6.3 Euclidean Algorithm and Associated Things

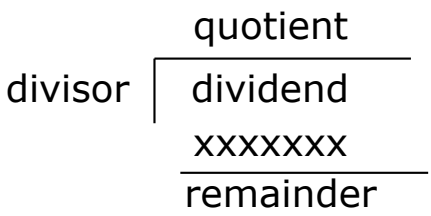
The Euclidean Algorithm for finding the greatest common divisor (also called the greatest common factor) is so named because it is found in Book 7 and Book 10 of Euclid's famous geometry book. Although Euclid's method is interesting we will develop the subject in a (seemingly) different way, more consistent with modern methods. The algorithm is closely connected to continued fractions and matrices and we want to look carefully at these connections.

The importance of the Euclidean Algorithm and the Greatest Common Divisor (GCD) goes far beyond the elementary applications, and we will see some of these. For example, did you ever wonder why $3\frac{1}{7} = \frac{22}{7}$ is a popular approximation for π? We explain this.

6.3.1 Basic Algorithm

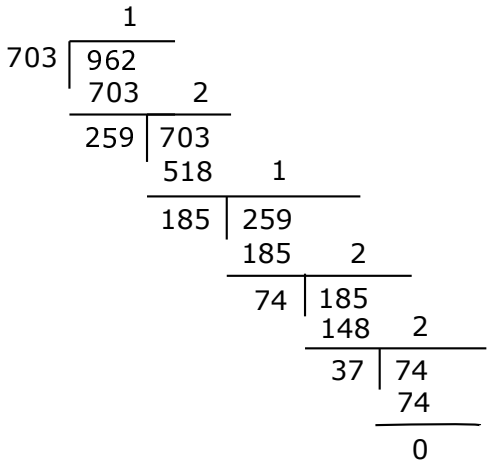
Let us say we wish to find the GCD of 962 and 703. Of course this could be done by factoring both numbers into primes, but for 1000 digit numbers this takes a few million years, so it a less than practical method, although of great theoretical importance. We have a much better method.

First we need some vocabulary to describe what is going on. In case you are a little rusty on the vocabulary this illustration will refresh your memory.



Vocabulary of Division

Now I show you the algorithm for finding the GCD. Examine it. Perhaps you can already see what is happening but if not I discuss it below. The first step is to divide the smaller number 703 into the larger number 962.



Euclidean Algorithm

We want the GCD of 962 and 703. We begin by dividing the smaller number 703 into the larger number 962. The quotient is 1 and the remainder is 259. Now

each further step is made by dividing the remainder into the previous divisor. Since the remainders are always less than the divisors, the numbers get smaller as we go down the algorithm. Eventually a remainder of 0 emerges. The GCD of 962 and 703 is then the previous remainder 37. Indeed $962 = 26 \cdot 37$ and $703 = 19 \cdot 37$. Thus 37 is a common divisor of 962 and 703. But what does it mean to be greatest?

Def $d > 0$ is the Greatest Common Divisor of a and b if and only if
1. d divides a and d divides b (abbreviated $d|a$ and $d|b$).
2. If $e|a$ and $e|b$ then $e|d$.

Thus "greatest" here means "greatest in the sense of division" though this amounts to the same thing as ordinary greatest. We will now show that indeed $37 = GCD(962, 703)$ We already showed 1. To show 2 we must rewrite the Euclidean algorithm in a slightly different form. Remember

DIVIDEND = QUOTIENT TIMES DIVISOR PLUS REMAINDER

We now write this out for each division in the Euclidean Algorithm to get the sequence

$$\begin{aligned} 962 &= 1 \cdot 703 + 259 \\ 703 &= 2 \cdot 259 + 185 \\ 259 &= 1 \cdot 185 + 74 \\ 185 &= 2 \cdot 74 + 37 \\ 74 &= 2 \cdot 37 + 0 \end{aligned}$$

Recall that if $d|r, s$ then $d|r \pm s$. Suppose now that $e|962, 703$. Then the first equation tells us that e divides 259. Knowing that e divides 703 and 259, the second equation tells us that e divides 185. Knowing that e divides 259 and 185, the third equation tells us that e divides 74. Knowing that e divides 185 and 74, the fourth equation tells us that e divides 37. Thus we have shown the second condition from the definition of the GCD.

We have done this with specific numbers but it should be clear that the same thing will work for any other numbers.

Note that our results allow us to reduce fractions in one step.

$$\frac{962}{703} = \frac{26 \cdot 37}{19 \cdot 37} = \frac{26}{19}$$

Although we will not stop to prove it here, this always happens in one step.

There is one more thing that we can squeeze out of the present material. We can run the above sequence of numbers backwards and after an annoying amount of calculation find two numbers r and s so that $37 = r \cdot 961 + s \cdot 703$. We find r and s to have different signs, and I can predict that r will be negative and s will be positive in this case. But following this thread is not productive; there are much better ways to find r and s. I only mention this because occasionally teachers do not present the simpler way and teach students to use this crude method.

6.3.2 Matrix Form of Euclidean Algorithm

We want to find the r and s mentioned at the end of the previous section. To do this it is convenient to construct a matrix representation of the Euclidean algorithm based on the sequence used in the last section. We do this by

$$\begin{aligned}
962 &= 1 \cdot 703 + 259 & \begin{pmatrix} 962 \\ 703 \end{pmatrix} &= \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 703 \\ 259 \end{pmatrix} \\
703 &= 2 \cdot 259 + 185 & \begin{pmatrix} 703 \\ 259 \end{pmatrix} &= \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 259 \\ 185 \end{pmatrix} \\
259 &= 1 \cdot 185 + 74 & \begin{pmatrix} 259 \\ 185 \end{pmatrix} &= \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 185 \\ 74 \end{pmatrix} \\
185 &= 2 \cdot 74 + 37 & \begin{pmatrix} 185 \\ 74 \end{pmatrix} &= \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 74 \\ 37 \end{pmatrix} \\
74 &= 2 \cdot 37 & \begin{pmatrix} 74 \\ 37 \end{pmatrix} &= \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 37 \\ 0 \end{pmatrix}
\end{aligned}$$

Now we put these equations together to get

$$\begin{pmatrix} 962 \\ 703 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 37 \\ 0 \end{pmatrix}$$

The 2×2 matrices all have determinant -1. They will be referred to as (left to right order) Q_1, Q_2, Q_3, Q_4, Q_5 and the numbers in the upper left corner of each, the quotients from the Euclidean Algorithm, as q_1, q_2, q_2, q_4, q_5

We will now start multiplying the matrices getting matrices R_1, R_2, R_2, R_4, R_5 defined by $R_{i+1} = R_i Q_i$ and $R_1 = Q_1$. The R_i are are the matrices at the front of the rows.

$$\begin{aligned}
\begin{pmatrix} 962 \\ 703 \end{pmatrix} &= \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 37 \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 37 \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} 4 & 3 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 37 \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} 11 & 4 \\ 8 & 3 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 37 \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} 26 & 11 \\ 14 & 8 \end{pmatrix} \begin{pmatrix} 37 \\ 0 \end{pmatrix}
\end{aligned}$$

Since $R_5 = Q_1 Q_2 Q_2 Q_4 Q_5$ and the Q_i all have determinant -1 we have $\det R_5 = (-1)^5 = -1$. We now have

$$\begin{pmatrix} 37 \\ 0 \end{pmatrix} = R_5^{-1} \begin{pmatrix} 962 \\ 703 \end{pmatrix} = \begin{pmatrix} 26 & 11 \\ 14 & 8 \end{pmatrix}^{-1} \begin{pmatrix} 962 \\ 703 \end{pmatrix} = (-1)^5 \begin{pmatrix} 8 & -11 \\ -14 & 26 \end{pmatrix} \begin{pmatrix} 962 \\ 703 \end{pmatrix}$$

from which we get, as we desired,

$$37 = (-1)^5(8 \cdot 962 - 11 \cdot 703) = -8 \cdot 962 + 11 \cdot 703$$

Thus $r = -8$ and $s = 11$

Now you are perhaps thinking that is an awful lot of work to come up with r and s, and that is correct; it is a lot of work. Moreover, the apparatus I used here is essential for proofs of certain results, as we will see. However for *computational* purposes a large part of the above calculation is redundant, and can be eliminated, so that in the end we can get r and s very quickly using a much abbreviated form of the above calculation. The trick is to use the following table.

$\cdots$	q_5	q_4	q_3	q_2	q_1		
$\cdots$	A_5	A_4	A_3	A_2	A_1	1	0
$\cdots$	B_5	B_4	B_3	B_2	B_1	0	1

The rules for forming the table (from right to left) are simple. The two rows are done totally separately. To find A_i one multiplies q_i times the previous A, that is A_{i-1} and adds the entry to the right of that, namely A_{i-2}. So it's multiply down and to the right, and then add the next one over. The same thing works for the B row. So

$$\begin{aligned} A_i &= q_i A_{i-1} + A_{i-2} \\ B_i &= q_i B_{i-1} + B_{i-2} \end{aligned}$$

These are called the *recursive relations*. With little practice this becomes quite automatic. For the obsessive, we will set $A_0 = 1$, $A_{-1} = 0$, $B_0 = 0$, and $B_{-1} = 1$. This makes the recursion rule above work for q_1 and q_2. It is not necessary to bother with this if you just remember the verbal description.

Now let's do an example. We'll use the q_i that we got from GCD$(962, 703)$.

2	2	1	2	1		
26	11	4	3	1	1	0
19	8	3	2	1	0	1

The perceptive among you will now see that the pairs of columns in the last two lines of the table are the R_i from earlier in the section, the last vertical pair (on the left) is the reduced fraction for $962/701$, and the next to last are the absolute values of r and s. The reason this works is that the recursive relations are simply counterfeiting the matrix multiplications, which you can see if you go back and look. We simply have dropped out all the redundant parts of the matrix calculations.

To get from the column $\{11, 8\}$ to the r, s, write the determinant of the matrix on the right:

$$\det \begin{pmatrix} 26 & 11 \\ 19 & 8 \end{pmatrix} = 26 \cdot 8 - 19 \cdot 11 = -1$$

You will always get 1 or -1 here. Adjust the sign in the next step. Multiply by -37 to get

$$\begin{aligned}(-37)\cdot 26\cdot 8-(-37)\cdot 19\cdot 11 &= (-37)(-1)\\ -962\cdot 8+703\cdot 11 &= 37\end{aligned}$$

and you have r and s. Be careful which is which.

6.3.3 Two-Variable Linear Diophantine Equations

Diophantes of Alexandria was probably born between 201 CE and 215 CE and an epigram in THE GREEK ANTHOLOGY implies that he died at age 89. The birth date has been the subject of argument for a long time, but this is the best guess right now. Even the man's name is not known for sure, Diophantês (accent on the a) or Diophantos (accent on the first o) are the two guesses. He was probably a Hellenistic Greek but there is a small chance he was a Hellenized Babylonian. He wrote 13 books (we would say chapters) on algebra and number theory, of which we have 1,2,3,8,9,10 in Greek and 4,5,6,7 in 'Arabic. (That's the majority view anyway.) One of the preoccupations of the books is integer solutions of equations, and thus the current usage of the term Diophantine equation is to find solutions that are integers. (Diophantês himself often was content with solutions that are fractions, but we don't let this slight inconsistency worry us.)

Two copies of the books in Greek survived the middle ages, brought to Italy by Greek scholars fleeing the fall of Constantinople to the barbarians in 1453. It can hardly be emphasized how important this was; it gave number theory an important leg up right at the beginning of the rebirth of mathematics in Europe, and number theory and algebraic geometry (the other subject Diophantês treated) have been ahead of the rest of mathematics ever since.

We are going to treat the Diophantine Equation $ax+by=c$, the "Diophantine" meaning that we seek solutions in *Integers*. Naturally a, b, c are integers. This is a nice problem because we can completely solve it, and there are not so many Diophantine problems where this is true. It's also fairly simple.

Before we begin the solution of our problem, we need two easy results.

Lemma 1 Suppose $d=\mathrm{GCD}(a,b)$. Then $\mathrm{GCD}(\frac{a}{d},\frac{b}{d})=1$.
Proof: we know there exist r, s so that

$$ra+sb=d$$

Since $d|a$, $d|b$ we can divide this by d to be

$$r\frac{a}{d}+s\frac{b}{d}=1$$

Trivially $1|\frac{a}{d},\frac{b}{d}$. Now suppose $e|\frac{a}{d},\frac{b}{d}$. Then by the last equation we have $e|1$. Thus $\mathrm{GCD}(\frac{a}{d},\frac{b}{d})=1$.

Lemma 2 Suppose $\text{GCD}(a,b) = 1$ and $a|bc$. Then $a|c$.
Proof: We know there exist r and s so that $ra + sb = 1$. Multiply this by c to get $rac + sbc = c$. Since $a|rac$ and $a|sbc$ (since $a|bc$) we have $a|c$.

So the problem is, find all solutions where x and y are integers of $ax+by = c$, where a, b, c are integers. Let $d = \text{GCD}(a,b)$. If there is a solution then we must have $d|c$ since $d|a, b$. So assume now that $d|c$ and we must show solutions exist and find them all. Since $d|c$ we have $c = gd$ for some integer g. From the previous section we know there exist integers r and s so that

$$ra + sb = d$$

If we multiply the equation by g we get

$$gra + gsb = gd = c$$

so setting $x_0 = gr$ and $y_0 = gs$ we have

$$ax_0 + by_0 = c$$

So we have shown $ax + by = c$ has a solution if and only if $d|c$ where $d = \text{GCD}(a,b)$.

The next job is to find *all* solutions. To this end let (x_1, y_1) be a second solution. Then we have

$$\begin{aligned} ax_1 + by_1 &= c \\ ax_0 + by_0 &= c \end{aligned}$$

and subtracting we have

$$a(x_1 - x_0) + b(y_1 - y_0) = 0$$

and then, dividing by $d = \text{GCD}(a,b)$ we have

$$\begin{aligned} \frac{a}{d}(x_1 - x_0) + \frac{b}{d}(y_1 - y_0) &= 0 \\ \frac{a}{d}(x_1 - x_0) &= -\frac{b}{d}(y_1 - y_0) \end{aligned}$$

We thus see that $a/d \,|\, b/d(y_1 - y_0)$ By lemma 1 we know $\text{GCD}(a/d, b/d) = 1$ and thus by lemma 2 we have

$$\frac{a}{d} \,|\, (y_1 - y_0)$$

so

$$\begin{aligned} y_1 - y_0 &= k\frac{a}{d} \\ y_1 &= y_0 + k\frac{a}{d} \end{aligned}$$

Then

$$\begin{aligned}\frac{a}{d}(x_1 - x_0) &= -\frac{b}{d}(y_1 - y_0) \\ &= -\frac{b}{d}k\frac{a}{d}\end{aligned}$$

Dividing out a/d we have

$$\begin{aligned}x_1 - x_0 &= -k\frac{b}{d} \\ x_1 &= x_0 - k\frac{b}{d}\end{aligned}$$

It is trivial to verify that x_1 and y_1 given by these formulas indeed do satisfy the Diophantine equation $ax + by = c$, so we have the

Theorem The Diophantine equation $ax + by = c$ has solutions if and only if $d|c$ where $d = \text{GCD}(a, b)$. If x_0, y_0 is one solution then all solutions (x_1, y_1) have the form

$$\begin{aligned}x_1 &= x_0 - k\frac{b}{d} \\ y_1 &= y_0 + k\frac{a}{d}\end{aligned} \qquad k \text{ any integer}$$

We now give an example. Find all the solutions of $962x + 703y = 111$.

We found in the last section that $\text{GCD}(962, 703) = 37$ and that

$$(-8) \cdot 962 + 11 \cdot 703 = 37$$

We have $111 = 3 \cdot 37$ so $37|111$ and solutions exist. Multiplying the previous equation by 3 we have

$$962 \cdot (-24) + 703 \cdot 33 = 3 \cdot 37 = 111$$

So $x_0 = -24$ and $y_0 = 33$. Then

$$\frac{a}{d} = \frac{962}{37} = 26 \quad \text{and} \quad \frac{b}{d} = \frac{703}{37} = 19$$

giving us

$$\begin{aligned}x &= x_0 - k\frac{b}{d} = -24 - 19k \\ y &= y_0 + k\frac{a}{d} = 33 + 26k\end{aligned}$$

Now let, for example, $k = -13$ and we get for (x, y)

$$\begin{aligned}x &= -24 - 19(-13) = 223 \\ y &= 33 + 26(-13) = -305\end{aligned}$$

We check: $962 \cdot 223 + 703 \cdot (-305) = 111$.

Now if the ability to find all solutions in integers to a problem like $962x + 703y = 111$ doesn't make you feel powerful, I can't guess what would.

6.3.4 Continued Fractions

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

What you see above is called a continued fraction. These have been studied for centuries and have many important applications. This one terminates at the fourth term but they can also go on forever, like decimals. They are closely associated with the Euclidean Algorithm as we now discuss. This time we will not introduce the notion with an example because I think it is a bit clearer to look at it in general. We will start by restructuring the Euclidean Algorithm as follows.

$$\begin{aligned}
a &= q_1 b + r_1 & \frac{a}{b} &= q_1 + \frac{r_1}{b} \\
b &= q_2 r_1 + r_2 & \frac{b}{r_1} &= q_2 + \frac{r_2}{r_1} \\
r_1 &= q_3 r_2 + r_3 & \frac{r_1}{r_2} &= q_3 + \frac{r_3}{r_2} \\
r_2 &= q_4 r_3 + r_4 & \frac{r_2}{r_3} &= q_4 + \frac{r_4}{r_3} \\
&\vdots & &\vdots \\
r_{k-3} &= q_{k-1} r_{k-2} + r_{k-1} & \frac{r_{k-3}}{r_{k-2}} &= q_{k-1} + \frac{r_{k-1}}{r_{k-2}} \\
r_{k-2} &= q_k r_{k-1} + 0 & \frac{r_{k-2}}{r_{k-1}} &= q_k
\end{aligned}$$

Recall that since the remainder is always smaller than the divisor we have

$$b > r_1 > r_2 > r_3 > \ldots > r_{k-1}$$

which causes the process to terminate because we have a decreasing sequence of positive integers which cannot go on forever. Hence a zero will eventually appear in the position for r_k for some k. Now we can rewrite the sequence on the right as

$$\frac{a}{b} = q_1 + \frac{1}{\frac{b}{r_1}} = q_1 + \cfrac{1}{q_2 + \cfrac{1}{\frac{r_1}{r_2}}} = q_1 + \cfrac{1}{q_2 + \cfrac{1}{q_3 + \cfrac{1}{\frac{r_2}{r_3}}}} = \ldots$$

$$\ldots = q_1 + \cfrac{1}{q_2 + \cfrac{1}{q_3 + \cfrac{1}{\ddots \; q_{k-1} + \cfrac{1}{\frac{r_{k-2}}{r_{k-1}}}}}} = q_1 + \cfrac{1}{q_2 + \cfrac{1}{q_3 + \cfrac{1}{\ddots \; q_{k-1} + \cfrac{1}{q_k}}}}$$

So the Euclidean algorithm can be rewritten in continued fractions, which will be handy. Our interest is in their use in approximation, which we will take up after the following example. We will start with our old friend 962/703. We run the Euclidean Algorithm again and also put it in the corresponding fractions.

$$\begin{aligned} 962 &= 1 \cdot 703 + 259 & \frac{962}{703} &= 1 + \frac{259}{703} \\ 703 &= 2 \cdot 259 + 185 & \frac{703}{259} &= 2 + \frac{185}{259} \\ 259 &= 1 \cdot 185 + 74 & \frac{259}{185} &= 1 + \frac{74}{185} \\ 185 &= 2 \cdot 74 + 37 & \frac{185}{74} &= 2 + \frac{37}{74} \\ 74 &= 2 \cdot 37 + 0 & \frac{74}{37} &= 2 + 0 \end{aligned}$$

Making this into a continuted fraction we have

$$\frac{962}{703} = 1 + \frac{1}{\frac{703}{259}} = 1 + \frac{1}{2 + \frac{1}{\frac{259}{185}}} = 1 + \frac{1}{2 + \frac{1}{1 + \frac{1}{\frac{185}{74}}}} = 1 + \frac{1}{2 + \frac{1}{1 + \frac{1}{2 + \frac{1}{\frac{74}{37}}}}}$$

Since $74/37 = 2$ and there is no remainder in this last division we have the complete continued fraction for 962/703:

$$\frac{962}{703} = 1 + \frac{1}{2 + \frac{1}{1 + \frac{1}{2 + \frac{1}{2}}}}$$

The convergents for a continued fraction are formed by lopping off all but the first k of the terms in the continued fraction. The naive way to calculate a continued fraction is to start at the bottom and work your way up. We will later find a better way. Thus we have:

$$C_1 = 1 \qquad C_2 = 1 + \frac{1}{2} = \frac{3}{2} \qquad C_3 = 1 + \frac{1}{2 + \frac{1}{1}} = \frac{4}{3} \qquad C_4 = 1 + \frac{1}{2 + \frac{1}{1 + \frac{1}{2}}} = \frac{11}{8}$$

and finally

$$\frac{962}{703} = C_5 = 1 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{2}}}} = \frac{26}{19}$$

Recall that

$$\frac{962}{703} = \frac{26 \cdot 37}{19 \cdot 37} = \frac{26}{19}$$

so the continued fraction always gives you the reduced form of the fraction. Now you might think the C_i numbers look familiar. Recall our table

2	2	1	2	1		
26	11	4	3	1	1	0
19	8	3	2	1	0	1

and we see that the table gives us the convergents. This is very handy. Moreover, the convergents are good approximations to 962/703 in the sense that 4/3 is the best approximation to 962/703 with denominator less than 8 and 11/8 is the best approximation to 962/703 with denominator less than 19 and with denominator 19 we get the exact value 26/19.

This is pretty much the whole story for rational numbers. The standard notation for the continued fraction is $[1, 2, 1, 2, 2]$. Each rational number has two continued fractions. We have been finding the "official" variant. The second one subtracts 1 from the last element and puts an extra 1 at the end. Thus

$$[1, 2, 1, 2, 2] \qquad \text{and} \qquad [1, 2, 1, 2, 1, 1]$$

This is not very important. You may recall that a rational number whose decimal terminated also had two sequences: $\frac{1}{2} = .5 = .499999\ldots$. Is this a coincidence or is there meaning to it? Nobody knows.

Rational Approximation of Irrationals

Our next subject is to use the continued fraction to get a rational approximation to an *irrational* number. The continued fraction for π is

$$\pi = [3, 7, 15, 1, 292, 1, 1, 1, 2, 1, 3\ldots]$$

This is an abbreviation for

$$\pi = 3 + \frac{1}{7.06251\ldots} = 3 + \cfrac{1}{7 + \cfrac{1}{15.9966\ldots}} = 3 + \cfrac{1}{7 + \cfrac{1}{15 + \cfrac{1}{1.00342\ldots}}}$$

$$= 3 + \cfrac{1}{7 + \cfrac{1}{15 + \cfrac{1}{1 + \cfrac{1}{292 + \cfrac{1}{1 + \cfrac{1}{\ddots}}}}}}$$

How did I get this terms? We will discuss that in a moment. First let's look at the convergents which I get from the table:

1	1	292	1	15	7	3		
208341	104348	103993	355	333	22	3	1	0
66317	33215	33102	113	106	7	1	0	1

Now we want to look at how good the estimates are. How good are they guaranteed to be? Well if the denominator is, for example 100, and the number to be approximated is 207/123 then if we take the nearest fraction with denominator 100 to $207/123 = 1.68293\ldots$ the closest fraction is 168/100 and the error in the approximation is about .003. What is the *worst* it could be. Well if the number were halfway between 168/100 and 169/100 then the error would be $.005 = 1/100 \div 2$ so the biggest possible error is .005. In general, if the denominator is n then the largest error for the closest fraction m/n is $1/m \div 2 = \frac{1}{2m}$.

Now let's see how good the approximations from the table are. Let's take 355/113. The worst the error could be (assuming the fraction is the best with denominator 113) is half of 1/113 which is .0044.... The actual error is

$$\pi - 355/113 = .000000266764\ldots$$

which is fantastically better than we expected from the analysis above. (If you used 355/113 instead of π to compute the circumference of the earth the error would be a little over 3 meters.) The reason is that the continued fraction algorithm gets to choose which denominator does the best job, and this gives us a much better approximation. In fact, a deeper analysis shows that the error is less than $1/(113)^2 = 0000783147\ldots$ but this estimate is crude and the actual error will usually be quite a bit less.

There is more to the story. One can prove that to get a better fraction than 355/113 you must go to 103993/33102. Put another way, no fraction with denominator 33101 or less will give you a better approximation than 355/113. Thus the table gives you the best rational estimates as you go along.

Another point of interest is the 292 in the continued fraction for π. If you think about it for a while and look at the continuted fraction, you can see that the 1/292 is a very small change in the previous term which is 1. We can interpret this to mean that at this point π was trying very hard to be the rational number 355/113, since that is what we would get if we chopped the

continued fraction off at the 1 just before the 292. A large number as a term in a continuted fraction means this, intuitively, and thus we could spectulate that the continued fraction with the smallest numbers, that is $[1, 1, 1, 1, 1, \ldots]$, would give a number as far from being a rational number as possible. This is in fact true, but would take a small book to prove. The general area here is called *Diophantine Approximation*.

One might ask if interesting irrational numbers have interesting continued fractions. Certainly this is not true for π. However, $e = 2.7181818\ldots$ does have an interesting continued fraction $[2, 1, 2, 1, 1, 4, 1, 1, 6, 1, 1, 8, 1, 1, 10, \ldots]$, or, written out

$$e = 2 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{4 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{6 + \cfrac{1}{1 + \cfrac{1}{\ddots}}}}}}}}}}$$

Since the terms are predictable, one can get rational approximations to e as accurate as one likes.

Computing the Terms in Continued Fractions

How did I get the numbers $[3, 7, 15, 1, 292, 1, 1, \ldots]$ for the continued fraction for π. One can do this in a computer algebra program like Mathematica or Matlab or one can do it on a calculator. The basic process is the same. On a calculator one does the following
pi (or 3.1415926535) store x
1/(x- integer(x)) store x
and then write down the integer part of what you see in the window and hit the enter key again. You'll have to figure out how your calculator stores numbers and how it takes the integer part of a number (which I symbolized by integer(x)). You should be able to get about 8 terms of the continued fraction before round off error destroys the accuracy.

Some computer algebra programs, for example Mathematica, have continued fraction algorithms built into them, and this has the advantage that some protections are built in to help with the roundoff error problem so the results have greater accuracy. They are also easy to use. Mathematica is expensive but for the mathematically inclined probably worth it. Explain to your parents how

important it is for your schoolwork. Matlab is a cheaper alternative and many people prefer it since they claim it is easier to use. Since I've never taken the time to learn it well, I can't give an informed opinion about this.

You can use the programming language PYTHON to find continued fraction terms for maybe 8 terms. You have to download PYTHON (it's free) to your computer and then either find a continued fraction program for PYTHON on the internet or use the following very crude program.

```
CF = [ ]
x = float(input("Decimal whose continued fraction you want "))
for i in range(0, 13):
    y = int(x)
    CF.append(y)
    if x-int(x):
        x = 1/(x-int(x))
    else:
        break
print(CF)
input("Input any letter ")
```

Input to this program must be a decimal real number, like 3.88822; do not try to enter a fraction. Float in the second line converts character data from the input into the decimal number you put in. To use this you will have to learn a little about PYTHON but everyone needs to know at least one programming language and there is nothing easier than PYTHON. If you have a MAC you already have PYTHON. If you have something else you can download Python. (Type download Python into your browser window for instructions.) In either case do a tutorial which will be fun. Python is very useful, usually fun, and does not require a compilation step as C++ and Java do. There is a helpful IDE (Integrated Development Environment) called Spyder which makes running it a little easier. Again, do the tutorial.

The program can be run (usually) by just clicking the icon. However, the window it works in will disappear before you can read it. I stop this by the last line in the program which makes the window stay open until you put in a letter, and gives you a chance to read the output. For example if you put in 8.3232323232 the output will be $[8, 3, 10, 1, 1, 1, 3156485, 1, 63, 164, 1, 5, 1]$. So the reliable part of the continued fraction will be $8, 3, 10, 1, 1, 1$. Then you type a, enter, and the window will disappear.

Whether Calculator or computing machine, all devices are subject to round off error. In this program round off error is often indicated by a large number in the output. The calculation is probably accurate till just before the large number, which is due to the round off error. You will have to practice for a while before you get good with either calculator or computer. Keep in mind that Mathematica has a built in continued fraction command so that you get many many terms before round off error screws it up.

Repeating Continued Fractions

Previously we noted that the continued fraction $[1,1,1,1,1,1,\ldots]$ might be a difficult number to approximate. But what number *is* it? We will answer this question. Let's call it ϕ pronounced fee. (Some people pronounce it fy but this is less cultured. Note there are differences between math pronunciation and physics pronunciation for some Greek letters.) Anyway we'll let

$$\phi = 1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{\ddots}}}}}$$

Then

$$\phi - 1 = \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{\ddots}}}}}$$

and

$$\begin{aligned} \frac{1}{\phi - 1} &= 1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{\ddots}}}} \\ &= \phi \end{aligned}$$

Thus

$$\begin{aligned} 1 &= \phi(\phi - 1) = \phi^2 - \phi \\ \phi^2 - \phi - 1 &= 0 \\ \phi &= \frac{1 \pm \sqrt{5}}{2} \qquad \text{using the quadratic formula} \end{aligned}$$

Now $1 - \sqrt{5} < 0$ and the continuted fraction will give a positive outcome so we know

$$\phi = \frac{1 + \sqrt{5}}{2}$$

This is a famous number although you may not have met it before. It is called the *Golden Section* or *Golden Ratio* and given a choice of rectangles with various

ratios between the sides most people will select the one with ratio ϕ as the most aesthetically pleasing. The ancient Greeks knew this and designed buildings, for example the Parthenon, so that the rectangle at the front was the designed with ratio the golden section. Most modern copies of the Parthenon preserve this tradition. The decimal value is 1.61803, but let us use the exact value to start calculating the continued fraction.

$$\begin{aligned}\phi &= \frac{1+\sqrt{5}}{2} = 1+\frac{-1+\sqrt{5}}{2} \quad \text{separating off the integer part} \\ &= 1+\frac{-1+\sqrt{5}}{2}\frac{-1-\sqrt{5}}{-1-\sqrt{5}} = 1+\frac{1-5}{2(-1-\sqrt{5})} \\ &= 1+\frac{2}{1+\sqrt{5}} = 1+\frac{1}{\phi}\end{aligned}$$

so

$$\phi = 1+\frac{1}{\phi} = 1+\frac{1}{1+\frac{1}{\phi}} = 1+\frac{1}{1+\frac{1}{1+\frac{1}{\phi}}} = 1+\frac{1}{1+\frac{1}{1+\frac{1}{1+\frac{1}{\phi}}}}$$

and that's how you make a continued fraction without a calculator. This same process will produce the continued fraction for any quadratic irrationality, for example

$$\frac{14+\sqrt{31}}{29}$$

will have a repeating continued fraction although with a longer period than ϕ and any repeating continuted fraction will be a number of this form.

There are many more fish in this lake but we must move on to other subjects.

6.4 Prime Numbers

6.4.1 Introduction

What it means to be a prime depends on the system one is in. In this section we will consider three rings, $\mathbb{Z}$, $\mathbb{Z}[i]$, and $\mathbb{Z}[\sqrt{-5}]$. The first two will be very similar, but in the third there will be some oddities. We will first discuss units, which interfere with our desire to have numbers factor uniquely into primes. Then we will define primes, and the official definition is different from the one you are used to, and we will discuss the difference. Then we will give a short introduction to the modern way of looking at these things.

The rings we will be discussing are all subsets of the complex numbers $\mathbb{C}$, which we mention because it means that the cancellation law works, since it works in $\mathbb{C}$ or in any field for that matter. Thus in our rings if $ab = ac$ then $b = c$ (provided of course that $a \neq 0$, which we will usually pass over without mention.)

6.4.2 Units

Units are divisors of 1, or to put it another way they are elements of the ring which have inverses. If the ring is a field then every non-zero element is a unit. Thus the concept is not too useful in a field. In $\mathbb{Z}$ there are only two, 1 and -1. This makes it possible to mostly ignore them which is what we have been doing. Here is the problem. We have two ways to factor 6 into primes, namely

$$6 = 2 \cdot 3 = (-2) \cdot (-3)$$

These are not different enough to constitute a real problem, but even in $\mathbb{Z}[i]$ it begins to become more of a problem. Is

$$5 = (1 + 2i)(1 - 2i) = (2 + i)(2 - 1)$$

one factorization of 5 or two. It is one, because the units of $\mathbb{Z}[i]$ are $\{1, i, -1, -i\}$ and so

$$5 = (1 + 2i)(1 - 2i) = (1 + 2i)1(1 - 2i) = (1 + 2i)(-i)(i)(1 - 2i) = (2 - i)(2 + i)$$

so the seemingly two different factorizations are in fact a single factorization differing only by units. We call numbers a, b that differ by units associates. Let's make official definitions now.

Def Let R be a ring. An element u in R is a *unit* if and only if there is a v in R for which $uv = 1$, which is the same thing as saying that u has an inverse.

Def Let R be a ring. Two elements a and b are *associates* if and only if there is a unit u so that $b = ua$. We write this as $a \sim b$.

Since $\mathbb{Z}$ has only two units each element has two associates, itself and its negative.

Since $\mathbb{Z}[i]$ has four units $\{1, i, -1, -i\}$ each element has four associates. For example

$$2 + i \sim i(2 + i) = -1 + 2i \sim (-1)(2 + i) = -2 - i \sim (-i)(2 + i) = 1 - 2i$$

Notice the important fact that the conjugate $2 - i$ of $2 + i$ is *not* among the associates. There is one important exception:

$$1 + i \sim (-i)(1 + i) = 1 - i \quad \text{so} \quad \overline{1 + i} \sim 1 + i$$

This has the interesting consequence that for many purposes 2 is a square in $\mathbb{Z}[i]$

$$2 = (1 + i)(1 - i) = (1 + i)(-i)(1 + i) = -i(1 + i)^2 \sim (1 + i)^2$$

We say "2 is a square in $\mathbb{Z}[i]$ up to a unit". And we remark that just as 2 is the even prime in $\mathbb{Z}$, $1 + i$ is the even prime in $\mathbb{Z}[i]$ and 2 is not a prime at all in $\mathbb{Z}$ since it factors.

A useful theorem which is easy to prove is

Theorem If $a|b$ and $b|a$ then $a \sim b$ (a and b are associates).

Proof Suppose $a|b$ and $b|a$. Then there are elements c and d of the ring for which $b = ca$ and $a = db$ From this we get

$$\begin{aligned} a &= db = dca \\ 1 &= dc \end{aligned}$$

Hence c and d are units and thus $a \sim b$.

6.4.3 The Definition of Prime

Before I give you the definition of prime we must discuss some philoosophy. From the point of view of logic one can make definitions any way one likes. But mathematics is not just logic; it has a beauty that goes far beyond mere logic. Otherwise why would so many people do it and love it. It's true that mathematics solves real life problems but that, for many mathematicians (certainly not all though), is icing on the cake. Many mathematicians love mathematics for itself. And for that reason they want it to run smoothly.

Now a way to make it *not* run smoothly is to define a term at one level and then when we generalize the game to cover more cases we say "OK, wait, we used to call this a widget but now in more general circumstances we call something else a widget." This is surely a recipe for confusion. The proper way to proceed is to define a concept so that when we proceed to a generalization the old definition continues to work. Sometimes this means that a centuries old definition needs to be modified. There is often considerable resistance to this among those used to the old ways, but the new definitions always win because mathematicians are compulsively consistent. Funeral by funeral the old way of talking disappears.

This problem is especially acute when the concept is part of elementary mathematics because it is then more solidly embedded in the culture, and the training of elementary teachers is only very slowly influenced by what happens in the world of higher mathematics.

And this is the case with the concept of prime number.

We have known for over a century that the old definition of prime, a number whose only factors are itself and 1, will not generalize to more complex situations. That is, in the more general situations the theorems that were true for prime numbers, so defined, become false. We would like to set the system up so that theorems about prime numbers remain true in more general situations, and this requires a change in the definition of prime. Naturally we want the primes in $\mathbb{Z}$ to still be the primes in $\mathbb{Z}$ with the new definition, that is $\{2, 3, 5, 7, 11, 13, 17, \ldots\}$, and we'll take care of that.

In more general circumstances what we think of as "prime" splits into two separate concepts which do not necessarily coincide. There are the primes and the indecomposables. Now it's time for the definitions.

Def In a ring R, an element a is indecomposable if and only if $a \neq 0$ and $a \neq$ unit and

if $a = bc$ then one of b or c is a unit
(and then a is an associate of the other).

Def In a ring R, an element p is prime if and only if $p \neq 0$ and $p \neq$ unit and
if p divides a product ab then p divides a or p divides b.
In symbols $p \neq 0$ and $p \neq$ unit and
if $p|ab$ then $p|a$ or $p|b$.

It is easy to prove that all primes are indecomposibles, and this is true in any commutative ring without divisors of 0. We assume that all our rings satisfy these two conditions from here on out.

Theorem All primes are indecomposibles
Proof Let p be a prime and assume $p = ab$. Since $p|p$ and p is prime, we must have $p|a$ or $p|b$. Suppose $p|a$. Then since $p = ab$ we have $a|p$. Then we know $p \sim a$ and thus b is a unit. Similarly if $p|b$ we have a is a unit. Thus, by the definition, p is an indecomposible.

It is not true *in general* that all indecomposibles are prime but we can prove that this is the case for $\mathbb{Z}$ and thus the two concepts select the same numbers in $\mathbb{Z}$. We mentioned earlier that we really wanted this to be the case.

Theorem In the ring $\mathbb{Z}$ all indecomposibles are primes.
Proof Let a be an integer and an indecomposible. We wish to show it is prime. So suppose $a|bc$. We must show a divides b or a divides c. Suppose a does *not* divide b. Then since a is indecomposible, the $\mathrm{GCD}(a, b) = 1$. Hence we know there exists r and s so that $ra + sb = 1$. Multiplying by c we have

$$rac + sbc = c$$

Since $a|rac$ and $a|sbc$ we have $a|c$ as required.

The key to this proof is the existence of r and s. We know r and s will exist if we have a Euclidean Algorithm in our ring. We have it in $\mathbb{Z}$ and it is not very hard to show there is also a Euclidean Algorithm in $\mathbb{Z}[i]$. The tricky part is that for the Euclidean Algorithm to work the remainders must be "smaller" than the divisors and that implies we must have some sort of *size* function for the ring. In the case of $\mathbb{Z}[i]$ the size of $z = a + bi$ is $N(z) = z\overline{z} = a^2 + b^2$. Using the geometric picture of $\mathbb{Z}[i]$ as the integer points in $\mathbb{R}^2$ it is not very hard to show the remainder has smaller size than the divisor, but there are tricky aspects and I wouldn't like to ad-lib this proof in front of a class. Best to have prepared notes. Rings with this property are called *Euclidean Rings* and they are rather scarce, although they are among the most popular rings. The previous proof would show that if a ring has a Euclidean Algorithm then the concepts of prime and indecomposable coincide. However such rings are scarce, so mostly prime and indecomposable do *not* coincide. There are also a few rings for which prime and indecomposable coincide but there in no Euclidean Algorithm. All of these things are extensively handled in Hasse.

A ring that definitely positively *does not* have a Euclidean Algorithm is $\mathbb{Z}[\sqrt{-5}] = \mathbb{Z}[i\sqrt{5}]$. Hence in this ring we expect prime and indecomposable might *not* coincide. This is the classical example of this phenomenon.

We review for a moment the definition of integer in a ring which contains the integers $\mathbb{Z}$. An element x is in then integral (which means "is an integer") if and only if it is a root of a polynomial $p(x) = x^n + b_{n-1}x^{n-1} + \ldots + b_1 x + b_0$ where the b_i are all ordinary integers from $\mathbb{Z}$. For our examples $n = 2$. So, for example, $z = 4 + \sqrt{-5}$ is an integer in the ring $\mathbb{Z}[\sqrt{-5}]$ because it is a root of $(z-4)^2 + 5 = z^2 - 8z + 21 = 0$. In $\mathbb{Z}[\sqrt{-1}] = \mathbb{Z}[i]$ the integers are $a + bi$ where $a, b \in \mathbb{Z}$. These last are called the Gaussian Integers. Factorization works in these rings of integers like it works in $\mathbb{Z}$.

Before we can proceed to this example we need to define the Norm of an element $z = a + b\sqrt{-5}$ in $\mathbb{Z}[\sqrt{-5}]$, where of course a, b are integers. As we have seen before,

Def In the ring $\mathbb{Z}[\sqrt{-5}]$ we define for $z = a + b\sqrt{-5}$
$N(z) = z\overline{z} = N(a + b\sqrt{-5}) = (a + b\sqrt{-5})(a - b\sqrt{-5}) = a^2 + 5b^2$.

Note that $N(z) \in \mathbb{Z}$. As we have done before, we can easily prove $N(z_1 z_2) = N(z_1)N(z_2)$.

Now let us consider $z = 4 + \sqrt{-5}$. This z is an integer since it is a root of $z^2 - 8z + 21 = 0$ with leading coefficient 1. We then have

$$(4 + \sqrt{-5})(4 - \sqrt{-5}) = 21 = 3 \cdot 7$$

This shows that $4 + \sqrt{-5}$ is not prime, because if it were prime it would have to divide either 3 or 7. But

$$\frac{3}{4 + \sqrt{-5}} = \frac{3}{4 + \sqrt{-5}} \frac{4 - \sqrt{-5}}{4 - \sqrt{-5}} = \frac{3(4 - \sqrt{-5})}{21} = \frac{4 - \sqrt{-5}}{7}$$

This last quantity has the equation $7z^2 - 8z + 3 = 0$ with leading coefficient 7 and hence is *not* an integer. A similar calculation shows that $7/(4 + \sqrt{-5})$ is also not an integer. So $4 + \sqrt{-5}$ does not divide either 3 or 7 and hence cannot be prime.

However, $4 + \sqrt{-5}$ is decomposable. For suppose that

$$\begin{aligned} 4 + \sqrt{-5} &= (a + b\sqrt{-5})(c + d\sqrt{-5}) \\ N(4 + \sqrt{-5}) &= N(a + b\sqrt{-5})N(c + d\sqrt{-5}) \\ 21 &= (a^2 + 5b^2)(c^2 + 5d^2) \end{aligned}$$

Since the right side is positive, the only possible factorizations of 21 that come into play here are $1 \cdot 21$, $3 \cdot 7$, $7 \cdot 3$, $21 \cdot 1$. Now $a^2 + 5b^2$ cannot be either 3 or 7 when $a, b \in \mathbb{Z}$, so we must have $a^2 + 5b^2 = 1$ or $a^2 + 5b^2 = 21$ in which case $c^2 + 5d^2 = 1$ In the first case we must have $a = \pm 1$ and $b = 0$ and $a + b\sqrt{-5}$ is a unit. In the second case $c = \pm 1$ and $d = 0$ so $c + d\sqrt{-5}$ is a unit. So in any factorization of $4 + \sqrt{-5}$, one of the factors is a unit. This is the definition of indecomposable.

So we have shown that in the ring $\mathbb{Z}[\sqrt{-5}]$ the element $4 + \sqrt{-5}$ is an indecomposable which is *not* a prime, which shows that in general the two concepts indecomposable and prime are different.

It is easy to show, now that you see how the equipment works, that neither 3 nor 7 is a prime in $\mathbb{Z}[\sqrt{-5}]$ for 3 divides $(4+\sqrt{-5})(4-\sqrt{-5}) = 21$ but divides neither factor. We saw above the both 3 and 7 are indecomposable. We have shown that in the ring $\mathbb{Z}[\sqrt{-5}]$ the number 21 cannot be factored into primes, and that it can be factored into indecomposables in two essentially different ways. This is a bit traumatic.

There is a famous story (which is false) that the mathematician Kummer thought that he had a proof of Fermat's last theorem (that $x^n + y^n = z^n$ has no solutions with integer x, y, z for $n > 2$ except the trivial ones) and brought it to his advisor Dirichlet. Dirichlet pointed out that the proof depended on unique factorization into primes in algebraic number fields, which we have just seen doesn't always work. Kummer went home and worked for forty years to fix the little deficiency, but was never able to do it. However, in the process he invented what he called ideal numbers (we call them ideals or divisors today) to try to fix the problem. In a later section we will briefly describe these. (This story always had suspicious aspects, and eventually it was traced to a mathematician named Study, who hated Kummer.)

6.5 The Fundamental Theorem of Arithmetic

The Fundamental Theorem of Arithmetic states that all integers in $\mathbb{Z}$ factor into a unit and primes and that the factorization is unique up to units. We will give a more accurate statement of the theorem when we are ready to prove it. However, it turns out that with very little extra work we can get the same theorem in a number of other rings, most importantly the Gaussian Integers $\mathbb{Z}[i]$. The critical requirement is that the ring have a size function, or Euclidean function, for use with the Euclidean algorithm.

6.5.1 Euclidean Functions

A Euclidean Function for a commutative ring R with no divisors of 0 (often called a Euclidean domain) is a function $S : R - \{0\} \to \mathbb{N} = \{1, 2, 3, 4, 5, 6, \ldots\}$ satisfying two properties.

1. For all a, b in the ring R with $b \neq 0$ there exist q and r so that

$$a = qb + r \qquad \text{with } r = 0 \text{ or } S(r) < S(b)$$

2. If a, b are non-zero elements of R then $S(a) \leq S(ab)$

You want to think of $S(a)$ as giving the size of a. There are several important examples of the concept.

1. $\mathbb{Z}$ where $S(a) = |a|$. This is the classical Euclidean algorithm.

2. $\mathbb{Z}[i]$ where $S(a+bi) = N(a+bi) = (a+bi)(a-bi) = a^2 + b^2$. The Gaussian Integers

3. $K[x]$, the ring of polynomials $a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$ where $n \geq 1$ and the a_i are in a field K and where $a_n \neq 0$. $S(a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0) = n$; the size is the degree of the polynomial.

An important *nonexample* is the ring $K[x, y]$ of polynomials with two variables with coeficients in a field K. There is no usable Euclidean function in $K[x, y]$. This partially accounts for the great differences between $K[x]$ and $K[x, y]$.

Def A Euclidean Domain is a commutative ring without divisors of 0 that has a Euclidean Function S.

The various methods that we used based on the Euclidean Algorithm, like the tables and continued fractions, work here as well, and provide the GCD and the r and s which we used for integers. If you recall, when showing that all irreducibles are primes the critical piece of the proof was that we could find r and s for which $ra + sb = \text{GCD}(a, b)$ and this works in any Euclidean Ring. Hence by repeating the proof given there in the section on primes one has

Theorem In a Euclidean Ring the primes and irreducibles coincide.

6.5.2 Units Again

The concept of Euclidean Domain which we introduced above has a slightly simpler variant that will make our life much easier. What we will do is replace item 2 in the definition by a stronger requirement: For nonzero a and b

$$S(ab) = S(a)S(b)$$

If a ring satisfies this requirement it will satisfy the item 2 requirement but this stronger definition allows us to prove the useful $S(1) = 1$. A Euclidean function that satisfies this instead of item 2 is called a multiplicative Euclidean function. All our examples have multiplicative Euclidean functions $S(a)$.

We will now, assuming a multiplicative Euclidean function (MEF), prove a number of things we will find useful.

In a Euclidean Domain with a MEF, we have S(1)=1.
We have $S(1) = S(1 \cdot 1) = S(1) \cdot S(1)$. But by definition $S(1) \neq 0$ Hence dividing out $S(1)$ we have $1 = S(1)$.

In a Euclidean Domain with a MEF, we have u is a unit if and only if $S(u) = 1$. If u is a unit then there is a v with uv=1. Then $1 = S(1) = S(uv) = S(u)S(v)$. Since $S(u)$ and $S(v)$ are integers, $S(u) = 1$. On the other hand, if $S(u) = 1$ then $1 = qu + r$ with $r = 0$ or $S(r) < S(u)$. The second alternative is impossible since $S(u) = 1$. Hence $r = 0$, $1 = qu$ and u is a unit.

A unit u in an Integral Domain with a MEF divides every element of the ring. Since u is a unit, it has an inverse v; $uv = 1$. For any a, $a = a \cdot 1 = auv$ so $u|a$.

If a in an Integral Domain with a MEF divides a unit u then a is a unit.
Let $a|u$. Since $u|a$ we have $a \sim u$ so a is a unit.

6.5.3 The Fundamental Theorem

We are now in a position to prove the fundamental Theorem that any element in a Euclidean Domain with a multiplicative Euclidean function(MEF) factors into irreducibles (or primes, same thing) in essentially exactly one way. We will do this in two steps. First we will prove existence: that any non-zero non-unit factors into irreducibles. Second we wil prove the uniqueness (up to units) of the factorization. The first is essentially easy but there is one small hassle. Also, it is true for a wide variety of rings. The second is more sophisticated but we have layed the groundwork pretty well.

Existence

In the existence section we will talk in terms of irreducibles since it is their properties which allow us to find a decomposition. We first define a term solely for use in this section.

Def We will say we have *factored* an element a of our Euclidean Domain R when we have found b and c which are non-units such that $a = bc$

Lemma If a factors so that $a = bc$ with neither b nor c units then $S(b)$, $S(c) < S(a)$.
Proof The fact that b and c are not units means that $S(b), S(c) \geq 2$. Since $S(a) = S(b)S(c)$ we then have $S(b)$, $S(c) < S(a)$

Lemma In a Euclidean Ring R (with a MEF) every non-zero non-unit element factors into indecomposibles.
Proof Let a be an non-zero non-unit element of R. If a is an indecomposible we are done. If a is not an indecomposible then there exist b,c which are not units and with $a = bc$. If b and c are irreducible we are done. If not, at least one, say b is not irreducible, so $b = de$, neither d nor e a unit. Thus $a = bde$. We can continue doing this until finally $a = e_1e_2e_3 \cdots e_{n-1}e_n$. How do we know the process does not go on forever? The previous lemma says that every time we break a factor not an indecomposible the S steps down. We cannot step down forever from the initial value $S(a)$ since we dealing with positive integers. There cannot be infinitely many factors since $S(a) = S(e_1)S(e_2)S(e_3) \cdots S(e_{n-1})S(e_n)$. Hence the process of factoring over and over must eventually stop.

The idea behind the last proof is "keep factoring until you can't factor any more" We need the MEF to keep the process from going on forever, which it could do in a more general situation. Consider the ring Ω of algebraic integers. In this ring

$$2 = \sqrt{2}\sqrt{2} = \sqrt{2}\sqrt[4]{2}\sqrt[4]{2} = \sqrt{2}\sqrt[4]{2}\sqrt[8]{2}\sqrt[8]{2} = \sqrt{2}\sqrt[4]{2}\sqrt[8]{2}\sqrt[16]{2}\sqrt[16]{2} = \ldots$$

We can clearly keep going in this vein forever. Thus, there are no indecomposibles in this ring.

The MEF is an adequate but slightly clumsy way to handle this problem. In more advanced treatments there is a concept called Noetherian[3] which does the

[3]Named after Emmy Noether, (1882-1935), the most important female mathematician of

same job an does it more elegantly, but would take too much time to develop here.

Uniqueness

In the uniqueness proof it is the prime quality that takes the center stage, as did the indecomposible property in the previous subsection. We will need the following lemma which says primes do not divide each other unless they are associates.

Lemma In the Euclidean Ring R, if p and if q are primes and $p|q$ then $p \sim q$.
Proof Let $p|q$. Then $q = pa$ for some a. Since q is prime, $q|p$ or $q|a$. in the first case we $p \sim q$ (p and q are associates). In the second case we have $a = qb$. Thus $q = pa = pqb$ so cancelling q we have $1 = pb$. Hence p is a unit which is a contradiction to the definition of prime, so this case is not possible.

It would also be useful to have the following lemma so it's proof will not clog up the proof of the main lemma.

Lemma If a prime divides a product $a_1a_2 \dots a_{s-1}a_s$ then it divides one of the factors.
Proof Let the prime p divide $a_1a_2 \dots a_{s-1}a_s$. Then since p is prime it will divide a_1 or $a_2 \dots a_{s-1}a_s$. If it divides a_1 we are done. Otherwise it divides $a_2a_3 \dots a_{s-1}a_s$. Then since p is prime it will divide a_2 or $a_3 \dots a_{s-1}a_s$. If $p|a_2$ we are done. Otherwise $p|a_3 \dots a_{s-1}a_s$. Continuing in this fashion we eventually are either done or $p|a_{s-1}a_s$. In that case $p|a_{s-1}$ or $p|a_s$ and we have proved the Lemma.

The phrase "Continuing in this fashion" conceals what would be more obsessively done with the technique of mathematical induction. Fortunately the proof is clear as it stands.

Now we come to the main result which I call a lemma since it will be put together with the lemma of the the previous subsection to get the Fundamental Theorem of Arithemetic.

Lemma In a Euclidean ring with a MEF an element can be factored into primes in essentially only one way.
Proof Let the element a have two factorizations into primes:

$$a = p_1p_2p_3 \dots p_{r-1}p_r = q_1q_2q_3 \dots q_{s-1}q_s$$

Without loss of generality we may assume that $r \leq s$ (otherwise rename the primes). Now $p_1|p_1p_2p_3 \dots p_{r-1}p_r$ so $p|q_1q_2q_3 \dots q_{s-1}q_s$. Hence by the lemma $p_1|q_j$ for some j, $1 \leq j \leq s$. Renumber the primes q_i so that q_j becomes q_1. This is just for convenience. Since $p_1|q_1$, our previous lemmas tell us that $p_1 \sim q_1$ and thus $q_1 = u_1p_1$ where u_1 is a unit. We then have

$$p_1p_2p_3 \dots p_{r-1}p_r = u_1p_1q_2q_3 \dots q_{s-1}q_s$$

the 20th Century. Noether's theorem about symmetries and conservations laws is fundamental to all modern physics. Noether is pronounced Nerter as spoken by a Kennedy.

and dividing out p_1 we have

$$p_2 p_3 \dots p_{r-1} p_r = u_1 q_2 q_3 \dots q_{s-1} q_s.$$

Since the prime p_2 cannot divide the unit u_1, we can repeat the process finding p_2 divides q_j for some j, $2 \leq j \leq s$. Renumber the q_i so that q_j becomes q_2 so $q_2 = u_2 p_2$. Divide out p_2 to get

$$p_3 \dots p_{r-1} p_r = u_1 u_2 q_3 \dots q_s$$

Repeat the process over and over (remembering $r \leq s$) until all the p_i have been divided out so we have

$$1 = u_1 u_2 \cdots u_r q_{r+1} \cdots q_s$$

But wait; this means all the q_i, $r+1 \leq i \leq s$ are all units. This is a contradiction; primes cannot be units. Hence q_i, $r+1 \leq i \leq s$ cannot exist, which means r must equal s and the equation becomes

$$1 = u_1 u_2 \cdots u_r$$

Thus the p_i and the q_j are associates of one another (in some order) and the two factorizations are the same except for units.

The Fundamental Theorem of Arithmetic

We now put the lemmas of the last two subsections together to get the

Theorem In a Eucliean Domain with a MEF all non-zero non-units factor into primes in a unique (up to order and units) way.

Proof We have lemmafied the proof so completely that there is nothing left to do. Remember for this kind of ring the indecomposables and the primes are the same elements.

It would be good to remember that we have not only proved this result for $\mathbb{Z}$ but also for $\mathbb{Z}[i]$ and for $K[x]$, the polynomials in one variable over a field K. Thus we know that polynomials with coefficients from $\mathbb{R}$ or $\mathbb{C}$ have unique factorization into indecomposable polynomials. This is even true for the ring of polynomials over the field $\mathbb{Z}$ mod 5 with it's 5 elements. This is a fine example of the power of abstraction. In the 19^{th} Century these would all be proved separately, but because we noticed that they all had MEFs we were able to prove all these cases with one proof. I'll even throw in one more important case. Recall that

$$\omega = \frac{-1 + i\sqrt{3}}{2} \qquad \text{cube root of 1}$$

is a root of $x^2 + x + 1$, so in spite of the denominator 2, ω is an integer. If we add ω to the ring $\mathbb{Z}$ to get $\mathbb{Z}[\omega]$, it turns out that $\mathbb{Z}[\omega]$ also has an MEF. Elements of the ring look like $a + b\omega$ with a and b in $\mathbb{Z}$. Then, recalling that $\omega^2 = \overline{\omega}$ we have $\omega + \omega^2 = \omega + \overline{\omega} = -1$ and $\omega\omega^2 = \omega^3 = 1$ and we can define S by

$$S(a + b\omega) = (a + b\omega)(a + b\overline{\omega}) = a^2 + ab(\omega + \overline{\omega}) + b^2\omega\omega^2 = a^2 - ab + b^2$$

One more comment before we leave this section. Recall we showed that in $\mathbb{Z}[\sqrt{-5}]$ the indecomposables were *not* prime. Hence there is no way to make $\mathbb{Z}[\sqrt{-5}]$ into a Euclidean ring.

6.6 Number Theory and Geometry

This section requires a little more background, specifically a little bit of Analytic Geometry. Problems like these were treated by Diophantês. This example lies in the border region between number theory and algebraic geometry.

We probably remember a few Pythagorean triples of integers, such as

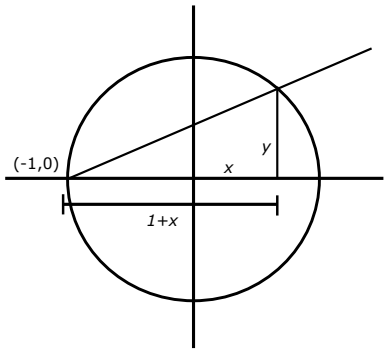
$$3^2 + 4^2 = 5^2 \quad \text{and} \quad 7^2 + 12^2 = 13^2$$

You probably think these are pretty rare but there are infinitely many of them as we are going to show, and indeed we will show you how to find them all.

A critical part of the derivation is the fact that if x_1 and x_2 are the roots of the quadratic equation $ax^2 + bx + c = 0$ then

$$x_1 + x_2 = -\frac{b}{a} \quad \text{and} \quad x_1 \cdot x_2 = \frac{c}{a}$$

You can easily prove this by adding the two roots as given by the quadratic formula or comparing coefficients of $a(x - x_1)(x - x_2)$ with $ax^2 + bx + c$.



Line through $(-1, 0)$ intersects circle at (x, y)

The analytic geometry part of the derivation consists of finding the intersection of a line through the point $(-1, 0)$ with the circle of radius 1 which is given by the equation $x^2 + y^2 = 1$. The equation of the line through $(-1, 0)$ with slope t is found from the the definition of slope as

$$\frac{y}{x+1} = t \quad \text{which simplifies to} \quad y = t(x+1)$$

Substituting this into $x^2 + y^2 = 1$ gives

$$\begin{aligned} x^2 + t^2(x+1)^2 &= 1 \\ (1+t^2)x^2 + 2t^2x + t^2 - 1 &= 0 \end{aligned}$$

The intersection of the line and circle occurs at two points, call their x coordinates x_1 and x. However we know one of these points (say x_1); it is $(-1, 0)$. Hence using the property $x_1 + x = -\frac{b}{a}$ we have

$$\begin{aligned} -1 + x &= -\frac{2t^2}{1+t^2} \\ x &= 1 - \frac{2t^2}{1+t^2} \\ &= \frac{1+t^2}{1+t^2} - \frac{2t^2}{1+t^2} \\ &= \frac{1-t^2}{1+t^2} \end{aligned}$$

To get the y coordinate of the second interesection point we use

$$\begin{aligned} y &= t(x+1) \\ &= t(\frac{1-t^2}{1+t^2} + 1) \\ &= t(\frac{1-t^2}{1+t^2} + \frac{1+t^2}{1+t^2}) \\ &= t(\frac{2}{1+t^2}) \\ &= \frac{2t}{1+t^2} \end{aligned}$$

Every choice of t will give us a point $(1-t^2)/(1+t^2), 2t/(1+t^2)$ on the unit circle, and because the coefficients in the formula are all rational, (x, y) will also be rational if you choose a rational t. The choicest points are in the first quadrant, $x > 0,\ y > 0$, for which we must choose t with $0 < t < 1$. For example

$$\begin{aligned} t = \frac{1}{2} \qquad & x = \frac{\frac{3}{4}}{\frac{5}{4}} = \frac{3}{5}, \quad y = \frac{1}{\frac{5}{4}} = \frac{4}{5} \\ t = \frac{1}{4} \qquad & x = \frac{\frac{15}{16}}{\frac{17}{16}} = \frac{15}{17}, \quad y = \frac{\frac{1}{2}}{\frac{17}{16}} = \frac{8}{17} \end{aligned}$$

Notice that each such point on the unit circle gives us a Pythogorean Triple (integer sided right triangle). The two above give us

$$\begin{aligned} \left(\frac{3}{5}\right)^2 + \left(\frac{4}{5}\right)^2 = 1 \quad &\longrightarrow \quad 3^2 + 4^2 = 5^2 \\ \left(\frac{15}{17}\right)^2 + \left(\frac{8}{17}\right)^2 = 1 \quad &\longrightarrow \quad 15^2 + 8^2 = 17^2 \end{aligned}$$

and you can make as many more as you like. If you want to cut the unit circle

out of the game you can substitue $t = v/u$ and simplify to get the formulas

$$\begin{aligned} x &= u^2 - v^2 \\ y &= 2uv \\ z &= u^2 + v^2 \end{aligned}$$

This will always get you Pythagorean triples if you put in integer u and v with $u > v$. It will also get you *all* such triples without a common factor, but the proof is a little technical and we won't go into it. Caution: the formula will get you all triples without common factor but it will also give you some *with* common factor, but not all of them. We have shown that there are infinitely many Pythagorean triples.

This method will work on all conic sections (ellipse, hyperbola, parabola) which are the quadratic curves. It will also work on a few cubic curves too, for example $y^2 = x^3$, but for cubic curves it is necessary to select just the right base point, whereas for quadratic curves you can use any point on the curve you like instead of the $(-1, 0)$ that we used above.

6.7 The Distribution of Primes

Mathematicians have been fascinated by the disribution of prime numbers in $\mathbb{Z}$ for hundreds of years. It turns out they are pretty randomly distributed as Dirichlet's theorem (see below) suggests. The first question natural to ask is how many primes are there below a certain upper limit, for example how many primes are there less than a million? We can give a pretty simple approximate answer to this question which is reasonably accurate.

The material in the next few sections is too difficult to give proofs because it requires Calculus, and not just ordinary real number Calculus but the Calculus of complex numbers, a subject of astounding beauty and subtlety. Don't fail to take the complex number Calculus course usually called Complex Variables or Complex Analysis (Junior or Senior level). Although I cannot give proofs the material is resonably easy to understand. It is typical of number theory that easy to understand theorems can have remarkably difficult proofs.

6.7.1 The Prime Number Theorem

Going all the way back to Euler (1707-1783) there were conjectures that $\pi(x)$, the number of primes less than x, was approximately

$$\pi(x) \approx \frac{x}{\ln(x)}$$

Early attempts to improve the accuracy of the formula were unsuccessful but we now have methods that work. However, we are not interested in those details. The $\ln(x)$ function is called the natural logarithm in contrast to the base 10 logarithm used extensively until the advent of hand calulators. The 10

logarithm has the button labled "log" on a hand calculator, whereas the natural logarithm has the button labled "ln". The two are proportional;

$$\log(x) = \frac{\ln(x)}{2.30259\ldots} \qquad \ln(x) = 2.30259\ldots \cdot \log(x)$$

and thus have the same mathematical properties but there are grubby constants associated with the use of $\log(x)$ in Calculus, so it's use should now be discouraged. There are only four kinds of people who should hit the log button on their calculators. The are: 1) Chemists calculating ph, 2)Astronomers calculating stellar brightness 3) Geologists calculating the number on the Richter scale for the strength of an earthquake and 4) Audiologists calculating decibels. If you are not one of those people you should not push the log button ever. Push the ln button[4].

Before you come to Calculus it doesn't really matter which logarithm you use. But for calculus the derivatives are different (d/dx is a symbol for the derivative):

$$\frac{d}{dx}\ln(x) = \frac{1}{x} \qquad \frac{d}{dx}\log(x) = \frac{1}{2.30259\ldots}\frac{1}{x}$$

Which would you prefer? If your finger should stray towards the log button *just say no.*

Let us look at how well the approximation works.

x	10	100	1000	10,000	100,000	1,000,000
$\pi(x)$	4	25	168	1229	9592	78498
$\frac{x}{\ln(x)}$	4.34	21.7	144.7	1085.74	8685.89	72382.4

For $x = 1,000,000$ the error is about 8%. The approximation is not as good as we might have hoped for but, on the other hand, it's not too bad. And it can be handy: how many primes are less than 10^{20}? The approximation gives us 2.17×10^{18}. Certainly beats counting them personally and the approximation is good enough for government work.

Also worth knowing is that as x increases the ratio of $\pi(x)$ and $x/\ln(x)$ approaches one. The technical notations for this are

$$\frac{\pi(x)}{\frac{x}{\ln(x)}} \to 1 \qquad \text{as } x \to \infty$$

$$\lim_{x\to\infty} \frac{\pi(x)}{\frac{x}{\ln(x)}} = 1$$

After over a hundred years of effort the prime number theorem was finally proved by Jacques Hadamard and Charles Jean de la Vallée-Poussin (separately) in 1896 using the Complex Calculus methods first introduced by Bernhard Riemann in 1859. Chebyshev had also come very close a little before Riemann's work. Chebyshev used less complicated methods than later workers.

[4] ln stands for logarithmus naturalis, latin for natural logarithem. log stands for Lower Our Grade.

6.7.2 Primes in Arithmetic Progressions

Preliminaries

An *arithmetic progression* starts at some integer and then jumps by the same amount for each step. Here is one that starts at $a_0 = 5$ and jumps by $d = 12$

$$5, 17, 29, 41, 53, 65, 77, 89, 101, \ldots$$

The formula for the sequence (new word for progression) is $a_i = 5 + 12i$ and the general formula is $a_i = a_0 + di$.

Note that this is closely connected to the numbers congruent to 5 mod 12, which are

$$\{\ldots -55, -43, -31, -19, -7, 5, 17, 29, 41, 53, 65, 77, 89, 101, \ldots\}$$

This is called the *residue class of 5 mod 12*. Thus an arithemetic progression is a rightward marching subset of a residue class, starting at a_0. The notation for the residue class is $[5] = [17] = [77] = [-19]$ and the residue class contains all integers congruent to 5 mod 12, or congruent to -19 mod 12, etc. Remember the test: a is congruent to 5 mod 12 if and only if $12|(a-5)$.

We are interested in primes in an arithmetic progression which we will always take to have positive terms as in the example. Here is another arithemetic progression much like the first, where $a_0 = 3$ and $d = 12$

$$3, 15, 27, 39, 51, 63, 75, 87, 99, 111, \ldots$$

Now for us the big difference between the two arithmetic porgressions is that there are lots of primes in the first one and exactly one prime in the second. Why is this. It is because in the second progression $3|a_i = 3 + 12i$, $i \in \mathbb{Z}$, $i \geq 1$ so only the prime 3 can be in progression. So which progressions, or which residue classes, can contain no or just one prime? For mod 12 only the classes [1],[5],[7],[11] can contain more than 1 prime. The other residue classes are [0]=[12],[2],[3],[4],[6],[8],[9],[10] and there is no more than 1 prime in each of these classes. We looked at 3. Let's look at [10]. All the elements of this class look like $10 + 12i$, so they are all even and can contain only the single even prime 2. (It doesn't actually contain 2, but it does have -2). Mentally check a couple of others.

Now if you think about which can and which can't have primes, the ones that can't are the ones whose GCD with 12 is not 1. GCD(5,12)=1 so the residue class [5] can have primes and the GCD(10,12)=2 so [10] can have no more than 1 prime. The classes $[a]$ with $\text{GCD}(a, 12) = 1$ are called *relatively prime residue classes to 12*. The same reasoning applies to the arithmetic progressions since the ones we are interested in are just the positive numbers in a residue class.

Euler's ϕ Function

Next question is, how many residue classes have many primes or how many residue classes are relatively prime to n. For $n = 12$ there are 4 such residue

classes. For 13 there are 12 such residue classes. For 6 there are only two, [1] and [5]. There is a formula for the number of numbers j in the sequence $0, 1, 2, \ldots, n-1$ which have GCD$(j, n) = 1$. It is called Eulers fee function. Let n be factored into primes where $p_1, p_2, \ldots, p_r$ are all distinct (different) and

$$n = p_1^{e_1} p_2^{e_2} \cdots p_{r-1}^{e_{r-1}} p_r^{e_r}$$

where $e_1, e_2 \ldots e_r$ are all positive integers. Then the number m of numbers in $0, 1, 2, \ldots, n-1$ relatively prime to n is $\phi(n)$ and thus there are $\phi(n)$ relatively prime residue classes to n. The formula for $\phi(n)$ is

$$\phi(n) = n(1 - \frac{1}{p_1})(1 - \frac{1}{p_2}) \cdots (1 - \frac{1}{p_{r-1}}(1 - \frac{1}{p_r})$$

Thus

$$\begin{aligned}
\phi(12) &= 12(1 - \frac{1}{2})(1 - \frac{1}{3}) = 12 \cdot \frac{1}{2} \cdot \frac{2}{3} = 4 \\
\phi(6) &= 6(1 - \frac{1}{2})(1 - \frac{1}{3}) = 6 \cdot \frac{1}{2} \cdot \frac{2}{3} = 2 \\
\phi(15) &= 15(1 - \frac{1}{3})(1 - \frac{1}{5}) = 15 \cdot \frac{2}{3} \cdot \frac{4}{5} = 8 \\
\phi(p) &= p(1 - \frac{1}{p}) = p \cdot \frac{p-1}{p} = p - 1
\end{aligned}$$

How hard is it to compute $\phi(n)$? For small numbers it is very easy, but it depends very essentially on factoring n into primes. As we have mentioned for large numbers, say numbers of 200 digits, factoring into primes is very difficult. No efficient method is known for finding the factorization. (In contrast, checking if a 200 digit number *is* a prime is very easy.) Since it is difficult to find the factorization, it is difficult to find $\phi(n)$ using the formula above. However, no other method for finding $\phi(n)$ is known. For bankers and secretive government agencies this is fortunate, because the coding methods they use depend on the difficulty of finding $\phi(n)$. (These coding methods are called public key cryptosystems.) An easy method of calculating $\phi(n)$ would attract the not necessarily benign interest of a great many secretive agencies.

Dirichlet's theorem

Let's think for a moment about what could happen here. The primes are distributed mod 12 among $\phi(12) = 4$ residue classes [1],[5],[7],[11]. It is conceivable that half of them primes are in [5], and a sixth of the are distributed evenly in each of [1],[7],[11]. Does this sound likely. It seems a little unfair doesn't it. And in fact it never happens.

In 1837 Johann Peter Gustav Lejeune Dirichlet[5] proved that there were infinitely many primes in each of the arithmetic progressions mod 12 in which

[5]The family was of Belgian origin (not French Huguenot as has been claimed) but had resided in western Germany for some time.

primes could live and moreover, there were approximately the same number in each arithmetic progression. Generalizing this we have

Dirichlet's theorem on Primes in Arithmetic Progressions. Let $n \geq 2$ be a positive integer and $a_1 = 1, a_2, \ldots a_{\phi(n)}$ be the numbers from 1 to n for which $\text{GCD}(a_i, n) = 1$. Then if we consider the $\pi(x)$ primes p between 1 and x there will be approximately $\pi(x)/\phi(n)$ in each of the arithmetic progressions beginning with a_i and with $d = n$.

Once again, this means that the primes do not bunch up in any of the residue classes but are equally distributed among the classes where they could live. Since this is true for every n, it suggests that the primes are fairly randomly distributed.

To prove this theorem Dirichlet had to invent the concept of L-function (too complex to describe here) and develop the theory of L-functions far enough to show that certain products of them do not vanish. This was a great contribution, and to this day there are several mathematicians who devote at least part of their lives to L-functions. It turned out that L-functions could be generalized in many ways, so Dirichlet kicked off the development of many subsequent strands of mathematics.

Computational Remarks

Look back at the formula for Euler's ϕ function, and you will notice that to find $\phi(n)$ it was necessary first to factor n. Contrary to one's intuition perhaps, it turns out that factoring integers is hard. It is pretty easy when the integers are small but for a 200 digit number I have seen the estimate of 16 million years. That was a while ago and probably today it would be shorter, perhaps 5 million years. Still that's a long wait for the factorization into primes, which you need to compute $\phi(n)$.

This is in contrast to other superficially similar problems, like is a 100 digit number a prime, which can be tested in a very short time with certain trickery. It's hard to tell in advance which problems are going to be hard.

The difficulty of factoring a number into primes, or the difficulty of finding $\phi(n)$, is closely connected with the problem of making unbreakable codes. A code called the RSA code was developed to make use of this. You can crack the code if you can find $\phi(m)$ for a two hundred digit number, but this is so difficult no one can do it, and so banks happily send cash back and forth and spies happily send reports back and forth using the RSA code, and nobody can crack the transmissions.

There is a whole class of these hard problems called NP-complete problems, and I have heard that if one of them is ever solved, all the others will be too. Nobody has yet proved there is no quick method for solving any of them, but nobody has solved one either. Stay tuned.

6.8 Indices

Before we get into the main subject of this section this is a convenient moment to prove Fermat's little theorem[6]. This was the first really important number theory to be proved in modern times that was not known to the ancients[7]. It is truly surprising in how many different circumstances this theorem comes to the rescue, so try to memorize it now like it was a treasure map.

Theorem Let p be an odd prime and $a \not\equiv 0 \bmod p$. Then

$$a^{p-1} \equiv 1 \bmod p$$

Proof (Leibniz (unpublished) and Euler) Form the numbers $\{1a, 2a, 3a, \ldots, (p-1)a\}$ and reduce them modulo p. Since all the numbers $\{1, 2, 3, \ldots, p-1\}$ have inverses mod p, if $ja \equiv ka$ then $j \equiv k$. Thus the reduced mod p numbers of $\{1a, 2a, 3a, \ldots, (p-1)a\}$ must all be different, and since there are $p-1$ of them they must be the numbers $\{1, 2, 3, \ldots, p-1\}$ in a different order. When we multiply the two sets together the results must be the same mod p; that is

$$\begin{aligned} (1a)(2a)(3a)\cdots((p-1)a) &\equiv 1 \cdot 2 \cdot 3 \cdot \cdots \cdot (p-1) \bmod p \\ 1 \cdot 2 \cdot 3 \cdot \cdots \cdot (p-1)a^{p-1} &\equiv 1 \cdot 2 \cdot 3 \cdot \cdots \cdot (p-1) \bmod p \\ a^{p-1} &\equiv 1 \bmod p \end{aligned}$$

We get to the last line by multiplying both sides of the previous congruence by the inverses of $\{1, 2, 3, \ldots, p-1\}$, which leaves us with the last line.

The theorem is so important and the results so amusing that I am going to do a numerical example of the proof, using $p = 7$ and $a = 3$. Then

$$\begin{aligned} (1 \cdot 3)(2 \cdot 3)(3 \cdot 3)\cdots((7-1) \cdot 3) &\equiv 3 \cdot 6 \cdot 9 \cdot 12 \cdot 15 \cdot 18 \equiv 3 \cdot 6 \cdot 2 \cdot 5 \cdot 1 \cdot 4 \\ (1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6) \cdot 3^6 &\equiv (3 \cdot 6 \cdot 2 \cdot 5 \cdot 1 \cdot 4) \cdot 1 \\ 3^{7-1} &\equiv 1 \bmod 7 \end{aligned}$$

We now return to the main subject of this section. We recall that the systems mod p are *finite fields*. There is a theorem which is just a little too hard for me to prove here that says

Theorem Let F be the finite field mod p that we get using $0, 1, 2, \ldots, p-1$ where multiplication is done mod p. Then there is a number g with $1 \le g \le p-1$ so that for any number a with $1 \le a \le p-1$ there is a unique α with $0 \le \alpha \le p-2$ for which $a \equiv g^\alpha$. Moreover, if g has this property then all other such elements h (called generators or primitive roots) can be found from g by $h \equiv g^\beta$ where $0 \le \beta \le p-2$ and $\mathrm{GCD}(\beta, p-1) = 1$. Since there are $\phi(p-1)$ such β there are $\phi(p-1)$ primitive roots.

[6] The name sounds funny because it is a not exactly accurate translation from German.

[7] We think. The destruction wrought on the cultural holdings of the ancient world by religious fanatics of various trivially different stripes was vast, so that we don't always know exactly what the ancients knew.

So let's now do an example. Let $p = 13$ and we want to find a primitive root g. If g is a primitive root then the powers of g will be the numbers $\{1, 2, 3, \ldots, 12\}$. We try $g = 3$

$$3^0 \equiv 1, \quad 3^1 \equiv 3, \quad 3^2 \equiv 9, \quad , 3^3 \equiv 1$$

and then further powers repeat the sequence 1,3,9 over and over so 3 is *not* a primitive root. Next we try $g = 2$ and the results are shown in the table:

i	0	1	2	3	4	5	6	7	8	9	10	11
2^i	1	2	4	8	3	6	12	11	9	5	10	7

and we see that the powers of 2 give $\{1, 2, 3, \ldots, 12\}$ mod 13. Hence $g = 2$ is a primitive root. The other primitive roots will be 2^j where $\mathrm{GCD}(j, 12) = 1$, so the j are $\{1, 5, 7, 11\}$ and all the $\phi(12) = 12\big(1 - \frac{1}{2}\big)\big(1 - \frac{1}{3}\big) = 4$ primitive roots are

$$2^1 \equiv 2, \quad 2^5 \equiv 6, \quad 2^7 \equiv 11, \quad 2^{11} \equiv 7$$

I checked these and indeed they are primitive roots.

Part of the reason I introduced this subject was to have a chance to explain the following. How did I find the primitive root. I tried 3 and it failed and then tried 2 and it succeeded. Is there a better way? It is embarassing to state that essentially no better way is known than trial and error. There is trickery to shorten the test for failure, but we are still doing trial and error. For 2 and a half centuries we have been trying to find a better way but so far no luck.

You can use the index table to do arithmetic mod 13 but there is a tricky aspect. The indices (= the exponents) work mod 12 not mod 13. So for example

$$\begin{aligned} 11^{-1} &\equiv (g^7)^{-1} \equiv g^{-7} \equiv g^5 \equiv 6 \qquad \text{since } -7 \equiv 5 \bmod 12 \\ 6/5 &\equiv g^5/g^9 \equiv g^{-4} \equiv g^8 \equiv 9 \\ \sqrt{3} &\equiv (g^4)^{\frac{1}{2}} \equiv \pm g^2 \equiv \pm 4 \equiv 4, 9 \end{aligned}$$

The last was included because in the next section we are going to obsess a lot about squares. Note that you can pick out the squares mod 13 easily from the table; they have even indices so they are 1,4,3,12,9,10, and for each of them, for example 12, the equation $x^2 \equiv 12$ is solvable. A glance at the table tells you that $x \equiv \pm 8 \equiv 8, 5$. If you look back at the multiplication table mod 13 you'll see that these are the elements on the diagonal. The elements with odd indices, 2,8,6,11,5,7 are not squares; the equation $x^2 \equiv 11$ is not solvable. An important fact that we see from this is

Corollary There are exactly $\frac{p-1}{2}$ non-zero squares mod p and the same number of non-squares.

We can use the index table to find a criterion for a mod p being a square. Since the squares have even indices, the squares are those elements with $a \equiv g^{2j}$ and thus

$$a^{\frac{p-1}{2}} \equiv (g^{2j})^{\frac{p-1}{2}} \equiv g^{p-1} \equiv 1 \bmod p$$

by Fermat's little theorem. You can prove by counting numbers of roots that for non-squares $a^{\frac{p-1}{2}} \equiv -1$. This is called Euler's criterion. It is not too handy to use but works really well for $a \equiv -1$.

Computational Note: Since there is no known way of finding a generator g mod p a computer algebra program like Mathematica or Matlab must either go through the trial and error process (slow) or have a big table which gives a g for each prime p (up to some large p). While inelegant in some ways, this method, called a *lookup table* is very very fast and it is mostly what is used in these circumstances. In fact, even when it's easy and fast to compute the result, like for sin(.434), it can be even faster to use lookup tables, and some programs now do use lookup tables for sin. Another natural place for a lookup table is for factoring numbers. While it impresses people no end to watch Mathematica factor some large number the truth is it has factorizations of the difficult numbers already hidden inside it; it's not actually doing anything but looking through a table.

6.9 Quadratic Reciprocity

This section in the book treats one of the great theorems of mathematics, the law of Quadratic Reciprocity. It is much easier to prove than the prime number theorem or Dirichlet's theorem but no one has ever come up with a quick and comfortable proof; all of them require some work.

I'll mention a personal experience before we get into the details. One day I was sitting in my office and the professor in the adjoining office came to me with a problem that arose in his research into card shuffling. He explained how it boiled down to a problem involving primes and squares, and I said immdiately "Oh, what you need is the law of Quadratic Reciprocity" and I explained how it worked and he went back to his office and finished his paper. Certainly an unlikely place to use a deep result from number theory.

An interesting feature of the theorem is that the first few people who studied it did not manage to prove it, in spite of the fact that they proved much harder things. When Gauss did prove it, he proved it in rather a complicated way. He eventually gave six proofs of the law, each from a slightly different angle, and then working off these eventually relatively simple proofs were found, although it's not easy to say what simple means in this circumstance. Proving the law became a shared hobby among ecclesiastical Englishmen in the last half of the 19th Century.

However, it was not until the development of the theory of algebraic integers that the true significance of the law emerged, and it was then that proofs were found that were more than just verifications but gave some hint as to what was really going on. Gauss had started this process with his proof using Gaussian integers ($\mathbb{Z}[i]$) and his proof of biquadratic reciprocity also using Gaussian Integers. Generalizations of the law of quadratic reciprocity went hand in hand with developments and generalizations of algebraic integers and their arithmetic. Far reaching theorems about the structure of rings of algebraic integers led to more

and more general forms of the law, until in the end it took serious effort to get from the most general form of the law back to the original form. There is a large book by Lammermeier on reciprocity laws for those who want to follow up on the subject, or who would like some inspiration for further studies. There are few laws or theorems which have had so great an influence on the subject in which they live.

It seems that on some level Gauss was aware of this, at least in part. In his book Disquitiones Arithmeticae (Arithmetic Investigations written 1798 published 1801) he not only proved the theorem for the first time but made it central to the development of the rest of the book. There was no obvious reason at that time to put it in such a central position, and it shows the depth of Gauss's genius that at the age of 21 he knew this was the correct path. This was the first comprehensive treatment of number theory which up to that time had been a hodgepodge of results published piecemeal. Gauss's book turned it into a real branch of mathematics. Unfortunately Gauss was far ahead of his time and his treatement, while perfectly readable today, was rather abstract for the era. It took a good while before his results were translated into forms readily understood by mathematicians of the time, but fortunately there were several great mathematicians, (e.g. Eisenstein, Dirichlet) who were able to do this.

There is a story that Gauss's book Disquitiones Arithmeticae has always been rare because of a fire at the publishing company. Certainly the original edition is extremely difficult and expensive to obtain today, and it is known that Gauss's favorite student Eisenstein was never able to obtain a copy. Dirichlet did have his own copy and it was his most prized possession. Nowadays it is easy to obtain a copy in English and copies are also available in German and Russian but I am not aware of copies available in the original Latin unfortunately.

So now it is time to discuss what the original form of the Law of Quadratic Reciprocity says. First let us build up some vocabulary for discussing the matter. As we have seen before, for a positive integer m it is possible to define residue classes mod m by putting numbers b into the class (called $[a]$) when b and a have the same remainder when divided by m. This is a clumsy criterion, but it is the same thing as

Def An integer b is in the class $[a]$ mod m if and only if $m|b-a$

and this is much easier to work with. Recall that we made tables of the residue classes mod 5 and mod 6 early in the book, but for convenience we just wrote 2 in the tables instead of [2]. We will continue to do this. In this section we are pretty excludively focused on tables mod p for primes $p > 2$ and only multiplication is of interest. I will present now the multiplication tables for 5, 7, 11 which is enough to illustrate the point.

5	1	2	3	4
1	1	2	3	4
2	2	4	1	3
3	3	1	4	2
4	4	3	2	1

7	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	4	6	1	3	5
3	3	6	2	5	1	4
4	4	1	5	2	6	3
5	5	3	1	6	4	2
6	6	5	4	3	2	1

11	1	2	3	4	5	6	7	8	9	10
1	1	2	3	4	5	6	7	8	9	10
2	2	4	6	8	10	1	3	5	7	9
3	3	6	9	1	4	7	10	2	5	8
4	4	8	1	5	9	2	6	10	3	7
5	5	10	4	9	3	8	2	7	1	6
6	6	1	7	2	8	3	9	4	10	5
7	7	3	10	6	2	9	5	1	8	4
8	8	5	2	10	7	4	1	9	6	3
9	9	7	5	3	1	10	8	6	4	2
10	10	9	8	7	6	5	4	3	2	1

Now to see how this works, use the 7 table and multiply 5 and 6. We want to know what residue class $5 \cdot 6$ is in, that is we want to know $[5 \cdot 6] = [30] = [2]$ because $30 = 4 \cdot 7 + 2$ and the remainder 2 is what interests us. We then put the [2] in the table but omitting the brackets to ease the crowding and the typesetting.

We are very interested in the numbers on the diagonals of these tables, which are the squares. The equation $x^2 \equiv a \bmod 7$ is solvable if and only if a is in one of the classes $[1], [4], [2]$. Thus $x^2 \equiv 32 \bmod 7$ is solvable because 32 is in [4] mod 7, and from the table we see the solutions are [2] and [5] so $2^2 \equiv 32 \bmod 7$ and $5^2 \equiv 32 \bmod 7$.

Note that in multiplication tables $(-a)^2 \equiv a^2 \bmod p$ but for $p \neq 2$ we have $-a \not\equiv a$ for $1 \leq a \leq p-1$. This $-a \equiv p - a$ and a are two unequal solutions to $x^2 \equiv a^2 \bmod p$. Since the residue classes mod p form a field, there can be only two solutions. So if we square the classes $[i], 1 \leq i \leq p-1$ we will get each square twice, once from $[i]$ and once from $[p-i] = [-i]$. Also note that we get all the squares by squaring $[1], [2], [3], \ldots, [(p-1)/2]$. Thus $(p-1)/2$ of the residue classes are squares. These are called *quadratic residues QR*. The other $(p-1)/2$ residue clesses, which are not on the diagonal and are not squares, are

called *quadratic nonresidues QN*[8]. Thus, with brackets omitted,

For	mod 3	the quadratic residues are 1
For	mod 5	the quadratic residues are $1, 4$
For	mod 7	the quadratic residues are $1, 4, 2$
For	mod 11	the quadratic residues are $1, 4, 9, 5, 3$
For	mod 13	the quadratic residues are $1, 4, 9, 3, 12, 10$
For	mod 17	the quadratic residues are $1, 4, 9, 16, 8, 2, 5, 13$
For	mod 19	the quadratic residues are $1, 4, 9, 16, 6, 17, 11, 7, 5$
For	mod 23	the quadratic residues are $1, 4, 9, 16, 2, 13, 3, 18, 12, 8, 6$

The QRs for 5, 7 and 11 are taken from the tables above. The rest I just computed by hand. Now we want to search for patterns in the above lists. We take pairs of primes from the lists. Using a slight modification of Gauss's notation we write 5 is a quadratic residue mod 11 as 5 QR 11, and 2 is *not* a quadratic residue mod 11 as 2 QN 11. Now we take pairs of primes from the list like 5, 11 and ask how the quadratic characters (QR and QN) correlate. Here is the data we can extract from the list. Note that in the second entry in the list it is necessary to reduce the first of the pair mod the second. In 13 QN 5 we must reduce $13 \equiv 3 \bmod 5$ and examine 3 in the 5 list where it is not found, so 3 QN 5 so 13 QN 5.

pair (p,q)	p Q? q	q Q? p
(5,13)	5 QN 13	13 QN 5
(5,17)	5 QR 17	17 QR 5
(5, 7)	5 QN 7	7 QN 5
(5, 11)	5 QR 11	11 QR 5
(5, 19)	5 QR 19	19 QR 5
(5, 23)	5 QN 23	23 QN 5
(7, 13)	7 QN 13	13 QN 7
(7, 17)	7 QN 17	17 QN 7
(11,13)	11 QN 13	13 QN 11
(11,17)	11 QN 17	17 QN 11
(13,17)	13 QR 17	17 QR 13
(13,19)	13 QN 19	19 QN 13
(13,23)	13 QR 23	23 QR 13
(17,19)	19 QR 19	19 QR 17
(17,23)	17 QN 23	23 QN 17

Well it looks like the two always come out with the same character. However, That is because I cherry picked the pairs of primes. Here are a few more pairs.

[8]The teminology seems odd to me. I would have called them nonquadratic residues. But the terminology as it stands is traditional.

pair (p,q)	p Q? q	q Q? p
(3,7)	3 QN 7	7 QR 3
(3,11)	3 QR 11	11 QN 3
(3,19)	3 QN 19	19 QR 3
(3,23)	3 QR 23	23 QN 3
(7,11)	7 QN 11	11 QR 7
(7,19)	7 QR 19	19 QN 7
(7,23)	7 QN 23	23 QR 7
(11,19)	11 QR 19	19 QN 11
(11,23)	11 QR 23	23 QN 11
(19,23)	19 QN 23	23 QR 19

In this list the pairs of primes come out with opposite character. There is a not too subtle difference between the pairs in the two lists. See if you can figure out what is different about the pairs of primes in the first list and the pairs of primes in the second list. (The lack of 3's in the first list is irrelevant.)

Legendre[9] studied this sort of material and was thus able to invent a fine notation to describe what is going on which is still used today. He was not able to prove the law which he so elegantly stated, although he thought he had proved it. Essentially he assumed that there are primes in certain arithmetic progressions (as we discussed in the section on Dirichlet's theorem) but it was decades later that this was proved, so at the time his proof was incomplete. He tried revising it several times but was never quite able to push it over the top.

Before we give Legendre's statement of the law let's spend a moment with some stuff we need to understand to appreciate it. Every thing turns on whether $p \equiv 1 \bmod 4$ or $p \equiv 3 \bmod 4$.

Let $p \equiv 1 \bmod 4$. Then $4|p-1$ so $p-1 = 4k$ and

$$\frac{p-1}{2} = 2k$$

so

$$(-1)^{\frac{p-1}{2}} = (-1)^{2k} = +1$$

Now let $p \equiv 3 \bmod 4$. Then $4|p-3$ so $p-3 = 4k$ and $p-1 = 4k+2$ and

$$\frac{p-1}{2} = 2k+1$$

so

$$(-1)^{\frac{p-1}{2}} = (-1)^{2k+1} = -1$$

Next we introduce Legendre's symbol for a being a quadratic residue mod p or not. The symbol is

$$\left(\frac{a}{p}\right)$$

[9] Adrien-Marie Legendre 1752-1833, French. A great mathematician who made many mistakes in his proofs. There is no portrait or photograph of him in spite of his having lived into the age of photography. All that is known is a caricature. The portrait shown in many history of mathematics books is actually that of the politician Louis Legendre. Although the French mint made medals for many mathematicians they unfortunately never got around to Legendre.

(*this is NOT a fraction*) and the definition is

$$\left(\frac{a}{p}\right) = \begin{cases} +1 & \text{if } a\,QR\,p \quad a \text{ is a quadratic residue mod } p \\ -1 & \text{if } a\,QN\,p \quad a \text{ is NOT a quadratic residue mod } p \end{cases}$$

We are now ready to state the law of quadratic reciprocity.

Theorem Let p and q be odd primes. Then

$$\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}} \left(\frac{p}{q}\right)$$

The modern way of writing this is more symmetrical although the old way is closer to how you use it. We can multiply both sides by $\left(\frac{p}{q}\right)$ to get

$$\left(\frac{q}{p}\right)\left(\frac{p}{q}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}$$

since $\left(\frac{p}{q}\right)^2 = 1$.

Notice that $(p-1)/2 \cdot (q-1)/2$ is even if either one or both of p and q is congruent to 1 mod 4 but $(p-1)/2 \cdot (q-1)/2$ will be odd if both p and q are congruent to 3 mod 4. Thus (as we found when looking at the examples above)

$$\left(\frac{q}{p}\right) = \left(\frac{p}{q}\right) \text{ if either } p \text{ or } q \text{ is congruent to 1 mod 4}$$

and

$$\left(\frac{q}{p}\right) = -\left(\frac{p}{q}\right) \text{ if both } p \text{ and } q \text{ are congruent to 3 mod 4}$$

There is another way to write the law of quadratic reciprocity which is also interesting. We define for odd numbers p

$$p^* = (-1)^{\frac{p-1}{2}} p$$

Note that $p^* \equiv 1 \bmod 4$. Then

$$p^* = \begin{cases} p & \text{for } p \equiv 1 \mod 4 \\ -p & \text{for } p \equiv 3 \mod 4 \end{cases}$$

Personally I suspect that the proper list of primes in $\mathbb{Z}$ should be the p^*, that is

$$\{2, -3, 5, -7, -11, 13, 17, -19, -23, 29, -31, 37, 41, -43, \ldots\}$$

but this is just my opinion.

Using this notation, the law of quadratic reciprocity could be restated, for odd primes p and q as

$$\left(\frac{q}{p}\right) = \left(\frac{p^*}{q}\right)$$

Using the material below you will prove this in the problems.

This is a convenient place to restate Euler's criterion from the end of the previous section. Recall Fermat's Little Theorem that for odd primes p we have if $p \not| a$ then $a^{p-1} \equiv 1 \bmod p$. Now recall that if a is a square, say $a \equiv b^2$ for some b, so then $a^{\frac{p-1}{2}} \equiv b^{2\cdot\frac{p-1}{2}} \equiv b^{p-1} \equiv 1$. Counting solutions as we did at the end of the previous section we have

Euler's criterion

$$\left(\frac{a}{p}\right) = \begin{cases} +1 & \text{if } a^{\frac{p-1}{2}} \equiv 1 \mod p \quad a \text{ is a quadratic residue mod } p \\ -1 & \text{if } a^{\frac{p-1}{2}} \equiv -1 \mod p \quad a \text{ is NOT a quadratic residue mod } p \end{cases}$$

From this it is obvious that

$$\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}} \qquad \bmod p$$

Letting $a = -1$ we have

$$\left(\frac{-1}{p}\right) \equiv (-1)^{\frac{p-1}{2}} \qquad \bmod p$$

but since both sides of this congruence are +1 or -1, congruence forces equality (since $-1 \not\equiv 1 \pmod p$ for $p > 2$) and thus we have the

First Complementary Theorem

$$\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}}$$

Using Euler's criterion, it is easy to prove the important property of the Legendre symbol

$$\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$$

With this and the first complementary theorem you will have no difficulty proving $\left(\frac{q}{p}\right) = \left(\frac{p^*}{q}\right)$ from the ordinary form of the quadratic reciprocity law.

To calculate $\left(\frac{a}{p}\right)$ for any integer a we also need the

Second Complementary Theorem

$$\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}} = (-1)^{\frac{p^*-1}{4}}$$

The first formula is traditional but the second is superior for almost all purposes.

We can now find $\left(\frac{a}{p}\right)$ for any integer a and prime p.

The proofs of the law of quadratic reciprocity and the second complementary theorem are not truly difficult but they are rather long for a book of this nature, so sadly I must omit them. The law of quadratic reciprocity is probably the deepest result discussed in this book, and is the jumping off point for much of the further development of number theory. It was first conjectured by Euler and first correctly proved by Gauss in his book *Disquisitiones Arithmeticae* published in 1801. The publishers warehouse burned down, explaining why the book was somewhat rare until it was reprinted many years later.

Perhaps it is time to give some examples. We will use $p = 31$ for our examples and will illustrate the sort of things that occur. We will also discuss while doing the examples various aspects that relate to the information. We will put the examples in the form of questions, like is -1 a square mod 31 or is -1 a square in the field $\mathbb{F}_{31}$. ($\mathbb{F}_{31}$ is the field with 31 elements represented by the integers i $0 \leq i \leq 30$. It can be proved that there is exactly one field for each prime power p^n.) I will switch off between the two descriptions so you will get used to both modes of speaking.

Is -1 a square mod 31? To do this we use the first complementary theorem:

$$\left(\frac{-1}{31}\right) = (-1)^{\frac{31-1}{2}} = (-1)^{15} = -1$$

which means that -1 is *not* a square in the field $\mathbb{F}_{31}$. This means that we can form $\mathbb{F}_{31}[\sqrt{-1}]$ which will consist of

$$\{a + b\sqrt{-1} \mid 0 \leq a, b \leq 30\}$$

in exactly the same way we form the Complex Numbers $\mathbb{C}$ from the Real Numbers $\mathbb{R}$. This gives us a field $\mathbb{F}_{961}$ with $961 = 31^2$ elements. We say that we have *adjoined* $\sqrt{-1}$ to the field $\mathbb{F}_{31}$.

Is 2 a square mod 31? To do this we use the second complementary theorem:

$$\left(\frac{2}{31}\right) = (-1)^{\frac{31^*-1}{2}} = (-1)^{\frac{-31-1}{4}} = (-1)^{-8} = +1$$

Thus we see that indeed the the congruence

$$x^2 \equiv 2 \pmod{31} \qquad \text{is solvable}$$

Notice, and this is important, that though we know there is a solution to $x^2 = 2$ in the field $\mathbb{F}_{31}$ we have no idea what the solution *is*. To find the solution we must mess around with our calculators or a computer algebra program. My calculator, which has a mod function on it, easily finds that

$$1^2 \equiv 1,\ 2^2 \equiv 4,\ 3^2 \equiv 9,\ 4^2 \equiv 16,\ 5^2 \equiv 25,\ 6^2 \equiv 5,\ 7^2 \equiv 18,\ 8^2 \equiv 2$$

and thus we have a solution 8. It would be easier if we had a table of indices. There are some techniques for solving this problem but they are not trivial. Notice that since 2 has a square root, adding $\sqrt{2}$ to $\mathbb{F}_{31}$ would produce a beast

quite unlike anything we have talked about in this book, which would definitely not be a field.

Notice that the other solution to $x^2 \equiv 2$ would be $-8 \equiv 23$. Since this is a field situation there are only the two solutions 8 and 23.

Next we ask if 3 is a square mod 31. We calculate

$$\left(\frac{3}{31}\right) = (-1)^{\frac{3-1}{2}\frac{31-1}{2}}\left(\frac{31}{3}\right) = (-1^{1\cdot 15}\left(\frac{1}{3}\right) = (-1)\cdot 1 = -1$$

(since $31 \equiv 1 \pmod 3$,) so 3 is not a square mod 31 and

$$x^2 = 3 \pmod{31} \qquad \text{is insolvable.}$$

We could, as before, add $\sqrt{3}$ to $\mathbb{F}_{31}$ to make

$$\mathbb{F}_{31}[\sqrt{3}] = \{a + b\sqrt{3} \mid 0 \leq a, b \leq 30\}$$

to get a field with 961 elements. However, more advanced theory tells us that there is exactly *one* field with 961 elements, so this is the same field as $\mathbb{F}_{31}[\sqrt{-1}]$ that we got before, with the elements in a different order. So we have gotten nothing new. This is the usual situation with finite fields.

Is 6 a square mod 31? Answer is quick;

$$\left(\frac{6}{31}\right) = \left(\frac{2}{31}\right)\left(\frac{3}{31}\right) = 1\cdot(-1) = -1$$

so 6 is not a square mod 31. This is our first chance to use the multiplicativity of the Legendre symbol.

Is 5 a square mod31?

$$\left(\frac{5}{31}\right) = (-1)^{\frac{5-1}{2}\frac{31-1}{2}}\left(\frac{31}{5}\right) = (-1)^{2\cdot 15}\left(\frac{1}{5}\right) = 1\cdot 1 = 1$$

Notice in the above list of squares we see that $6^2 \equiv 5 \bmod(31)$.

The following two calculations show how useful it is to remove squares from the numerator of a Legendre symbol, since $\left(\frac{\text{square}}{p}\right) = 1$.

$$\left(\frac{8}{31}\right) = \left(\frac{4}{31}\right)\left(\frac{2}{31}\right) = 1\cdot 1 = 1$$

$$\left(\frac{12}{31}\right) = \left(\frac{4}{31}\right)\left(\frac{3}{31}\right) = 1\cdot(-1) = -1$$

We get something new when we ask if 15 is a square mod mod 31. Using things we have already calculated, we get

$$\left(\frac{15}{31}\right) = \left(\frac{3}{31}\right)\left(\frac{5}{31}\right) = (-1)\cdot 1 = -1$$

so 15 is not a square mod 31. Do you see what is *new* here? To do this we had to factor 15 into primes. This was so easy you probably missed it, but there is an important point here, one that always rears it head when we factor things into primes. Factoring into primes for large numbers is *difficult.* Hence we always ask, is there a way to do this without factoring into primes. This was answered by Jacobi[10]

To simplify the calculation of the Legendre symbol, Jacobi introduced a new symbol (which looks exactly the same) but where the denominator is not constrained to be prime. Here is the definition.

Def Let a be any integer and let b be a positive *odd* integer. Let b factor into primes thus: $b = p_1 p_2 \cdots p_r$ where the p_i *need not be distinct.* Then

$$\left(\frac{a}{b}\right) = \left(\frac{a}{p_1}\right)\left(\frac{a}{p_2}\right)\cdots\left(\frac{a}{p_r}\right)$$

where the symbols on the right are Legendre symbols.

One can prove, and indeed I have included a problem set where *you* get to do the proofs, that the same laws work for the Jacobi symbol as for the Legendre symbol. First we have

$$\left(\frac{ac}{b}\right) = \left(\frac{a}{b}\right)\left(\frac{c}{b}\right)$$

for all integers a and c and all odd positive b. Next we have

JACOBI SYMBOL MATERIAL HERE

6.10 The Chinese Remainder Theorem

This last section in the book is important for many reasons, and also relatively easy. One reason for the section is it gives me an opportunity to talk about mathematics in general and connect this theorem to two important principles.

In my opinion, and it is just my opinion, mathematics is about two things. The first is symmetry. We are used to seeing symmetry in geometry where it is intuitively clear. In the more algebraic areas of mathematics it is less clear, but remember every time you see some pattern in numerical or algebraic data you are seeing some kind of symmetry, so in this sense it is everywhere.

For example we saw that if either p or q is congruent to 1 (mod 4) then

$$\left(\frac{q}{p}\right) = \left(\frac{p}{q}\right)$$

[10] 1804-1851. Born in Potsdam, he was the first important mathematician of Jewish ancestry in modern times. He became a professor at the University of Königsberg where he remained until 1843. He had health problems and moved to Berlin where he was given a royal pension for (most of) the remainder of his life. He made contributions to many areas of mathematics, the most important being number theory, elliptic and theta functions, and mechanics. Jacobi and the author have very similar interests which is not related to the fact that they have the same birthday.

We might say that p and q enter into the formula symmetrically. However, when p and q are both congruent to 3 (mod 4) the formula changes and is now symmetric in a different way. We say the original symmetry is broken.

$$\left(\frac{q}{p}\right) = -\left(\frac{p}{q}\right)$$

We like symmetry, and when symmetry is broken as in the second formula we want to know why. Investigating this broken symmetry was one of the most important investigations mathematicians ever undertook, and they have found many many ways to look at the problem.

The second thing that mathematics is about is collecting a lot of information in restricted circumstances and then integrating that information to find out something about a general situation. To give a more concrete example think of electrically sophisticated ants living on a very large surface. They want to know the general shape of their surface bit they are unable to see the entire surface at once, due to it's size. They put a lot of points on the surface, which they call vertices, and draw lines from the points (on the surface) to nearby points. They then count the number of edges between the points, and count the number of regions that have one vertex near their point. They make arrangements with the ants of the nearby vertices about which vertex gets to count the region as its region, for counting purposes only, and similarly for the edges. The ants at each vertex then report into the central office by telephone how many edges and regions their vertex has alloted to it. The central office then calculates a certain number and knows from the value of that number that the surface is a two-hole donut. That is, they have taken the local information from each vertex and they have used it to find the global structure of the surface, a two hole donut. I call this a local-global correlation. These occur all over mathematics and where ever they occur mathematicians find them interesting.

Now for an example from number theory. Suppose we are interested in squares mod 935. How should we attack this problem? We see $935 = 5 \times 11 \times 17$ We investigate the situation with squares mod 5, mod 11 and mod 17. This is the *local* information. We already know a lot about how to attack these three problems. Perhaps we can put this local information together to get information mod 935 which is *global* information. Now replace 935 by some number a with twenty seven distinct prime factors. We can investigate squares mod a by investigating squares mod each of the twenty seven primes and then putting the information together. I hope you see that this looks a bit like the problem the ants faced in the previous paragraph.

Before giving a more complex numerical example we will state the basic theorem for this kind of work.

Def Two integers m and n are called relatively prime if and only if $\mathrm{GCD}(m, n) = 1$.

Def A set $\{m_1, m_2, \ldots, m_r\}$ of positive integers is called *pairwise relatively prime* if and only if $\mathrm{GCD}(m_i, m_j) = 1$ for any distinct i, j with $1 \leq i, j \leq r$.

The phrase *pairwise relatively prime* is often abbreviated to *pairwise prime* for convenience.

Theorem (Chinese Remainder Theorem[11])
Let $\{m_1, m_2, \ldots, m_r\}$ be pairwise relatively prime positive integers and let $\{a_1, a_2, \ldots, a_r\}$ be any integers. Let $M = m_1 m_2 \cdots m_r$. Then there is an integer a for which

$$\begin{array}{rcll} a & \equiv & a_1 & \bmod m_1 \\ a & \equiv & a_2 & \bmod m_2 \\ \ldots & \ldots & \ldots & \ldots \\ a & \equiv & a_r & \bmod m_r \end{array}$$

Moreover, the solution a is unique mod M.

The final sentence in the theorem means if a and b both satisfy the system of congruences $a, b \equiv a_i \bmod m_i$ then $a \equiv b \bmod M$

The proof of this theorem is just like the following example, where our numbers are just replaced by letters. This was the standard method of proving things among the Sumerians and Babylonians, where a numerical example in which the numbers had no special properties was given, and many other cultures have followed this pattern of proof. We might call this *proof by example.* The student recognizes that the same pattern would work when the numbers are changed. Much high school algebra is taught in exactly this way, continuing a 5000 year old tradition.

In many cases of this kind, we are given several sets of values $\{a_1, a_2, \ldots, a_r\}$ and we need to find the a for each set. Thus we would like a systematic way to find a. To do this, we find what the professionals call *a set of r orthogonal idempotents* $e_1, e_2, \ldots, e_r$. Once found, it is trivial to find an a for any set of data.

We will now illustrate how to find the e_i when $r = 3$ and $m_1 = 5$, $m_2 = 11$, $m_3 = 17$. The e_i are solutions to the following congruences.

$$\begin{array}{rcll} e_1 & \equiv & 1 & \bmod 5 \\ e_1 & \equiv & 0 & \bmod 11 \\ e_1 & \equiv & 0 & \bmod 17 \end{array} \qquad \begin{array}{rcll} e_2 & \equiv & 0 & \bmod 5 \\ e_2 & \equiv & 1 & \bmod 11 \\ e_2 & \equiv & 0 & \bmod 17 \end{array} \qquad \begin{array}{rcll} e_3 & \equiv & 0 & \bmod 5 \\ e_3 & \equiv & 0 & \bmod 11 \\ e_3 & \equiv & 1 & \bmod 17 \end{array}$$

The three congruences for e_1 can be collapsed into

$$\begin{array}{rcll} e_1 & \equiv & 1 & \bmod 5 \\ e_1 & \equiv & 0 & \bmod 187 \end{array}$$

since $11|e_1$ and $17|e_1$ are equivalent to $187|e_1$ since 11 and 17 are primes and $187 = 11 \times 17$. This amounts to

[11] The earliest known formulation of this theorem is in the book Sūnzǐ Suànjīng, roughly the mathematical manual of Master Sung, written sometime in the 3rd to 5th Centuries CE. Sadly, nothing is known about Master Sung.

$$\begin{aligned} e_1 - 1 &= 5\ell \\ e_1 &= 187k \end{aligned}$$

and subtracting the first from the second we have

$$187k - 5\ell = 1$$

This is a linear Diophantine equation which we learned to solve long ago by using the Euclidean algorithm. Running the Euclidean Algortithm on 5 and 187 we get, after some trivial manipulations, that

$$-2 \cdot 187 + 75 \cdot 5 = 1$$

This equation tells us that if we set $e_1 = -2 \cdot 187$ then

$$\begin{aligned} e_1 &\equiv 1 \mod 5 \\ e_1 &\equiv 0 \mod 187 \end{aligned}$$

We are almost done. It is convenient (but not necessary) to have the e_i satisfy $0 \leq e_i \leq M = 935$ Thus $e_1 = -2 \cdot 187 = -374 \equiv 561 \mod 935$. We can solve for e_2 and e_3 using the congruences

$$\begin{aligned} e_2 &\equiv 1 \mod 11 \\ e_2 &\equiv 0 \mod 85 \end{aligned} \qquad \begin{aligned} e_3 &\equiv 1 \mod 17 \\ e_3 &\equiv 0 \mod 55 \end{aligned}$$

and we then find $e_2 = -4 \cdot 85 = -340 \equiv 595$ and $e_3 = -4 \cdot 55 = -220 \equiv 715 \mod 935$. So to sum up we have

$$e_1 \equiv 561 \qquad e_2 \equiv 595 \qquad e_3 \equiv 715 \qquad \mod 935$$

Now to solve the system of congruences

$$\begin{aligned} a &\equiv a_1 \mod m_1 = 5 \\ a &\equiv a_2 \mod m_2 = 11 \\ a &\equiv a_3 \mod m_2 = 17 \end{aligned}$$

it suffices to set

$$a \equiv a_1 e_1 + a_2 e_2 + a_3 e_3 \qquad \mod 935$$

For example the system of congruences

$$\begin{aligned} a &\equiv 3 \mod m_1 = 5 \\ a &\equiv 7 \mod m_2 = 11 \\ a &\equiv 15 \mod m_2 = 17 \end{aligned}$$

has solution

$$a \equiv 3e_1 + 7e_2 + 15e_3 \equiv 3 \cdot 561 + 7 \cdot 595 + 15 \cdot 715 = 16573 \equiv 678 \qquad \mod 935$$

Why does it work. Consider $m_2 = 11$ The congruences above show that $e_1 \equiv e_3 \equiv 0 \bmod 11$ and $e_2 \equiv 1 \bmod 11$. Hence mod 11 we have

$$a \equiv 3e_1 + 7e_2 + 15e_3 \equiv 3 \cdot 0 + 7 \cdot 1 + 15 \cdot 0 \equiv 7 \qquad \bmod 11$$

and the same for the other two congruences. We leave it to the problems to work out why any solution of the system will be congruent to 678 mod 935.

The e_i have a couple of interesting properties. First, $e_i^2 \equiv e_i \bmod M$. For example $e_2^2 = 595^2 = 354025 \equiv 595 \bmod 935$. A ring element satisfying $e^2 = e$ is called an idempotent. Second the product of any two distinct e_i will be 0. For example $e_1 e_2 = 561 \cdot 595 = 333795 \equiv 0 \bmod 935$. Thus we say e_1 and e_2 are *orthogonal* and thus the e_i form a set of *orthogonal idempotents*. This is important in the theory of rings because it allows one to break a ring down into a product of smaller rings, but we will not go into this here. We do a bit more of this in the problems.

Now we will use our local-global connection to solve the congruence

$$x^2 \equiv 586 \qquad \bmod 935$$

Clearly if there is a solution then it is necessary that the following three congruences have a solution:

$$x^2 \equiv 586 \equiv 1 \mod 5 \qquad x^2 \equiv 586 \equiv 3 \mod 11 \qquad x^2 \equiv 586 \equiv 8 \mod 17$$

The first is clearly solvable. For the other two we have

$$\left(\frac{3}{11}\right) = (-1)^{\frac{3-1}{2}\frac{11-1}{2}} \left(\frac{11}{3}\right) = (-1)^{1\cdot 5} \left(\frac{2}{3}\right) = 1$$

$$\left(\frac{8}{17}\right) = \left(\frac{4}{17}\right)\left(\frac{2}{17}\right) = 1 \cdot (-1)^{\frac{17^*-1}{4}} = 1$$

Hence all three are solvable, so the original congurence $x^2 \equiv 586 \pmod{935}$ is solvable as we now demonstrate using the Chinese remainder theorem.

I get out my calculator and punching the buttons (see problems) I find that the solutions of the three congruences are

$$x \equiv \pm 1 \equiv 1, 4 \mod 5 \qquad x \equiv \pm 5 \equiv 5, 6 \mod 11 \qquad x \equiv \pm 5 \equiv 5, 12 \mod 17$$

These can be combined in 8 ways to give 8 solutions. They are (all mod 935)

$$\begin{array}{rcrcrcrcrcrcrcr}
1e_1 & + & 5e_2 & + & 5e_3 & = & 1\cdot 561 & + & 5\cdot 595 & + & 5\cdot 715 & = & 7111 & \equiv & 566 \\
1e_1 & + & 5e_2 & + & 12e_3 & = & 1\cdot 561 & + & 5\cdot 595 & + & 12\cdot 715 & = & 12,116 & \equiv & 896 \\
1e_1 & + & 6e_2 & + & 5e_3 & = & 1\cdot 561 & + & 6\cdot 595 & + & 5\cdot 715 & = & 7706 & \equiv & 226 \\
1e_1 & + & 6e_2 & + & 12e_3 & = & 1\cdot 561 & + & 6\cdot 595 & + & 12\cdot 715 & = & 12711 & \equiv & 556 \\
4e_1 & + & 5e_2 & + & 5e_3 & = & 4\cdot 561 & + & 5\cdot 595 & + & 5\cdot 715 & = & 8794 & \equiv & 379 \\
4e_1 & + & 5e_2 & + & 12e_3 & = & 4\cdot 561 & + & 5\cdot 595 & + & 12\cdot 715 & = & 13,799 & \equiv & 709 \\
4e_1 & + & 6e_2 & + & 5e_3 & = & 4\cdot 561 & + & 6\cdot 595 & + & 5\cdot 715 & = & 9389 & \equiv & 39 \\
4e_1 & + & 6e_2 & + & 12e_3 & = & 4\cdot 561 & + & 6\cdot 595 & + & 12\cdot 715 & = & 14,394 & \equiv & 369
\end{array}$$

These are all solutions of $x^2 \equiv 586 \pmod{935}$. Lets check one at random. To check if $709^2 \equiv 586 \pmod{935}$ we compute $709^2 - 586$ and see if 935 divides it. Indeed it does:

$$709^2 - 586 = 502795 = 537 \times 935$$

One can check the other 7 solutions in the same way. A little thought will also convince you that there can be no other solutions. There are not always 8 solutions to $x^2 \equiv a \pmod{935}$. We saw there were 8 solutions to $x^2 \equiv 586 \mod 935$ but there are only 4 solutions to $x^2 \equiv 485 \pmod{935}$ and only two to solutions to $x^2 \equiv 880 \pmod{935}$. One would like a formula for the number of solutions. To do this we will have to modify slightly the Legender (or Jacobi) symbol for use in counting. The additional rule is

$$\text{If } p \text{ is an odd prime and } p|a \text{ then } \left(\frac{a}{p}\right) = 0$$

$$\text{If } b \text{ is odd and GCD}(a,b) > 1 \text{ then } \left(\frac{a}{b}\right) = 0$$

We will not be using the second rule; I just include it for information.

We see immediately that the formula for the number of solutions of $x^2 \equiv a \pmod{p}$ is

$$N(x^2 \equiv a \bmod p) = 1 + \left(\frac{a}{p}\right)$$

for if $p \not| a$ and a is a square we get 2, which is correct. If $p \not| a$ and a is not a square we get 0, which is correct. And if $p|a$ then $a \equiv 0 \pmod{p}$ and we get one solution $0 \pmod{p}$ which is also correct, where we have used our new rule for the Legendre symbol. This illustrates the use of the new rule in counting.

Now for the number of solutions of $x^2 \equiv a \bmod 935$ we can write the formula for the number of soltions as

$$N(x^2 \equiv a \bmod 935) = \left(1 + \left(\frac{a}{5}\right)\right)\left(1 + \left(\frac{a}{11}\right)\right)\left(1 + \left(\frac{a}{17}\right)\right)$$

There is one other fact that is useful when working with $x^2 \equiv a \bmod b$ *where* b *is odd*. It is this

For odd primes p with $p \not| a$

$x^2 \equiv a \bmod p^n$ is solvable if and only if $x^2 \equiv a \bmod p$ is solvable

In both cases there are two solutions

For example $x^2 \equiv 99 \mod 5^3 = 125$ is solvable and has two solutions because $x^2 \equiv 99 \equiv 4 \mod 5$ is solvable. It is even possible to start with a solution mod 5 and compute the solution mod 125 by a well defined procedure which involves solving certain linear congruences, but we will not go down this path.

And finally we note that the situation with squares mod powers of 2 is anomalous.

For odd a

$x^2 \equiv a \bmod 2^n$ $n \geq 3$ is solvable if and only if $x^2 \equiv a \bmod 8$ is solvable

In both cases there are four solutions

This ends our brief tour of number theory. If you work through the problems for this chapter you will know more than most students who have taken a college course in number theory. I apologize for failing to provide proofs for everything but I wanted to keep things relatively elementary and fun. There are many fine elementary books on number theory if you want to follow up on your experience here. (There are also many less wonderful books. I suggest you check the book to find out if they prove the law of quadratic reciprocity. If they don't, find another book.) Here are a few random suggestions.

FRIENDLY BOOKS

Davenport, H. THE HIGER ARITHMETIC. (Many editions, Very friendly approach.)

Andrews, George E. NUMBER THEORY. (Dover Paperback, cheap. Also does some combinatorics.)

Hardy, G. H. & Wright, Edward M. AN INTRODUCTION TO THE THEORY OF NUMBERS (Well loved classic.)

Burton, David M. ELEMENTARY NUMBER THEORY (Modern college textbook)

Hurwitz & Kritikos LECTURES ON NUMBER THEORY.

HEAVIER BOOKS

Hasse, H. LECTURES ON NUMBER THEORY. Wonderful book but out of print.

Hasse, H. NUMBER THEORY Moderately difficult introduction to the modern Theory.

Serre, Jean-Pierre A COURSE IN ARITHMETIC Compact introduction to modern ideas and methods.

REALLY REALLY HARD BOOKS

Weil, Andre BASIC NUMBER THEORY. Much mathematical knowledge required to read this.

Bibliography

[Rovelli] Quantum Gravity, Cambridge Univ Press, 2005

[Rovelli & Vidotto] Covariant Loop Quantum Gravity, Cambridge Univ Press, 2015

[Cantor] Contributions to the Founding of the Theory of Transfinite Numbers, Dover, New York.

[Dawson] Dawson, C. Bryan, Calculus Set Free; Infinitesimals to the Rescue, vol I,II 2018

[Hermes] Hermes, Hans, Enumerability, Decidability, Computability: An Introduction to the Theory of Recursive Functions (Grundlehren der mathematischen Wissenschaften) Berlin 1969

[Keisler] Keisler, H. Jerome, Elementary Calculus, Dover Reprint, N.Y. 2012

[Quine] Quine, W.V.O. Methods of Logic, 4th ed, Harvard University Press 1982.

[Needham] Needham, Tristan, Visual Complex Analysis, Oxford University Press 1997.