Preface: Leaving the Sixties

I left for college in 1967: the sixties, when everything was changing. My dorm was co-ed, the first in the country; Vietnam war protests shut down the University and students smoked weed in the lounge. But also: an entering physics major, I picked up Feynman’s *Lectures On Physics*. Feynman wrote,

The special problem we tried to get at with these lectures was to maintain the interest of the very enthusiastic and rather smart students coming out of the high schools . . . . They have heard a lot about how interesting and exciting physics is – the theory of relativity, quantum mechanics, and other modern ideas. [Instead] they were made to study inclined planes, electrostatics, and so forth, and after two years it was quite stultifying.

Today, most students take an AP Calculus course in high school; they repeat the course in college, sometimes with the same book. They’ve heard about amazing discoveries in science, engineering, medicine and technology that change the world, but when they enter a calculus classroom, they’re magically transported back to 1967, as though the last fifty years of progress never happened.

This book is for people who have had calculus and want to know more, or who may need to use it or understand it in their work. It’s not about precise proofs, but about understanding in depth why and when and how to use calculus as a professional.

It is unlike other books; I want to reconnect calculus with with all the world: history, religion, philosophy, literature, psychology and especially the modern scientific and technical world.

Calculus is an incredible intellectual adventure. Let’s start.
Chapter 0: Introduction

Why Calculus?

On Jody Foster’s first day at Yale she wrote a friend: ‘My Calculus book is three inches thick. I can’t survive three inches of calculus.’ My students are more direct: ‘I’m going to be a doctor. Why do I need calculus?’ Or, ‘I’m gonna be an engineer: we use computers for everything. Why do I need this class?’ Instead of a direct answer, I’ll tell a story about a doctor, and an engineer, and the ECG, a graph of the electrical currents in a heartbeat, used to evaluate the health of a heart (see p5).

The story begins in the 1790’s, when the wife of an Italian doctor, Luigi Galvani, noticed that the muscles of some frog legs she’d hung out for dinner twitched when a spark of electricity touched them. Galvani followed up, experimenting for many years, concluding that the nervous system operated through electricity. But in 1790 there were no instruments to measure electricity; scientists detected electrical impulses in the body using frog legs. In the 1830’s, Carlo Matteucci, a neurophysiologist, laid a frog leg on a beating heart, and the leg twitched in rhythm with the heartbeats, showing that heart muscles also generate an electric current when they contract.

Opening the chest of a patient and laying a frog leg across the beating heart isn’t considered a safe and easy diagnostic technique, but if the chest weren’t opened up, the electrical impulses from the heart would be too weak to measure. It took another seventy years before the Dutch physiologist Willem Einthoven developed a device to measure those weak currents, and then did the extra engineering to record them.

Figure 2 shows an early ECG machine. The electric currents from the heart cause a very thin wire to move; on the right, a light beam shines on the wire, and on the left, a motorized strip of photograph paper records the movements. And . . . the world’s first human ECG recording, Figure 3.
Technology advances: Einthoven’s moving strip of photograph paper was replaced by a pen moving up and down, writing on a moving strip of paper. The paper has lines and squares to make reading the recordings accurate. And the graph is sharper, showing more detail. (Figure 4).

Now the diagnosis: the patient is suffering from an erratic heart beat, and an MD can tell by measuring the distance between peaks. In Figure 4 each little box is .04 sec wide, so an old-fashioned MD would count the squares between the peaks, and compute how many beats per minute there were. In this (short) graph, an erratic heartbeat shows up as a varying number of beats per minute (for more details, see the note on p5).

That’s what a 1950’s MD would do; even in 2014, in a visit to a hospital, my ECG was on a piece of paper a lot like Figure 4. But it doesn’t help in an emergency room situation; there isn’t time to sit down with a nice piece of paper and count squares. Fortunately, we’re not in the 1950’s; we have computers.

Technology advances: now it’s the engineer’s turn. To keep records of a patient’s heart condition, there’s no need to store dozens of sheets of paper in a file cabinet. Instead of paper, write the electrical signals directly into a computer.

Once the data is in a computer—well, instead of drawing the graph on paper and counting squares, a microchip draws the graph on a LCD screen . . . and it counts the squares, it computes the beats-per-minute.

Today you can buy a cell-phone size ECG unit, as in Figure 5. An MD can get a quick sense of whether a heart is behaving normally, and then, if needed, run a full ECG (which takes 10 minutes or so, requires a nurse, square counting, and $$). Technology is great.

Imagine you’re the engineer programming the chip, working with an MD who also didn’t need calculus,. Before the chip can count squares, it has to somehow ‘see’ the peaks. Calculus tells you peaks happen at critical points, where a derivative is zero. The software to find critical points, using a microchip, was written by Jiapu Pan in Shanghai and Willis Tompkins in Wisconsin, using a derivative-based algorithm. Figure 6 shows an example of the technique; it relies on successive mathematical operations to strip off the irrelevant parts of the beat.

So this engineer and doctor did need calculus for their work, and their paper is now one of the most often-read papers in biomedical engineering.
Notes

Notes for Chapter 0


p3 An records the electrical activity of the heart.

The heart is a collection of chambers to hold blood, muscles to pump blood, and veins and arteries to channel the blood. Muscle contractions are associated with electrical currents (we’ll go into more detail later in the book) and the current is measured by placing electrodes on the body; the record of the current flow is the ECG. A printout gives a visual record of how the different parts of the heart contract, the size of the contraction, and the timing between contractions; over a hundred years of work allows us to diagnose heart problems using this information. The ECG serves as a kind of microscope into the heart, with the advantage that the doctor doesn’t have to do surgery to inspect the internal workings of the heart. It’s a simple, cheap way of monitoring basic health.

p4 The erratic heartbeat in Figure 4 is called fibrillation. As we said, the heart is a series of chambers and pumps, and a working heart chamber pumps only after it gets blood from the preceding chamber. In fibrillation, the pumping happens at random, and the blood can slosh around the heart, never getting to the body. Depending upon the chamber that has the fibrillation, this can lead to death within minutes.
Chapter 1: Numbers

Section 1: Background & History

In the Introduction, I asked a medical question about an erratic heart-beat, and wound up with some numbers and a graph. No-one blinks, because numbers are already everywhere: time, speed, area, temperature, pressure, heart rate, GPA . . .

The issues for this section are: What are numbers? How can they be such a central part of science, technology and even of our lives?

Numbers aren’t things, like spoons, that you can touch, but if numbers are ideas or other kinds of mental construct, how did the idea of number get into our minds? Philosophers have examined these for millennia (see p 13), but this book is about the contributions of modern science. We’ll start with a study of animal intelligence.

A scientist has made holes in a log and randomly places worms in those holes. A robin watches, and when the scientist leaves, the robin immediately goes to the log to munch on worms (Figure 7), and it selects the holes with the most worms.

The study shows that these birds can “count” at least to the extent of perceiving ‘more’ and ‘less.’ Other studies show this ability is cross-species: robins, rats, newly hatched chicks, new-born babies. They won’t pass an algebra test, but something in these tiny brains ‘gets’ number. For more details, see p 13.

‘Number’ seems to be like ‘time’ or ‘space’ – mysterious because each is built into the way our eyes and brains perceive the world (for the research, see p 12). A traditional way to talk about these mental processes is to say that objects in the world have qualities – color and weight are examples of qualities. Our senses perceive these qualities, and our minds form mental versions called qualia (Compare Figure 8). By analogy, when we look at a scene, we perceive a certain number of objects: the quality we perceive is ‘numerosity’. It seems that we, and other animals, perceive the world as having numerosity (For more on numerosity, see p 13).

This may answer what number are, but doesn’t answer out second question: we share numerosity with rats, robins and baby chicks, but they don’t count out change at the local Starbucks, check how fast they’re jogging or buy 4G phones. Humans use numerosity differently than animals do: how is this?
Part of the answer relates to information processing in our brain: we’re not very good at it. Someone on the interstate driving at 50 in a 65 zone and wandering out of their lane is probably on the phone, and they can’t handle two tasks at once.

It isn’t just driving; studies show we can only do a small number of things at once: we can only focus on a small number of features in the environment and can keep only a very small number of facts in our immediate memory. Compared to our computers, our hardware is slow, limited and obsolete. (A good read on these limitations is David Eagleman’s *Incognito: The Secret Lives of the Brain*, Vintage; Reprint edition 2012)

This extends to numerosity: in Figure 9, it’s hard to get the answer by just looking; numerosity fails for large quantities. Recent research suggests the part of our brain that detects numerosity is limited: it runs out of room (see p.14).

But we can answer the question in Figure 9 by counting. One view is that counting is a form of external storage: both counting on fingers and using number words are examples. These are symbolic representations of numbers. Figure 10 shows a third storage trick: making marks (tallies) on wood or paper.

There’s another dimension to this: all external counting can become a public act, entering public memory or becoming stored records. This greatly extends the uses of numbers (see p.14).

Finger counting is very old; some research suggests that numerosity is localized in the intraparietal sulcus, an area of the brain also associated with hand-eye co-ordination. Though finger counting could be an evolutionary device, the problem with fingers and words is the same as the problem with attention: we run out very quickly. What we do then is almost universal: anthropologists have found many societies in which people count higher than ten by referring to parts of their bodies (see p.14).

For example, we can count to five using the fingers of the left hand; at five, the right hand uses one finger to indicate we’ve counted to five. The left hand recycles its fingers and counts to five again; the right hand uses a second finger to indicate we’ve counted two fives; this would be base five counting. Figure 11 shows a counting system, used in markets in the Middle East, which uses two hands to count to sixty, thus, using base sixty.
A body counting system like that in Figure 11 groups certain quantities together. In the example of using one hand to count fives, the numbers we count are naturally grouped into fives. A fully developed system based on counting with the five fingers of the hand would have markers for five (bending the first finger of the right hand; for $5 \cdot 5 = 25$ (bending all five fingers of the right hand), $5 \cdot 5 \cdot 5 = 125$ (perhaps by touching the left ear), etc. We’d call this a base five system.

While systems like this work for personal transactions, they’re of no use for long distance trade – or, for collecting taxes or ruling a country. It’s been argued that public number words or signs make numerosity more useful. Figure 12 gives an excellent example. Denise Schmandt-Besserat of the University of Texas at Austin established the early origins of writing and the use of written numbers in Mesopotamia. Figure 12 depicts clay objects used for record keeping, from Susa, Iran ca 3300 BCE. She writes (see p15):

_The early city states still used tokens to control the levy of dues. When individuals could not pay, the tokens representing the amount of their debts were kept in a round clay envelope. In order to be able to verify the content of the envelope without breaking it, the tokens were impressed on the surface before enclosing them. A cone left a wedge-shaped mark and a disc a circular one. It was the invention of writing._

Our question was, why is numerosity so important an aspect of the human way of life? A partial answer is that the transition to symbolic representations changed numerosity from an event in the brain, to a public act. Counting become part of culture.

There’s a bit more, though: figure 13 shows an early Mesopotamian symbolic system, partially based on tens (see p14). Each round dot represents a ten; each vertical line a one; it would be analogous to our writing $10, 10, 10, 1, 1$. This notation has its advantages; we could just as easily write it $1, 10, 1, 10, 10$. Order doesn’t matter.
The disadvantage comes when we want to write something like 99. One could do what the Romans did: introduce new symbols like V for five and L for fifty, then use what is called a subtractive system, writing IV to mean 5−1. But this gets complicated: you need extra symbols like I, V, X, L, C, M...

Figure 14, from the Chinese Qin and Han dynasties, shows an improvement over the Mesopotamian notation. Scribes used vertical and horizontal lines for numbers; there are basic symbols for one through nine, and small boxes partition the numbers. The position of each box from the left determines 9 vs. 90 vs. 900. In contrast to Figure 13, Figure 14 demonstrates positional notation. And one further development, the use of empty boxes to denote zero.

This system, known as rod arithmetic, allowed the operations of addition, subtraction and multiplication to be performed by arranging numbers on a rod, aligning boxes, and performing the operations as we'd do to day:

\[
\begin{array}{c}
54 \\
23
\end{array}
\]

\[
31
\]

It’s a very fast method for doing arithmetic, and scholars in the Qin Dynasty had need of it: the dynasty was characterized by many monumental building projects, including the Great Wall of China. This was a massive engineering project, requiring thousands and thousands of complex computations.

This showcases a new side of numbers. An official planning a vast engineering project like the Great Wall needs to think to the future: how many bricks, and laborers and jugs of wine and rice will I need for the project? And what taxes will I have to raise? These computations are projections of present knowledge into the future. It is a different kind of thinking than merely accounting for how many bricks or provisions I can see in front of me.

So this is another factor in the advance of numerosity: the development of good notation, and fast algorithms for performing computations, and using numbers to predict the future. All of these are important in modern uses of mathematics, and we’ll see much more of them.

The only difference is how we write our numbers. The writing at the top of Figure 17 shows how numbers were written in early Indian mathematics. The bottom shows the same numbers in Hindu-Arabic notation, the system used in modern Western science and technology.
Historically, though, much of the Chinese numeration system remained in China, which was largely closed off to contact with contemporary civilizations. Thus, modern Western scientific numbers are referred to as ‘Hindu-Arabic’. As Figure 17 shows, the signs themselves were developed in India. Arabic-Islamic civilization made several important contributions; by 662CE use of the Indian notational system had spread west to ancient Iran/Iraq. A mathematician of the Caliph’s court in Baghdad, Muhammad ibn Musa al-Khwarizmi (see Figure 18) wrote The Book of Addition and Subtraction according to the Indian Calculations (the original Arabic manuscript is lost, so we only know the book by its Latin title).

Arabic mathematicians took the Hindu system and added the decimal point, use of decimals to write fractions, a notation for the zero, as well as the algorithms needed to add, subtract, divide and multiply efficiently. The term ‘algorithm’ is, in fact, a European mistranslation of al-Khwarizmi’s name. In al-Khwarizmi’s time, the Caliphate ruled a vast commercial and political empire; ”Arabs ...had been using [Hindu numbers] for centuries to calculate interest, convert currencies, and solve other problems of trade” (see The Crest of the Peacock: Non-European Roots of Mathematics by George Gheverghese Joseph, Princeton University Press; Third edition, 2010.)

The Hindu-Arabic system came to Europe though Leonardo of Pisa. The son of an Italian merchant who traded in the Islamic world, he worked in customs, and to do his work, learned the Hindu-Arabic algorithms. He wrote, The Book of the Abacus, which popularized these techniques for commerce.

Commerce, again? We’ve ignored all the mathematics in modern science; the origins go very far back in the human story. Figure 19 is a copy of an inscribed antler bone, with marking showing the phases of the moon: a lunar calendar (see p15). The bone may date to 32,000 BCE; however accurate the date, humans were clearly exploring the cycles of the heavens for a very long time.

The constellations – Scorpio, Gemini, Capricorn, etc, come to us from Sumeria (modern southern Iraq), about 2000-3000 BCE. Similarly, the Babylonian base sixty counting system; this gave the 360 degrees of a circle.

The origins of these are disputed and possibly lost in time, but written work transmits culture. There is a vast collection of cuneiform texts relating positions of the planets and constellations with earthly events. In other words, the cultures of the fertile crescent practiced divination, the art of reading the future from omens. These omens...
could be patterns in the livers of sacrificed animals, or they could be the patterns of the heavens – that is, astrology. Professor Ulla Koch-Westenholz writes (see p15):

stars and planets were the celestial manifestations of gods, but also seem to have been gods in their own right . . . Sometimes evil omens from a planet were seen as the expression of anger of the god whose celestial image the particular planet was (e.g. Jupiter = Marduk), so that that particular god had to be appeased. In this way, messages could be sent directly from a god to the king, . . . auspicious Venus omens [are seen] as an expression of the love Istar holds for the king.

But early astrologers couldn’t predict the next appearance of a constellation, or an eclipse, or any planetary sign: the best they could do was interpret events after they’d happened. Once again, for a culture that possessed writing and record keeping, this too was useful. Centuries of omens correlated with astronomical observations were recorded in a great compendium, the *Enâma Anu Enlil*, (named after the first words of the tablet, ”the day when the gods Anu and Enlil…”).

For example, Figure 20 records twenty-one years of the rising and setting of the planet Venus; modern astronomers have even worked backwards, to date the tablet; one going back to 1581 BCE (see p15).

Beginning in the Babylonian period (about 1800 BCE), astronomers used the position of planets over many years to develop two systems for making predictions. Asger Aaboe (see p15) called these Systems A and B. Each divided the sky into several regions, and had the planets moving with different speeds through the different parts of the sky. The ratios of speeds was always that of numbers like 3:2, small consecutive integers

Babylonian astronomy may have given the idea ’music of the spheres’. Greek theories of musical harmony also relied on ratios of small consecutive integers. The Greek knew Babylonian tables (see p15); if they knew Systems A and B, they may have believed that the motions of the planets also played out celestial harmonies.

The answer to the central question of this chapter? By 2000 BCE, numbers had become an essential part of trade, government, early scientific astrology, and even religion. And much more was to come.
Supplement to Section 1: Vision

We’ve said numerosity is a perception of qualities in the environment, but can we assume that our senses perceive everything, just as it is? Ethology, the study of animal behavior, shows that many kinds of animals can perceive some aspects of number. That makes it easier to study number, because we can perform experiments on some kinds of animals that would be unethical on cats, chimps and humans. And a very good animal is the frog. Scientists can stabilize a frog, locate the various portions of the eye, the nerves and the centers in the brain that receive the nerve impulses. They then show the frog various kinds of scenes, and see which nerves fire.

The article ‘What the Frog’s Eyes Tells the Frog’s Brain’ (see p16) discusses this research. Figure 21 shows that even the eye has several cell layers; these take the raw response from the rods and cones and shape it. The eye and the optic nerve system decides to pass certain kinds of information on to the brain, and suppress other kinds. We say the eye filters visual information before passing it to the brain (see p16 on filters).

The brain too filters information. Layers and layers of cells respond only to certain kinds nerve firings. These layers determine the qualities a frog can perceive in a scene. What the frog can see is: differences in contrast (an insect standing out from the background?), convexity (shape of a bug?), a moving edge (moving insect?) and dimming of light (predator in back?). See p16 for further reading.

Figure 22 gives an interpretation of what a pond might look like to a frog. It’s nothing like what we would see. Research has shown how very much of the world we see that the frog cannot. It’s equally likely that there’s much of the world we cannot see. Figure 23 shows a picture of a flower in normal sunlight and also in ultraviolet light.

Bees can see ultraviolet, and it’s conjectured that the flower has evolved markings for bees to locate nectar, to pollinate the plant. Locating nectar is not particularly important to us, so our eyes simply don’t see this pattern.

Charles Darwin’s brother once reminded him of Plato’s philosophy, that we understand abstract ideas (like numbers) because our soul has experienced these abstracts before birth. Darwin wrote in his notebooks that instead of a preexisting soul, we have pre-existing monkey ancestors. That is, Darwin contended perception evolved to help organisms survive; in evolutionary terms, our brains compute numerosity because as our species evolved, numerosity helped to survive. And: what we didn’t need, we can’t see.
Notes for Chapter 1 Section 1: Introduction


Augustine wrote of the importance of the anagogic, passages of scripture referring to the afterlife. The author has re-interpreted this to mean predicting the future, in particular, the dates for Christian holy days, e.g., Easter.

p6 Two themes run through this book
i) "How can we know?" about the world, universe, God? The study of this question is epistemology.
ii) "How can math, a purely mental activity, tell us about the physical world?"

These questions pervade the Western intellectual tradition. In 1960 the Nobel physicist Eugene Wigner wrote a paper, The Unreasonable Effectiveness of Mathematics in the Natural Sciences: theoretical physicists can predict effects to eight decimal places of accuracy. How?

p6 Information on the robin experiment, and Figure 7 are from Simon Hunt, Jason Low and K. C. Burns, Adaptive numerical competency in a food-hoarding songbird, Proc. R. Soc. B (2008) 275, 2373–2379


p6 What is ‘numerosity’? One approach is to say that perceiving numerosity is like perceiving color. Neuroscience gives a different approach:

Humans and many other animal species have evolved a capacity to represent approximate number. This ‘number sense’ is at the heart of the preverbal ability to perceive and discriminate large numerosities and relates to the intraparietal sulcus, a brain area which contains neurons tuned to approximate number . . . and which is functionally active already at 3 months of age in humans. Children discriminate numerosity long before language acquisition and formal education, as early as at 3 hours after birth.

Humans view a scene and perceive numerosity, but in some cases, the perception of numerosity is weak; again, see Manuela Piazza et. al., in Cognition 116(2010) p33. Many such people also have difficulty with ordinary calculations, and are said to have dyscalculia. We assume that this is due to genetic differences, but we can also find the area of the brain where numerosity is processed. Processing seems to be related to the intraparietal sulcus, shown in Figure 24; see Brian Butterworth et. al., Dyscalculia: From Brain to Education, Science 332, 1049 (2011).

Wikipedia tells us that the IPS "principal functions are related to perceptual-motor coordination (e.g., directing eye movements and reaching) and visual attention, which allows for visually-guided pointing, grasping, and object manipulation that can produce a desired effect. The IPS is also thought to play a role in other functions, including processing symbolic numerical information, visuospatial working memory . . . .

The research examining the portion of the brain that processes numerosity is in Harvey, B.M., it et. al, Topographic representation of numerosity in human parietal cortex. Science, 341, p. 1123- (2013). Figure 25 shows how the localization is arranged. The higher numbers seem to have less and less area of the brain devoted to their recognition, suggesting this is why number perception fails at high numbers.

The distinction between the internal perception of numerosity, and the public naming of numbers, is at the core of the philosopher Ludwig Wittgenstein’s private language argument: is it possible for a human to have a language that no-one else has? Or does the idea of language necessarily involve rules, rules which only make sense if they are public? See http://plato.stanford.edu/entries/private-language/ or the Wikipedia article on private language.

The attempt to reconstruct an ancestor to all languages is called the ProtoWorld hypothesis; see the Wikipedia article. Reconstructed works are preceded by an asterisk, as in *tik. *tik means a finger, but is also related to indicate, to point, and also to digit, and to the number one.


Figure 13 is from T. Cuyler Young, Jr., of the Royal Ontario Museum, Toronto, Canada, who excavated Godin Tepe in the 1960’s. It is
provided from Denise Schmandt-Besserat of the University of Texas at Austin, in See How Writing Came About, University of Texas Press; Abridged edition (1997).

p8 The quotation is from Schmandt-Besserat’s book From Accounting To Writing, https://sites.utexas.edu/dsb/tokens/from-accounting-to-writing/. In a series of articles, Professor Schmandt-Besserat established the origin of written number systems, leading to written language. These systems evolved from commercial accounting needs in the ancient Near East, from 8000 BCE to 2000 BCE. It seems the use of writing for taxes and government record-keeping followed later.

In Trust In Numbers (Princeton University Press, 1996), Theodore M. Porter argues:

My approach here is to regard numbers, graphs, and formulas first of all as strategies of communication. . . . only a very small proportion of the numbers and quantitative expressions loose in the world today make any pretense of embodying laws of nature, or even of providing complete and accurate descriptions of the external world. They are printed to convey results in a familiar, standardized form, or to explain how a piece of work was done in a way that can be understood far away. They conveniently summarize a multitude of complex events and transactions.

p10 This interpretation is due to Alexander Marshack; see The Roots of Civilization: The Cognitive Beginnings of Man’s First Art, Symbol and Notation, Moyer Bell Ltd, December 1991.


Notes for Supplement to Section 1: Vision


The idea of filtering our senses, to pick out the parts we need to survive is the same kind of filtering as Pan and and Tompkins used to reduce a complicated heart beat, to a simple series of peaks, thus finding the main part of the heartbeat; see Figure 26. Pan and Tompkins are applying one filter after another to get the exact time when the heartbeat happens.

The frog work is from the late 1960’s; we now know a great deal more about vision. The article *The Movies In Our Eyes*, Scientific American, April 2007, discusses some of this research. The authors, Frank Werblin and Botond Roska write:

*The retina actually performs a significant amount of preprocessing right inside the eye and then sends a series of partial representations to the brain for interpretation. We came to this surprising conclusion after investigating the retinas of rabbits, which are remarkably similar to those in humans.*

The retina sends a series of images to the brain; the authors call these ‘movies’ as the images change over time (see Figure 27). They identify twelve different kinds of moves that the retina generates and sends on: some show the edges of a scene, some show brightness, or reflectance, and some show information we have no name for. The brain then integrates these movies into the vision we perceive.

Once again, we are limited and perceive only what our eyes pass on.
Chapter 1: Numbers

Section 2: Building The Numbers

We’ve looked at numbers from the perspectives of history, neuroscience and frogs; it’s time to see how mathematicians have thought about numbers through the millenia.

We’ll start with the numbers \{1, 2, \ldots\}, also called the counting numbers. As we’ve seen, these appeared early human history, which makes them seem – well, so obvious that they don’t require further thought. Mathematically, they’re undefined, and it took serious work to get at the ideas behind the ‘obvious’.

The counting numbers are called the natural numbers, denoted by \mathbb{N}, and intuitively given as a list:

\[ \mathbb{N} = \{1, 2, 3, \ldots\} \]

If we want to include the negatives and zero, we have the integers,

\[ \mathbb{Z} = \{0, \pm 1, \pm 2, \pm 3, \ldots\} \]

The \mathbb{Z} here is from the German word Zählen, or ‘numbers’. None of this, though, tells us what numbers are, and none of this tells us what’s actually on the infinite list \{1, 2, 3, \ldots\}.

One approach was to think about the meaning of ‘counting’. Gottlob Frege (a German mathematician working in the mid nineteenth century) thought back to a child pointing at objects one at a time and saying the numbers ‘one, two, three’: the process of counting is that of matching one collection, say toys, with another, perhaps fingers, or a list of number words. If we match the two collections \{\Sigma, \Delta, \Gamma\} and \{\infty, +, \int\} we’d say they have the same number of items. This process partitions the world into collections which have the same number, and we could define three to be the collection of all collections that have the same number as \{\infty, +, \int\}; see p29.

Another approach follows a very old idea; we know Aristotle wrote of it, and it was common in early Greek mathematics. The number ‘1’ was considered the father of all numbers, because all numbers could be generated from him. And we wouldn’t consider ‘1’ as a number at all, as a father is not one of his own children.

A thousand years after Aristotle, the mathematician al-Khwarizmi wrote:

Because one is the root of all numbers; number is nothing but a collection of ones.
The Italian mathematician Giuseppe Peano used this approach to define the natural numbers through axioms $P_1, P_2, \ldots P_4$:

$P_1$: If $n$ is a natural number, then $n$ has a successor. Peano denoted it as $n'$; for us, it means $n + 1$.

$P_2$: If $n + 1 = m + 1$, then $n = m$.

$P_3$: There’s a unique natural number $n$ which is not the successor of any other natural number. We denote this as the number ‘1’.

These axioms aren’t enough; a collection like \{\ldots, −\frac{1}{2}, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots\} has the same property, so do many others. These all contain extra numbers; we eliminate the extras by defining the natural numbers as the smallest collection with the ‘plus one’ property. The trick to define ‘smallest’ is to state that there are no smaller collections with the ‘plus one’ property:

$P_4$: Assume $S$ is a collection of natural numbers that contains 1. Assume it also has the ‘plus one’ property: whenever $n$ is in $S$, $n + 1$ is also in $S$. Then $S$ is actually all of $\mathbb{N}$.

Axiom $P_4$ has another name, the Axiom of Induction. But it isn’t enough: are there any collections, where all four axioms are true? Sure there are: take $\mathbb{N} = \{1, 1 + 1, 1 + 1 + 1, \ldots\}$ and so on. But it’s the ‘\ldots’ and the ‘and so on’ that trip us up (see p29).

‘\ldots’ means, of course, ‘continue this forever,’ but we’d have to understand ‘forever,’ we’d have to understand infinity. That’s the fifth axiom:

$P_5$: An infinite set exists.

This is called the Axiom of Infinity; see p29 for more details.

Now we know what the counting numbers are, what do we really know? From Frege’s construction, we’d have the number 2 as the super-collection of all "two" sets: \{1, 3\}, \{\pi, e\}, \{\Gamma, \sigma\} \ldots. Oh, that really helps. Or, we have it as \{1, 1 + 1, 1 + 1 + 1, \ldots\}. Again, not helping.

Base ten notation organizes these infinite lists into comprehensible chunks, as seen in daily life: cash a check for $432$, you get four hundreds, three tens and two singles, rather like in Figure 28. Instead of repeating three tens, we write repeated addition as multiplication: \[30 = 10 + 10 + 10 = 3 \cdot 10\]. Thirty three is then \[33 = 3 \cdot 10 + 3\].
To write ten tens, we think of repeated multiplication as exponentiation: \(10 \cdot 10 = 10^2\). This series of notations is the decimal system; written as a decimal, our cashed check becomes

\[4 \cdot 100 + 3 \cdot 10 + 2 = 6 \cdot 10^2 + 3 \cdot 10^1 + 2 \cdot 10^0\]

Decimals make it easier to write large numbers, but they also group quantities by size. For example, a burger is less than ten dollars, local concerts less than a hundred, books for a semester less than a thousand, and so on. These gradations are called cognitive reference points, and research shows we use these reference points naturally (see p30).

But we need to prove that every natural number can be written with decimals; that’s actually tricky. The Axiom of Induction suggests a way: to show that the collection of decimals is the same as the collection of natural numbers, we’d show the number ‘1’ is a decimal (easy: \(1 = 1 \cdot 10^0\)). We’d also show that if you can write every number \(n\) as a decimal, you can write \(n + 1\) as a decimal.

That’s easy for numbers like 43 and 44: if we know \(43 = 10 \cdot 4 + 3\), then

\[44 = 43 + 1 = (10 \cdot 4 + 3) + 1 = 10 \cdot 4 + 4\]

The problem is with numbers like 99, or 999, even 1276899: you have to deal with carrying a one, and each of the numbers 99, 999, 1276899 is different; there’s no general pattern. So it’s all a little harder: what we actually need to do is rethink induction. The right technique here is called strong induction; see p30.

Once you have that, you see that the powers of 10 measure all the natural numbers: for example, \(100 \leq 632 \leq 1000\). This is called

**The Archimedean Principle:** If \(n\) is a natural number, then there is always a power of ten, \(10^m\), with \(10^m \leq n < 10^{m+1}\)

Now that we (finally) know the counting numbers are decimals, the next issue is fractions. Easy: fractions are quotients of integers, \(\frac{p}{q}\), so we write the collection of all fractions as \(Q\) (Quotients). Getting to decimals turns out to be difficult. Once again, a journey through history.

As we saw, Mesopotamian scribes used written numbers for accounting (see p30). Workers (or forced laborers) were paid in standard-sized units of grain, or bread, or beer. The difficulty was in dividing: say three standard loaves divided among four workers. Each worker gets \(\frac{3}{4}\) of a loaf, but how: do you cut each loaf in half, then half again, to get fourths? Then each worker gets three little slices? Let’s not even try jugs of beer. The scribes needed more than the ability
to write fractions like \( \frac{2}{3} \); they needed to compute with fractions, and relate those computations to other kinds of numbers.

Figure 29 is a papyrus which showed archaeologists how the Egyptians did exactly that: all fractions were rewritten as sums of unit fractions. Unit fractions are fractions with one as the numerator, so, instead of writing \( \frac{3}{4} \), the Egyptian system was to write \( \frac{3}{4} = \frac{1}{2} + \frac{1}{4} \). This method also makes the division of the loaves of bread more practical: everyone gets a half loaf instead of just small pieces.

In Mesopotamia, scribes used base sixty, so they wrote a fraction like \( \frac{1}{12} \) as \( \frac{5}{60} \), and \( \frac{4}{45} = \frac{1}{12} + \frac{1}{120} = \frac{5}{60} + \frac{20}{60^2} \). Or, in their notation, 5, 20.

They also knew that when you divide 7 into 1, the division never ends; in modern notation, \( \frac{1}{7} = 0.142857142857142857 \) and so on. The "and so on" can be taken to mean that if you continue to divide, you will continue to get blocks of 142857.

The decimal is infinite – whatever that means. To avoid even thinking about ‘infinite’ we say \( \frac{1}{7} \) is a repeating decimal (see p30). The standard notation for a repeating decimal is \( \frac{1}{7} = 0.\overline{142857} \).

There: problem solved. You don’t have to think about infinity, just draw lines over some numbers. You can even say every fraction \( \frac{p}{q} \) is a repeating decimal – at least if you’re willing to write \( \frac{12}{10} = 1.2 = 1.2\overline{0} \). This follows from the

The Archimedean Principle for Rationals: If \( r = \frac{p}{q} > 1 \) is a rational number number, then there is always a power of ten, \( 10^m \), with \( 10^m \leq n < 10^{m+1} \). If in addition \( 0 < r < 1 \), there is always a negative power of ten, \( 10^{-m} \), with \( 10^{-(m+1)} < n \leq 10^{-m} \).

See p31 for the idea of the proof.

Mesopotamian scribes also had techniques for these fractions; they’d write something like \( \frac{1}{7} = 0.142857 \) but warn "approximation given since 7 does not divide". They were running a government, so it was a practical world; the scribes needed quick answers and techniques they could use. Infinity was just a nuisance.

But a thousand years later, an unknown scribe wrote

\[ 8, 34, 16, 59 < \frac{1}{7} < 8, 34, 18 \]

The decimal version is

\[ 0.14285640... < \frac{1}{7} < 0.1428611 \]

There’s a message to decipher in this progression of ideas: when we write \( \frac{1}{7} = 0.142857 \), we don’t know how accurate this is. For example,
if \( \frac{1}{7} \) is the amount of tax on a piece of land, and you’re a government, you want the largest number you can get away with (rounding up). If you’re the one paying that tax, you want the smallest (rounding down). Where does 0.142857 fit? The number all alone doesn’t say.
Writing \(0.14285640... < \frac{1}{7} < 0.1428611\) does say: it tells you the largest and smallest value you could take, but it doesn’t leave you with just one number. Our practical scribes had a solution: take the average, 0.14285875. Now we have a number to use, and we know the largest and smallest variations.

Moderns write it differently; we’d say 0.14285875 is an approximation to the real value of \(\frac{1}{7}\), but that it isn’t the real value. The way to talk about approximation versus real value is to introduce the idea of error: error = real value - approximation. Here, the error is \(\frac{1}{7} - 0.14285875\).

This doesn’t seem to help, because we don’t know the real value of \(\frac{1}{7}\). But we do know how large and how small \(\frac{1}{7}\) could be:

\[
0.14285640... < \frac{1}{7} < 0.1428611
\]

So

\[
0.14285640... - 0.14285875 < \frac{1}{7} - 0.14285875 < 0.1428611 - 0.14285875
\]

And, rewriting, \(-0.00000235 < error < 0.00000235\), or as we’d write it today, \(|error| < 0.00000235\) (actually, \(|error| < 2.35 \times 10^{-6}\)).

How does this help? Imagine some lowly scribe presenting the taxes to his boss. The boss remarks, “You have taken the seventh part; perhaps the tax is too small.” But now our scribe can bow low and say, “Oh Shining One, the tax on this land is ten bushels of rice, and the error is but a part of one grain of rice”.

This is an important idea: we don’t really know what numbers mean, until we know the error in computing them. Controlling error is central to modern science, technology, and industry, and it was invented thousands of years ago in Mesopotamia.

The Chinese scholar Lui Hui (Figure 30) expressed similar ideas when calculating the value of \(\pi\): he used the approximation \(\pi \approx 3\) and added the comment that this was not the true value, but was good enough for most practical purposes (see p31).

Lui Hui also added a comment: he said that he’d estimated \(\pi\) by computing the area of a 96-agon inscribed inside a circle (a 96-agon is a 96-sided figure, so triangles, squares, are 3-agons, 4-agons). He gave a formula for going from one approximation to a better one: see Figure 31, where he goes from a 6-agon to a 12-agon. He then goes to a 24-agon, a 48-agon, and finally a 96-agon.

A new idea: not just an approximation, but a way to get better and better approximations; smaller and smaller errors. It’s the beginnings of the idea of a limit.
So we’re good: numbers act as public symbols making them useful for commerce and government; many cultures developed efficient methods of writing and computing with numbers. What more do we need?

We perceive numerosity, but also qualities like size, position, length, area, angle, volume, weight, ... We’re still on the question of why numbers are prevalent in the modern world; one reason is that they can be linked with these other kinds of qualities. That link creates a set of associations that leads to an entirely new kind of number.

Geometric qualities: length, distance, area, volume ... Why? Because bureaucracies always need taxes, but how do you compute the tax? If you just take part of the harvest from a farmer, the farmer can hide the harvest before the tax collector visits. Instead, compute the area of the land planted, then compute how much produce the land should yield, then take part of that. But ancient inheritance involved subdividing land amongst many children, so taxable lands had complicated shapes. Figure 32 shows an Egyptian computation of the area of a fairly complicated figure. Figure 32 shows the state’s interest in geometry.

Areas: the area of a triangle is half the base times the height, so geometers had to understand the connection between numbers and lengths of lines. This links counting with measuring; quantity and geometry. We use yardsticks all the time, but the link was less obvious in ancient times: Figure 33 shows surveyors’ tools as gifts from the Gods. People took their lengths seriously – and the state taxation depended on public trust in those tools (not exactly the first time we’ve run across the idea of public standards).

Figure 34 shows a problem in geometry from Mesopotamia about 1700 BCE. It’s a right triangle, with two equal sides, each side length one. The Pythagorean theorem tells us the square of the hypotenuse is $1^1 + 1^2 = 2$, but what’s the length of the hypotenuse? On the tablet, in base 60, it’s written 1, 24, 51, 10; in decimal notation, 1.41421. We’d write the answer as $\sqrt{2}$. The Mesopotamian answer in Figure 34 looks like the kind of approximations we used for $\sqrt{2}$: nothing new or surprising here.

How did the Mesopotamian scribes get their approximations? No one knows, but here’s the best idea historians have come up with; at least it gives the answer the Mesopotamian scribe wrote.
Start with a first guess for $\sqrt{2}$: say, $\frac{3}{2} = 1.5$. That’s too big, but $\frac{2}{\frac{3}{2}} = \frac{4}{3} = 1.33\ldots$ is too small. Take the average; that’ll be in-between, so it will be closer to the true value than either guess. So, if $g_1$ is the first guess, then you get a second, better guess with

$$g_2 = \frac{1}{2} \left( g_1 + \frac{2}{g_1} \right)$$

With $g_1 = \frac{3}{2}$, then

$$g_2 = \frac{1}{2} \left( \frac{3}{2} + \frac{2}{\frac{3}{2}} \right) = \frac{1}{2} \left( \frac{3}{2} + \frac{4}{3} \right) = \frac{17}{12}$$

And again:

$$g_3 = \frac{1}{2} \left( \frac{17}{12} + \frac{2}{\frac{17}{12}} \right) = \frac{1}{2} \left( \frac{17}{12} + \frac{24}{17} \right) = \frac{577}{408}$$

How good are these approximations? Square them, and compare with $2$:

$$g_1^2 = \frac{9}{4} = 2.25; \quad g_2^2 = \frac{289}{144} = 2.00694; \quad g_3^2 = \frac{332929}{166464} = 2.0000060073\ldots$$

We’re back in the Lui Hui situation: not just an approximation, but a way to get better and better approximations. And as before, you carry out as many decimal places as you need, get on with your job, report to the Chief Scribe. Get your beer ration.

Written history tells us that Greek mathematicians discovered $\sqrt{2}$ is not like $\frac{1}{7}$: it cannot be written as a quotient $p/q$, and it is not a repeating decimal (for the proof, see p31 ). As the number is not a rational, it came to be called ir-rational (the same construction is used in the word ‘irrelevant’; it means not relevant. Using ‘ir’ to mean ‘not’ is a left over from Latin). For some of the history of irrationals, see see p32.

So: what is $\sqrt{2}$? We could still say it’s like $\frac{1}{7}$, because you can get more and more digits of $\sqrt{2}$. While you can’t get them by long division, at least you have a process, giving the numbers $g_1, g_2, \ldots$ whose squares approximate $2$.

But this is a kind of fraud. When you do the long division for $\frac{1}{7}$, you can see exactly where the decimal starts to repeat and why. With the $g_1, g_2, \ldots$ you don’t know what the $g_1, g_2, \ldots$ are going to do, or why. Could $g_4$ be $\frac{17}{12}$ again? Then $g_5$ would be $\frac{577}{408}$ again, and you wouldn’t get better approximations. How can you rule that out? And even if it does get ‘better and better’, exactly what is getting better, and what is it getting better to?

These were difficult questions (the final answer came something like 2000 years later) and the most significant response, historically, was
to avoid the question. The Greek mathematician Eudoxus of Cnidos (408 - 355 BCE) undid the link between quantity and geometry by developing a consistent theory of magnitude. A magnitude was an undefined term that could express geometric ideas such as lengths, areas and volumes of geometric figures; Eudoxus showed how to manipulate magnitudes as ratios of line segments, analogous to manipulation of numbers. For example, we can define $\frac{m}{n} = \frac{p}{q}$ to mean $mq = np$, and for much of Greek mathematics, fractions were thought of as ratios. Hence, our quotients $\frac{p}{q}$ to the Greeks were the rational numbers, because they were ratios (‘rational’ is from the word ‘ratio,’ meaning in Latin ‘to compute’ and in Proto Indo-European, *reh-, to ‘put in order’; see Wiktionary). Ratios were enough to do the geometry Eudoxus and most Greek mathematicians wanted, for example, the construction of figures using a ruler and a compass (see p 32).

The Eudoxian theory was influential for centuries; even Newton, in his Arithmetica Universalis of 1707 defined numbers as ratios of line segments. The prevailing opinion was stated by the German mathematician Michael Stifel (1486-67), who was critical of using approximations to define an irrational:

\[ \ldots \text{considerations compel us to deny that irrational numbers are numbers at all. To wit, when we seek to subject them to [decimal representation] \ldots we find they flee away perpetually, so that not one of them can be apprehended precisely \ldots Now that cannot be called a true number which which is of such a nature that it lacks precision \ldots so an irrational number \ldots is hidden in a kind of cloud of infinity.} \]

In Europe there were no alternatives to Eudoxus for over a thousand years. But the mathematician Leonardo of Pisa, also known as Fibonacci, was aware of Arabic work on algebra; in 1225 CE, he published the solution to a problem mentioned by Omar Khayyam, in his book Al-jabr: solve the equation $x^3 + 2x^2 + 10x = 20$ (see Figure 35).

Fibonacci showed there were no integer solutions, no rational solutions, and that the solution could not be constructed by ruler and compass. So the number was irrational, but of some unknown kind.

The Mesopotamian scribe on p 30 would probably shrug his shoulders; what did it matter, as long as he could compute three or four digits of these numbers, and keep the Chief Scribe happy? But for us it’s more difficult: am I going to run into new kinds of irrationals each time I solve a new equation?

Figure 35: Solving The Cubic
Five hundred years after Omar Khayyam, the Italian mathematician Girolamo Cardano found a formula for finding roots of cubics. Here, the formula is applied to Khayyam’s equation by the Wolfram Alpha computer program.
In an attempt to create some kind of order, mathematicians began to rethink their irrationals. Fractions like $\frac{1}{7}$, and irrationals like $\sqrt{2}$, $\frac{1+\sqrt{5}}{2}$, ... and even Fibonacci’s irrational, are all solutions of equations:

$$7x - 1 = 0, \quad x^2 - 2 = 0, \quad x^2 - x - 1 = 0, \quad x^3 + 2x^2 + 10x - 20 = 0$$

Solutions were designated algebraic numbers, since they could be obtained by solving algebra equations (note the shift from numbers that come from geometry, to numbers that come from algebra).

$\pi$, $e$ didn’t seem to be like algebraic numbers; the mathematician Euler remarked (1744) that these two seemed to go beyond the techniques of algebra. As the Latin for ‘go beyond’ is ‘transcend’, Euler suggested that these two were transcendental numbers. In 1878 $e$, and in 1882 $\pi$, were each shown to be transcendental.

Of course $\pi$ is a solution to the equation $\cos(\theta) = -1$, and $e$ to the equation $\ln(x) = 1$. Again, new equations, new irrationals. By this time, European mathematicians knew many new kinds of functions, for example the Bessel function $J_0(x)$, which describes the pattern of rings when light is diffracted through a small hole (see Figure 36). All light through microscopes and telescopes become diffracted, and the presence of the first dark ring determines how far apart two objects have to be to seem distinct through the lens. The dark rings occur when the intensity is zero, that is, $J_0(x) = 0$. Is this going to involve totally new kinds of numbers?

Just how many kinds of irrational are there?

There’s another issue: we know we can get more and more decimal places of $\frac{1}{7}$, and we believe we can do that for $\sqrt{2}$ and $\pi$, but what about these new numbers? Do we even know they’re decimals? Answering this question leads us to visit a night-mare world: chains of numbers with no decimal expansion, ascending and descending into the infinitely large and infinitely small.

To enter this world, all you have to do is imagine .9999 ... doesn’t equal 1. Intuitively, the decimal ‘never gets there.’ Then there’s a gap: the number $\gamma$ measures the size of the gap, with $\gamma = 1 - .9999 ... > 0$. What kind of things are inside that gap? We’re going to discover that there’s a lot of numbers inside there – and that none of those numbers are decimals. This is the problem: the world of numbers might be strange beyond imagining.

Let’s look a bit at $\gamma$. How big is it? What’s in its decimal expansion?

If the decimal expansion of $\gamma$ starts with something like .276 ..., then that first decimal place makes $\gamma > .1$. But $\gamma = 1 - .9999 ... = \ldots$
$1 - .9 - (.09999\ldots) < 1 - .9 = .1$ We can’t have $1 < \gamma < .1$, so $\gamma$ has to start with something like $\gamma = .0276\ldots$ The problem is that $\gamma = 1 - .9999\ldots < 1 - .99 = .01$ and again, we can’t have $0.01 < \gamma < .01$, so $\gamma$ has to start with something like $\gamma = .00276\ldots$

The problem is, actually, that this never stops: $0 < \gamma < \frac{1}{10^k}$ for all $k$. $\gamma$ has the decimal expansion $\gamma = 0.0000\ldots$ but still $\gamma$ isn’t zero. So we have what we feared: a number that doesn’t have a decimal expansion at all. So much for our scribe and his beer ration.

We could say $\gamma$ is an infinitely small non-zero number. Then $\frac{\gamma}{2} > \frac{\gamma}{3} > \frac{\gamma}{4} > \ldots$ are also infinitely small numbers with no decimal expansion. And so are $\gamma > \gamma^2 > \gamma^3\ldots$.

So we don’t have just one infinitely small number, we have whole chains of them, getting smaller and smaller. We’ll see it get’s worse – infinitely worse.

Since $\gamma < \frac{1}{10^n}$ for all $n$, then $\omega = \frac{\gamma}{10^n}$ for all $n$. Again, this means $\omega$ can’t start with $\omega = 1.374\ldots$, and it can’t start with $\omega = 1384.732\ldots$ because $1384.732\ldots < 10,000 = 10^4$ but $\omega > 10^4$. So $\omega$ can’t start with any numbers before the decimal point: $\omega$ is infinitely large. And so are $\omega^2 < \omega^3\ldots$.

And as if those aren’t large enough, $\mu = \omega^\omega$ is larger than all of them. Now we start again and do $\nu = \mu^\omega\ldots$ and we get chains of larger and larger infinities.

Now let’s get more infinitely small numbers:

\[ \gamma > \gamma^2 > \gamma^3 > \ldots > \gamma^\omega > \gamma^\mu > \ldots \]

The gap contains a nightmare of infinities!

There’s supposed to be a way out, if you know the theory of limits. We’re supposed to know $0.999\ldots = 1$ because the collection \{.9, .99, .999\ldots\} has 1 as a limit. Let’s try that. First, a little notation:

\[
.9 = \frac{9}{10}; .99 = \frac{9}{10} + \frac{9}{10^2}; .999 = \frac{9}{10} + \frac{9}{10^2} + \frac{9}{10^3} \ldots
\]

Now we can talk about the limit: to say the limit of the \{.9, .99, .999\ldots\} is 1, is to say that for every $\epsilon > 0$ there’s a point after which the sequence is at least $\epsilon$ close to 1:

\[
|1 - \left(\frac{9}{10} + \frac{9}{10^2} + \frac{9}{10^3} \ldots + \frac{9}{10^k}\right)| < \epsilon
\]

But

\[
1 - \left(\frac{9}{10} + \frac{9}{10^2} + \frac{9}{10^3} \ldots + \frac{9}{10^k}\right) = \frac{1}{10^k}
\]
So we’re saying there’s a point after which \( \frac{1}{10^k} < \epsilon \). But, this is the whole thing about numbers like \( \gamma \): if \( \epsilon \) is one of our infinitely small numbers, it’s the other way around: \( \epsilon < \frac{1}{10^k} \).

Not only do numbers like \( \gamma \) mess up ideas about decimals, they mess up the whole theory of limits.

Maybe our logic is bad? We can find a mistake in one of those computations? It would tell us no such \( \gamma \) could ever exist?

No.

The British mathematician John Conway actually invented a system, called the surreal numbers, \( S \), which have these infinitely small and infinitely large numbers. There is no simple way to eliminate the surreals; you have to define real numbers in a way that will exclude them right from the start. Not to be a broken record: we are making a choice; we are constructing real numbers to be the way we think they should be. The next section shows how it was done.
Notes for Chapter 1 Section 2: Building The Numbers

p17 Frege’s construction is based on the idea of equivalence of sets. Sets $A$ and $B$ are equivalent if there exists a mapping $f : A \rightarrow B$ which is one to one and onto (for each element $b \in B$, there is one element $a \in A$ with $f(a) = b$ (onto), and there is only one such $a$ (one to one). The collection of all sets equivalent to $A$ is called the equivalence class of $A$, denoted $[A]$, and the collection of equivalence classes defines the numbers.

Soon after Frege’s work was published, as Die Grundlagen der Arithmetik (The Foundations of Arithmetic), the British philosopher Bertrand Russell noted that phrases such as ‘the collection of all collections’ lead to paradoxes. See Russell’s Paradox on Wikipedia.


p18 At the time European mathematicians were thinking about the construction of the natural numbers by “continuing in the same way’, the Austrian philosopher Ludwig Wittgenstein began his examination of the foundations of mathematics, by thinking on the idea of ‘following a rule’, or ‘going on in the same way’. From the Stanford Encyclopedia of Philosophy (http://plato.stanford.edu/entries/wittgenstein/)

One of the issues most associated with the later Wittgenstein is that of rule-following. [...] Wittgenstein begins his exposition by introducing an example: ‘... we get [a] pupil to continue a series (say +2) beyond 1000 – and he writes 1000, 1004, 1008, 1012Ö. What do we do, and what does it mean, when the student, upon being corrected, answers “But I did go on in the same way”? Wittgenstein proceeds [to ask] How do we learn rules? How do we follow them? [...] Are they in the mind, along with a mental representation of the rule? Do we appeal to intuition in their application? Are they socially and publicly taught and enforced?


p18 The axiom that infinite sets exist is phrased a bit oddly. It states that exists a set $S$ and a map $f : S \rightarrow S$ such that $f$ is one to one but not onto. Here’s the idea: let $S = \{1, 1 + 1, 1 + 1 + 1, \ldots\}$, and define $f$ as $f(n) = n + 1$. By the Peano axiom P1, every $n$ has a successor $n + 1$, so $f$ is defined for every $n$. To show $f$ is one to one, let $f(n) = f(m)$; then $n + 1 = m + 1$. By Peano P2, $n = m$. To show $f$ is not onto, assume there’s an $n$ with $f(n) = 1$. But this means
$n + 1 = 1$, or, in the terms of the Peano axioms, ‘$1$’ is the successor of $n$. But Peano $P3$ says ‘$1$’ is not the successor of any $n$.

The Axiom of Infinity was introduced by the Hungarian mathematician John von Neumann, as part of a very concrete construction of the natural numbers; he used set theory to construct standard meanings for 1, 2, etc. If $\phi$ denotes the empty set, then $\phi$ is a good candidate to be zero, as $\phi$ has zero elements inside. Then $\{\phi\}$ is a candidate for the number one, as it has one element in it. We could also write:

$$1 = \{0\} = \{\phi\}; \hspace{0.5cm} 2 = \{0, 1\} = \{\phi, \{\phi\}\}; \hspace{0.5cm} 3 = \{0, 1, 2\} = \{\phi, \{\phi\}, \{\phi, \{\phi\}\}\}$$

Now we see how to continue, and the Axiom of Infinity allows us to ‘keep on going’ giving us the infinite set $\mathbb{N}$. In the study of infinities, $\mathbb{N}$ can be proved to be the ‘smallest’ infinite list. It’s a bit tricky to define both what an infinite list might mean, and how one infinity could be smaller than another. These matters were addressed by the German mathematician Richard Dedekind, who wrote the influential book *Was sind und was sollen die Zahlen?* (roughly, "What are numbers and what should they be?"). The individual who did most to clear up issues about infinities of different sizes was the mathematician Georg Cantor, who developed the theory of transfinite numbers in the late nineteenth century.

In cognitive science, numbers like like 10, 100, 1000... are called cognitive reference points. Eleanor Rosch did a series of experiments suggesting that these numbers were preferentially referred to when subjects thought about a collection of random numbers such as 102, 173. See Eleanor Rosch, *Cognitive Reference Points*, Cognitive Psychology 7, 1974 p532.


Phrases like "$1/7$ does not divide" go back to the Old Babylonian Period (3200BCE-1600 BCE); the use of approximations for $1/7$ appear in the Seleucid Period (312 BCE- 62 BCE). See The Crest of the
The Archimedean Principle for rational numbers begins with the division algorithm (not surprising, as \( \frac{p}{q} \) is a division!). Roughly, a number like \( \frac{80}{9} \) can be written as \( 8 + \frac{8}{9} \); the second term is a fraction less than one. This means that when \( 10^0 \leq 8 < 10^1 \), it’s still true that \( 10^0 \leq 8 + \frac{8}{9} < 10^1 \).

The second half, dealing with \( 0 < r < 1 \), follows by applying the the Archimedean Theorem to \( \frac{1}{r} > 1 \) and then inverting the inequalities.

The work of the mathematician Lui Hui appeared in commentaries and solutions to the Chinese text *The Nine Chapters on the Mathematical Art*, written in 263 CE.

A photograph of the Yale Babylonian Collection’s Tablet YBC 7289 (c. 1800 to 1600 BCE), showing a Babylonian approximation to the square root of 2 (1,24,51,10 base 60). Babylonian mathematicians knew Pythagoras’ Theorem relating the sides of a right triangle. The photo is by Yale professor Bill Casselman; see http://www.math.ubc.ca/~cass/Euclid/ybc/ybc.html).

The result that \( \sqrt{2} \) is not a rational is contained in the works of the mathematician Euclid; the following proof is a modernized version of the one he gave.

The first result we need is that every integer is either even (has a factor of 2: 2, 4, 6, 8, …) or odd (has no factor of 2: 1, 3, 5, 7, …). This is clear enough from the list, and is easy to prove rigorously, using induction. It’s equally easy that the squares of even numbers are even (4, 16, 36, 64, …), and the squares of odd numbers are odd ((1, 9, 25, 49, …), and proving this is just a little bit of algebra. But it’s important, because it allows us to do square-root things: normally, if \( n \) had a factor of 2, all we could say that \( n = \sqrt{n^2} \) has a factor of \( \sqrt{2} \). But, using even/odd, we can do more: if \( n^2 \) has a factor of 2, then \( n \) also has a factor of 2.
The last result we need is that you can simplify fractions by canceling out common factors: \( \frac{60}{36} = \frac{30}{18} = \frac{15}{9} = \frac{5}{3} \) and now there’s no longer any common factors. This is a bit harder to prove; it needs our strong induction.

Now we’re ready to start: If you can find \( \sqrt{2} \) as a fraction \( \frac{p}{q} \), then

\[
\frac{p}{q} = \sqrt{2} \quad \text{so} \quad \frac{p^2}{q^2} = 2 \quad \text{so} \quad p^2 = 2q^2
\]

This shows \( p^2 \) has a factor of 2, so \( p \) also has a factor of 2; write it \( p = 2r \). Then \( p^2 = 2q^2 \) becomes \( (2r)^2 = 2q^2 \) or \( 4r^2 = 2q^2 \). Divide by 2 to get \( 2r^2 = q^2 \); now \( q^2 \) has a factor of 2, so \( q \) also has a factor of 2. But we already cancelled all the 2’s, and so, no such \( p \) and \( q \) can exist.

\( \sqrt{2} \) was not the first irrational discovered by the Greeks, and it seems likely that the original proof was more a geometric sketch than the kind of very careful proof Euclid wrote.

p24 The discovery of irrational numbers is attributed to the philosopher Hippasus of Metapontum; the irrational in question was likely \( \frac{1 + \sqrt{5}}{2} \), derived from a pentagram, but much of this is lost to history. He lived in the late fifth century BC (that is, from 500 BCE to 401 BCE, closer to 401), and was a member of the Pythagoreans, an ascetic and mystic sect who believed that certain numbers represented maleness, others justice, etc. Ratios of integers explained musical harmony and the ‘music of the spheres’. They explained all the world by integers, so the discovery of irrationals was a serious challenge to their beliefs.

p25 Irrationals appeared in Book X of Euclid’s Geometry (around 300 BCE). These came from ‘ruler and compass’ constructions: for example, a right triangle with side lengths one has hypotenuse \( \sqrt{2} \). Continuing on, you can get \( \sqrt{1 + \sqrt{2}} \), \( \sqrt{5 + \sqrt{3}} \), \( \sqrt{\sqrt{5} \pm \sqrt{3}} \). Euclid showed that the theory of magnitudes could generate these; his irrationals all came from geometry.

Both Archimedes and Hero of Alexandria worked on finding the value of \( \pi \), the ratio of the circumference of a circle to its diameter, or the area to the square of the radius. They approximated the circle by \( n \)-agons, dissected these into triangles, as in Figure 40, and computed areas or lengths of their sides; either led to irrationals. But they needed numbers; both used variations on the estimates

\[
p + \frac{r}{2p + 1} \leq \sqrt{p^2 + r} \leq p + \frac{r}{2p}
\]

While the estimate is reminiscent of the Mesopotamian computations for \( \frac{1}{7} \), the computation was motivated, once more, by geometry.
Figure 41 gives a geometric proof that \((p + q)^2 = p^2 + 2pq + q^2\): it shows that if you start with a square of area \(p^2\), you can get a square of larger area \((p + q)^2\) by adding two rectangles of area \(pq\), and a square of area \(q^2\).

Here’s how this gives us our approximation: \(\sqrt{(p + q)^2} = p + q\), so 
\[\sqrt{p^2 + 2pq + q^2} = p + q.\] If \(q\) is a small number (think \(q = .01\)) then \(q^2\) is even smaller (think \(q^2 = .0001\)). In our approximation, we can ignore it; then \(\sqrt{p^2 + 2pq} \approx p + q\). If we let \(r = 2pq\), then \(q = \frac{r}{2p}\) and we get Archimedes’ approximation, \(\sqrt{p^2 + r} \approx p + \frac{r}{2p}\). This is also the result we’d get from a tangent line approximation (later in the book).
Chapter 1: Numbers

Section 3: The Real Numbers

In the mid-1800’s, many European mathematicians worked on irrationals (see 38). There were two issues; we’ve discussed the first, the kinds of irrationals – were they all infinite decimals? But there was a second, trickier problem.

Limits allowed mathematicians to answer questions like the computation of $\sqrt{2}$: when we take \{g_1, g_2, \ldots\} defined by $g_2 = \frac{1}{2} \left( g_1 + \frac{2}{g_1} \right)$, \etc, and if we believe that \{g_1, g_2, \ldots\} has a limit, $g$, then the limit has to be $g = \frac{1}{2} \left( g + \frac{2}{g} \right)$. Solving this gives $g^2 = 2$. So the theory of limits tells us that ‘better and better’ means the \{g_1, g_2, \ldots\} become progressively better approximations to $\sqrt{2}$; with a theory of limits, we can say the limit of the approximations is $\sqrt{2}$.

But we don’t know anything like $\sqrt{2}$ exists, and we also don’t know the \{g_1, g_2, \ldots\} even have a limit.

The mathematician Georg Cantor resolved both these issues, giving a construction of the real numbers. We said construction – it wasn’t as though everyone suddenly hit on the one idea that was out there waiting to be discovered. Nor, when the ideas were published, did everyone say ‘Why of course, that’s what I was thinking all along’; several constructions competed for acceptance.

Cantor first took an idea from from Cauchy, who had shown how to talk about convergence without mentioning the limit (see (Figure 43; Cauchy was working on a different issue; see 38). Cauchy’s idea starts with a list of numbers (a sequence) \{a_1, a_2, a_3, \ldots\} abbreviated as $a_k$; it’s understood that $k$ goes through all the natural numbers, sequentially.

If a sequence $a_k$ converges after a while all the $a_k$ have to be close to the limit, so they have to be close to each other. The phrase ‘after a while’ translates to the existence of a number $N$ specifying when that closeness happens; ‘all’ translates to ‘all numbers bigger than $N$.’ In symbols, $j \geq N, k \geq N$.

That the numbers in the sequence are close to each other means that the distance between them is small; that translates into saying that $|a_n - a_m|$ is small. How small? Well, to be a limit, it has to get smaller and smaller. Put another way, if you tell me how small you want it, I can make it that small. ‘How small you want it’ translates to ‘for any $\epsilon > 0$’; and ‘I can make it that small’ now translates to ‘$|a_i - a_k| < \epsilon$.’ When you put it all together, you have a definition of convergence

Figure 42: Georg Cantor
Cantor (1845-1918) solved the riddle of the irrationals, and opened up the theory of the infinite.

Figure 43: Augustin-Jean Cauchy
The French mathematician Cauchy (1789-1857) and a few of his many mathematical accomplishments. He developed the modern theory of limits, and published the first modern calculus book, basic all results on limits, as textbooks do today.
that doesn’t mention the actual limit:

**Cauchy Condition for Convergence:** If the sequence \(a_n\) converges, then for every \(\epsilon > 0\), there is a number \(N\) such that, whenever both \(j \geq N, k \geq N\), then \(|a_j - a_k| < \epsilon\).

Cantor needed a second idea: we believe that the sequence 3, 3.1, 3.14, 3.145, \ldots converges to \(\pi\), but we don’t know what \(\pi\) is. Why not define \(\pi\) to be the sequence 3, 3.1, 3.14, 3.145, \ldots? In general, a real number is defined to be any sequence of rational numbers. Well, not any sequence: the sequence 1, 0, 4, 0, 9, 0, 16, 0, \ldots doesn’t converge, so we don’t want that. We want only the convergent sequences. Wait a second: that’s be Cauchy sequences. Go back to 3, 3.1, 3.14, 3.145, \ldots.

\[|a_1 - a_2| = .1 = \frac{1}{10}; \ |a_2 - a_3| = .04 = \frac{4}{100}; \ |a_3 - a_4| = .005 = \frac{5}{1000},\] etc. These are getting progressively smaller (well, not quite: see 38).

There’s that pesky \(\epsilon\) being infinitely small issue, again. Cantor avoids it by restating the Cauchy condition: instead of saying ‘for every \(\epsilon > 0\)’, he says ‘for every rational number \(\epsilon > 0\)’. No more worries about infinitely small epsilons.

Cantor calls the collection of Cauchy sequences the **real numbers**, \(\mathbb{R}\). There are lots of problems:

**Question:** i) So what are rational numbers now?

**Answer:** A rational number, say \(\frac{1}{2}\), would be the sequence \(\{\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \ldots\}\).

**Question:** ii) How am I supposed to add, subtract, etc these things?

**Answer:** This turns out to be amazingly easy: if \(R = \{r_1, r_2, \ldots\}\) and \(S = \{s_1, s_2, \ldots\}\) then \(R + S = \{r_1 + s_1, r_2 + s_2, \ldots\}\). Same for multiplication, etc. The \(R + S, \) etc are still sequences of rational numbers, and it takes only a little work to show they’re also Cauchy sequences.

**Question:** iii) What happens to the whole theory of limits?

**Answer:** We’d like to have the usual definitions, but tricky part is defining statements like \(|R_k - R| < \epsilon\), more generally \(R < S\), which leads to or \(0 < S - R\). In the end, we just need to know what it means for a real number \(R = \{r_1, r_2, \ldots\}\) to be positive.

Here’s where Cantor was sneaky: just as with the definition of Cauchy sequences, he used rational numbers to define inequalities between the reals. We’d like to just say that \(R\) is positive if the rational numbers making up \(R\) are positive, but that doesn’t work: take \(R = \{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots\}\). Intuitively, \(R\) is the ‘limit’, and that ‘limit’ is zero, not positive. So Cantor defined \(R \succ 0\) as follows:

\[R = \{r_1, r_2, \ldots\} \succ 0\] means there is a rational number \(r \succ 0\) and an \(N\) such that if \(k \geq N\), then \(r_k \geq r\): the rational \(r \succ 0\) keeps the ‘limit’
We also put on the extra words ‘if \( k \geq N \). Again, it’s about limits: the sequence \( \{-1, -0.5, 0, 0.5, 0.59, 0.599, 0.5999, \ldots\} \) has limit .6, which is positive: what the first three terms do is irrelevant. That what \( N \) does: if I take \( N = 4 \), then I’m saying the first three are irrelevant. Now with \( r = .5 \), and \( N = 4 \) for all \( k \geq N, r_k \geq r \).

**Question:** iv) There are lots of sequences of rationals converging to \( \sqrt{2} \); which one are we supposed to use?

**Answer:** We can use any sequence we like. All you need to do is check: sequences \( a_k, b_k \) define the same real number if, intuitively, they have the same limit; see 38 for the technical details.

**Question:** v) What if I took sequences of reals? Do I get some weird new kind of number?

**Answer:** Actually, no. If you took a sequence of real numbers \( \{R_1, R_2, \ldots\} \) with limit \( R \), you can find rational numbers \( \{r_1, r_2, \ldots\} \) with \( R = \{r_1, r_2, \ldots\} \). Very roughly, the idea is this: if

\[
R_1 = \{a_1, a_2, a_3, \ldots\} \\
R_2 = \{b_1, b_2, b_3, \ldots\} \\
R_3 = \{c_1, c_2, c_3, \ldots\}
\]

Then

\[
R = \{a_1, b_2, c_3, \ldots\} \text{ etc.}
\]

Cantor now is able to prove three easy results:

**One:** Every Cauchy sequence of real numbers converges to a real number. This is called the **completeness** of the real numbers.

**Two:** Every real number is a limit of rational numbers; in fact, if \( R = \{r_1, r_2, \ldots\} \), then, with sloppy notation, \( \lim r_k = R \). This is called the **density** of the rational numbers.

**Three:** All the known irrationals were limits of rational numbers, so all the known irrationals were real numbers.

Finally, Cantor’s definition of \( R > 0 \), using rational numbers, allows us to show the real numbers \( \mathbb{R} \) don’t contain infinitely small numbers. Here’s why:

Take a supposed infinitely small real number \( \gamma = \{r_1, r_2, \ldots\} > 0 \). \( \gamma > 0 \) means there’s a rational number \( r > 0 \) and an \( N \) such that for all \( k \geq N, r_k \geq r \).
Now we have our rational number \( r = \frac{p}{q} \leq \gamma < \frac{1}{10^k} \) for all \( k \). But this contradicts the Archimedean Principle for Rationals (p20), that for a rational \( r, 0 < r < 1 \), there’s an \( m \) with \( \frac{1}{10^m} < r \). No infinitely small rational numbers, no infinitely small real numbers. And, BTW, .9999\ldots actually is equal to 1.

We used decimal expansions to get an idea of what is in the infinite collection \( \mathbb{N} \); we’d like something similar for \( \mathbb{R} \). We know each real number is a limit of rational numbers, and each rational number is a repeating decimal, but it might be that the limit is some very weird kind of object.

Ideally, we’d have \( \pi = \{3, 3.1, 3.14, \ldots\} \). This is actually true; we’ll look at it in the next section. But, all we have right now is that the real numbers are all infinite decimals, and amongst those infinite decimals, the irrationals are the non-repeating decimals. It’s a very simple and very useful way to think about real numbers (it’s worth noting, though, that these aren’t necessarily the best way to approximate real numbers and to compute them; see p39).

But Cantor’s theory also opened questions contemporaries preferred to not ask, questions which are still now troubling. For example: we can actually compute the decimal approximations for \( \sqrt{2} \) and \( \pi \) with the approximations given by the Mesopotamians and by Lui Hui. We can even write computer programs to do the computations. Call numbers like \( \pi, \sqrt{2} \) the \textit{computable} numbers. Are all numbers computable? Cantor showed they are not: most real numbers are not computable (see p38).
Notes for Chapter 1 Section 3: The Real numbers


p34 Cauchy’s work on convergence wasn’t motivated by the status of the irrationals. He was trying to give the methods of calculus a logical justification supplying rigorous proofs for all the results. This type of work is ‘working on the foundations’ of the calculus; the analogy being that without good foundations, buildings could collapse.

After Newton and Leibniz developed calculus, mathematicians worked to extend the ideas and give applications to physics, mechanics and engineering; meaning and proof were low priority. This began to change in the mid-eighteen hundreds; Cauchy was one of several mathematicians working to separate out the true results from the false. We’ll meet some of these later on in the course.

p36 Two Cauchy sequences in Cantor’s construction are equivalent if they have the same limit. Since Cantor was trying to develop a theory of limits, he used a Cauchy sequence idea: Cauchy sequences $a_k$, $b_k$ are equivalent if for every $\epsilon > 0$, there is a number $N$ such that, whenever $k \geq N$, then $|a_k - b_k| < \epsilon$. The collection of all sequences equivalent to $a_k$ is denoted $[a_k]$, and the real number corresponding to $a_k$ is actually the equivalence class $[a_k]$.

p35 For the Cauchy condition it isn’t enough to show $|a_3 - a_4| < \epsilon$; you also have to show $|a_3 - a_5| < \epsilon$ and $|a_3 - a_6| < \epsilon$ etc. But for some sequences there’s a cheat: if you can show $|a_n - a_{n+1}| < \frac{\epsilon}{10^n}$ for all $n$, then the sequence is Cauchy. We certainly have that with 3, 3.1, 3.14, 3.145, …; in fact it’s easy to show any infinite decimal gives a Cauchy sequence. Say you have .8769539…; make a sequence out of it as we did with $\pi$: .8, .87, .876, ….. Then $|a_n - a_{n+1}| < \frac{9}{10^n}$, because decimal digits can’t be greater than 9.

p37 What are real numbers that are or are not ‘computable?’ To understand ‘computable’, look at our basic example, the computation of $\sqrt{2}$. We start with a guess $g_1$, then set $g_2 = \frac{1}{2} \left( g_1 + \frac{2}{g_1} \right)$ The sidenote shows how to make a computer program for this. We talked, pre-Cantor, about defining $\sqrt{2}$ to be whatever this procedure gives. We could go one further; we could say that the computable numbers are the numbers that you can get from computer programs. Why not?

A Program For $\sqrt{2}$

```
REAL g
G = 1
FOR j = 1 TO 100
  g = \frac{1}{2} \left( g + \frac{2}{g} \right)
END
PRINT g
```
It’s a very practical definition of computing!

This means that the collection of computable numbers will be found in the collection of all computer programs (along with a lot of junk, like ’hello world’). But the collection of all programs is contained in the collection of all finitely long collections of words. However, the collection of all real numbers, in Cantor’s theory, is the same as all infinite lists. Cantor was able to show that this collection is much larger than the collection of computable numbers. Put another way, if you randomly picked a number from all the reals, the probability that the number is computable is zero.

This argument is part of Cantor’s theory of transfinite numbers. (For a popular, non-technical presentation, see Rudy Rucker’s Infinity and the Mind, Princeton University Press 2004. A basic technical introduction is in Michael J. Schramm’s Introduction to Real Analysis, Dover Publications 2008.

We have a method to approximate numbers like $\sqrt{2}$ by decimals: the method gives the successive approximations 1, 1.4, 1.41, 1.414, . . . . If you square these numbers, you get 1, 1.96, 1.9881, 1.999396, . . . , but there are more general ways to approximate $\sqrt{2}$.

There’s a very old technique, called continued fractions that gives a very general way to approximate a large class of irrationals. Here’s the idea: $\sqrt{2}$ satisfies the quadratic $x^2 - 2 = 0$. I can’t factor that, but I can factor $x^2 - 1 = 1$ as $(x - 1)(x + 1) = 1$, hence

$$x - 1 = \frac{1}{1 + x} \quad \text{or} \quad x = 1 + \frac{1}{1 + x}$$

Oops, there are $x$’s on both sides. That’s okay, I already solved for $x$; let’s plug that into the right side:

$$x = 1 + \frac{1}{1 + x} = 1 + \frac{1}{1 + \left(1 + \frac{1}{1 + x}\right)} = 1 + \frac{1}{2 + \frac{1}{1 + x}}$$

and, if we continue, we get

$$x = 1 + \frac{1}{2 + \frac{1}{2 + \frac{1}{2 + \ldots}}}$$

Now let’s compute $\sqrt{2}$. If we start with $x = 1$, we get $x = \frac{3}{2}$, then $\frac{7}{5}$, $\frac{17}{12}$, $\frac{41}{29}$, . . . . The squares are 2.25, 1.96, 2.006944 . . . , 1.99888109 . . .

See https://en.wikipedia.org/wiki/Continued_fraction
Chapter 1: Numbers

Section 4: Thinking In Decimals

We saw that every real number is a limit of rational numbers, and therefore a limit of decimals. It would be nice to have a more specific characterization; an example would be that \( \pi \) is the limit of \{3, 3.1, 3.14, \ldots\}. Where are we with this?

On p19 we showed all natural numbers can be written with decimals: it followed from the Archimedean Principle:

**The Archimedean Principle:**
If \( x \) is a positive integer, then there is always a power of ten, \( 10^m \), with \( m \geq 0 \) and \( 10^m \leq x < 10^{m+1} \).

We showed a similar principle for rationals. Because Cantor showed each real is limit of rationals, we also get:

**The Archimedean Principle for Reals:** If \( x > 1 \) is a real number, then there is always a power of ten, \( 10^m \), with \( m \geq 0 \) and \( 10^m \leq x < 10^{m+1} \).

Let’s see what this tells us about a real number, say \( r = 632.014159 \ldots \). With \( m = 2 \), \( 100 \leq 632.14159 \ldots < 1000 \). Then \( 1 \leq \frac{632.014159\ldots}{100} < 10 \), and therefore \( r \) is a leading decimal digit plus a fraction.

Now subtract the leading decimal digit (6) from \( r \), and continue until you’re left with just the fraction (.014159 \ldots). We also have an Archimedean principle for fractions:

For positive fractions \( y \) less than one, \( \frac{1}{y} > 1 \), so \( 10^m \leq \frac{1}{y} < 10^{m+1} \), and flipping the inequalities: there is always a power of ten, \( 10^m \), with \( m \geq 0 \) and \( \frac{1}{10^{m+1}} < y \leq \frac{1}{10^m} \).

In our case, \( m = 2 \) and \( .01 < 014159 \ldots \leq .1 \). Now instead of dividing by a power of ten, we multiply by 100: \( 1 < 1.4159 \ldots \leq 10 \) and once again we have a leading decimal digit plus a fraction. Subtract the 1, and apply the same procedure to .4159 \ldots.

What we’ve done is to generate a sequence, \{600, 630, 632, 632.01, 632.014, \ldots\}. After the second term, we get:

\[
\begin{align*}
0 & \leq 632.014159 \ldots - 632 < 1 \\
0 & \leq 632.014159 \ldots - 632.0 < .1 \\
0 & \leq 632.014159 \ldots - 632.01 < .01 
\end{align*}
\]

and so on. These inequalities are enough to show that 632.014159 \ldots is the limit of \{600, 630, 632, 632.01, 632.014, \ldots\}. We have exactly what we wanted: real numbers are limits of very simple finite decimals.
Now that we know what numbers are, we can reclaim the connection between numbers and geometry: we want to show that every point on a line corresponds to exactly one real number (this idea was known to Aristotle, and inspired an entirely different method of constructing the real numbers). Intuitively, the connection will be like taking a ruler to measure lengths; rulers associate numbers to positions.

To make a ruler, take a line, pick a starting point to correspond to the number zero, and choose some fixed length to represent the natural number 1. Mark off the line in units of 1, moving to the right. By the Peano Axioms, p18, this gives us all the natural numbers, arranged along the right side of zero (we can put the negative numbers on the left side of zero; the ideas will be the same).

To account for fractions, a ruler in English units would start subdividing the space between integers: first, cut the space in half, then each side is cut in half again, etc. This marks off the ruler in halves, eighths, sixteenths, and so on. Again, as we move right we go from $\frac{1}{8}$ to $\frac{1}{4}$ to $\frac{1}{2}$ and so on.

Metric rulers, though, use decimals: we’d divide the space between units into ten pieces, then each of those into ten again, etc. It’s like a meter stick, marked off in centimeters, micrometers millimeters and so on.

So far what we have are just ruler markings on the line; this only gives us finite decimals like $\frac{1}{10} + \frac{3}{10} + \frac{4}{10^2}$: we don’t even have $\frac{1}{7}$ yet. For that we’ll need a little more.

Let’s take $\pi$ as an example. Since it’s a real number, it’s a limit of known decimals: $\pi = 3.1459 \ldots$. To locate this with a ruler, we’d the initial digit tells us it’s between 3 and 4. The second digit tells us it’s between the markings for 3.1 and 3.2, and so on: each decimal digit gives us a range of markings on the ruler, and $\pi$ is between all of those. Does this locate $\pi$ as a single point on the line? Let’s try it.

Imagine there were a second number, $\phi$ sharing the space with $\pi$. Then $\pi, \phi$ would each be in the interval between 3 and 4, so they’d have to be no further apart than the length of that interval:

$$-1 \leq \pi - \phi \leq 1$$

And, as $3.1 \leq \pi \leq 3.2$, we have

$$-0.1 \leq \pi - \phi \leq 0.1$$

And, as $3.14 \leq \pi \leq 3.15$, we have

$$-0.01 \leq \pi - \phi \leq 0.01$$
We’ll cut this short: \( \pi, \phi \) share the same location, so for every \( m \),

\[
-\frac{1}{10^m} \leq \pi - \phi \leq \frac{1}{10^m}
\]

By the Achimedean Principle for reals, neither \( \pi - \phi \) nor \( \phi - \pi \) can be a positive number; hence \( \pi - \phi = 0 \). There’s only one number at these successively accurate ruler markings.

If you run this process in reverse, any point on the line sits between markings on the ruler, and these give the decimal expansion.

This construction lets us use geometric and decimal intuitions consistently together:

1. As we move right on the line, we go from 2 to 32 to 632, and we’d write \( 2 < 32 < 632 \): moving right corresponds to growing larger.
2. Inequalities like these are easy: we just check 632 has more decimal places to the left of the decimal point than 32.
3. Decimals work well with our notions of limits to infinity: as we move right along the line, we move towards infinity, and, correspondingly, as we add more zeroes immediately to the left of the decimal points, numbers go to infinity (see p.42): 98, 870, 7600, 65000 . . .
4. For decimal fractions, moving right goes from \( .125 \) to \( .25 \) to \( .5 \); this is easier to see as moving from \( .125 \) to \( .250 \) to \( .500 \), and again the leading decimal place determines the size.
5. Again, decimals work well with the idea of a limit to zero. As we move left to zero, we add more and more zeroes immediately to the right of the decimal point: \( .1, \ .02, \ .003, \ldots \) (see p.42).

Decimals work well with geometric and analytic intuitions, but can be inconvenient for modern science. Two examples:

1. One of the fastest chemical reactions is the absorption of light by a photosynthetic molecule; this occurs in \( .000000000000015 \) seconds. Oddly enough, this is about the same amount of time that stock market transactions take place in electronic trading.
2. While we know the amount of \( \mathrm{CO}_2 \) in the atmosphere is increasing, temperatures on earth haven’t increased accordingly. A new study shows that from 1865 to 1997, the worlds oceans have absorbed about \( 150000000000000000000000 \) Joules of energy (see Figure 46). For comparison, one Joule is about the amount of energy released by dropping a tomato from three feet.

There are several ways to deal with the zeroes; the first is scientific notation, which counts the number of decimal places to the right of the leading decimal, and puts that number into a power of ten.
Thus, \(100 = 1 \times 10^2\), \(21060 = 2.106 \times 10^4\); \(1500000000000000000000000000\) Joules is \(1.5 \times 10^{23}\) Joules. Scientific notation requires a single (non-zero) digit to the left of the decimal point.

For a fractional numbers like .0231, scientific notation counts the number of decimal places to the right of the non-zero leading decimal, so \(.0231 = 2.31 \times 10^{-2}\) and \(.00000000000015\) seconds is \(1.5 \times 10^{-16}\) seconds. Alternatively, \(.0231 = \frac{231}{10^2} = 2.31 \times 10^{-2}\). In either case, the effect is still to have only a single (non-zero) digit to the left of the decimal point.

Powers of ten are reference points for daily life, but powers of a thousand are more often used in science, engineering, and technology: we have a meter, \(10^0\) meters, a millimeter, \(10^{-3}\) meters, a micrometer \(10^{-6}\) meters, a nanometer, \(10^{-9}\) meters, etc. \(10^{-15}\) units is a femto-unit, so the photosynthetic reaction above takes place in 15 femtoseconds. Similarly, \(10^{21}\) units is a zetta-unit, so the amount of heat energy released into the atmosphere is 150 zettajoules.

A system using powers of \(10^3\) or \(10^{-3}\) is called engineering notation. In this notation, \(21060\) would be written as \(2.106 \times 10^4\), while \(.0231\) would be \(23.1 \times 10^{-3}\).

One example comes from computer engineering, measuring the speed of modern computers. Consumers look at processor speed: ‘wow that’s a 4ghz processor’. Scientists prefer to measure how many computations a computer can do. Additions or multiplications of natural numbers are called integer operations; those can be very fast. But most scientific computations are with decimals; these are called floating point operations. Scientists and engineers need to know how many floating point operations a computer can perform per second; this is referred to as FLOPS (the FLOP measurement is not easy to calculate; it’s often a measure of an ideal, rather than what a computer can do on real computations; see p.45).

To look at some examples of computer speeds, some terminology: A \textit{mega}- is \(10^6\) (a million); a \textit{giga}- is \(10^9\) (a billion); a \textit{tera}- is \(10^{12}\) (a trillion); a \textit{peta}- is \(10^{15}\) (a quadrillion).

The first commercial supercomputer was the CRAY-1 (Figure 47) introduced to Los Alamos in 1976; it computed at 100 megaflops. In comparison, the first Macintosh from 1984 (Figure 48) was about \(10,000\) times slower. Apple’s fastest computer in 2016, the MacPro (Figure 49) runs at a maximum of 436 gigaflops. And the 2016 world speed champion is the Chinese Sunway Taihu Light, (Figure 50) running at 93 petaflops, though America, Switzerland, Japan and Saudi Arabia also have (slower) petaflop machines.
Extremely fast computers are used to do aerodynamic computations to reduce air resistance in 18-wheelers; to compute Bitcoins; to forecast the weather a week in advance; to design nuclear weapons; to model the motions of molecules as they engage in chemical reactions (for example, to design more effective drugs or to find genetic components of disease).
Notes for Chapter 1 Section 4: Thinking In Decimals

More decimal places to the left of the decimal makes numbers larger. A decimal-based definition of limit would go something like this:

**Definition** We say \( x \to \infty \) if \( x \) contains more and more decimal places to the left of the decimal point.

We’ll see how this matches with standard definitions of \( x \to \infty \). We can use the same ideas of size in decimal numbers to see how to define \( x \to 0 \); the numbers .5, .05, .005 get smaller and smaller. So,

**Definition** We say \( x \to 0 \) if \( x \) contains more and more leading zeros immediately to the right of the decimal point.

Each of these definitions can be made to work because of the Archimedean Property: the powers of ten measure the size of real numbers, and powers of ten are expressed as a 1 followed or preceded by zeroes.

For some of the difficulties thinking about FLOPS computations, see https://devtalk.nvidia.com/default/topic/745504/comparing-cpu-and-gpu-theoretical-gflops/
Chapter 1: Numbers

Section 5: Units

Notes for Chapter 1 Section 5: Units
Chapter 1: Numbers

Section 6: Accuracy

Notes for Chapter 1 Section 6: Accuracy
Chapter 1: Numbers

Section 7: Computer Numbers

Notes for Chapter 1 Section 7: Computer Numbers
Chapter 1: Numbers

Section 8: Data

Notes for Chapter 1 Section 8: Data
List of Figures

Chapter 0

p3 Figure 2 ECG machine picture from http://en.wikipedia.org/wiki/Image: Willem_Einthoven_ECG.jpg.

p3 Figure 3 The first human ECG. From Einthoven W. The galvano-metric registration of the human electrocardiogram. Pfluger’s Arch f.d. ges Physiol 1903;99:472-480.

p4 Figure 4 Picture of ECG indicating irregular heartbeat, from http://www.bem.fi/book/19/19.htm.

p4 Figure 5 Picture of portable ECG unit from http://www.skymall.com/handheld-ecg-monitor/HDW107.html

p4 Figure 6 is from Alfonso, Valentino X ECG QRS Detection, in Tompkins, Ed. Biomedical Digital Signal Processing, C-Language Examples and Laboratory Experiments for the IBM PC, Prentice Hall PTR, 1995.


Chapter 1

p6 Figure 7 from Simon Hunt, Jason Low and K. C. Burns, Adaptive numerical competency in a food-hoarding songbird, Proc. R. Soc. B (2008) 275, 2373-2379.

p6 Figure 8, Portrait of David Hume, by by Allan Ramsay, from Scottish National Portrait Gallery.

p7 Figure 11 is from http://www.gregapodaca.com/numerography/files/020.html

p7 Figure 10 is from http://afrolegends.com/2013/08/29/the-ishango-bone-craddle-of-mathematics/

p8 The image in Figure 12 is by Ben M. Harvey, Utrecht University.

p8 Figure 13 is from T. Cuyler Young, Jr., of the Royal Ontario Museum, Toronto, Canada.

p8 Figure 14 is licensed under Public Domain via Commons https://commons.wikimedia.org/wiki/File:Zero_in_Rod_Calculus.png#/media/File:Zero_in_Rod_Calculus.png

p9 Figure 15 is by Gisling, licensed under Public Domain via Com-

Figure 20 is from http://www.livius.org/articles/person/kidinnu-the-chaldaeans-and-babylonian-astronomy

Figure 21 Retina diagram from http://hubel.med.harvard.edu/book/b9.htm

Figure 22 is a still from the movie 'What the Frog Sees' by Skip Battaglia. Video at http://www.visualnews.com/2012/10/16/the-frog-does-not-see-what-does-not-move-to-the-frog-to-move-is-to-exist/

Figure 23 is from http://www.howplantswork.com/2008/11/30/flowers-what-you-see-versus-what-the-bees-see

Figure 24 is from "ParietCapts lateral" by Sebastian023. Licensed under CC BY-SA 3.0 via Commons - https://commons.wikimedia.org/wiki/

The image in Figure 25 is by Ben M. Harvey, Utrecht University.

Figure 27 is from Schmandt-Besserat, From Accounting to Writing, cited there as permission from MusŒe du Louvre, DŒpartement des AntiquitŒs Orientales.

Figure 29 is from the University College, London, http://www.ucl.ac.uk/museums-static/digitalegypt/lahun/uc32159.html

Figure 30 is from http://www.slideshare.net/sbrierton/liu-hui1

Figure 31 is from CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=4322998

Figure 33 is from https://commons.wikimedia.org/wiki/File:Table_of_Shamash_relief.jpg licensed under Creative Commons. The original is in the British Library, room 55.

Figure 34 is from the personal collection of Bill Casselman; see http://www.math.ubc.ca/~cass/Euclid/ybc/ybc.html. The original tablet is in the Yale Babylonian Collection; see http://nelc.yale.edu

Figure 36 is from https://commons.wikimedia.org/wiki/File:Laser_Interference.JPGdiffract


Figure 39 is by Moritz NŒhr, an Austrian photographer also known for his portraits of the composer Gustav Mahler and the artist Gustav Klimt. From the archives of the Austrian National Library, and is public domain.
Figure 37 is by Lukas Lansky

Figure 38 is from Princeton University Office of Communications, by Denise Applewhite

Figure 40 is from http://mrstevesscience.blogspot.com/2011/03/help-6th-graders-figure-out-for.html

Figure 41 is from http://www.mathaware.org/mam/oo/master/essays/B3D/2/binomial.html

Figure 43 is from http://www-groups.dcs.st-and.ac.uk/history/PictDisplay/Cauchy.html

Figure 44 is from a portrait in the Museo del Prado, Spain. This is from https://www.cs.drexel.edu/crorres/Archimedes/contents.html

Figure 45 is taken from http://sputniknews.com/business/20150903/1026525213.html

Figure 47 is from http://www.macobserver.com/tmo/article/The_Fastest_Mac_Compared_to_Todays_Supercomputers

Figure 48 is from http://www.oldcomputers.net/macintosh.html

Figure 49 is from http://corespeed3d.com/tag/apple-mac-pro/

Figure 50 is from http://www.gizmag.com/supercomputer-china-taihulight-processor-petaflops/43982/
## Index

### A
- abacus, 10
- accounting, 15, 19, 30
- aerodynamics, 44
- algebraic numbers, 26
- algorithm, 10
- anagogic, 13
- anthropologists, 7
- approximation, 22
- Arabic, 10
- Archimedean Principle, 19, 20, 40
- Archimedean Principle for Reals, 40
- Archimedes, 32
- Aristotle, 17, 41
- astrology, 11
- Augustine, 13

### B
- base sixty, 7
- Babylonian, 31
- Babylonian Astronomy, 11
- Bakhshali Manuscript, 9
- base five, 7
- bee, 12
- Bessel function, 26
- binomial expansion, 33
- Bitcoins, 44

### C
- calender, 10
- Cantor, 34, 37–39
- Cantor, Georg, 30, 38
- Cauchy Condition for Convergence, 35
- Cauchy, Augustin, 34
- chemical reactions, 44
- Chinese, 9, 22, 43
- cognitive reference points, 30
- cognitive science, 30
- completeness, 36
- computable, 37, 38
- continued fractions, 39
- Conway, 28
- counting, 17
- Cray, 43
- cuneiform, 10

### D
- Darwin, Charles, 12
- Dedekind, Richard, 38
- Dehaene, Stanislaw, 13, 14
- density of rationals, 36
- diffraction, 26
- digital filter, 4
- disease, genetic, 44
- divination, 10
- drug design, 44
- dyscalculia, 14
- Darwin, Charles, 12
- Dedekind, Richard, 38
- Dehaene, Stanislaw, 13, 14
- density of rationals, 36
- diffraction, 26
- digital filter, 4
- disease, genetic, 44
- divination, 10
- drug design, 44
- dyscalculia, 14

### E
- Easter, 13
- ECG, 3, 5
- eighteen-wheelers, 44
- Einthoven, 3
- Engineering, 9
- Enma Anu Enlil, 11
- equivalence, 29
- equivalence class, 29, 38
- error, 22
- ethology, 12
- Euclid, 32
- euler, 31
- Eudoxus, 25
- Euler, 26

### F
- fertile crescent, 10
- Fibonacci, 25, 26
- fibrillation, 5
- filter, 12, 16
- Foster, Jody, 3
- foundations of calculus, 38
- fractions, 19
- fractions, unit, 20
- Frege, Gottlob, 17
- frog, 12
- frog brain, 12

### G
- Galvani, 3
- geometric, 23
- gigaflap, 43
- Godin Teke, 8, 18
- Godin Tepe, 14
- Great Wall of China, 9
- Greek, 24, 25
- Gwalior Numerals, 9

### H
- Han, 9
- heartbeat, 16
- Hero of Alexandria, 32
- Hindu-Arabic numerals, 10
- Hippasus, 32
- Hume, David, 6

### I
- induction, 31
- Induction, Axiom of, 18
- infinite, 18, 30
- infinite decimal, 20
- integers, 17
- intraparietal sulcus, 7, 13
- IPS, 14
- Ishango Bone, 7
- Islam, 10
K
- al-Khwarizmi, 10, 17, 29

L
- Leibnitz, Gottfried, 38
- Leonardo of Pisa, 10, 25
- limit, 22
- Lu Hui, 37
- Lui Hui, 22

M
- Mac, 43
- magnitude, 25
- Marshack, Alexander, 15
- Mesopotamia, 19, 30
- Mesopotamian, 23, 37
- modeling, 44
- music of the spheres, 11, 32

N
- natural numbers, 17
- Newton, 25
- Newton, Isaac, 38
- nightmare, 27
- numerosity, 6, 12, 13
- von Neumann, John, 30

O
- omens, 10

P
- Pan, Jiapu, 50
- Pan-Tompkins, 16, 50
- Peano, 41
- Peano, Giuseppe, 18
- petaflop, 43
- Philosophy, 6
- photosynthesis, 42
- Plato, 12
- positional notation, 9
- projection, 9
- ProtoWorld, 14
- Pythagorean Theorem, 23
- Pythagoreans, 32

Q
- qualities, 6
- Qin, 9
- qualia, 6

R
- rational numbers, 25
- real numbers, 35
- repeating decimal, 20
- Robin, 6
- Roman numerals, 9
- Rosch, Eleanor, 30
- ruler and compass, 32

S
- Schmandt-Besserat, Denise, 15
- scientific notation, 42
- Scots, 6
- Seleucid, 31
- sequence, 34
- Shamash, 23
- Shining One, 22
- stock market trading, 42
- strong induction, 19, 30, 32
- surreal numbers, 28
- symbolic, 7
- System A and B, 11, 15

T
- tally, 7
- taxes, 23
- Tompkins, Willis, 50
- transcendental numbers, 26
- transfinite numbers, 30, 39

U
- ultraviolet light, 12

W
- warming, global, 42
- weapons, nuclear, 44
- weather forecasting, 44
- Wigner, Eugene, 13
- Wittgenstein, Ludwig, 29