The Art of Proof

A Concrete Gateway to Mathematics

Matthias Beck & Ross Geoghegan

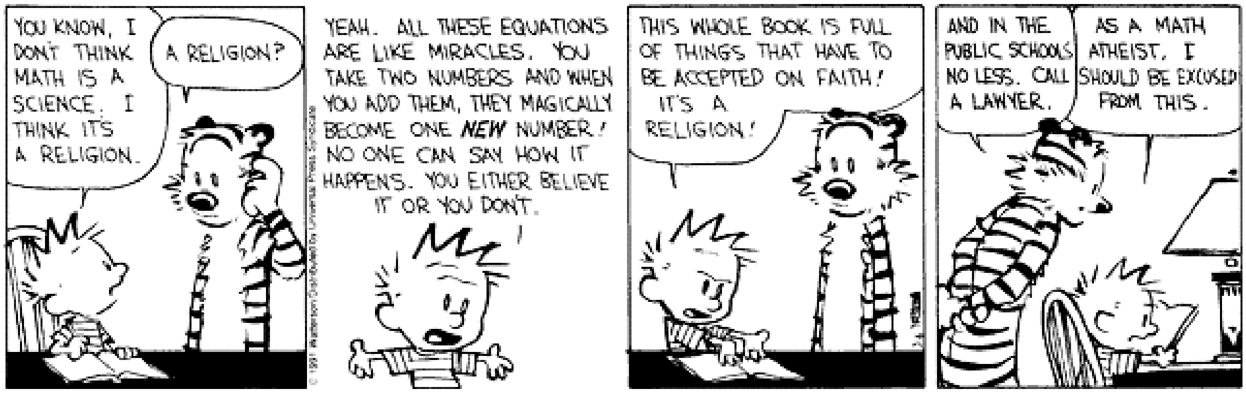
I SIGNED UP FOR A NEW COURSE TODAY
IT'S CALLED "INTRODUCTION TO MATH"
IS IT HARD?
NOT AT ALL..
I JUST WALKED RIGHT IN, AND SAID, "HOW DO YOU DO, MATH!"

Pure mathematics is, in its way, the poetry of logical ideas.

Albert Einstein (1879–1955)

Contents

Notes for the Student


YOU KNOW, I DON'T THINK MATH IS A SCIENCE. I THINK IT'S A RELIGION.
A RELIGION?
YEAH. ALL THESE EQUATIONS ARE LIKE MIRACLES. YOU TAKE TWO NUMBERS AND WHEN YOU ADD THEM, THEY MAGICALLY BECOME ONE NEW NUMBER! NO ONE CAN SAY HOW IT HAPPENS. YOU EITHER BELIEVE IT OR YOU DON'T.
THIS WHOLE BOOK IS FULL OF THINGS THAT HAVE TO BE ACCEPTED ON FAITH! IT'S A RELIGION!
AND IN THE PUBLIC SCHOOLS NO LESS. CALL A LAWYER.
AS A MATH ATHEIST, I SHOULD BE EXCUSED FROM THIS.

You have been studying important and useful mathematics since the age of three; most likely, the body of math you know can be described as Sesame-Street-through-calculus. This is all good and serious math—from the beautiful algorithm for addition, which we all learned in grade school, through high-school algebra and geometry, and on to calculus.

Now you have reached the stage where the details of what you already know have to be refined. You need to understand them from a more advanced point of view. We want to show you how this is done. We'll take apart what you thought you knew (adding some new topics when it seems natural to do so), and reassemble it in a manner so clear that you can proceed with confidence to more difficult mathematics—algebra, analysis, combinatorics, geometry, number theory, statistics, topology, etc.

Actually, we won't be looking at everything you know—that would take too long. We concentrate here on numbers: integers, fractions, real numbers, decimals, complex numbers, and cardinal numbers. We wish we had time to do the same kind of detailed examination of high-school geometry, but that would be another semester-long course, and, as mathematical training, it would only teach the same things over again. To put that last point more positively: once you understand what we're teaching in this book—in this course—you'll be able to apply these methods and ideas to other parts of math in future courses.

On first sight you may find this book unusual, maybe even alarming. Here is one comment we received recently from a student who used a test version:

> The overall feel of the book is that it is very "bare-bones", there isn't much in the way

> of any additional explanations of any of the concepts. While this is nice in the sense that the definitions and axioms are spelled right out without anything getting in the way, if a student doesn't initially understand the concepts underlying the sentence, then they're screwed. As it stands, the book seems to serve as a supplement to a lecture, and not entirely as a stand-alone learning tool.

This student has got it exactly right. We intend this book to be supplemented by discussion in a professor's class. If you think about what's involved in writing any book of instruction you'll realize that the authors had better be clear on the intended readership and the way they want the book to be used. While we do believe that *some* students can use this book as a stand-alone, our experience in using this material (experience stretching over twenty-five years) tells us that this will work for only a few. So please regard your professor as Part 3 of this book (which comes in two Parts). He/she is the source for getting all the insights we didn't—indeed, couldn't—write down.

We (the two authors) are active research mathematicians, and we know, for ourselves as well as for our students, that learning mathematics through oral discussion is usually easier than learning mathematics through reading, even though reading is necessary in order to get the details right. So we have written a kind of manual or guide for a semester-long discussion—inside and outside class.

Please read the *Notes for the Instructor* on the following pages. There's much there that's useful for you too. And good luck. Mathematics is beautiful, satisfying, fun, deep and rich.

San Francisco *Matthias Beck*
Binghamton *Ross Geoghegan*
May 2008

Notes for the Instructor

Logic moves in one direction, the direction of clarity, coherence and structure. Ambiguity moves in the other direction, that of fluidity, openness, and release. Mathematics moves back and forth between these two poles. [...] It is the interaction between these different aspects that gives mathematics its power.
William Byers (*How Mathematicians Think*, Princeton University Press, 2007)

This book is intended primarily for students who have studied calculus or linear algebra and who now wish to take courses that involve theorems and proofs in an essential way. The book is also for students who have less background but have strong mathematical interests.

We have written the text for a one-semester course; typically such a course has a title like "Gateway to Mathematics" or "Introduction to proofs" or "Introduction to higher mathematics." Our book is shorter than most texts designed for such courses. Our belief, based on many years of teaching this type of course, is that the roles of the instructor and of the textbook are less important than the degree to which the student is invited/requested/required to do the hard work. So we have omitted many of the proofs.

Here is what we are trying to achieve:

1. To show the student some important and interesting mathematics.

2. To show the student how to read and understand statements and proofs of theorems.

3. To help the student discover proofs of stated theorems, and to write down the newly discovered proofs correctly, and in a professional way.

4. To foster for the student something as close as is feasible to the experience of doing research in mathematics. Thus we would want the student to actually discover theorems and write down correct and professional proofs of those discoveries. This is different from being able to write down proofs of theorems that have been pre-certified as true by us (in the text) or by the instructor (in class).

Once the last of these has been achieved, the student is a mathematician. We have no magic technique for getting the student to that point quickly, but this book might serve as a start.

We should add that real success takes time, usually much more than one semester. But we hope to put the basics in place so that skills may grow as the student takes other courses.

Many books intended for a gateway course are too abstract for our taste. They focus on the different types of proofs and on developing techniques for knowing when to use each method. We prefer to start with useful mathematics on day one, and to let the various methods of proof, definition, etc., present themselves naturally as they are needed in context. Here is a quick indication of our general philosophy:

On Teaching Methods: We do not start with customary dry chapters on “Logic” and “Set Theory.” Rather we take the view that the student is intelligent, has considerable prior experience with mathematics, and knows, from common sense, the difference between a logical deduction and a piece of nonsense (though some training in this may be helpful!). To defuse fear from the start, we tell the student: “A theorem is simply a sentence expressing something true; a proof is just an explanation of why it is true.” Of course, that opens up many other issues of method, which we gradually address as the course goes on.

On Choice of Material: We say to the student something like the following: “You have been studying important and useful mathematics since the age of three; the body of math you know is Sesame-Street-through-calculus. Now it’s time to revisit (some of) that good math and to get it properly organized. The very first theorem most of you ever heard proved was when you asked some adult: Is there a biggest number? (What answer were you given? What would you answer now if a four-year-old asked you that question?) Later on, you were taught to represent numbers base 10, and to add and multiply them. Did you realize how much is buried behind that (number systems, axioms, algorithms, ...)? We will take apart what you thought you knew, and we’ll reassemble it in a manner so clear that you can proceed with confidence to more difficult mathematics.”

On Grading Homework—The Redline Method: It is essential that the student regularly hand in written work and get timely feedback. One method of grading that we have found successful lessens the time-burden on the instructor and puts the responsibility on the shoulders of the student. It works like this:

Certain theorems in the book are assigned by the instructor: proofs are to be handed in. The instructor reads a proof until a (real) mistake is found—this might be a sentence that is false or a sentence that has no meaning. The instructor draws a red line under that sentence and returns the proof to the student at the next meeting. No words are written on the paper by the instructor: it is the student’s job to figure out why the red line was put there. Pasting as necessary so as not to have to rewrite the correct part (the part above the red line), the student then hands in a new version and the process of redlining is repeated until the proof is right. It takes discipline on the instructor’s part not to write helpful hints in the margin, but it is best not to. Instructors may wish to limit the number of rewrites, but anything up to six should be acceptable. Once the proof is correct, the student gets credit for that problem—the same credit whether it is right the first time or the sixth time. There is a safety valve: the student who cannot understand why a red line was drawn can go to office hours: in that case, we advise the instructor not to tell the student the reason right away but to ask him or her to read the entire proof aloud, sentence by sentence. Almost always, when the student reaches the redlined sentence it becomes clear what the issue is.

In all this we are not looking for perfection of expression—that will come with time. We start with the attitude that a proof is just an explanation of why something is true, and the student should come to understand that a confused explanation is no more acceptable in mathematics than in ordinary life. But it is counterproductive for the instructor to have a “gotcha” mentality. The red line should be reserved for real mistakes of thought. Minor errors of expression can be overlooked, especially in the early weeks.

To put this another way, the student needs to believe that writing correct mathematics is not an impossible task. We should be teaching rigor, but not rigor mortis.

We sometimes say in class that we will read the proof rather as if it were a computer program: if the program doesn't run, there must be some first line where the trouble occurs. That's where the red line is.

We can say from experience that this method of grading is psychologically taxing on the student. That is good. But it is tough love.

One Semester: The material covered in this book should be more than enough for a 15-week, 3-hours-per-week course. It is structured in parts of equal size, namely a discrete part (integers, induction, modular arithmetic, finite sets, etc.) and a continuous part (real numbers, limits, decimals, infinite cardinal numbers, etc.) We recommend that both parts should get equal "airtime" in a mathematical gateway course. Thus the instructor should resist the temptation to let class discussion of Part 1 slide on into the eighth week. Some discipline concerning homework deadlines is needed at that point too, so that students will give equal time and attention to the second half. (The instructor who ignores this advice will probably come under criticism from colleagues: this course is often a prerequisite for real analysis.)

Problems: There are two kinds of exercises for students in this book:

1. The main body of the text consists of *propositions* (called *theorems* when they are particularly important), in which the mathematics is developed. In principle, these propositions are meant to be proved by the students; however, proving *all* of them is likely to be overwhelming, so the instructor must exercise judgment. Besides this, some of the propositions are proved in the text, to give the student a feel for how to approach a proof of a certain statement (and also to introduce different proof methods). Of the remaining propositions, we tend to prove roughly half in class on the blackboard and give the other half as homework problems. (In a really small class, the instructor may ask students to go to the blackboard and present their proofs.)

2. There are also exercises called *projects*. These are more exploratory, sometimes open-ended, problems for the students to work on. They vary greatly in difficulty—some are more elementary than the propositions, some concern unsolved conjectures, and some are writing projects intended to entice exploration by the students. We would encourage students to do these in groups. Some could be the basis for an outside-class pizza party, one project per party.

Collaboration: We encourage collaborative work subject to some constraints. First, it works best if no collaboration is allowed until everyone has done some exercises on their own. Secondly, collaboration becomes an impediment if a strong student is constantly "carrying" a weaker student: the instructor needs to be alert for this. Just as in mathematical research, collaborative work should be clearly labeled as such.

Copying: Most students understand that copying serves no purpose. In a class of thirty or less, the instructor should have no difficulty in catching copying. One solution is to ask the suspect to explain his or her work verbally in the office.

Note from Ross Geoghegan: This book is a development of class notes I have been using and altering for the past twenty-five years. I experienced a version of the Moore Method as a graduate student, and its influence has stuck. Since then I have been influenced by some of my colleagues at Binghamton. For example:

1. When I started teaching this course, Louis McAuley advised me to emphasize decimals in introducing the real numbers, since that is how students have previously met them. I have tried many variants of this over the years, seeking a workable compromise between actually constructing the reals and stating axioms for them.

2. It was the late Craig Squier who convinced me that a course of this kind should be split evenly between the discrete and the continuous, with special emphasis on induction and recursion in the discrete part.

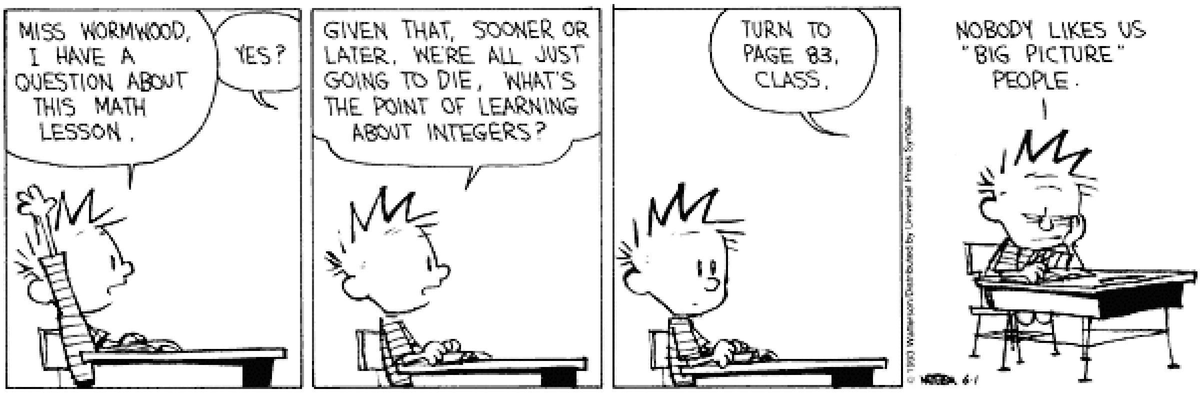
We thank Laura Anderson, Fernando Guzmàn, and Eric Hayashi, who tried out variants on these notes and gave useful feedback.

San Francisco *Matthias Beck*
Binghamton *Ross Geoghegan*
May 2008

Part I: The Discrete

Chapter 1

Integers


MISS WORMWOOD, I HAVE A QUESTION ABOUT THIS MATH LESSON.
YES?
GIVEN THAT, SOONER OR LATER, WE'RE ALL JUST GOING TO DIE, WHAT'S THE POINT OF LEARNING ABOUT INTEGERS?
TURN TO PAGE 83, CLASS.
NOBODY LIKES US "BIG PICTURE" PEOPLE.
© 1993 Watterson/Distributed by Universal Press Syndicate
6-1

You already have an informal understanding of the integers—they are numbers like $0, 1, -3, 34, \ldots$ Here we begin by writing down a list of properties of the integers that your previous experience will tell you ought to be considered to be true, things you always believed anyway. We call these properties *axioms*. Axioms are statements and definitions that form the starting point of a mathematical discussion; items which are assumed (by an unspoken agreement between author and reader) without question or deeper analysis. Advanced mathematics normally begins with a statement of axioms.[1] Once the axioms are settled, we then explore how much can be logically deduced from them. A mathematical theory is rich if a great deal can be deduced from a few primitive (and intuitively acceptable) axioms.

In short, we have to start somewhere. The axioms in one course may in fact be theorems in a deeper course whose axioms are more primitive. The list of axioms is simply a clearly stated starting point.

[1] If you open a randomly chosen mathematics book in the library, you may not see a list of axioms on the first page, but they are present implicitly; usually the author is assuming prerequisite knowledge of more basic mathematics which rests on axioms well known to the reader.

1.1 Axioms

We assume the existence of a set[2], denoted $\mathbb{Z}$, whose members are called **integers**. This set $\mathbb{Z}$ is assumed to be equipped with binary operations[3] called **addition**, $+$, and **multiplication**, $\cdot$, satisfying the following five axioms, as well as Axiom 3.1 to be introduced in Chapter 3:

Axiom 1.1. *For all integers m, n, and p:*

(i) $m + n = n + m\,.$ *(commutative property of addition)*

(ii) $(m + n) + p = m + (n + p)\,.$ *(associative property of addition)*

(iii) $m \cdot (n + p) = m \cdot n + m \cdot p\,.$ *(distributive property)*[4]

(iv) $m \cdot n = n \cdot m\,.$ *(commutative property of multiplication)*

(v) $(m \cdot n) \cdot p = m \cdot (n \cdot p)\,.$ *(associative property of multiplication)*

Axiom 1.2. *There is an integer $0 \in \mathbb{Z}$ such that for all $m \in \mathbb{Z}$, $m + 0 = m$.* *(identity element for addition)*

Axiom 1.3. *There is an integer $1 \in \mathbb{Z}$ such that $1 \neq 0$ and for all $m \in \mathbb{Z}$, $m \cdot 1 = m$.* *(identity element for multiplication)*

Axiom 1.4. *For each $m \in \mathbb{Z}$, there exists an integer, denoted $-m$, such that $m + (-m) = 0$.* *(additive inverse)*

Axiom 1.5. *For all integers m, n, and p, if $m \neq 0$ and $m \cdot n = m \cdot p$, then $n = p$.* *(cancellation)*

The $\in$ symbolizes **is an element of**—that is, $0 \in \mathbb{Z}$ means "0 is an element of the set $\mathbb{Z}$." The symbol "=" means **equals**. It is a verb, and to say $m = n$ means that m and n are the same element. Thus m can be substituted for n in any expression without changing the meaning of that expression. For instance, if $m = n$ then we can conclude that $m + p = n + p$ for any p. We note for future reference some properties that the symbol "=" satisfies:

(i) For all m, $m = m$. *(reflexivity)*

(ii) For all m and n, if $m = n$ then $n = m$. *(symmetry)*

(iii) For all m, n, and p, if $m = n$ and $n = p$ then $m = p$. *(transitivity)*

Students often ask: do we have to say "by reflexivity," "by symmetry," and "by transitivity" each time we use "="? The answer is *no*. But you should be aware of these matters. Indeed, there's a fourth common use of "=" not covered by (i), (ii) and (iii). This is:

[2] You might ask: how is a **set** defined? For now, let's use the word intuitively: a set is a collection of "things" or **elements** or **members**.

[3] A **binary operation** on a set S is a procedure which takes two elements of S as input and gives a third as output.

[4] Strictly speaking, the right-hand side should read $(m \cdot n) + (m \cdot p)$. It is a useful convention to always multiply before adding, whenever an expression contains both $+$ and $\cdot$.

(iv) Think of mathematical objects as "words" in an "alphabet" [Example: x^2+2x-1] which can be put together to make longer words [Example: u is x^2+2x-1 and v is $y^2 + 2x - 7\pi$ so that the symbol uv is used to denote the compound word $(x^2+2x-1)(y^2+2x-7\pi)$]. If abc and adc are compound words, and if $b=d$ then $abc=adc$. That is: if $b=d$ you may replace b by d whenever b occurs. *(replacement)*

The symbol "$\neq$" means **is not equal to**. To say $m \neq n$ means m and n are different. Note that "$\neq$" satisfies symmetry always, transitivity sometimes but not always, and never satisfies reflexivity. Similarly, the symbol $\notin$ means **is not a member of**.

1.2 First Consequences

At this point, the only facts we *know* to be true about the integers are Axioms 1.1–1.5. In the language of mathematics, the axioms are **true**[5] or are **facts**. Every time we prove that some statement follows logically from the axioms we are proving that it too is true, just as true as the axioms, and from then on we may add it to our list of facts. Once we have established that the statement is a fact (i.e., is true) we may use it in later logical arguments: it is as good as an axiom because it follows from the axioms.

From now on, we will use the common notation mn to denote $m \cdot n$. We start with some propositions that show our axioms still hold when we change the orders of some terms:

Proposition 1.1. *For all integers m, n, and p, $(m+n)p = mp+np$.*

Let's prove Proposition 1.1. The left-hand side $(m+n)p$ equals $p(m+n)$ by Axiom 1.1(iv). Now we may use Axiom 1.1(iii) to deduce that $p(m+n) = pm+pn$. Finally, we use Axiom 1.1(iv) again: $pm = mp$ and $pn = np$. In summary we have proved:

$$(m+n)p \overset{\text{Axiom 1.1(iv)}}{=} p(m+n) \overset{\text{Axiom 1.1(iii)}}{=} pm+pn \overset{\text{Axiom 1.1(iv)}}{=} mp+np\,,$$

that is, $(m+n)p = mp+np$. □ [6]

This proof is well described as a *direct proof*: all we have used are some statements that we know to be true (Axioms 1.1(iii) and (iv)), and we mixed those together in a way that provided us with the statement in Proposition 1.1. You can prove the next propositions in a similar way; try it!

Proposition 1.2. *For all integers m, $0+m=m$ and $1 \cdot m = m$.*

Proposition 1.3. *For all integers m, n, and p, if $m+n=m+p$ then $n=p$.*

Proposition 1.4. *If $m, x_1, x_2 \in \mathbb{Z}$ satisfy the equations $m+x_1=0$ and $m+x_2=0$, then $x_1=x_2$.*

(This means that, given $m \in \mathbb{Z}$, the element $-m$ mentioned in Axiom 1.4 is the unique solution of $m+x=0$.)

Proposition 1.5. *For all integers m, n, p, and q:*

[5] What is truth? That is for the philosophers to discuss. Mathematicans avoid such matters by the *axiomatic method*: in mathematics a statement is true if and only if it follows logically from the agreed axioms.

[6] The □ signifies the end of a proof.

(i) $(m+n)(p+q) = (mp+np) + (mq+nq)$.

(ii) $m + (n + (p+q)) = (m+n) + (p+q) = ((m+n)+p)+q$.

(iii) $m + (n+p) = (p+m) + n$.

(iv) $m(np) = p(mn)$.

(v) $m(n + (p+q)) = (mn+mp) + mq$.

(vi) $(m(n+p))\,q = (mn)q + m(pq)$.

Next are some propositions that refine our knowledge about 0 and 1:

Proposition 1.6. *If $x \in \mathbb{Z}$ has the property that for all $m \in \mathbb{Z}$, $m + x = m$, then $x = 0$.*

(This means that the element 0 mentioned in Axiom 1.2 is the unique solution of $m + x = m$.)

Proposition 1.7. *If $x \in \mathbb{Z}$ has the property that for some $m \in \mathbb{Z}$, $m + x = m$, then $x = 0$.*

These last two propositions look suspiciously alike; we only have to switch the phrases "for all" and "for some." The outcome could hardly be more different. When we say some statement is true *for all*, say, $m \in \mathbb{Z}$, we talk about a *universal* property of the set $\mathbb{Z}$. On the other hand, when some statement is true *for some* $m \in \mathbb{Z}$, we make an *existence* statement: there is (at least) one $m \in \mathbb{Z}$ that satisfies the statement.

Proposition 1.8. *For all $m \in \mathbb{Z}$, $m \cdot 0 = 0 = 0 \cdot m$.*

Proposition 1.9. *If $x \in \mathbb{Z}$ has the property that for all $m \in \mathbb{Z}$, $mx = m$, then $x = 1$.*

(Thus the element 1 mentioned in Axiom 1.3 is the unique solution of $mx = m$.)

Proposition 1.10. *If $x \in \mathbb{Z}$ has the property that for some nonzero $m \in \mathbb{Z}$, $mx = m$, then $x = 1$.*

The propositions in this chapter are meant to be proven in the order they are presented here. Let's try to prove Proposition 1.10. This proposition contains a typical if-then statement: **if** statement ♡ is true **then** statement ♣ is true as well. Statement ♡ here is "$x \in \mathbb{Z}$ has the property that for some nonzero $m \in \mathbb{Z}$, $mx = m$," and statement ♣ is "$x = 1$." We prove this if-then statement in the most direct way: *assume* ♡ is true; then try to show that ♣ follows. In our case, we assume (in addition to what we already know from previous propositions and the axioms) that somebody gives us an $x \in \mathbb{Z}$ and says there is some nonzero $m \in \mathbb{Z}$ for which $mx = m$. We first use Axiom 1.3:

$$m \cdot x = m = m \cdot 1\,,$$

and then apply Axiom 1.5 to the left- and right-hand side of this last equation (note that $m \neq 0$) to deduce that $x = 1$. In summary, assuming $x \in \mathbb{Z}$ has the property that $mx = m$ for some nonzero $m \in \mathbb{Z}$, we concluded that $x = 1$, and this proves our if-then statement. □

Here are some more propositions about inverses and cancellation:

Proposition 1.11. *For all $m, n \in \mathbb{Z}$, $(-m)(-n) = mn$.*

Corollary 1.12. $(-1)(-1) = 1.$[7]

Proposition 1.13.

(i) *For all* $m \in \mathbb{Z}$, $-(-m) = m$.

(ii) $-0 = 0$.

Proposition 1.14. *Given* $m, n \in \mathbb{Z}$ *there is one and only one* $x \in \mathbb{Z}$ *such that* $m + x = n$.

(Later, we will call this solution $n - m$, but at this stage we have no subtraction operation.)

In mathematics, when we say "one of the statements ... is true" we mean "at least one of the statements ... is true." If we mean that *exactly one* statement is true we must say so: we use a phrase like "exactly one," "one and only one," "precisely one."

Proposition 1.15.

(i) $1 \cdot 1 = 1$.

(ii) *If* $x \in \mathbb{Z}$ *and* $x \cdot x = x$ *then* $x = 0$ *or* 1.

Proposition 1.16. *For all integers* m *and* n*:*

(i) $-(m + n) = (-m) + (-n)$.

(ii) $-m = (-1)m$.

(iii) $(-m)n = m(-n) = -(mn)$.

Proposition 1.17. *If* $mn = 0$ *then* $m = 0$ *or* $n = 0$.

This last proposition contains the innocent-looking word **or**. In every-day language, the meaning of "or" is not always clear. It can mean an *exclusive or* (as in "either ... or ... but not both") or an *inclusive or* (as in "either ... or ... or both"). In mathematics, the word "or", without further qualification, is always inclusive. For example, in Proposition 1.17 it might well happen that *both* m and n are zero.

This is so important that we'll say it again: In mathematics, "$\heartsuit$ or $\clubsuit$" always means either $\heartsuit$, or $\clubsuit$, or both $\heartsuit$ and $\clubsuit$.

Proof of Proposition 1.17. Again we have an if-then statement, so let's assume that the integers m and n satisfy $mn = 0$. We need to prove that either $m = 0$ or $n = 0$ (or both). One idea you might have is to rewrite 0 on the right-hand side of the equation $mn = 0$ as $m \cdot 0$ (using Proposition 1.8):

$$m \cdot n = m \cdot 0\,. \tag{1.1}$$

This new equation suggests that we use Axiom 1.5 to cancel m on both sides. We have to be careful here: we can only do that if $m \neq 0$. But that's no problem: if $m = 0$ we are done, since then the statement "$m = 0$ or $n = 0$" is true (note that in that case it might still happen that $n = 0$). If $m \neq 0$, we cancel m in (1.1) to deduce $n = 0$, which again means that the statement "$m = 0$ or $n = 0$" holds. In summary, we have shown that if $mn = 0$ then $m = 0$ or $n = 0$. □

Here's something you may try to prove: Assuming Axioms 1.1–1.5 we proved Proposition 1.17. On the other hand, if we assume Axioms 1.1–1.4 *and* the statement of Proposition 1.17, we can prove the statement of Axiom 1.5. In other words, we could have taken Proposition 1.17 as an axiom in place of Axiom 1.5. This illustrates the arbitrariness of axioms.

[7] We call this result a *corollary* because it follows immediately from the previous result.

1.3 Subtraction

We now define a new operation on $\mathbb{Z}$, called $-$ and known as **subtraction**:

$$m - n \qquad \text{is defined to be} \qquad m + (-n)\,.$$

Proposition 1.18. *For $m, n, p, q \in \mathbb{Z}$:*

(i) $(m - n) + (p - q) = (m + p) - (n + q)$.

(ii) $(m - n) - (p - q) = (m + q) - (n + p)$.

(iii) $(m - n)(p - q) = (mp + nq) - (mq + np)$.

(iv) $m - n = p - q$ *if and only if* $m + q = n + p$.

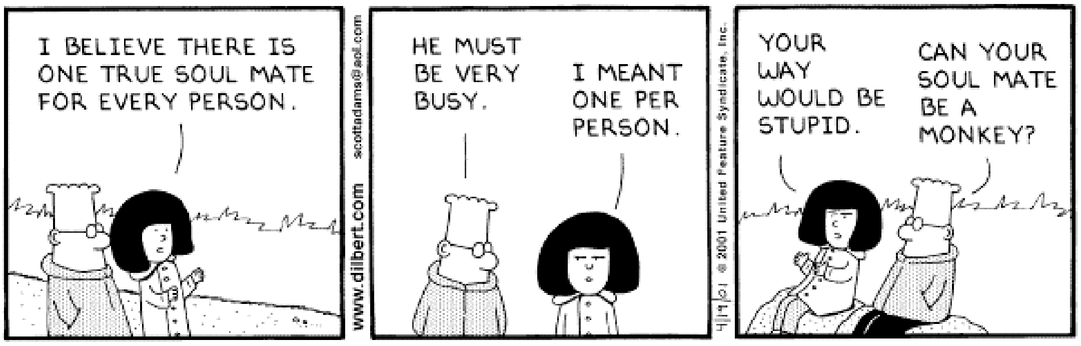
(v) $(m - n)p = mp - np$.

This chapter has been an illustration of the axiomatic method. We avoided philosophical discussions about the integers—questions like "what *is* an integer?" and "in what sense do integers exist?" We simply agreed that, for our purposes, a set satisfying some axioms is assumed to exist. And we explored the consequences. Everything in this book is introduced on that basis. The first axioms we use are found in this chapter. Later, we will need other axioms as we enrich our theory. (The only other axioms used in this book are found in Chapters 3 and 10.)

Chapter 2

Quantifiers

Noel sing we, both all and some.
A line in a 15th century English Christmas carol


I BELIEVE THERE IS ONE TRUE SOUL MATE FOR EVERY PERSON.
scottadams@aol.com
www.dilbert.com
HE MUST BE VERY BUSY.
I MEANT ONE PER PERSON.
© 2001 United Feature Syndicate, Inc.
4/19/01
YOUR WAY WOULD BE STUPID.
CAN YOUR SOUL MATE BE A MONKEY?

2.1 The Universal and Existential Quantifiers

We have already made extensive use of the phrases *for all* and *there exists.* We now take a closer look at these and, while we're at it, we introduce some useful notation.

The symbol $\forall$ means *for all* or *for each* or *for every* or *for any* or *whenever.* Whether or not you think these five phrases mean the same thing in ordinary conversation, they do mean the same thing in mathematics.

The symbol $\exists$ means *there exists* or (in the plural) *there exist.* It is always qualified by a property: i.e., one always says "$\exists$ ... such that..." Another translation of $\exists$ is *for some.* Let's look at some examples.

(i) Axiom 1.2 could be written: $\exists\, 0 \in \mathbb{Z}$ such that $\forall\, m \in \mathbb{Z}\ m + 0 = m$.

(ii) Axiom 1.4 could be written: $\forall\, m \in \mathbb{Z}\ \exists\, n \in \mathbb{Z}$ such that $m + n = 0$.

The symbol $\forall$ is the **universal quantifier** and the symbol $\exists$ is the **existential quantifier**. It is instructive to break up the two sentences in the examples:

(i) $(\exists\, 0 \in \mathbb{Z}$ such that$)$ $(\forall\, m \in \mathbb{Z})$ $m + 0 = m$.

(ii) $(\forall\, m \in \mathbb{Z})$ $(\exists\, n \in \mathbb{Z}$ such that$)$ $m + n = 0$.

Here are some features to note:

- Each statement consists of segments of the form ($\exists$ ... such that) and ($\forall$ ...), and then a final statement: $m + 0 = m$ in the first, and $m + n = 0$ in the second example.
- The order is important. In (ii) n depends on m.
- The informal phrase *for some* really means $\exists$; for example, informally we might write Axiom 1.2 as:

  $$\text{for some } 0 \in \mathbb{Z},\ m + 0 = m \text{ for all } m \in \mathbb{Z}.$$

  The clumsiness of this last sentence should explain why *for some* is not a good phrase for beginners to use. However, you will see *for some* in books and it usually (always?) hides an existence statement "$\exists$ ... such that."

For a more complicated example, let's look again at (ii) above (i.e., at Axiom 1.4) and at the statement that we get by switching the two quantifiers:

(ii) $(\forall\, m \in \mathbb{Z})$ $(\exists\, n \in \mathbb{Z}$ such that$)$ $m + n = 0$.

(iii) $(\exists\, n \in \mathbb{Z}$ such that$)$ $(\forall\, m \in \mathbb{Z})$ $m + n = 0$.

The key fact to note is that in (ii), n depends on m; change m and you expect you'll have to change n accordingly. Axiom 1.4 asserts the existence of the additive inverse for a given number m; it is indeed the case that different m's have different additive inverses.

Now let's look at (iii): Here we claim the existence of an n that will work as the "additive inverse" for *any* m—a statement that you know is wrong. This is one illustration that

$$(\forall\, m \in \mathbb{Z})\, (\exists\, n \in \mathbb{Z} \text{ such that})\ \ldots$$

and

$$(\exists\, n \in \mathbb{Z} \text{ such that})\ (\forall\, m \in \mathbb{Z})\ \ldots$$

mean quite different things.

You can have several $\forall$-phrases in a row, but they can always be reorganized into one:

$$(\forall\, \heartsuit)\, (\forall\, \clubsuit) \qquad \text{has the same meaning as} \qquad (\forall\, \heartsuit \text{ and } \clubsuit)\,.$$

You can also find yourself saying

$$(\exists\ \heartsuit \text{ and } \clubsuit \text{ such that})\ \ldots\ .$$

Then you are asserting the existence of two things which in combination have some property ... .

Once you organize a sentence this way, its *negation* is easily found. After meditating a bit on how to negate statements, you should arrive at the following rules for finding the negation:

(1) Change each ($\forall$ ...) into ($\exists$ ... such that).

(2) Change each ($\exists$ ... such that) into ($\forall$ ...).

(3) Negate the final phrase.

So the statement "Axiom 1.4 does not hold" means:

$$(\exists\, m \in \mathbb{Z} \text{ such that})\ (\forall\, n \in \mathbb{Z})\ m + n \neq 0\,.$$

One more piece of notation: The phrase ($\exists!\, n \in \mathbb{Z}$ such that ...) means that there is a *unique* $n \in \mathbb{Z}$ with the given property. There are two statements here:

(i) existence ($\exists\, n \in \mathbb{Z}$ such that ...) and

(ii) uniqueness (if $n_1 \in \mathbb{Z}$ and $n_2 \in \mathbb{Z}$ both have the given property, then $n_1 = n_2$).

To keep this text as readable as possible, we will avoid the notations $\forall$ and $\exists$ and instead write out "for all" and "there exist(s)." Nevertheless, we recommend using the shortcut notations $\forall$ and $\exists$ when taking notes, working on proofs, etc.

2.2 Implications

We've already talked a little about if-then statements. We use the symbol $\Rightarrow$ to abbreviate an if-then statement; that is, $\heartsuit \Rightarrow \clubsuit$ has the same meaning as "if $\heartsuit$ then $\clubsuit$." For example, $\heartsuit$ could be the statement "it is raining" and $\clubsuit$ the statement "the street is wet." Then $\heartsuit \Rightarrow \clubsuit$ says "if it is raining, then the street is wet." There are two more commonly-used phrases equivalent to $\heartsuit \Rightarrow \clubsuit$, namely "$\heartsuit$ implies $\clubsuit$" and "$\heartsuit$ only if $\clubsuit$." The last phrase can be confusing and is mostly used in *double implications*: we say "$\heartsuit$ if and only if $\clubsuit$" when we want to express the two if-then statements $\heartsuit \Rightarrow \clubsuit$ and $\clubsuit \Rightarrow \heartsuit$; notationally we abbreviate this by $\heartsuit \Leftrightarrow \clubsuit$.[1]

An implication $\heartsuit \Rightarrow \clubsuit$ can be rewritten in terms of the *negatives* of the statements $\heartsuit$ and $\clubsuit$; namely, $\heartsuit \Rightarrow \clubsuit$ has the same meaning as saying (not $\clubsuit$) $\Rightarrow$ (not $\heartsuit$). This latter statement is the *contrapositive* of $\heartsuit \Rightarrow \clubsuit$.

Project 2.1. *This equivalence is something that causes people trouble. Do you really believe that $\heartsuit \Rightarrow \clubsuit$ has the same meaning as (not $\clubsuit$) $\Rightarrow$ (not $\heartsuit$)? Write down three examples. Come up with one that will persuade your English-literature-major roommate that $\heartsuit \Rightarrow \clubsuit$ has indeed the same meaning as (not $\clubsuit$) $\Rightarrow$ (not $\heartsuit$).*

What is the *negation* of the if-then statement $\heartsuit \Rightarrow \clubsuit$? Let's use an example: say we tell you "On mondays we have lunch at the student union" (mathematically speaking: *if* it's monday, *then* we have lunch at the student union). You'd like to prove us wrong (i.e., you feel that the negation of this statement is true), then you'll probably hang out at the student union on mondays checking whether we actually show up. In other words, to prove us wrong, you would show us that "today is monday but you didn't have lunch at the union." So the *negation* of the if-then statement $\heartsuit \Rightarrow \clubsuit$ is the statement $\heartsuit$ and (not $\clubsuit$).

[1] For completeness, we should also mention the statement "$\heartsuit$ if $\clubsuit$," which means $\clubsuit \Rightarrow \heartsuit$, but this is not commonly used.

This has a consequence that might seem counterintuitive at first sight. Namely, the statement $\heartsuit$ and (not $\clubsuit$) is *false* when $\heartsuit$ is *false* (whether $\clubsuit$ is true or false). But this means that $\heartsuit \Rightarrow \clubsuit$ is *true* in these cases, no matter if $\clubsuit$ is true or false! Let's reiterate this: when the statement $\heartsuit$ is *false*, the statement $\heartsuit \Rightarrow \clubsuit$ is *always true.*

This is a good moment to discuss the words *true* and *false.* In ordinary life, it is often not easy to say that a given sentence is true or false; maybe it is neither, maybe it was written down to suggest imprecise ideas. Even deciding what propositions make sense can be difficult. For example:

(1) She loves me, she loves me not.

(2) Colorless green ideas sleep furiously.

(3) What did the professor talk about in class today? Actually it was like totally confusing because he just went on and on about all this stuff about integers, you know, and things like that, and I was like totally not there.

The first of these is poetry, and is expressing a thought entirely different from what the words actually say. The second, a famous example due to the linguist Noam Chomsky (b. 1928) is grammatically correct but meaningless. The third—well, what can we say?

These types of sentence are not allowed in mathematics. People doing mathematics usually prefer the word *statement* rather than *sentence.* It's not easy to say precisely what a mathematical statement is, but this much we can say: it should be a sentence in the ordinary sense, and it should be clearly either true or false, and definitely not both. You may not know whether a particular statement is true or false; in fact, much of mathematics is concerned with trying to decide whether a statement is true or false. But it should be clear on first reading that this is the kind of sentence which by its very structure has to be either true or false.

A sentence consists of words, so we should discuss what kinds of words belong in a mathematical statement. It should be the case that at least the nouns, verbs, adjectives, and adverbs are words whose meanings we already know.[2] So in mathematics we have to build a dictionary before we can even write a mathematical proposition. The words in our dictionary are often called *terms* and they are usually introduced once and once only as *definitions.* There is a huge logical problem here; where do we start? How could we possibly write down the first definition? But if you think about it, you'll realize that the same problem arises in the learning of languages. If we are learning a second language, we can build up the whole dictionary by using some words from our first language at the beginning. But how did each of us learn his or her first language? That's a deep problem in psychology. What we can all say for sure from our own experience is that we didn't learn our first language by building a formal dictionary. Somehow, we knew the meanings of some words and that's what got us started. So, in the same way, we have to start our mathematics by honestly admitting that there are going to be some undefined terms whose meanings we know intuitively. This may seem like a logical mess but it is real life and we are not able to disentangle it. As a practical matter, none of this will cause us much trouble. . .

Project 2.2. *We will soon find out that it is important to practice negating statements. Note that you don't need to know the meaning of the terms below to negate the following statements.*

(i) *G is normal and H is regular.*

[2] Let's ignore words like "the" or "if".

(ii) *Any cubic polynomial has a real root.*

(iii) *The newspaper article was neither accurate nor entertaining.*

(iv) *A sequence of real numbers is convergent only if it is bounded.*

(v) *If x is a real number then x^2 is positive or zero.*

(vi) *H/N is a normal subgroup of G/N if and only if H is a normal subgroup of G.*

(vii) *For all $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that $n \geq N$ implies $|a_n - L| < \epsilon$.*

Chapter 3

Natural Numbers and Induction

Suppose that we think of the integers lined up like dominoes. The inductive step tells us that they are close enough for each domino to knock over the next one, the base case tells us that the first domino falls over, the conclusion is that they all fall over. The fault in this analogy is that it takes time for each domino to fall and so a domino which is a long way along the line won't fall over for a long time. Mathematical implication is outside time.
Peter J. Eccles (*An Introduction to Mathematical Reasoning*, p. 41)

Before getting to the point, we recall standard notation used in describing sets. You will have seen this before.

3.1 Subsets and Set Equality

Let A and B be sets. We write $A \subseteq B$ ("A is a **subset** of B") whenever it is true that every member of A is a member of B:

$$A \subseteq B \qquad \text{means that for all } x, \qquad x \in A \ \Rightarrow \ x \in B\,.$$

The symbol $\supseteq$ is also used: $B \supseteq A$ means $A \subseteq B$.

The **empty set** is the set with no elements. It is denoted $\varnothing$ or $\{\}$.

Proposition 3.1. *The empty set is a subset of any set, that is: $\varnothing \subseteq S$ for any set S.*

Proof. Given a set S, we need to show that if $x \in \varnothing$ then $x \in S$. However, $\varnothing$ contains no element, so the statement "$x \in \varnothing$" is always false. As we discussed in Section 2.2, an *if* ♡ *then* ♣ statement in which ♡ is false is always true; thus $x \in \varnothing \Rightarrow x \in S$ is a true statement, i.e., $\varnothing \subseteq S$. □

Proposition 3.2. *Let A, B, C be sets.*

(i) *$A \subseteq A$.*

(ii) *If $A \subseteq B$ and $B \subseteq C$ then $A \subseteq C$.*

Proof. (i) $A \subseteq A$ means "if $x \in A$ then $x \in A$", which is apparently a true statement.

(ii) Suppose $A \subseteq B$ and $B \subseteq C$. We need to show that if $x \in A$ then $x \in C$. Given $x \in A$, $A \subseteq B$ implies that $x \in B$. This, in turn, implies with $B \subseteq C$ that $x \in C$. □

Project 3.3. *Read through this proof for the case where A is empty. Then there is no such x. Do you see why the proof still holds?*

Next comes **equality**[1] of sets: two sets are equal if they have precisely the same elements. Another way of saying this is that each is a subset of the other: in symbols, $A = B$ when $A \subseteq B$ and $B \subseteq A$. When you want to prove that two sets A and B are equal, it is usually best to prove the two statements $A \subseteq B$ and $B \subseteq A$ separately. Here are some facts about set equality:

Proposition 3.4. *Let A, B, C be sets.*

(i) *$A = A$.*

(ii) *If $A = B$ then $B = A$.*

(iii) *If $A = B$ and $B = C$ then $A = C$.*

These three properties should look familiar—we mentioned them already when we talked about equality of two integers. Back then we called them *reflexivity*, *symmetry*, and *transitivity*, respectively. We'll see these properties again in Section 6.1.

3.2 Axioms for the Natural Numbers

Now we're ready to define the set $\mathbb{N}$ of **natural numbers** axiomatically.

Axiom 3.1. *There is a subset $\mathbb{N} \subseteq \mathbb{Z}$ with the following properties:*

(i) $1 \in \mathbb{N}$.

(ii) *If $n \in \mathbb{N}$ then $n + 1 \in \mathbb{N}$.*

(iii) $0 \notin \mathbb{N}$.

(iv) *For every $n \in \mathbb{Z}$ such that $n \neq 0$, we have $n \in \mathbb{N}$ or $-n \in \mathbb{N}$.*

(v) *If a subset $A \subseteq \mathbb{Z}$ satisfies* (i) *and* (ii) *then $\mathbb{N} \subseteq A$.*

So far, the only integers with names are 0 and 1. Let's name some more:

We'll use the symbol	2	to denote	$1 + 1$
	3		$2 + 1$
	4		$3 + 1$
	5		$4 + 1$
	6		$5 + 1$
	7		$6 + 1$
	8		$7 + 1$
	9		$8 + 1$.

The integers 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 are called **digits**. All except 0 belong to $\mathbb{N}$. This follows from (i) and (ii) of Axiom 3.1. From previous experience, you know what we mean by symbols like

[1] By now you've probably become accustomed to the fact that we use boldface phrases to *define* a term (such as set equality in this case). Contrary to some other disciplines, in mathematics it is important we define each term only *once*.

63, 721, -2719. They are names of other integers. While we may use these names informally (for example, to number the pages of this book) we'll give a proper treatment of the "base 10" system later in Chapter 8.

Proposition 3.5. *Suppose $B \subseteq \mathbb{N}$ is such that:* (1) $1 \in B$ *and* (2) *if* $n \in B$ *then* $n+1 \in B$*. Then* $B = \mathbb{N}$.

Proof. The hypothesis says that $B \subseteq \mathbb{N}$. By Axiom 3.1(v), $\mathbb{N} \subseteq B$. Therefore $B = \mathbb{N}$. □

This proposition immediately gives us a famous principle—important enough to justify the word "theorem."

Theorem 3.6 (Principle of mathematical induction—first form)**.** *Suppose $P(k)$ is a statement depending on a variable $k \in \mathbb{N}$. In order to prove the statement "$P(k)$ is true for all $k \in \mathbb{N}$" it is sufficient to prove: $P(1)$ is true, and for a given $n \in \mathbb{N}$, $P(n)$ implies $P(n+1)$.*

Proof. Let $B = \{k \in \mathbb{N} : P(k) \text{ is true}\}$[2], and suppose that we can prove $P(1)$ is true, and $P(n)$ implies $P(n+1)$. This means $1 \in B$; and if $n \in B$ then $n+1 \in B$. By Proposition 3.5, $B = \mathbb{N}$, so $P(k)$ is true for all $k \in \mathbb{N}$. □

Proofs that use Theorem 3.6 are called *proofs by induction.* Let's look at an example.

Let m and n be integers. We say m **is divisible by** n (or, alternatively, n **divides** m) if there exists $j \in \mathbb{Z}$ such that $m = jn$.

Proposition 3.7.

(i) *For any integer $m \neq 0$, m is not divisible by 0.*

(ii) *For any $k \in \mathbb{N}$, $k^3 + 2k$ is divisible by 3. (By k^3 we mean $k \cdot k \cdot k$.)*

(iii) *For any $k \in \mathbb{N}$, $k^4 - 6k^3 + 11k^2 - 6k$ is divisible by 4.*

(iv) *For any $k \in \mathbb{N}$, $k^3 + 5k$ is divisible by 6.*

Proof of (ii). We will use induction on k. Let $P(k)$ denote the statement "$k^3 + 2k$ is divisible by 3." The induction principle states that we first need to check $P(1)$ (the **base case**), that is, the statement "$1^3 + 2 \cdot 1$ is divisible by 3," which is certainly true: $1^3 + 2 \cdot 1 = 1 + 2 = 3$, and 3 is divisible by 3 by definition (the equation $3 = 3j$ has the solution $j = 1 \in \mathbb{Z}$).

Next comes the **induction step**, that is, we assume that $P(n)$ is true for some n and show that $P(n+1)$ holds as well. So assume that $n^3 + 2n$ is divisible by 3, that is, there exists $y \in \mathbb{Z}$ such that $n^3 + 2n = 3y$. Our goal is to show that $(n+1)^3 + 2(n+1)$ is divisible by 3, that is, we need to show the existence of $z \in \mathbb{Z}$ such that $(n+1)^3 + 2(n+1) = 3z$. But the left-hand side of this equation can be rewritten as

$$\begin{aligned}(n+1)^3 + 2(n+1) &= n^3 + 3n^2 + 3n + 1 + 2n + 2 \\ &= \left(n^3 + 2n\right) + 3n^2 + 3n + 3 \\ &= 3y + 3n^2 + 3n + 3 \\ &= 3\left(y + n^2 + n + 1\right).\end{aligned}$$

[2] The set notation $S = \{x \in A : x \text{ satisfies statement } \clubsuit\}$ means that the elements of the set S are precisely those $x \in A$ that satisfy statement ♣. Many people also use the notation $S = \{x \in A \mid x \text{ satisfies statement } \clubsuit\}$ for this.

So we can set $z = y + n^2 + n + 1$, which is an integer (because y and n are in $\mathbb{Z}$). Thus we have proved that there exists $z \in \mathbb{Z}$ such that $(n+1)^3 + 2(n+1) = 3z$, as desired. □

Project 3.8. *Come up with (and prove) other divisibility statements.*

Proposition 3.9.

(i) *If $m \in \mathbb{N}$ and $k \in \mathbb{N}$ then $m + k \in \mathbb{N}$.*

(ii) *If $m \in \mathbb{N}$ and $k \in \mathbb{N}$ then $mk \in \mathbb{N}$.*

(iii) *For $m \in \mathbb{Z}$, one and only one of the following is true: $m \in \mathbb{N}$, $-m \in \mathbb{N}$, $m = 0$.*

(*Hint for a possible proof of* (i) and (ii): Given $m \in \mathbb{N}$, prove the statements "$m + k \in \mathbb{N}$" and "$mk \in \mathbb{N}$" by induction on $k \in \mathbb{N}$.)

Proof of (iii). Axiom 3.1(iv) already tells us that if $m \in \mathbb{Z}$ is not zero, $m \in \mathbb{N}$ or $-m \in \mathbb{N}$. In other words, for any $m \in \mathbb{Z}$, one of the three statements $m \in \mathbb{N}$, $-m \in \mathbb{N}$, $m = 0$ is true. The hard part is show that *only one* of the three statements applies. If $m = 0$, Axiom 3.1(iii) tells us that $m = 0 \notin \mathbb{N}$, and Proposition 1.13 then implies $-m = -0 = 0 \notin \mathbb{N}$.

Now it remains to prove that if $m \neq 0$, then m and $-m$ cannot both be in $\mathbb{N}$. We use a technique called **proof by contradiction**. The idea is simple: say we want to prove that some statement ♣ is true. Then we start our argument by assuming that ♣ is *false* and we show that this leads to a contradiction. Typically we deduce from it some other statement that is obviously false, such as $0 = 1$ or the negation of one of our axioms.[3]

To prove Proposition 3.9(iii) we need to prove that, given an $m \neq 0$, m and $-m$ are not both in $\mathbb{N}$. The negation of this conclusion is the statement that both $m, -m \in \mathbb{N}$. So we assume (hoping to arrive at a contradiction) that m and $-m$ are both in $\mathbb{N}$. Part (i) tells us that then $m + (-m) \in \mathbb{N}$. But we also know, by Axiom 1.4, that $m + (-m) = 0$. Combining these two statements yields

$$0 = m + (-m) \in \mathbb{N}\,.$$

But $0 \in \mathbb{N}$ contradicts Axiom 3.1(iii)—the statement $0 \in \mathbb{N}$ is precisely the negation of that axiom. This contradiction means that our assumption that both m and $-m$ are in $\mathbb{N}$ must be false, that is, at most one of m and $-m$ is in $\mathbb{N}$. □

This is a good place to say a bit more about proof by contradiction. Let ♡ be a mathematical statement. Then ♡ and "not ♡" both make sense. To say "♡ is false" is the same thing as saying "not ♡ is true." In other words, we assume the **Law of the Excluded Middle**, that ♡ cannot be both true and false—there is no middle ground. This assumption in our logic lies behind proof by contradiction. If you want to prove ♡ is true you can do so directly, or you can prove it by contradiction, i.e.:

[3] The validity of proof by contradiction depends on there being no contradictions built into our axioms. It is the authors' job to make sure our axioms are consistent (i.e., lead to no contradictions). To be really honest we should tell you more: you probably believe (as we do) that the axioms we've introduced so far don't contradict each other, but an amazing theorem of Kurt Gödel (1906–1978) says that *we can't prove this.* In fact, for any "reasonable" system of axioms for arithmetic, it is impossible to prove their consistency without moving to a "higher" theory—roughly speaking, one with more axioms.

(1) Suppose ♡ is false.

(2) Show this leads to a conclusion directly contradicting our axioms or contradicting something that we have already proved to follow from our axioms.

You may then conclude that ♡ must be true, since assuming ♡ false leads to a contradiction, and there is no middle ground. This method of proof is particularly useful for statements that begin "There does not exist . . . " (Suppose it did exist . . .)[4]

3.3 Ordering the Integers

From previous mathematics, you are accustumed to the symbol $<$. If we write $7 < 9$ you read it as "7 is less than 9," and if we write $m < n$ you read it as "m is less than n." If you look back over what we have done so far, you will notice that we have not ordered the integers, and until we do so, phrases like "m is less than n" make no sense. Now that we have defined the positive integers $\mathbb{N}$, we can order the integers. Here is how we do it:

Let $m, n \in \mathbb{Z}$. The statements $m < n$ (m **is less than** n) and $n > m$ (n **is greater than** m) both mean that $n - m \in \mathbb{N}$. The notation $m \leq n$ (m **is less than or equal to** n) and $n \geq m$ (n **is greater than or equal to** m) means that $m < n$ or $m = n$.

Some propositions for inequalities follow below.

Proposition 3.10. *For all $m, n, p \in \mathbb{Z}$, if $m < n$ and $n < p$ then $m < p$.*

Proposition 3.11. $\mathbb{N} = \{n \in \mathbb{Z} : n > 0\}$.

Proposition 3.12. *For all $n \in \mathbb{N}$ there exists $m \in \mathbb{N}$ such that $m > n$.*

The next proposition states that every number n is either positive ($n > 0$), negative ($n < 0$)[5], or zero. Note that 0 is neither positive nor negative.

Proposition 3.13. *For any $n \in \mathbb{Z}$ exactly one of the following holds: $n > 0$, $n < 0$, $n = 0$.*

Proposition 3.14. *For any $n \in \mathbb{N}$, $n \geq 1$.*

Proposition 3.15. *If $m, n \in \mathbb{Z}$ satisfy $m \leq n \leq m$ then $m = n$.*

Proposition 3.16. *There is no integer x such that $0 < x < 1$.*

(*Hint for a possible proof:* This is a good test case for you to construct a proof by contradiction.)

Proposition 3.17. *For all $m, n, p \in \mathbb{Z}$:*

(i) *If $m < n$ then $m + p < n + p$.*

[4] In the early 20th century there was controversy in the mathematical world as to whether a theorem is really proved if it is only proved by contradiction. There was a feeling that a proof is stronger and more convincing if it is direct. With the rise of computer science and interest in computability this is a serious issue in certain circles. We can say, however, that today proof by contradiction is accepted as valid by all but a tiny number of mathematicians.

[5] It is unfortunate that people use the term *negative* to describe two very different phenomena: namely, for the additive inverse of a number and to state that a number is less than zero. For example $-(-3)$, the "negative" of -3, is positive. Be on the watch for confusion of this sort.

(ii) *If $m < n$ and $0 < p$ then $mp < np$.*

(iii) *If $m < n$ and $p < 0$ then $np < mp$.*

Proposition 3.18. *For any $m, n \in \mathbb{Z}$, exactly one of the following is true: $m < n$, $m = n$, $m > n$.*

Proposition 3.19. *Let $n \in \mathbb{Z}$. There is no integer x such that $n < x < n + 1$.*

Proposition 3.20. *If $m \in \mathbb{Z}$ and $m \neq 0$ then $m^2 \in \mathbb{N}$.*

Proposition 3.21. *The equation $x^2 = -1$ has no solution in $\mathbb{Z}$.*

(*Hint for a possible proof:* Another good test case for a proof by contradiction.)

Proposition 3.22. *For all $m, n, p, q \in \mathbb{Z}$:*

(i) *$m - p < m - q$ if and only if $p > q$.*

(ii) *If $p > 0$ and $mp < np$ then $m < n$.*

(iii) *If $p < 0$ and $np < mp$ then $m < n$.*

(iv) *If $m \leq n$ and $0 \leq p$ then $mp \leq np$.*

Proposition 3.23. *If $m, n \in \mathbb{N}$ and n is divisible by m then $m \leq n$.*

It is sometimes useful to start inductions at an integer other than 1:

Theorem 3.24 (Principle of mathematical induction—first form revisited)**.** *Let $P(k)$ be a statement, depending on a variable $k \in \mathbb{Z}$, that makes sense for all $k \geq m$, where m is a fixed integer. In order to prove the statement "$P(k)$ is true for all $k \geq m$" it is sufficient to prove: $P(m)$ is true, and for $n \geq m$, $P(n)$ implies $P(n + 1)$.*

Proof. Let $Q(k)$ be the statement $P(k + m - 1)$, that is, $Q(1) = P(m), Q(2) = P(m + 1)$, etc. The original Theorem 3.6 on induction stated that to prove the statement "$Q(k)$ is true for all $k \geq 1$" it is sufficient to prove $Q(1)$ is true, and for $n \geq 1$, $Q(n)$ implies $Q(n + 1)$. But this is equivalent to saying that to prove the statement "$P(k)$ is true for all $k \geq m$" it is sufficient to prove $P(m)$ is true, and for $n \geq m$, $P(n)$ implies $P(n + 1)$. □

Theorem 3.24 is useful, for example, to prove statements like the following. (One can also prove this without induction—try both ways!).

Proposition 3.25. *For all integers $k \geq 2$, $k^2 < k^3$.*

Chapter 4

More on Sets

Phyllis explained to him, trying to give of her deeper self, “Don’t you find it so beautiful, math? Like an endless sheet of gold chains, each link locked into the one before it, the theorems and functions, one thing making the next inevitable.”
John Updike (*Village*)

4.1 Intersections and Unions

The **intersection** of two sets A and B is

$$A \cap B = \{x : x \in A \text{ and } x \in B\} .$$

The **union** of A and B is

$$A \cup B = \{x : x \in A \text{ or } x \in B\} .$$

The set operations $\cap$ and $\cup$ allow us alternative ways of writing certain sets. Here are two examples.

Proposition 4.1. $\{3x + 1 : x \in \mathbb{Z}\} \cap \{3x + 2 : x \in \mathbb{Z}\} = \varnothing$.

Proposition 4.2. $\{2x : x \in \mathbb{Z}, 3 \le x\} = \{x \in \mathbb{Z} : 5 \le x\} \cap \{x \in \mathbb{Z} : x \text{ is divisible by } 2\}$.

For two sets A and B, we define the **set difference**

$$A - B = \{x : x \in A \text{ and } x \notin B\} .$$

Given a set $A \subseteq X$, we define the **complement** of A in X to be $X - A$. If the bigger set X is clear from the context, we often write $\overline{A}$ or A^c for the complement of A (in X).

Example. We define the **even integers** to be those elements of $\mathbb{Z}$ that are divisible by 2. The **odd integers** are defined to be the complement of the set of even integers (in $\mathbb{Z}$).

Theorem 4.3 (DeMorgan’s laws)**.** *Given two subsets $A, B \subseteq X$,*

$$(A \cup B)^c = A^c \cap B^c \qquad \textit{and} \qquad (A \cap B)^c = A^c \cup B^c .$$

Project 4.4. *Determine which of the following set identities are true; prove your assertions.*

$$A - (B \cup C) = (A - B) \cup (A - C) .$$
$$A \cap (B - C) = (A \cap B) - (A \cap C) .$$

The empty set $\varnothing$ is "extreme" in that it is the smallest possible set. $S \neq \varnothing$ if and only if there is an x such that $x \in S$. One would like to go to the other extreme and define a set that contains "everything"; however, we know since at least the days of Bertrand Russell (1872–1970) that such a set is hard to come by:

Project 4.5. *Consider the set $S = \{X : X \text{ is a set and } X \notin X\}$. Is the statement $S \in S$ true or false?*

4.2 Cartesian Products and Functions

Let A and B be sets. From them we obtain a new set

$$A \times B = \{(a, b) : a \in A \text{ and } b \in B\} .$$

A symbol of the form (a, b) is called an **ordered pair**. The set $A \times B$ is called the **(cartesian) product** of A and B. It is the set of all ordered pairs whose first entry is a member of A and whose second entry is a member of B.

Example. $\mathbb{Z} \times \mathbb{Z}$ is the set of all ordered pairs of integers.

Cartesian products allow us to define the arguably most important concept in all of mathematics. A **function** f with **domain** A and **codomain** B is a rule that assigns to each $a \in A$ one and only one element $f(a)$ of B. We write $f : A \to B$. The **graph** of this function is

$$\Gamma(f) = \{(a, b) \in A \times B : b = f(a)\} .$$

The graph is a subset of $A \times B$ that has exactly one member with first entry a for each $a \in A$.

Example. A binary operation on a set A is a function $f : A \times A \to A$. For example, Axiom 1.1 could be restated as: There are two functions plus: $\mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ and times: $\mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ such that for all integers m, n, and p,

$$\begin{aligned} \text{plus}(m, n) &= \text{plus}(n, m) \\ \text{plus}\left(\text{plus}(m, n), p\right) &= \text{plus}\left(m, \text{plus}(n, p)\right) \\ \text{times}\left(m, \text{plus}(n, p)\right) &= \text{plus}\left(\text{times}(m, n), \text{times}(m, p)\right) \\ \text{times}(m, n) &= \text{times}(n, m) \\ \text{times}\left(\text{times}(m, n), p\right) &= \text{times}\left(m, \text{times}(n, p)\right) . \end{aligned}$$

Some people object to the above definition of *function* on the grounds that the word *rule* is too vague. They prefer to define a function by defining its graph. Thus they would say: a function with domain A and codomain B is a subset Γ of $A \times B$ of the following kind: for each $a \in A$ there is one and only one element of Γ whose first entry is a. If $(a, b) \in \Gamma$, they might write $b = f(a)$ and $\Gamma_f = \Gamma$, informally calling the function f and calling Γ the *graph* of f.

Sometimes mathematicians ask whether a function is *well defined*. What they mean is: "Does the rule you propose really assign to each element of the domain one and only one value in the codomain?" Talking about a well-defined function is a bit like asking if a fact is true—as if there was such a thing as a false fact. But the word is used, so you should know what it means.

Chapter 5

Recursive Definitions

Induction makes you feel guilty for getting something out of nothing, and it is artificial, but it is one of the greatest ideas of civilization.
Herbert S. Wilf

In this chapter we will use the concept of induction to *define* things. Let's start with some examples.

5.1 Examples

Example (Sums)**.** Suppose that for each $j \in \mathbb{N}$, we are given some $x_j \in \mathbb{Z}$.[1] For each $k \in \mathbb{N}$, we want to *define* an element of $\mathbb{Z}$ called $\sum_{j=1}^{k} x_j$. We do so in a way that resembles the principle of induction:

(i) Define $\sum_{j=1}^{1} x_j$ to be x_1.

(ii) Assuming $\sum_{j=1}^{n} x_j$ already defined, we define $\sum_{j=1}^{n+1} x_j$ to be $\left(\sum_{j=1}^{n} x_j\right) + x_{n+1}$.

(We still need to make sure this is a legitimate way to define things.) We sometimes write $x_1 + x_2 + \cdots + x_k$ when we really mean $\sum_{j=1}^{k} x_j$ as defined above.

Example (Products)**.** Similarly, we define $\prod_{j=1}^{k} x_j$ as follows:

(i) Define $\prod_{j=1}^{1} x_j = x_1$.

(ii) Assuming $\prod_{j=1}^{n} x_j$ defined, define $\prod_{j=1}^{n+1} x_j = \left(\prod_{j=1}^{n} x_j\right) \cdot x_{n+1}$.

One can also write $x_1 x_2 \cdots x_k$ for $\prod_{j=1}^{k} x_j$.

Let's denote the nonnegative integers by $\mathbb{Z}_{\geq 0} = \mathbb{N} \cup \{0\} = \{n \in \mathbb{Z} : n \geq 0\}$.[2]

Example (Factorials)**.** As a third example, we define $k!$ ("k factorial"):

[1] In other words, we are given a function $f : \mathbb{N} \to \mathbb{Z}$ where we write x_j instead of $f(j)$.

[2] This is somewhat controversial: about half of all mathematicians would include 0 in the set of natural numbers (by changing our Axiom 3.1 slightly). Fortunately, this is all a matter of definition...

(i) Define $0! = 1$.

(ii) Assuming $n!$ defined (where $n \in \mathbb{Z}_{\geq 0}$), define $(n+1)! = (n!) \cdot (n+1)$.

In these three examples, a function is being defined step by step: the number $f(n+1)$ can only be written down when you already know the number $f(n)$, so it may take serious calculation to actually write down $f(1,000,000)$. The definition of a function given in the previous chapter uses the word *rule*... Sometimes the rule that assigns a value $f(x)$ to each x is given by a formula, e.g., $f(x) = x^2 + 3$. You can see at a glance what answer the rule gives, for any choice[3] of x. Other times, as in these examples, the rule is given recursively, and one could ask if such a rule truly defines a function. The answer is *yes.* In fact, the legitimacy of this method of defining functions can be deduced from our existing axioms, i.e., it is a theorem:

Theorem 5.1. *Let $m \in \mathbb{Z}$ and let A be a set. Then a function $f : \{k \in \mathbb{Z} : k \geq m\} \to A$ is well defined by giving the following:*

(i) *An explicit definition of $f(m)$ (i.e., an assignment of one and only one element of A to be $f(m)$), and*

(ii) *For each $n \geq m$, an explicit definition of $f(n+1) \in A$ in terms of $f(n)$.*

Such a definition is called a *recursive definition.* In the above examples, Theorem 5.1 is being used in the special case where $A = \mathbb{Z}$. In the first two examples, $m = 1$; in the third, $m = 0$.

Proof. Let $P(k)$ be the statement "$f(k)$ is well defined." We will use induction on k (Theorem 3.24) now. For the base case ($k = m$) we know $P(m)$ holds true because of (i). For the induction step, assume that $P(n)$ is true, i.e., $f(n)$ is well defined, for a particular n. Then (ii) allows us to deduce that $f(n+1)$ is also well defined, and that finishes our induction. □

Example (Exponents)**.** Let b be a fixed integer. We define b^k for all integers $k \geq 0$ by:

(i) $b^0 = 1$.

(ii) Assuming b^n defined, let $b^{n+1} = b^n \cdot b$.

Note that this doesn't yet define b^k for $k < 0$. In calculus books a^x is defined to be $e^{x \log a}$ (where log is the natural logarithm, sometimes written ln or $\log_e$). But a is required to be positive because the domain of the function log is the set of positive real numbers. On the other hand, in high school we were accustomed to writing, for example, $(-3)^2 = 9$ or $(-3)^3 = -27$. At our present stage, we only have integers, not real numbers. Our definition of b^k when $b < 0$ agrees with the high-school definition.

Proposition 5.2. *For $b \in \mathbb{N}$ and $k, m \in \mathbb{Z}_{\geq 0}$,*

(i) $b^k \in \mathbb{N}$.

(ii) $b^m b^k = b^{m+k}$.

[3] Well, not quite: the number x might be so huge that it would take the biggest computer in the world to write it down, and even that computer might not be able to handle the number $x^2 + 3$; however in a mathematical sense our sentence is correct.

(iii) $(b^m)^k = b^{mk}$.

Proof. We prove (i) by induction on $k \ge 0$; let $P(k)$ be the statement "$b^k \in \mathbb{N}$." The base case $P(0)$ follows with $b^0 = 1 \in \mathbb{N}$ (by definition). For the induction step, suppose we know that $b^n \in \mathbb{N}$ for some n; our goal is to conclude that $b^{n+1} \in \mathbb{N}$ also. By definition, $b^{n+1} = b^n \cdot b$; and since we know that both b and b^n are in $\mathbb{N}$, we conclude by Proposition 3.9 that their product is also in $\mathbb{N}$.

For part (ii), suppose we are given $b \in \mathbb{N}$ and $m \in \mathbb{Z}_{\ge 0}$. We will prove the statement $P(k) : b^m b^k = b^{m+k}$ by induction on $k \ge 0$. The base case $P(0)$ follows with $b^0 = 1$ and $m + 0 = m$, and so $P(0)$ simply states that

$$b^m \cdot 1 = b^m,$$

which holds by Axiom 1.3. For the induction step, assume we know that $b^m b^n = b^{m+n}$ for some n; our goal is to conclude that $b^m b^{n+1} = b^{m+n+1}$. The left-hand side of this equation is, by definition,

$$b^m b^{n+1} = b^m \cdot b^n \cdot b\,, \tag{5.1}$$

whereas the right-hand side is, again by definition,

$$b^{m+n+1} = b^{m+n} \cdot b\,. \tag{5.2}$$

That the right-hand sides of (5.1) and (5.2) are equal follows with our induction assumption.

Part (iii) follows in a similar fashion. Namely, suppose we are given $b \in \mathbb{N}$ and $m \in \mathbb{Z}_{\ge 0}$. We will prove $P(k) : (b^m)^k = b^{mk}$ by induction on $k \ge 0$. The base case $P(0)$ states that

$$(b^m)^0 = b^{m\cdot 0}.$$

However, both sides of this equation are 1—the left-hand side by definition, and the right-hand side by Proposition 1.8. For the induction step, assume we know that $(b^m)^n = b^{mn}$ for some n; our goal is to prove $(b^m)^{n+1} = b^{m(n+1)}$. The left-hand side of this equation is, by definition,

$$(b^m)^{n+1} = (b^m)^n (b^m)\,, \tag{5.3}$$

and the right-hand side is, by definition and part (ii),

$$b^{m(n+1)} = b^{mn+m} = b^{mn} b^m. \tag{5.4}$$

That the right-hand sides of (5.3) and (5.4) are equal follows with our induction assumption. □

The recursive definition of exponents allows us to make more divisibility statements; here are a few examples.

Proposition 5.3. *For any $k \in \mathbb{N}$:*

(i) $5^{2k} - 1$ *is divisible by* 24.

(ii) $2^{2k+1} + 1$ *is divisible by* 3.

(iii) $10^k + 3 \cdot 4^{k+2} + 5$ *is divisible by* 9.

Project 5.4. *Determine for which natural numbers $k^2 < 2^k$ and prove your answer.*

As another example of a recursive construction, we invite you to further explore unions and intersections of sets.

Project 5.5 (Unions and intersections)**.** *Given sets $A_1, A_2, A_3, \ldots$, develop recursive definitions for $\bigcup_{j=1}^k A_j$ and $\bigcap_{j=1}^k A_j$. Find and prove the extension of DeMorgan's laws (Theorem 4.3) for these unions and intersections.*

5.2 Sequences

A **sequence** in the set A is a function $f : \mathbb{N} \to A$.

Example. $A = \mathbb{Z}$ and $f(k) = k^3 + k$. Then $f(1) = 2$, $f(2) = 10$, $f(3) = 30$ etc. $f(k)$ is called the k^{th} *term* of the sequence. Sometimes we use a different notation: $x_k = k^3 + k$. Then $x_1 = 2$, $x_2 = 10$, $x_3 = 30$ etc. But this is the same concept: we have just renamed $f(k)$ to be x_k. Then the whole sequence is denoted by $(x_k)_{k=1}^{\infty}$ or $(x_k)_{k\in\mathbb{N}}$. Our example above could be written as $\left(k^3 + k\right)_{k\in\mathbb{N}}$.

Suppose $m, n \in \mathbb{Z}$ and m divides n. Then, by definition, there exists $j \in \mathbb{Z}$ such that $n = mj$. We denote[4] this integer j by $\frac{n}{m}$.

Example ("$3x + 1$ problem")**.** Pick your favorite natural number m, and define the following sequence:

(i) Define $x_1 = m$, that is, set x_1 to be your favorite number.

(ii) Assuming x_n defined, define $x_{n+1} = \begin{cases} \frac{x_n}{2} & \text{if 2 divides } x_n, \\ 3x_n + 1 & \text{otherwise.} \end{cases}$

For example, if your favorite natural number is $m = 1$ then the sequence $(x_k)_{k\in\mathbb{N}}$ starts with $1, 4, 2, 1, 4, 2, 1, 4, 2, \ldots$ It is a famous open conjecture that, no matter which $m \in \mathbb{N}$ one chooses as the starting point, the sequence eventually takes on the value 1 (from which point the remaining sequence looks like our above example).

Note that a sequence is a function. In the first example, above, we have defined our sequence by a formula. In the second example, the sequence is defined recursively.

Proposition 5.6.

(i) *Let $m \in \mathbb{Z}$ and let $(x_j)_{j\in\mathbb{N}}$ be a sequence in $\mathbb{Z}$. Then $m \cdot \left(\sum_{j=1}^{k} x_j\right) = \sum_{j=1}^{k} (mx_j)$.*

(ii) *If $x_j = 1$ for all $j \in \mathbb{N}$ then $\sum_{j=1}^{k} x_j = k$ for all $k \in \mathbb{N}$.*

(iii) *If $x_j = n \in \mathbb{Z}$ for all $j \in \mathbb{N}$ then $\sum_{j=1}^{k} x_j = kn$ for all $k \in \mathbb{N}$.*

Proposition 5.7.

(i) *Let a, b, c be integers such that $a < b < c$. Then $\sum_{j=a}^{c} x_j = \sum_{j=a}^{b} x_j + \sum_{j=b+1}^{c} x_j$.*[5]

(ii) *Let r be a natural number. Then $\sum_{j=a}^{b} x_j = \sum_{j=a+r}^{b+r} x_{j-r}$.*

(iii) $\sum_{j=a}^{b}(x_j + y_j) = \left(\sum_{j=a}^{b} x_j\right) + \left(\sum_{j=a}^{b} y_j\right)$.

Proposition 5.8. (i) $\sum_{j=1}^{k} j = \frac{k(k+1)}{2}$.

[4] Note that $\frac{n}{m}$ is an integer here; this definition does not describe rational numbers (i.e., fractions). For example, our number system does not yet include $\frac{1}{2}$.

[5] As a warm-up exercise, try to figure out a rigorous definition for the sum $\sum_{j=m}^{n} x_j$ for two integers $m \leq n$.

(ii) $\sum_{j=1}^{k} j^2 = \frac{k(k+1)(2k+1)}{6}$.
(In particular, $k(k+1)$ is divisible by 2 and $k(k+1)(2k+1)$ is divisible by 6, for all $k \in \mathbb{N}$.)

Project 5.9. *Find (and prove) a formula for $\sum_{j=1}^{k} j^3$.*

Setting $x_k = \sum_{j=1}^{k} j$ illustrates the fact that we can think of a sum like $\sum_{j=1}^{k} j$ as a sequence with parameter k. Such sequences, defined as sums, are called **series**. Here is another example, the **geometric series**:

Proposition 5.10. *For $x \neq 1$ and $k \in \mathbb{N} \cup \{0\}$, $\displaystyle\sum_{j=0}^{k} x^j = \frac{1 - x^{k+1}}{1 - x}$.*

5.3 The Binomial Theorem

Theorem 5.11. *If $k, m \in \mathbb{Z}_{\geq 0}$ and $m \leq k$ then $m!(k-m)!$ divides $k!$.*

Thus $\frac{k!}{m!(k-m)!}$ is an integer; it is customary to denote the integer by the symbol $\binom{k}{m}$ and call it the **binomial coefficient**.[6]

Proof. We prove the statement $P(k)$: "for all $m \in \mathbb{Z}_{\geq 0}$ such that $m \leq k$, there exists $j \in \mathbb{Z}$ such that $k! = j\, m!(k-m)!$" by induction on k.

The base case $k = 0$ implies $m = 0$, and we can choose $j = 1$ to obtain $P(0)$: $0! = 1 \cdot 0! \cdot 0!$.

For the induction step, we assume $P(n)$ is true for a particular n. To prove $P(n+1)$ for a given m, we would like to use the induction hypothesis $P(n)$ for m and $m-1$; we will first treat the cases $m = 0$ and $m = n+1$.

When $m = 0$, we choose $j = 1$, and obtain the statement $P(n+1)$: $(n+1)! = 1 \cdot 0!\,(n+1)!$. When $m = n+1$, we again choose $j = 1$, and obtain the statement $P(n+1)$: $(n+1)! = 1 \cdot (n+1)!\,0!$.

Now suppose $1 \leq m \leq n$. The induction hypothesis $P(n)$, applied to both m and $m-1$, implies that there exist integers a and b such that

$$n! = a\,(m-1)!\,(n-m+1)! \qquad \text{and} \qquad n! = b\,m!\,(n-m)!\,.$$

But then

$$\begin{aligned}(n+1)! &= n!\,(m+n+1-m)\\ &= n!\,m \;+\; n!\,(n+1-m)\\ &= a\,(m-1)!\,(n-m+1)!\,m \;+\; b\,m!\,(n-m)!\,(n+1-m)\\ &= (a+b)\,m!\,(n-m+1)!\end{aligned}$$

which completes our induction step. □

In our proof of Theorem 5.11—more precisely, in the last math line—we obtained the following recursive identity for the binomial coefficients:

[6] In other math classes you will come across the more general definition $\binom{k}{m} = \frac{k(k-1)(k-2)\cdots(k-m+2)(k-m+1)}{m!}$. This expression has the advantage that k is not required to be an integer—it can be a real or even a complex number. (The number m is still required to be a nonnegative integer.) You should convince yourself that the two definitions agree when $k \in \mathbb{Z}_{\geq 0}$.

Corollary 5.12. *For $1 \le m \le k$, $\binom{k+1}{m} = \binom{k}{m-1} + \binom{k}{m}$.*

You have most certainly seen this recursion in disguise; namely when you discussed binomial expansions in school: $(a+b)^2 = a^2 + 2ab + b^2$, $(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$, etc. Your teacher (hopefully) explained that one can obtain the coefficients from the **binomial triangle**:

$$\begin{array}{ccccccccccccc}
 & & & & & & 1 & & & & & & \\
 & & & & & 1 & & 1 & & & & & \\
 & & & & 1 & & 2 & & 1 & & & & \\
 & & & 1 & & 3 & & 3 & & 1 & & & \\
 & & 1 & & 4 & & 6 & & 4 & & 1 & & \\
 & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & \\
1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1 \\
 & & & & & & \vdots & & & & & &
\end{array}$$

Corollary 5.12 says how to compute a row from the previous one, and the following theorem is about the binomial expansion we've just talked about.

Theorem 5.13 (Binomial theorem for integers)**.** *If $a, b \in \mathbb{Z}$ and $k \in \mathbb{Z}_{\ge 0}$ then*

$$(a+b)^k = \sum_{m=0}^{k} \binom{k}{m} a^m \, b^{k-m}.$$

An immediate corollary of the binomial theorem, namely the special case $a = b = 1$, gives another relation among the binomial coefficients:

Corollary 5.14. *For $k \in \mathbb{Z}_{\ge 0}$, $\displaystyle\sum_{m=0}^{k} \binom{k}{m} = 2^k$.*

A slight variation of the binomial theorem is the general product formula of calculus—although this looks like a completely different topic at first sight.

Proposition 5.15 (Leibniz's formula[7])**.** *The formulas for differentiation of sums and products in calculus are:*

$$\begin{aligned}
(u+v)' &= u' + v' \\
(uv)' &= uv' + u'v \\
(ku)' &= ku' \quad \textit{where } k \textit{ is a constant.}
\end{aligned}$$

Define $w^{(k)}$ recursively by $w^{(0)} = w$, $w^{(n+1)} = (w^{(n)})'$. Then (formally—no calculus involved):

$$(uv)^{(k)} = \sum_{m=0}^{k} \binom{k}{m} u^{(m)} v^{(k-m)}.$$

[7] This formula was found by Gottfried Leibniz (1646–1716)—a codiscoverer (with Isaac Newton) of calculus—in 1678.

The binomial coefficients appear all over mathematics. We give one more property, which will become handy soon. An integer $n \geq 2$ is **prime** if it is only divisible by ± 1 and $\pm n$. The first few primes are 2, 3, 5, 7, 11, 13, 17, etc.

Proposition 5.16. *Let p be prime and $0 < m < p$. Then $\binom{p}{m}$ is divisible by p.*

Primes are of fundamental importance in the arithmetic of integers. One can prove (try it—e.g., you can use induction) that every integer > 1 can be factored into primes in an essentially unique way. There are many intriguing theorems about primes and just as many open questions. To give a flavor of the latter, we mention an example: **Twin primes** are two primes that differ by 2 (the smallest possible difference between two odd primes). Examples of twin primes include $(3, 5)$, $(17, 19)$, and $(41, 43)$. It is an open question whether there are infinitely many twin primes.[8] We will not discuss primes further in this book, except in Proposition 6.11.

[8] This open problem raises the much simpler question whether there are infinitely many primes. The answer (i) is not entirely obvious, (ii) is yes, and (iii) has probably been known for 2500 years. After discussing Proposition 6.5 you should be able to prove it yourself.

Chapter 6

Equivalence Relations and Integers Modulo n

Mathematics is the art of giving the same name to different things.
Jules Henri Poincaré (1854–1912)

6.1 Relations

A **relation** on a set A is a subset of $A \times A$. Given a relation $R \subseteq A \times A$, we often write $x \sim y$ instead of $(x, y) \in R$. In words, we say that x **is related to** y (by the relation R).

Example. Some familiar examples of relations $a \sim b$ in $\mathbb{Z}$ are:

- $a = b$
- $a < b$
- $a \leq b$
- a divides b.

What we mean here is that any of $=, <, \leq$, and *divides* can play the role of $\sim$.

Example. The graph of a function $f : A \to A$ is a special case of a relation (for which there is only one $(x, y) \in R$ for every $x \in A$).

The relation $R \subseteq A \times A$ is an **equivalence relation** if it has the following three properties:

(i) $a \sim a$ for all $a \in A$. *(reflexivity)*

(ii) $a \sim b$ implies $b \sim a$. *(symmetry)*

(iii) $a \sim b$ and $b \sim c$ imply $a \sim c$. *(transitivity)*

Given an equivalence relation $\sim$ on A, the **equivalence class** of $a \in A$ is $[a] := \{b \in A : b \sim a\}$.

Example. Of the above examples of relations $a \sim b$, only the one defined by $a = b$ is an equivalence relation.

Proposition 6.1. *Given an equivalence relation $\sim$ on a set A and $a, b \in A$.*

(i) $a \in [a]$.

(ii) $a \sim b$ *if and only if* $[a] = [b]$.

The equivalence classes defined by an equivalence relation on the set A chop up A in the following sense:

Proposition 6.2. *Given an equivalence relation on a set A.*

(i) *For $a_1, a_2 \in A$, either $[a_1] = [a_2]$ or $[a_1] \cap [a_2] = \varnothing$.*

(ii) *For every $a \in A$, there is an equivalence class containing a.*

A **partition** of A is a set Π of subsets of A satisfying:

(i) If $P_1, P_2 \in \Pi$, then either $P_1 = P_2$ or $P_1 \cap P_2 = \varnothing$.

(ii) Every $a \in A$ belongs to some $P \in \Pi$.

Proposition 6.2 states that the equivalence classes (with respect to a given equivalence relation on A) form a partition of A. The following proposition provides the converse:

Proposition 6.3. *Let Π be a partition of A. Define $\sim$ by: $a \sim b$ if and only if a and b lie in the same element of Π. Then $\sim$ is an equivalence relation.*

Project 6.4. *Determine which of the following is an equivalence relation. If it is, determine the equivalence classes.*

(1) *The relation $\sim$ on $\mathbb{Z}$ defined by $x \sim y$ if and only if $x < y$.*

(2) *The relation $\sim$ on $\mathbb{Z}$ defined by $x \sim y$ if and only if $x \leq y$.*

(3) *The relation $\sim$ on $\mathbb{Z}$ defined by $x \sim y$ if and only if $|x| = |y|$. Here $|x|$ denotes the* absolute value[1] *of $x \in \mathbb{Z}$, which is defined to be x if $x \geq 0$ and to be $-x$ if $x < 0$.*

(4) *The relation $\sim$ on the set of all lines (say, in a plane) defined by $L_1 \sim L_2$ if and only if L_1 is parallel to L_2.*

(5) *The relation $\sim$ on $\mathbb{Z}$ defined by $x \sim y$ if and only if $x|y$ or $y|x$.*

[1] We will discuss absolute value in more detail in Section 13.2.

6.2 The Integers Modulo n

The concept of an equivalence relation is best illustrated with an example. Here we use "clock arithmetic" (except that we do not require our clocks to have 12 numbers). For this we need a theorem that expresses something you always knew: that when you divide one positive integer into another, there will be a remainder (possibly 0) which is less than the dividing number.

Theorem 6.5 (Division Algorithm)**.** *Fix $n \in \mathbb{N}$. For every $m \in \mathbb{Z}$ there exist $q, r \in \mathbb{Z}$ such that $0 \le r \le n-1$ and*

$$m = qn + r\,.$$

(*Hint for a possible proof:* Consider first the case $m \ge 0$ by induction on m, and then the case $m < 0$.)

Corollary 6.6. *The integer m is odd if and only if there exists $q \in \mathbb{Z}$ such that $m = 2q + 1$.*

You thought you knew this corollary already? Try proving it with out the Division Algorithm (Theorem 6.5).

Proposition 6.7. *The integer m is even if and only if m^2 is even.*

Given a fixed $n \in \mathbb{N}$, we define the relation $\equiv$ on $\mathbb{Z}$ by $x \equiv y$ if and only if $x - y$ is divisible by n. When there is any possibility of ambiguity about n, we write this as $x \equiv y \bmod n$.

Example. Let's discuss the case $n = 2$. Here $x \equiv y \bmod 2$ means that $x - y$ is even, in other words, x and y have the *same parity.* In some sense, the relation $\equiv$ (for different n) generalizes the notion of parity.

Proposition 6.8.

(i) *$\equiv$ is an equivalence relation on $\mathbb{Z}$.*

(ii) *The equivalence relation $\equiv$ has exactly n distinct equivalence classes, namely $[0], [1], \ldots, [n-1]$.*

This set of equivalence classes is called the **set of integers modulo** n, written $\mathbb{Z}_n$ or $\mathbb{Z}/n\mathbb{Z}$. You might think of the numbers in $\mathbb{Z}_n$ as remainders when we divide integers by n, and this is essentially the content of the Division Algorithm (Theorem 6.5). We will define addition and multiplication operations on $\mathbb{Z}_n$ and see that $\mathbb{Z}_n$ satisfies many (but not all) of our axioms for integers.

Proposition 6.9. *If $a \equiv a'$ and $b \equiv b'$ then $a + b \equiv a' + b'$ and $ab \equiv a'b'$.*

This proposition allows us to make the following definition: For elements $[a]$ and $[b]$ of $\mathbb{Z}_n$, we define **addition** $\oplus$ and **multiplication** $\odot$ on $\mathbb{Z}_n$ via

$$[a] \oplus [b] = [a + b] \qquad \text{and} \qquad [a] \odot [b] = [ab]\,.$$

Proposition 6.10. *Addition $\oplus$ and multiplication $\odot$ on $\mathbb{Z}_n$ are commutative, associative, and distributive. $\mathbb{Z}_n$ has an additive identity and a multiplicative identity.*

The set $\mathbb{Z}_n$ is of fundamental importance in a variety of areas. For example, many computer encryption schemes are based on arithmetic in $\mathbb{Z}_n$. Among the different $\mathbb{Z}_n$'s there are distinguished members, namely those for which n is prime:

Theorem 6.11 (Fermat's little theorem[2]). *For an integer m and a prime p,*

$$m^p \equiv m \bmod p\,.$$

(*Hint for a possible proof:* Use induction on $m \geq 0$ and Proposition 5.16.)

$\mathbb{Z}_n$ is by no means the only meaningful example of a set of equivalence classes on which we can define new binary operations. Here is another one:

Project 6.12. *On $\mathbb{Z} \times (\mathbb{Z} - \{0\})$ we define the equivalence relation $(m_1, n_1) \sim (m_2, n_2)$ if and only if $m_1 n_2 = n_1 m_2$. (You should check that this is indeed an equivalence relation.) For two equivalence classes $[(m_1, n_1)]$ and $[(m_2, n_2)]$, we define addition and multiplication via*

$$[(m_1, n_1)] + [(m_2, n_2)] = [(m_1 n_2 + m_2 n_1, n_1 n_2)] \quad \textit{and} \quad [(m_1, n_1)] \cdot [(m_2, n_2)] = [(m_1 m_2, n_1 n_2)]\,.$$

What properties do these binary operations have?

Your instructor may or may not want to tell you why this project is important.

Here is one more example of an equivalence relation that you might be familiar with from a previous math class.

Example. In linear algebra let V be a finite-dimensional vector space and let W be a linear subspace. Define the equivalence relation $\sim$ on V by $v_1 \sim v_2$ if and only if $v_1 - v_2 \in W$. The set of equivalence classes is denoted by V/W and is again a vector space. In the language of linear algebra, the map $v \mapsto [v]$ is a surjective linear map $V \to V/W$ whose kernel is W.

[2] Where there is a *little* theorem there is a *big* theorem. In Fermat's case it states that for $n \geq 3$, the equation $x^n + y^n = z^n$ has no solutions $(x, y, z) \in \mathbb{N}^3$. This is known as Fermat's Last Theorem. Pierre de Fermat (1601–1665) wrote the statement in the margin of a book in 1637 and claimed he had a proof which was too long to fit in the margin. Many famous and would-be-famous mathematicians tried to give a proof, but mistakes were always found. It is now accepted that the proof given in 1995 by Andrew Wiles (b. 1953) is correct. The word "last" refers to the fact that all other theorems stated (often without proof) by Fermat were proved over the intervening 358 years, most of them long ago. (The methods used by Wiles are highly sophisticated and could not possibly have been known to Fermat.)

Chapter 7

The Well Ordering Principle, Strong Induction, and More Recursions

If things are nice there is probably a good reason why they are nice: and if you do not know at least one reason for this good fortune, then you still have work to do.
Richard Askey

7.1 The Well Ordering Principle and a Second Form of Induction

We would now like to establish the following simple-sounding fact: every nonempty subset of $\mathbb{N}$ has a smallest element. Let's first define things properly. Suppose $A \subseteq \mathbb{Z}$ is not empty.[1]

(i) We say that A is **bounded above** if there exists $b \in \mathbb{Z}$ such that for all $a \in A$, $a \leq b$. The number b is called an **upper bound** for A. If $b \in A$, we say that b is the **largest element** of A.[2]

(ii) We say that A is **bounded below** if there exists $b \in \mathbb{Z}$ such that for all $a \in A$, $b \leq a$. The number b is called a **lower bound** for A. If $b \in A$, we say that b is the **least element** of A.

(iii) We say that A is **bounded** if it is both bounded above and bounded below.

Example. The integers -2 and 1 are lower bounds for $\mathbb{N}$, and 1 is the least element of $\mathbb{N}$.

Theorem 7.1 (Well ordering principle)**.** *Every nonempty subset of* $\mathbb{N}$ *has a least element.*

Theorem 7.1 can be very useful and is well worth memorizing. We'll see it in action several times in what follows.

We now recall the principle of induction (Theorem 3.6) in what we called its first form: Suppose $P(k)$ is a statement depending on a variable $k \in \mathbb{N}$. In order to prove the statement "$P(k)$ is true for all $k \in \mathbb{N}$" it is sufficient to prove: $P(1)$ is true, and $P(n)$ implies $P(n+1)$. There is a second form of this, a trick that people call **strong induction**. It's just another way of stating the same idea, but it sounds different, and it can be useful:

[1] Later on, we'll use exactly the same definition for subsets of $\mathbb{R}$.

[2] This language suggests that the largest element of A, if it exists, is unique, which you may try to prove.

Theorem 7.2 (Principle of mathematical induction—second form)**.** *Suppose $P(k)$ is a statement depending on a variable $k \in \mathbb{N}$. In order to prove the statement "$P(k)$ is true for all $k \in \mathbb{N}$" it is sufficient to prove:*

(i) *$P(1)$ is true;*

(ii) *if $P(j)$ is true for all $1 \leq j \leq n$, then $P(n+1)$ is true.*

One way to prove this theorem is by the first form of induction (Theorem 3.6)—try proving the statement "$P(j)$ is true for all $1 \leq j \leq k$" by induction on k.

Project 7.3 (Principle of mathematical induction—second form revisited)**.** *State and prove a strong induction theorem for a statement "$P(k)$ is true for all $k \geq m$" for a given $m \in \mathbb{Z}$, paralleling Theorem 3.24.*

7.2 Higher-Order Recursions

The strong induction principle allows us to define a new kind of recursive sequence. Here is a famous example:

Example. The **Fibonacci numbers**[3] are defined by $f : \mathbb{N} \to \mathbb{N}$, $f(1) = 1$, $f(2) = 1$, and

$$f(n) = f(n-2) + f(n-1) \qquad \text{for } n \geq 3\,. \tag{7.1}$$

A **recurrence relation** for a sequence is a formula for computing the k^{th} element in terms of the lower-indexed elements. The equation (7.1) for the Fibonacci numbers is an example.

Project 7.4. *Calculate $f(13)$.*

The Fibonacci numbers obey a surprising formula—assuming for a moment that we know real numbers, such as $\sqrt{5}$.

Proposition 7.5. *The k^{th} Fibonacci number is given directly by the formula*

$$f(k) = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2} \right)^k - \left(\frac{1-\sqrt{5}}{2} \right)^k \right).$$

Note that this result implies that the strange expression on the right-hand side is always an integer.

Proof. Let $a = \frac{1+\sqrt{5}}{2}$ and $b = \frac{1-\sqrt{5}}{2}$. We prove $P(k) : f(k) = \frac{1}{\sqrt{5}} \left(a^k - b^k \right)$ by (strong) induction on $k \in \mathbb{N}$. For starters, we check $P(1)$ and $P(2)$, for which the formula gives $f(1) = 1 = f(2)$. For the induction step, assume that $P(j)$ is true for $1 \leq j \leq n$, for some n. Then, by definition of the

[3] The Fibonacci numbers are named after Leonardo of Pisa (c. 1170—c. 1250), also known as Leonardo Fibonacci.

Fibonacci sequence and induction assumption,

$$\begin{aligned}
f(n+1) &= f(n) + f(n-1) \\
&= \frac{1}{\sqrt{5}}\left(a^n - b^n\right) + \frac{1}{\sqrt{5}}\left(a^{n-1} - b^{n-1}\right) \\
&= \frac{1}{\sqrt{5}}\left(a^n + a^{n-1} - b^n - b^{n-1}\right) \\
&= \frac{1}{\sqrt{5}}\left(a^{n-1}(a+1) - b^{n-1}(b+1)\right) \\
&= \frac{1}{\sqrt{5}}\left(a^{n-1}a^2 - b^{n-1}b^2\right) \\
&= \frac{1}{\sqrt{5}}\left(a^{n+1} - b^{n+1}\right).
\end{aligned}$$

Here the penultimate step follows with

$$a+1 = \frac{3+\sqrt{5}}{2} = \frac{1+2\sqrt{5}+5}{4} = a^2 \qquad \text{and} \qquad b+1 = \frac{3-\sqrt{5}}{2} = \frac{1-2\sqrt{5}+5}{4} = b^2. \qquad \square$$

The Fibonacci numbers have numerous other properties; to give a flavor we present a few more here (try to prove them without using Proposition 7.5):

Proposition 7.6. *For all $k, m \in \mathbb{N}$, where $m \geq 2$, $f(m+k) = f(m-1)f(k) + f(m)f(k+1)$.*

Proposition 7.7. *For all $k \in \mathbb{N}$, $f(2k+1) = f(k)^2 + f(k+1)^2$.*

Proposition 7.8. *For all $k, m \in \mathbb{N}$, $f(mk)$ is divisible by $f(m)$.*

Chapter 8

Base Ten Representation of Integers

If people do not believe that mathematics is simple, it is only because they do not realize how complicated life is.
John von Neumann (1903–1957)

In our axioms two (distinct!) elements of $\mathbb{Z}_{\geq 0}$ were given names: 0 and 1. Later some more elements of $\mathbb{Z}_{\geq 0}$ were given names: 2, 3, 4, 5, 6, 7, 8, 9. Now we give the name 10 to the integer $9+1$.

Proposition 8.1. *If $n \in \mathbb{Z}_{\geq 0}$ then $n < 10^n$.*

Let us define a function $\nu : \mathbb{Z}_{\geq 0} \to \mathbb{N}$ as follows: $\nu(0) = 1$, and for all $n \in \mathbb{N}$, $\nu(n)$ equals the least element of $\{t \in \mathbb{N} : n < 10^t\}$. $\nu(n)$ is called **the number of digits of n with respect to base 10**. Our definition of ν makes sense because, by Proposition 8.1, $\{t \in \mathbb{Z}_{\geq 0} : n < 10^t\}$ contains n and is therefore nonempty. So the Well Ordering Principle (Theorem 7.1) guarantees that this set contains a unique least element.

Example. $\nu(d) = 1$ for all digits d. $\nu(10) = 2$.

Proposition 8.2. *For all $n \in \mathbb{N}$, $\nu(n) = k$ if and only if $10^{k-1} \leq n < 10^k$.*

The whole literate world has been taught that every nonnegative integer can be represented by a finite string of digits, and that different strings correspond to different integers. Since none of this is in our axioms, we must derive it from the axioms.

Theorem 8.3 (Existence of base ten representation for elements of $\mathbb{Z}_{\geq 0}$)**.** *Let $n \in \mathbb{Z}_{\geq 0}$. Then there exist digits $x_0, x_1, \ldots, x_{\nu(n)-1}$ such that $n = \displaystyle\sum_{i=0}^{\nu(n)-1} x_i\, 10^i$. Moreover, if $n > 0$ then there exists such digits with $x_{\nu(n)-1} > 0$.*

Proof. The trick is to prove this theorem by induction on $\nu(n)$ (rather than, e.g., by induction on n).

If $\nu(n) = 1$, then either $n = 0$ or $1 \leq n < 10$, by Proposition 8.2. In either case, the required digit x_0 is n, and if $n > 0$, $x_0 > 0$. Assume the theorem is true for all $n \in \mathbb{Z}_{\geq 0}$ such that $\nu(n) \leq m$, where $m \geq 1$. Now let $\nu(n) = m+1$. Then $10^m \leq n < 10^{m+1}$, by Proposition 8.2. Let

$S = \{t \in \mathbb{N} : n - 10^m < t \cdot 10^m\}$. We see that $S \neq \varnothing$ because $9 \in S$. By the well ordering principle (Theorem 7.1), S has a least element r. Since $9 \in S$, r is a digit ≥ 1. Since $n - 10^m < r \cdot 10^m$ we have $n < (r+1)10^m$. And we also know that $r \cdot 10^m \leq n$, because, if $r = 1$ we have seen $10^m \leq n$, and if $r > 1$ then this follows from the "least element" property of r. Let $s = n - r \cdot 10^m$. Then $s < 10^m$ (why?). So $\nu(s) \leq m$. By our induction assumption, there exist digits $x_0, \ldots, x_{\nu(s)-1}$ such that $s = \displaystyle\sum_{i=0}^{\nu(s)-1} x_i 10^i$. We have $n = r \cdot 10^m + s$. Thus $n = \displaystyle\sum_{i=0}^{m} y_i 10^i$ where

$$y_i = \begin{cases} x_i & \text{if } i \leq \nu(s) - 1, \\ 0 & \text{if } \nu(s) \leq i < m, \\ r & \text{if } i = m. \end{cases}$$

Since $r > 0$ and $m = \nu(n) - 1$, we are done. □

Proposition 8.4. *For any $r \in \mathbb{N}$, $\left(\displaystyle\sum_{i=0}^{r-1} 9 \cdot 10^i\right) + 1 = 10^r$.*

Proposition 8.5 (Uniqueness of base ten representation for elements of $\mathbb{Z}_{\geq 0}$)**.** *Let $n \in \mathbb{Z}_{\geq 0}$. Let $n = \sum_{i=0}^{p} x_i 10^i = \sum_{i=0}^{q} y_i 10^i$ where $p, q \in \mathbb{Z}_{\geq 0}$, each x_i and each y_i is a digit, $x_p \neq 0$, and $y_q \neq 0$. Then $p = q$, and $x_i = y_i$ for all i.*

Proposition 8.6. *If $\nu(n) > \nu(n-1)$ then n is a power of* 10.

Proposition 8.7. *Let $n \in \mathbb{Z}_{\geq 0}$. Then n is divisible by* 3 *if and only if the sum of its digits is divisible by* 3.

(*Hint for a possible proof:* Write $n = \sum_{i=0}^{\nu(n)-1} x_i \, 10^i$ as in Theorem 8.3. Define $\sigma(n) = \sum_{i=0}^{\nu(n)-1} x_i$. Prove that $n - \sigma(n)$ is divisible by 3.)

Project 8.8. *Come up with (and prove) other divisibility tests.*

Let $m \in \mathbb{Z}_{\geq 0}$. By Theorem 8.3 and Proposition 8.5, for each $i \in \mathbb{Z}_{\geq 0}$ such that $0 \leq i \leq \nu(m) - 1$ there is a unique digit x_i such that $m = \sum_{i=0}^{\nu(m)-1} x_i \, 10^i$. It is convenient (as we have all been taught since childhood) to represent m by the string of digits $x_{\nu(m)-1} x_{\nu(m)-2} \cdots x_2 x_1 x_0$. For example, $m = 3 \cdot 10^0 + 6 \cdot 10^1 + 8 \cdot 10^2 + 0 \cdot 10^3 + 7 \cdot 10^4$ is represented by 70863. This string is called the **base ten representation** of m. The following is a criterion for deciding when $m < n$:

Theorem 8.9. *Let $m, n \in \mathbb{Z}_{\geq 0}$. Let m have base ten representation $x_{\nu(m)-1} x_{\nu(m)-2} \cdots x_2 x_1 x_0$ and let n have base ten representation $y_{\nu(n)-1} y_{\nu(n)-2} \cdots y_2 y_1 y_0$. Then $m < n$ if and only if either*

(i) $\nu(m) < \nu(n)$

or

(ii) $\nu(m) = \nu(n)$ *and $x_j < y_j$, where j is the least element of $\{i \in \mathbb{Z}_{\geq 0} : x_k = y_k$ for all $k > i\}$.*

Here are the steps for a possible proof:

(1) $9\sum_{i=0}^{k} 10^i < 10^{k+1}$.

(2) If $m < n$ then $\nu(m) \leq \nu(n)$.

(3) If $m < n$ either (i) or (ii) holds.

(4) If (i) or (ii) holds, then $m < n$.

Any integer $k \geq 2$ can be used as a base for representing the integers. (Why should we not use base 1?) The base 2 is usually used in computer code. This section has been written so that the proof of existence and uniqueness can easily be imitated from the case $k = 10$ by making simple changes as follows.

(1) The definition of *digit* must be changed. For base k, the digits are $\{n \in \mathbb{Z}_{\geq 0} : n < k\}$. They must be named with symbols. E.g., for base 2 use 0 and 1. For base 12 use 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, t, e.

(2) In the definition of $\nu(n)$ replace 10 by k.

(3) In the proofs of the existence and uniqueness theorems replace 10 by k everywhere, and use the new ν.

Chapter 9

The Algorithm for Addition of Two Nonnegative Numbers (Base 10)

One and one and one is three.
The Beatles (*Come together*)

An **algorithm** is a procedure, e.g., a computer program, for doing something mathematical step by step. (That's not a formal definition: it's not an easy matter to write down the formal meaning of the word *algorithm.*)

We saw that each $m \in \mathbb{Z}_{\geq 0}$ has a unique representation as $\sum_{i=0}^{\nu(m)-1} x_i\, 10^i$ with $x_{\nu(m)-1} > 0$ when $m > 0$. If $n \in \mathbb{Z}_{\geq 0}$ and $n = \sum_{i=0}^{\nu(n)-1} y_i\, 10^i$, we want an algorithm for the digits $z_0, z_1, \ldots$ when $m + n$ is written as $\sum_{i=0}^{\nu(m+n)-1} z_i\, 10^i$. We are all familiar with this: for example, if $m = 332$ and $n = 841$ our previous knowledge of mathematics leads us to believe the statement $m + n = 1173$. What did we do to get 1173? Our goal is to describe the process rigorously.

The Algorithm: Given as input digits $x_0, \ldots, x_q$ and $y_0, \ldots, y_q$, the output of the algorithm is the ordered pairs $(z_0, i_0), \ldots, (z_{q+1}, i_{q+1})$ where each z_k is a digit and each i_k is 0 or 1. The z_k's and i_k's are defined in stages recursively:

Stage 0: The input consists of two digits x_0 and y_0. The output is (z_0, i_0) as follows: if $x_0+y_0 < 10$ then $z_0 = x_0 + y_0$ and $i_0 = 0$; if $10 \leq x_0 + y_0$ then $x_0 + y_0 = d_0 + 10$ where d_0 is a digit (why?) and in that case $z_0 = d_0$ and $i_0 = 1$.

Stage k: Here, $1 \leq k \leq q$. The input consists of two digits x_k and y_k as well as i_{k-1} which is either 0 or 1. The output is (z_k, i_k) as follows:

Case 1: If $x_k + y_k < 10$ and $i_{k-1} = 0$, define $z_k = x_k + y_k$ and $i_k = 0$.

Case 2: If $x_k + y_k < 9$ and $i_{k-1} = 1$, define $z_k = x_k + y_k + 1$ and $i_k = 0$.

Case 3: If $10 \leq x_k + y_k$ and $i_{k-1} = 0$ then there is a unique digit d_k such that $x_k + y_k = d_k + 10$; define $z_k = d_k$ and $i_k = 1$.

Case 4: If $9 \leq x_k + y_k$ and $i_{k-1} = 1$ then there is a unique digit d_k such that $x_k + y_k + 1 = d_k + 10$; define $z_k = d_k$ and $i_k = 1$.

Stage $q+1$: Define $z_{q+1} = i_q$ and $i_{q+1} = 0$.

We remark that the output (z_0, i_0) only depends on x_0 and y_0. If $k \geq 1$, the output (z_k, i_k) only depends on x_k, y_k and i_{k-1}. Now we show that the algorithm gives the right answer:

Theorem 9.1. *If $m = \sum_{i=0}^{q} x_i\, 10^i$ and $n = \sum_{i=0}^{q} y_i\, 10^i$ where each x_i and each y_i is a digit then $m + n = \sum_{i=0}^{q+1} z_i\, 10^i$, where the digits $z_0, \ldots, z_{q+1}$ are obtained from the algorithm.*

Note that we allow $x_q = 0$ or $y_q = 0$, so these sums might not be the standard base 10 representations of m and n. The reason for doing it this way is easily seen if you add 27 to 4641. We are in effect adding 0027 to 4641.

Proof. We proceed by induction on q. In the base case $q = 0$, the theorem says that $x_0 + y_0 = z_0 + i_0 \cdot 10$ which is true—just look at Stage 0.

For the induction step, assume the theorem is true whenever q is replaced by $q - 1$. When $x_0, \ldots, x_{q-1}$ and $y_0, \ldots, y_{q-1}$ are the input, let $(z'_0, i'_0), \ldots, (z'_q, i'_q)$ form the output. By induction $\sum_{i=0}^{q-1} x_i\, 10^i + \sum_{i=0}^{q-1} y_i\, 10^i = \sum_{i=0}^{q} z_i\, 10^i$. We already remarked that $z'_k = z_k$ when $k \leq q - 1$. The last stage of the algorithm gives $z'_q = i_{q-1}$. So

$$\begin{aligned} m + n &= \sum_{i=0}^{q-1} x_i\, 10^i + x_q\, 10^q + \sum_{i=0}^{q-1} y_i\, 10^i + y_q\, 10^q \\ &= \sum_{i=0}^{q-1} z_i\, 10^i + i_{q-1}\, 10^q + x_q\, 10^q + y_q\, 10^q \\ &= \sum_{i=0}^{q+1} z_i\, 10^i, \end{aligned}$$

since the algorithm gives $z_q\, 10^q + z_{q+1}\, 10^{q+1} = (x_q + y_q + i_{q-1})10^q$. □

Let $p = \max\{\nu(m), \nu(n)\}$. If $m = \sum_{i=0}^{\nu(m)-1} x_i\, 10^i$ and $n = \sum_{i=0}^{\nu(n)-1} y_i\, 10^i$ and if we define $x_k = 0$ for $\nu(m) \leq k \leq p - 1$ and $y_k = 0$ for $\nu(n) \leq k \leq p - 1$ (if there are such k's), then $m = \sum_{i=0}^{p-1} x_i\, 10^i$ and $n = \sum_{i=0}^{p-1} y_i\, 10^i$. The algorithm tells us how to find $z_0, \ldots, z_p$ in $m + n = \sum_{i=0}^{p} z_i\, 10^i$.

Proposition 9.2. *Let $p = \max\{\nu(m), \nu(n)\}$. Then $\nu(m + n) = p$ or $p + 1$.*

(*Hint for a possible proof:* Use Exercise 8.2.)

Example. (i) If $m = 332$, $n = 841$, $p = 3$, then $\nu(m + n) = 4$.

(ii) If $m = 32$, $n = 641$, $p = 3$, then $\nu(m + n) = 3$.

Proposition 9.3. *$z_p = 0$ or 1. If $z_p = 0$, $\nu(m + n) = p$. If $z_p = 1$, $\nu(m + n) = p + 1$.*

Proof. By the last stage of the algorithm, $z_p = i_{p-1}$, which is 0 or 1. If $z_p = 1$ then $\nu(m + n)$ is $p + 1$. If $z_p = 0$ then $\nu(m + n) \leq p$. By Proposition 9.2, $\nu(m + n) = p$. □

The same approach can be used to prove the correctness of the other grade school algorithms:

(1) Subtraction

$$\begin{array}{r} 461 \\ 29 \\ \hline 432 \end{array}$$

(2) Long addition

$$\begin{array}{r} 461 \\ 29 \\ 391 \\ \hline 881 \end{array}$$

(3) Long multiplication

$$\begin{array}{r} 461 \\ 29 \\ \hline 4149 \\ 9220 \\ \hline 13369 \end{array}$$

Project 9.4. *In the long addition* (2) *above, the digits* 5*,* 7*,* 8*, and* 0 *do not appear in the numbers being added, while* 1 *and* 9 *appear more than once. Is there a long addition as in* (2) *so that each digit is used exactly once and the total is* 100*? (Each number should be written in the usual way,* 7 *rather than* 07 *for example). Either find such a long addition, or prove there is none.*

Part II: The Continuous

Chapter 10

Real Numbers

Mathematical study and research are very suggestive of mountaineering. Whymper made several efforts before he climbed the Matterhorn in the 1860's and even then it cost the life of four of his party. Now, however, any tourist can be hauled up for a small cost, and perhaps does not appreciate the difficulty of the original ascent. So in mathematics, it may be found hard to realise the great initial difficulty of making a little step which now seems so natural and obvious, and it may not be surprising if such a step has been found and lost again.
Louis Joel Mordell (1888–1972)

Just like the integers, the real numbers—which you saw in calculus and before—will be defined through a set of axioms. They will look somewhat like the axioms of Chapters 1 and 3, with a few subtle twists which make all the difference.

10.1 First Axioms

We assume the existence of a set, denoted $\mathbb{R}$, whose members are called **real numbers**. This set $\mathbb{R}$ is assumed to be equipped with binary operations $+$ and $\cdot$ satisfying the following Axioms 10.1–10.7.

Axiom 10.1. *For all $x, y, z \in \mathbb{R}$:*

(i) $x + y = y + x$.

(ii) $(x + y) + z = x + (y + z)$.

(iii) $x \cdot (y + z) = x \cdot y + x \cdot z$.

(iv) $x \cdot y = y \cdot x$.

(v) $(x \cdot y) \cdot z = x \cdot (y \cdot z)$.

The product $x \cdot y$ is often written xy.

Axiom 10.2. *There is a real number $0 \in \mathbb{R}$ such that for all $x \in \mathbb{R}$, $x + 0 = x$.*

Axiom 10.3. *There is a real number $1 \in \mathbb{R}$ with $1 \neq 0$ such that for all $x \in \mathbb{R}$, $x \cdot 1 = x$.*

Axiom 10.4. *For all $x \in \mathbb{R}$, there exists a real number denoted $-x \in \mathbb{R}$, such that $x + (-x) = 0$.*

Axiom 10.5. *For all $x \in \mathbb{R} \setminus \{0\}$, there exists a real number, denoted x^{-1}, such that $x \cdot x^{-1} = 1$.*

Proposition 10.1. *Suppose $x, y, z \in \mathbb{R}$ and $x \neq 0$. If $xy = xz$ then $y = z$.*

Let's compare Proposition 10.1 with Axiom 1.5 in Chapter 1. The proposition asserts that the cancellation property described in Axiom 1.5 also holds in $\mathbb{R}$. It follows that $\mathbb{R}$ and $\mathbb{Z}$ both satisfy Axioms 1.1–1.5 of Chapter 1. Hence any proposition we discovered about $\mathbb{Z}$ using only Axioms 1.1–1.5 is also valid for $\mathbb{R}$. The same holds true for the propositions about subtraction, once we make the following definition.

As with $\mathbb{Z}$, we define **subtraction** in $\mathbb{R}$ by

$$x - y = x + (-y) .$$

Here is a definition that we could not make in $\mathbb{Z}$: We define a new operation $\div$ on $\mathbb{R}$, called **division**, through

$$y \div x = y \cdot x^{-1} .$$

Axiom 10.5 does not assert the existence of 0^{-1}; so division makes no sense when $a = 0$. In the language of Section 4.2, the division function is

$$\text{division} : \mathbb{R} \times (\mathbb{R} - \{0\}) \to \mathbb{R}, \ \text{division}(y, x) = y \cdot x^{-1} .$$

The number x^{-1} is usually written as $\frac{1}{x}$, and $y \div x$ is usually written[1] as $\frac{y}{x}$.

Project 10.2. *Think about why division by 0 does not make sense. Come up with some arguments that will convince your English-literature-major roommate.*

10.2 Ordering the Reals

Axiom 10.6. *There is a subset $\mathbb{R}_{>0} \subseteq \mathbb{R}$ (whose members are called* **positive real numbers***) satisfying:*

(i) *If $x, y \in \mathbb{R}_{>0}$ then $x + y \in \mathbb{R}_{>0}$.*

(ii) *If $x, y \in \mathbb{R}_{>0}$ then $xy \in \mathbb{R}_{>0}$.*

(iii) *For $x \in \mathbb{R}$, one and only one of the following is true: $x \in \mathbb{R}_{>0}$, $-x \in \mathbb{R}_{>0}$, $x = 0$.*

This does not look like Axiom 3.1 for $\mathbb{N}$. However, Axiom 10.6 very much resembles Proposition 3.9.

Analogous to the less-than definition in $\mathbb{Z}$, we write $x < y$ or $y > x$ if $y - x \in \mathbb{R}_{>0}$ and $x \leq y$ or $y \geq x$ if we additionally allow $x = y$. The analogy of the $<$ relation on $\mathbb{R}$ with the one defined on $\mathbb{Z}$ continues:

Proposition 10.3. *Let $x, y, z, w \in \mathbb{R}$.*

[1] Do not confuse the division symbol with the symbol | which describes the divisibility property of integers.

(i) *If $x < y$ and $y < z$ then $x < z$.*

(ii) *If $0 < x < y$ and $0 < z \leq w$ then $xz < yw$.*

(iii) *If $x \neq y$ then $x < y$ or $y < x$.*

(iv) *If $x < 0$ and $y < z$ then $xy > xz$.*

(v) *If $x \leq y \leq x$ then $x = y$.*

Proposition 10.4. *Suppose $x, y \in \mathbb{R}$ satisfy $x \geq 0$ and $y \geq 0$. Then $x = y$ if and only if $x^2 = y^2$.*

The next proposition states which of the axioms of $\mathbb{N}$ in Chapter 3 hold for $\mathbb{R}_{>0}$.

Proposition 10.5.

(i) $1 \in \mathbb{R}_{>0}$.

(ii) *If $x \in \mathbb{R}_{>0}$, then $x + 1 \in \mathbb{R}_{>0}$.*

(iii) $0 \notin \mathbb{R}_{>0}$.

(iv) *For every $x \in \mathbb{R}$ such that $x \neq 0$, either $x \in \mathbb{R}_{>0}$ or $-x \in \mathbb{R}_{>0}$.*

Proof. We prove (i) by contradiction. Assume that $1 \notin \mathbb{R}_{>0}$. Then Axiom 10.6(iii) implies that $-1 \in \mathbb{R}_{>0}$, since $1 \neq 0$. But then, by Axiom 10.6(ii), $1 = (-1)(-1) \in \mathbb{R}_{>0}$, which contradicts our initial assumption that $1 \notin \mathbb{R}_{>0}$.

(ii) follows immediately with (i) and Axiom 10.6(i), and (iii) and (iv) with Axiom 10.6(iii). □

It follows that everything deduced previously for $\mathbb{Z}$ from Axioms 1.1–1.5 and 3.1 holds automatically for $\mathbb{R}$, provided Axiom 3.1(v) is not used.[2] We will see shortly that Axiom 3.1(v) does not hold in $\mathbb{R}$. First we note:

Proposition 10.6. *If $x \in \mathbb{R}_{>0}$ then $\frac{1}{x} \in \mathbb{R}_{>0}$.*

Here is one way in which $\mathbb{R}_{>0}$ and $\mathbb{N}$ are very different. Note that the integer 1 is the least element of $\mathbb{N}$, by Proposition 3.14. By contrast:

Theorem 10.7. *$\mathbb{R}_{>0}$ does not have a least element.*

Proof. Define the real number $2 = 1 + 1$; by Proposition 10.5(i) and Axiom 10.6(i), $2 \in \mathbb{R}_{>0}$. Proposition 10.6 implies that $2^{-1} = \frac{1}{2}$ is also positive.

We claim further that $\frac{1}{2} < 1$; otherwise, Proposition 10.3(ii) (with $1 < 2$ and $0 < 1 \leq \frac{1}{2}$) would imply that $1 < 1$, a contradiction.

We have hence established $0 < \frac{1}{2} < 1$ and can start the actual proof of Theorem 10.7. We will use a proof by contradiction. Assume that there exists a least element $s \in \mathbb{R}_{>0}$. Then we can use Proposition 10.3(ii) (with $0 < \frac{1}{2} < 1$ and $0 < s \leq s$) to deduce $\frac{1}{2} \cdot s < s$. However, $\frac{1}{2} \cdot s \in \mathbb{R}_{>0}$ (by Axiom 10.6(ii)), which contradicts the fact that s is the least element in $\mathbb{R}_{>0}$. □

[2] $\mathbb{Z}$ and $\mathbb{R}$ are quite different objects—we have not yet embedded $\mathbb{Z}$ in $\mathbb{R}$ as your intuition might expect us to do (and as we will do later). The point of the axiomatic method is that if the axioms are the same, all conclusions from the axioms must also be the same. That is why we are pointing out which axioms for $\mathbb{Z}$ and $\mathbb{R}$ are the same and which are not.

We labeled Theorem 10.7 as a theorem rather than a proposition to emphasize its importance. An argument could be made (though we wouldn't dream of making it) that Theorem 10.7 is the most important theorem of Part II of this book. What we're really trying to say here is that in many of your advanced mathematics courses—courses with words like *analysis* and *topology* in their titles—the instructor will use Theorem 10.7 regularly. It may not be mentioned explicitly, but it will be used in "ϵ-δ arguments." We will discuss this in more detail in Chapter 13.

A closely related result, also of great importance, is:

Theorem 10.8. *Suppose $x, y \in \mathbb{R}$ with $x < y$. Then there exists $z \in \mathbb{R}$ such that $x < z < y$.*

There is one more axiom for $\mathbb{R}$, for which we first need another definition: An upper bound for $A \subseteq \mathbb{R}$ is a **least upper bound** for A if it is less than or equal to every upper bound for A. Least upper bounds are unique if they exist:

Proposition 10.9. *If x_1 and x_2 are least upper bounds for A then $x_1 = x_2$.*

The least upper bound of A is denoted $\sup A$ (an abbreviation for **supremum**) or $\operatorname{lub} A$.

Example. $-1 = \sup\{-n : n \in \mathbb{N}\}$.

The least upper bound of a set may not exist. For example:

Proposition 10.10. *$\mathbb{R}_{>0}$ does not have a least upper bound.*

Proposition 10.11. *Let A be a set. If $\sup A \in A$ then $\sup A$ is the largest element in A. Conversely, if A has a largest element b then $b = \sup A$.*

At this point it is useful to define **intervals**. They come in nine types: Suppose $x < y$, then

$$\begin{aligned}
[x, y] &= \{z \in \mathbb{R} : x \le z \le y\} \\
(x, y] &= \{z \in \mathbb{R} : x < z \le y\} \\
[x, y) &= \{z \in \mathbb{R} : x \le z < y\} \\
(x, y) &= \{z \in \mathbb{R} : x < z < y\} \\
(-\infty, x] &= \{z \in \mathbb{R} : z \le x\} \\
(-\infty, x) &= \{z \in \mathbb{R} : z < x\} \\
[x, \infty) &= \{z \in \mathbb{R} : x \le z\} \\
(x, \infty) &= \{z \in \mathbb{R} : x < z\} \\
(-\infty, \infty) &= \mathbb{R}\,.
\end{aligned}$$

Example. The intervals $(2, 3)$, $[-1, 3]$, and $(-\infty, 3]$ all have least upper bound 3. The first of these intervals, $(2, 3)$, does not have a largest element.

Project 10.12. *For $B \subseteq \mathbb{R}$, $B \neq \varnothing$ we can define the* greatest lower bound *(*$\inf B$*, for* infimum*, or* $\operatorname{glb} B$*) of B. Give the precise definition for* $\inf B$ *and prove that it is unique if it exists.*

Here is the final axiom for the real numbers.

Axiom 10.7. *Every non-empty subset of $\mathbb{R}$ that is bounded above has a least upper bound.*

This axiom, which concludes our definition of $\mathbb{R}$, is given here only because people referring back later might forget to include it in the list. It needs explanation, indeed a chapter of its own—Chapter 13.

Chapter 11

More on Functions

I believe that numbers and functions of analysis are not the arbitrary result of our minds; I think that they exist outside of us, with the same character of necessity as the things of objective reality, and we meet them or discover them, and study them, as do the physicists, the chemists and the zoologists.
David Hilbert (1862–1943)

We have now defined two number systems $\mathbb{Z}$ and $\mathbb{R}$, and intuitively, we think of the integers as a subset of the real numbers. However, nothing in our axioms tells us explicitly that $\mathbb{Z}$ can be viewed as a subset of $\mathbb{R}$. In fact, at the moment we have no axiomatic reason to think that the symbols 0 and 1 defined in the axioms for $\mathbb{Z}$ can be thought of as the same symbols 0 and 1 defined in the axioms for $\mathbb{R}$. Just for now, we will be more careful and write $0_{\mathbb{Z}}$ and $1_{\mathbb{Z}}$ for these special members of $\mathbb{Z}$, and $0_{\mathbb{R}}$ and $1_{\mathbb{R}}$ for the corresponding special members of $\mathbb{R}$. Of course, intuitively we are accustomed to identifying $0_{\mathbb{Z}}$ with $0_{\mathbb{R}}$ and identifying $1_{\mathbb{Z}}$ with $1_{\mathbb{R}}$. We will justify this identification in this chapter by giving an *embedding* of $\mathbb{Z}$ into $\mathbb{R}$, that is, a function whose image of each integer is the corresponding number in $\mathbb{R}$. In preparation, we must first discuss functions with special properties.

11.1 Injections and Surjections

Here is the concept that will allow us to embed $\mathbb{Z}$ into $\mathbb{R}$ (in Section 11.2): A function $f : A \to B$ is **injective** (or is **one-to-one**, or is an **injection**) if

$$\text{for all } a_1, a_2 \in A: \ a_1 \neq a_2 \text{ implies } f(a_1) \neq f(a_2)\,.$$

An equivalent definition (see Section 2.2) for a function $f : A \to B$ to be injective is

$$\text{for all } a_1, a_2 \in A: \ f(a_1) = f(a_2) \text{ implies } a_1 = a_2\,.$$

Example. The function $f_1 : \mathbb{Z} \to \mathbb{Z}$ defined by $f_1(n) = 3n$ is injective. The function $f_2 : \mathbb{Z} \to \mathbb{Z}$ defined by $f_2(n) = n^2$ is not injective.

The function $f : A \to B$ is **surjective** (or is **onto**, or is a **surjection**) if

$$\text{for all } b \in B \text{ there exists } a \in A \text{ such that } f(a) = b\,.$$

The **image** (or **range**) of a function $f : A \to B$, denoted $f(A)$, is the set $\{f(a) : a \in A\}$. Thus $f : A \to B$ is onto if and only if $f(A) = B$, or, in words, if its image equals its codomain. The function $f : A \to B$ is **bijective** (or a **bijection** or a **one-to-one correspondence**) if f is both injective and surjective.

Example. A seemingly trivial but important function is the **identity function** of the set A, namely the function $\mathrm{id}_A : A \to A$, defined by $\mathrm{id}_A(a) = a$ for all $a \in A$. This function is bijective.

Project 11.1. *Determine which of the following functions is injective, surjective, or bijective. Justify your assertions.*

(i) $f : \mathbb{Z} \to \mathbb{Z},\ f(n) = n^2$.

(ii) $f : \mathbb{Z} \to \mathbb{Z}_{\geq 0},\ f(n) = n^2$.

(iii) $f : \mathbb{Z}_{\geq 0} \to \mathbb{Z}_{\geq 0},\ f(n) = n^2$.

(iv) $f : \mathbb{R} \to \mathbb{R},\ f(x) = 3x + 1$.

(v) $f : \mathbb{R}_{\geq 0} \to \mathbb{R},\ f(x) = 3x + 1$.

(vi) $f : \mathbb{Z} \to \mathbb{Z},\ f(x) = 3x + 1$.

Project 11.2. *Construct (many) functions that are*

(i) *bijective;*

(ii) *injective, but not surjective;*

(iii) *surjective, but not injective;*

(iv) *neither injective nor surjective.*

Justify your claims.

The **composition** of two functions $f : A \to B$ and $g : B \to C$ is

$$g \circ f : A \to C \ \text{ given by } \ (g \circ f)(a) = g(f(a)) \text{ for all } a \in A\,.$$

Proposition 11.3.

(i) *If $f : A \to B$ is injective and $g : B \to C$ is injective then $g \circ f : A \to C$ is injective.*

(ii) *If $f : A \to B$ is surjective and $g : B \to C$ is surjective then $g \circ f : A \to C$ is surjective.*

(iii) *If $f : A \to B$ is bijective and $g : B \to C$ is bijective then $g \circ f : A \to C$ is bijective.*

A **left inverse** to a function $f : A \to B$ is a function $g : B \to A$ such that $g \circ f = \mathrm{id}_A$. A **right inverse** to a function $f : A \to B$ is a function $g : B \to A$ such that $f \circ g = \mathrm{id}_B$. A **(two-sided) inverse** to f is a function that is both a left inverse and a right inverse.

Proposition 11.4.

(i) *f is injective if and only if f has a left inverse.*

(ii) *f is surjective if and only if f has a right inverse.*

(iii) *f is bijective if and only if f has an inverse.*

Proof. (i) Suppose $f : A \to B$ is injective. Then fix an $a_0 \in A$ and define the function $g : B \to A$ through

$$g(b) := \begin{cases} a & \text{if } b \text{ is in the image of } f \text{ and } f(a) = b, \\ a_0 & \text{otherwise.} \end{cases}$$

Then g is a well-defined function, because f is injective.

Conversely, suppose $f : A \to B$ has a left inverse g, that is, $g : B \to A$ is a function such that $g \circ f = \mathrm{id}_A$. Suppose $a_1, a_2 \in A$ satisfy $f(a_1) = f(a_2)$; to prove that f is injective we will show that this equation implies $a_1 = a_2$. Because g is a function, $f(a_1) = f(a_2)$ implies that

$$(g \circ f)(a_1) = g(f(a_1)) = g(f(a_2)) = (g \circ f)(a_2) .$$

Comparing the left-hand side of this equation with the right-hand side yields $a_1 = a_2$, since $g \circ f = \mathrm{id}_A$.

(ii) Suppose $f : A \to B$ is surjective. We will construct a function $g : B \to A$ as follows: Given $b \in B$, choose an $a \in A$ such that $f(a) = b$ (we can possibly find more than one such a, in which case we choose one). We define this a to be the image of b under the function g, that is, we define[1] $g(b) = a$. With this definition, we obtain $(f \circ g)(b) = f(g(b)) = f(a) = b$, that is, g is a right inverse to f.

Conversely, suppose $f : A \to B$ has a right inverse g, that is, $g : B \to A$ is a function such that $f \circ g = \mathrm{id}_B$. We need to show that f is surjective, that is, given $b \in B$, we need to find $a \in A$ such that $f(a) = b$. Given such a $b \in B$, we define $a = g(b)$. Then by construction, $f(a) = f(g(b)) = (f \circ g)(b) = b$.

(iii) follows from (i) and (ii). □

Proposition 11.5. *Let A and B be sets. There exists an injection from A to B if and only if there exists a surjection from B to A.*

Proposition 11.6. *Suppose $f : A \to B$ is a function where A and B are finite sets with the same number of elements. Then f is injective if and only if f is surjective.*

Project 11.7. *Suppose $f : A \to B$ and $g : B \to C$. Decide if each of the following is true or false; in each case prove the statement or give a counterexample.*

(i) *If f is injective and g is surjective then $g \circ f$ is surjective.*

(ii) *If $g \circ f$ is bijective then g is surjective and f is injective.*

[1] The *Axiom of Choice* is the assertion that this way of defining the function g is legitimate. Lurking here is a deep issue in Set Theory.

11.2 The Relationship of $\mathbb{Z}$ and $\mathbb{R}$

We want to embed $\mathbb{Z}$ into $\mathbb{R}$. To do this, we define a function $e : \mathbb{Z} \to \mathbb{R}$ as follows:

(i) Define e on $\mathbb{Z}_{\geq 0}$ recursively: $e\left(0_{\mathbb{Z}}\right) = 0_{\mathbb{R}}$ and, assuming $e(n)$ defined for a fixed $n \in \mathbb{Z}_{\geq 0}$, define $e\left(n + 1_{\mathbb{Z}}\right) = e(n) + 1_{\mathbb{R}}$.[2]

(ii) If $n \in \mathbb{Z}$ and $n < 0$, define $e(n) = -e(-n)$.

Proposition 11.8. *For all $n \in \mathbb{Z}$,*

(i) $e\left(n + 1_{\mathbb{Z}}\right) = e(n) + 1_{\mathbb{R}}$.

(ii) $e\left(n - 1_{\mathbb{Z}}\right) = e(n) - 1_{\mathbb{R}}$.

(iii) $e(n) = -e(-n)$.

The point of this proposition is that the statements occurring in the definition of e hold for *all* $n \in \mathbb{Z}$.

Proof. (i) The first equation holds by definition when $n \in \mathbb{Z}_{\geq 0}$. If $n = -1$ then

$$e\left(n+1\right) = e\left(0\right) = 0 = -1 + 1 = e\left(n\right) + 1 \,.$$

If $n < -1$ then $n + 1$ is negative and

$$e\left(n+1\right) = -e\left(-(n+1)\right) = -e\left(-n-1\right) \overset{(\star)}{=} -\left(e(-n) - 1\right) = -e(-n) + 1 = e(n) + 1 \,.$$

Here $(\star)$ follows from $e(-n) = e\left((-n-1)+1\right) = e(-n-1) + 1$ (note that $-n - 1 > 0$).
(ii) The second equation follows from part (i) with

$$e(n) = e\left((n-1)+1\right) = e(n-1) + 1 \,.$$

(iii) The equation $e(n) = -e(-n)$ holds by definition for negative n. We will prove it now for $n \geq 0$ by induction. The base case $n = 0$ follows because $0 = -0$ (in $\mathbb{Z}$ as well as in $\mathbb{R}$!), and so $e(0) = -e(-0)$. For the induction step, assume that $e(n) = -e(-n)$. Then, by applying part (i) and (ii),

$$e(n+1) = e(n) + 1 = -e(-n) + 1 = -\left(e(-n) - 1\right) = -e(-n-1) = -e\left(-(n+1)\right) .$$

Proposition 11.9. *The function e is injective.*

(*Hint for a possible proof:* Use Proposition 11.4(i), i.e., construct a left inverse of e.)

Proposition 11.10. *The function e preserves addition: that is, temporarily using the notation $+_{\mathbb{Z}}$ for addition in $\mathbb{Z}$ and $+_{\mathbb{R}}$ for addition in $\mathbb{R}$, $e\left(m +_{\mathbb{Z}} n\right) = e(m) +_{\mathbb{R}} e(n)$.*

[2] Here the first addition takes place in $\mathbb{Z}$, whereas the second addition happens in $\mathbb{R}$.

Proof. Fix $m \in \mathbb{Z}$. We will prove that for all $n \in \mathbb{Z}$, $e\,(m+n) = e(m) + e(n)$ in three steps. First, the result holds for $n = 0$ since $e(0) = 0$.

Next, we prove $P(n) :\ e\,(m+n) = e(m) + e(n)$ by induction on $n \in \mathbb{N}$, which will establish the proposition for *positive* n. The base case $P(1)$ follows by definition. For the induction step, assume $P(n)$. Then, by Proposition 11.8(i),

$$e\,(m + (n+1)) = e\,(m+n) + 1 \overset{(\star)}{=} e(m) + e(n) + 1 = e(m) + e(n+1)\,.$$

Here $(\star)$ follows from the induction hypothesis.

Finally, we prove $Q(n) :\ e\,(m + (-n)) = e(m) + e(-n)$ by induction on $n \in \mathbb{N}$, which will establish the proposition for *negative* n. The base case $Q(1)$ follows from Proposition 11.8(ii). For the induction step, assume $Q(n)$. Then, again by Proposition 11.8(ii),

$$\begin{aligned} e\,(m + (-(n+1))) &= e\,(m + (-n) - 1) = e\,(m + (-n)) - 1 \overset{(\star)}{=} e(m) + e(-n) - 1 \\ &= e(m) + e(-n-1) = e(m) + e\,(-(n+1))\,, \end{aligned}$$

where $(\star)$ follows from the induction hypothesis. □

Proposition 11.11. *The function e preserves multiplication: that is, $e\,(m \cdot_{\mathbb{Z}} n) = e(m) \cdot_{\mathbb{R}} e(n)$, where $\cdot_{\mathbb{Z}}$ and $\cdot_{\mathbb{R}}$ are the multiplication operations on $\mathbb{Z}$ and $\mathbb{R}$, respectively.*

Proposition 11.12. *The function e preserves order, that is, if $m, n \in \mathbb{Z}$ satisfy $m <_{\mathbb{Z}} n$ then $e(m) <_{\mathbb{R}} e(n)$.*

Proof. Given $m \in \mathbb{Z}$, we will prove $P(n) : e(n) - e(m) \in \mathbb{R}_{>0}$ by induction on $n \geq m+1$. The base case $P(m+1)$ follows from the definition of e since $e(m+1) - e(m) = e(m) + 1 - e(m) = 1 \in \mathbb{R}_{>0}$. For the induction step, assume $P(n)$. Then, again by definition,

$$e(n+1) - e(m) = e(n) + 1 - e(m) = (e(n) - e(m)) + 1 \in \mathbb{R}_{>0}\,,$$

since, by induction hypothesis, $e(n) - e(m) \in \mathbb{R}_{>0}$. □

Thus $\mathbb{R}$ has a subset $e(\mathbb{Z})$ which behaves exactly like $\mathbb{Z}$ with respect to addition, multiplication, and order. It follows that $e(\mathbb{Z})$ behaves like $\mathbb{Z}$ with respect to every property of $\mathbb{Z}$ we have discussed in this book. *From now on* we will not distinguish between $n \in \mathbb{Z}$ and $e(n) \in \mathbb{R}$. We drop notational distinctions such as $0_{\mathbb{Z}}$, $1_{\mathbb{R}}$, $+_{\mathbb{Z}}$, $\cdot_{\mathbb{R}}$, $<_{\mathbb{Z}}$, $<_{\mathbb{R}}$, and write $\mathbb{Z} \subseteq \mathbb{R}$.

In this book we have studied $\mathbb{Z}$ and $\mathbb{R}$ separately as if they were apples and oranges. Now, in this chapter, we have embedded $\mathbb{Z}$ in $\mathbb{R}$. Usually, people don't think this way. They simply think of $\mathbb{Z}$ as a subset of $\mathbb{R}$ (as you have always done). We will do that from now on. However, it is useful training to distinguish between the kinds of mathematics that are done in $\mathbb{Z}$, in $\mathbb{N}$, and in $\mathbb{R}$. Mathematicians know that this distinction is important.

Chapter 12

Rational Numbers

5 out of 4 people have trouble with fractions.
Billboard in Danby, NY

Recall that when $x, y \in \mathbb{R}$ and $y \neq 0$ then there is a real number $\frac{x}{y}$. Now that we have embedded $\mathbb{Z}$ in $\mathbb{R}$ we define rational numbers by considering the special case in which x and y are integers. The real number $z \in \mathbb{R}$ is **rational** if $z = \frac{m}{n}$ where $m, n \in \mathbb{Z}$ and $n \neq 0$. Non-rational real numbers are **irrational**. The set of all rational numbers is denoted by $\mathbb{Q}$.

Irrational numbers will be the subject of later chapters; in particular, we will see that not all real numbers are rational. For now, we develop the machinery for fractions that you are used to since grade school.

Proposition 12.1. *If $x, y, z \in \mathbb{R}$ with $y \neq 0$ and $z \neq 0$ then $\frac{xz}{yz} = \frac{x}{y}$.*

We mentioned earlier that each positive integer can be factorized in a unique way as a product of primes. If $\frac{m}{n}$ is rational then, multiplying above and below by -1 if necessary, we may assume $n > 0$. Cancelling prime factors common to m and n, we can always represent a rational as $\frac{m}{n}$ where $n > 0$ and m and n do not have any common factors. This representation is *in lowest terms.*

Proposition 12.2. *Suppose m_1, n_1, m_2, n_2 are integers, and m_1 and n_1 do not have any common factors. Then $\frac{m_1}{n_1} = \frac{m_2}{n_2}$ implies that m_1 divides m_2 and n_1 divides n_2.*

Proposition 12.3. *For all $m_1, n_1, m_2, n_2 \in \mathbb{Z}$, where $n_1, n_2 \neq 0$, $\frac{m_1}{n_1} + \frac{m_2}{n_2} = \frac{m_1 n_2 + n_1 m_2}{n_1 n_2}$.*

Proposition 12.4. *In the following, all m's and n's are integers and all n's are non-zero.*

(i) *For all $\frac{m_1}{n_1}, \frac{m_2}{n_2}, \frac{m_3}{n_3} \in \mathbb{Q}$:*

(a) $\frac{m_1}{n_1} + \frac{m_2}{n_2} = \frac{m_2}{n_2} + \frac{m_1}{n_1}$.

(b) $\left(\frac{m_1}{n_1} + \frac{m_2}{n_2}\right) + \frac{m_3}{n_3} = \frac{m_1}{n_1} + \left(\frac{m_2}{n_2} + \frac{m_3}{n_3}\right)$.

(c) $\frac{m_1}{n_1}\left(\frac{m_2}{n_2} + \frac{m_3}{n_3}\right) = \frac{m_1}{n_1} \cdot \frac{m_2}{n_2} + \frac{m_1}{n_1} \cdot \frac{m_3}{n_3}$.

(d) $\frac{m_1}{n_1} \cdot \frac{m_2}{n_2} = \frac{m_2}{n_2} \cdot \frac{m_1}{n_1}$.

(e) $\left(\frac{m_1}{n_1}\,\frac{m_2}{n_2}\right)\frac{m_3}{n_3} = \frac{m_1}{n_1}\left(\frac{m_2}{n_2}\,\frac{m_3}{n_3}\right)$.

(ii) *For all* $\frac{m}{n} \in \mathbb{Q}$, $\frac{m}{n} + 0 = \frac{m}{n}$.

(iii) *For all* $\frac{m}{n} \in \mathbb{Q}$, $\frac{m}{n} \cdot 1 = \frac{m}{n}$.

(iv) *For all* $m, n \in \mathbb{Z}$, $n \neq 0$, $\frac{m}{n} + \frac{(-m)}{n} = 0$.

(v) *If* $\frac{m_1}{n_1} \neq 0$ *and if* $\frac{m_1}{n_1} \cdot \frac{m_2}{n_2} = \frac{m_1}{n_1} \cdot \frac{m_3}{n_3}$ *then* $\frac{m_2}{n_2} = \frac{m_3}{n_3}$.

Proposition 12.5. *The rational number* $\frac{m}{n} \in \mathbb{Q}$ *is positive (i.e.,* $\frac{m}{n} \in \mathbb{R}_{>0}$*) if and only if either* $m > 0$ *and* $n > 0$, *or* $m < 0$ *and* $n < 0$.

Proposition 12.6.

(i) *The sum of two positive rationals is a positive rational.*

(ii) *The product of two positive rationals is a positive rational.*

(iii) *For every* $\frac{m}{n} \in \mathbb{Q}$ *such that* $\frac{m}{n} \neq 0$, *either* $\frac{m}{n}$ *is positive or* $\frac{-m}{n}$ *is positive, and not both.*

You learned all this in grade school under the title *fractions*. . . As promised, we are systematically organizing some of the mathematics you previously knew.

If you compare $\mathbb{Q}$ as described here with $\mathbb{R}$ as described in Chapter 10, you will see that they have much in common. Indeed, $\mathbb{Q}$ satisfies Axioms 10.1–10.6 of Chapter 10, so anything proved for $\mathbb{R}$ using only those axioms holds for $\mathbb{Q}$ too. However, as we will see in Project 14.5, Axiom 10.7 does not hold in $\mathbb{Q}$.

We finish this chapter with a project that connects what we have been doing here with linear algebra:

Project 12.7. *In linear algebra you learned about vector spaces and linear maps.* $\mathbb{R}$ *itself is a one-dimensional vector space, and a linear map from* $\mathbb{R}$ *to* $\mathbb{R}$ *is a function of the form* $f(x) = cx$ *for some fixed* $c \in \mathbb{R}$. *Prove that if* $f : \mathbb{R} \to \mathbb{R}$ *is continuous and satisfies* $f(x+y) = f(x) + f(y)$ *for all* $x, y \in \mathbb{R}$, *then* f *is a linear map. Here is one possible way to proceed: Assuming that* $f : \mathbb{R} \to \mathbb{R}$ *satisfies* $f(x+y) = f(x) + f(y)$, *prove that*

(1) *For all* $n \in \mathbb{Z}$ *and* $x \in \mathbb{R}$, $f(nx) = nf(x)$.[1]

(2) *For all* $n \in \mathbb{Z}$, $m \in \mathbb{N}$, *and* $x \in \mathbb{R}$, $f\left(\frac{n}{m}x\right) = \frac{n}{m}f(x)$.

(3) *There exists* $c \in \mathbb{R}$ *such that for all* $x \in \mathbb{Q}$, $f(x) = cx$.

[1] *Hint:* prove this first for $n \in \mathbb{N}$.

Chapter 13

Completeness of $\mathbb{R}$

It is a pain to think about convergence but sometimes you really have to.
Sinai Robins

This chapter is devoted to consequences of Axiom 10.7: every non-empty subset of $\mathbb{R}$ that is bounded above has a least upper bound. One such consequence is the following.

13.1 The Integers are Unbounded

Theorem 13.1. *$\mathbb{N}$, considered as a subset of $\mathbb{R}$, is not bounded above.*

Proof. We will prove this by contradiction. Suppose $\mathbb{N}$ were bounded above. Then, by Axiom 10.7, $\mathbb{N}$ would have a least upper bound: call it u. The interval $\left(u - \frac{1}{2}, u\right] = \left\{x \in \mathbb{R} : \ u - \frac{1}{2} < x \leq u\right\}$ must contain some $n \in \mathbb{N}$ since otherwise $u - \frac{1}{2}$ would be an upper bound for $\mathbb{N}$ (contradicting the fact that u is the least upper bound). But if $u - \frac{1}{2} < n \leq u$ then $u + \frac{1}{2} < n + 1$, so $u < n + 1$ (note that $\frac{1}{2} > 0$). By Axiom 3.1, $n + 1 \in \mathbb{N}$, so u is not an upper bound for $\mathbb{N}$, which is a contradiction. □

Let's pause for a moment to think about Theorem 13.1. In Proposition 3.12 we saw that there is no largest natural number. If the real numbers are pictured by a horizontal line (draw one) and if the first few natural numbers are marked on that line (do it: $1, 2, 3, 4, \dots$), we have to rule out the possibility that the dots representing the natural numbers bunch up on one another in some finite part of the line. In mathematics we use the term *converge* rather than "bunch up." Whatever the word, Theorem 13.1 says that, in fact, there is no real number that is bigger than all the natural numbers.

Since $\mathbb{N} \subseteq \mathbb{Z}$ we immediately obtain the following.

Corollary 13.2. *$\mathbb{Z}$ is not bounded above.*

We proved earlier that 1 is the least element of $\mathbb{N}$. Note that $\mathbb{Z}$ does not have a least element:

Proposition 13.3. *$\mathbb{Z}$ is not bounded below.*

Another consequence of the unboundedness of $\mathbb{N}$ is the following innocent-looking but useful proposition.

Proposition 13.4. *For any $\epsilon > 0$, there exists $n \in \mathbb{N}$ such that $\frac{1}{n} < \epsilon$.*

Proof. By Proposition 13.1, there exists $n \in \mathbb{N}$ such that $n > \frac{1}{\epsilon}$ or, equivalently by Proposition 10.3(ii), $\frac{1}{n} < \epsilon$. □

In Theorem 10.8 we showed that between any two real numbers there is a third. We will now prove that between any two real numbers we can find a *rational* number.

Theorem 13.5. *Let $x, y \in \mathbb{R}$ with $x < y$. Then there is a rational number r such that $x < r < y$.*

Proof. Let $z = \frac{x+y}{2}$ and $\epsilon = \frac{y-x}{2}$; note that $\epsilon > 0$. We need need to prove that there exists $r \in \mathbb{Q}$ such that $z - \epsilon < r < z + \epsilon$.

By Proposition 13.4, there exists $m \in \mathbb{N}$ such that $\frac{1}{m} < \epsilon$. Furthermore, because $\mathbb{Z}$ is unbounded, there exists $n \in \mathbb{Z}$ such that $n > mz$. Choose this n minimal, that is, $n - 1 \le mz < n$. With Proposition 10.3, we can rewrite these inequalities as

$$\frac{n}{m} - \frac{1}{m} \le z < \frac{n}{m} \ .$$

The left-hand inequality implies $\frac{n}{m} \le z + \frac{1}{m} < z + \epsilon$, whereas the right-hand inequality implies $\frac{n}{m} > z > z - \epsilon$, that is,

$$z - \epsilon < \frac{n}{m} < z + \epsilon \ .$$ □

In Chapter 16, we will complement Theorem 13.5 by showing that there is also an *irrational* number between x and y.

Corollary 13.6. *There is no smallest positive rational number.*

13.2 Absolute Value and Distance

The **absolute value** of $x \in \mathbb{R}$, denoted $|x|$, is defined to be x if $x \ge 0$ and to be $-x$ if $x < 0$. This definition implies that $|x| \ge 0$ always.

Proposition 13.7. *For $x, y \in \mathbb{R}_{\ge 0}$, $x < y$ if and only if $x^2 < y^2$.*

Proposition 13.8. *For any $x \in \mathbb{R}$, $|x^2| = |x|^2 = x^2$.*

Proposition 13.9. *For any $x, y \in \mathbb{R}$:*

(i) $|x| = 0$ *if and only if* $x = 0$.

(ii) $|xy| = |x|\,|y|$.

(iii) $-|x| \le x \le |x|$.

(iv) $|x + y| \le |x| + |y|$.

Proof. (i) If $x = 0$ then, by definition, $|x| = x = 0$. Conversely, if $x \neq 0$ then either $x > 0$, in which case $|x| = x > 0$, or $x < 0$, in which case $|x| = -x > 0$ by Proposition 10.3(iv) (with $a = -1$). In both cases we conclude $|x| > 0$, so in particular, $|x| \neq 0$.

(ii) There are four cases:
If one of x, y is zero, both sides of $|xy| = |x||y|$ are zero.
If $x > 0$ and $y > 0$ then by Axiom 10.6 $xy > 0$, and $|xy| = xy = |x||y|$.
If $x > 0$ and $y < 0$ then by Proposition 10.3(iv) $xy < 0$, and $|xy| = -xy = x(-y) = |x||y|$. Since the equation $|xy| = |x||y|$ is symmetric in x and y, this also covers the case $x < 0, y > 0$.
Finally, if $x < 0$ and $y < 0$ then by Proposition 10.3(iv) $xy > 0$, and $|xy| = xy = (-x)(-y) = |x||y|$.

(iii) There are two cases:
If $x \geq 0$ then $|x| = x$ and $-x \leq 0$; hence by Proposition 10.3(i) $-|x| = -x \leq x = |x|$.
If $x < 0$ then $|x| = -x > 0$; hence, again by Proposition 10.3(i), $-|x| = x < -x = |x|$.
In both cases we obtain $-|x| \leq x \leq |x|$.

(iv) By Proposition 13.8,

$$|x + y|^2 = (x + y)^2 = x^2 + 2xy + y^2 = |x|^2 + 2xy + |y|^2.$$

By part (iii), $2xy \leq |2xy| = 2|x||y|$; here the last equality follows from part (ii). Hence

$$|x + y|^2 = |x|^2 + 2xy + |y|^2 \leq |x|^2 + 2|x||y| + |y|^2 = (|x| + |y|)^2\,,$$

from which we conclude $|x + y| \leq |x| + |y|$ with Proposition 13.7. □

Proposition 13.10. *For any $x, y \in \mathbb{R}$:*

(i) $|x - y| = 0$ *if and only if* $x = y$.

(ii) $|x - y| = |y - x|$.

(iii) $|x - z| \leq |x - y| + |y - z|$.

(iv) $|x - y| \geq ||x| - |y||$.

One of the beautiful things about mathematics is that in involves both algebra and geometry; in fact, there are times when one wants to express a geometrical statement using the language of algebra, and there are other times when one wants to express an algebraic statement in the language of geometry. Absolute value provides a good example. Mathematicians think of $|x-y|$ as the **distance** from the point x to the point y in the line $\mathbb{R}$. In fact, let us make this the definition of the word distance. This agrees with the everyday definition of that word: distance is never negative (try going for a walk -2 miles in length) and the distance from a point to a different point is never 0. In this language, we can rewrite Proposition 13.10:

(i) The distance from x to y is 0 if and only if x equals y.

(ii) The distance from x to y equals the distance from y to x.

(iii) The distance from x to z is at most the sum of the distances from x to y and from y to z.

Try putting (iv) in words: sometimes algebraic language is easier.

Proposition 13.11. *Let $x, y \in \mathbb{R}$. Then $x = y$ if and only if for any $\epsilon > 0$ we have $|x - y| < \epsilon$.*

13.3 Limits

Let $(x_k)_{k=1}^{\infty}$ be a sequence in $\mathbb{R}$ and let $L \in \mathbb{R}$. We say (x_k) **converges to** L if

$$\text{for all } \epsilon > 0 \text{ there exists } N \in \mathbb{N} \text{ such that for all } n \geq N,\ |x_n - L| < \epsilon\,.$$

When (x_k) converges to L, we call L the **limit** of the sequence (x_k), and we write $\lim\limits_{k\to\infty} x_k = L$. Here are two examples of convergent sequences.

Proposition 13.12.

(i) $\lim\limits_{k\to\infty} \dfrac{k-1}{k} = 1\,.$

(ii) $\lim\limits_{k\to\infty} \left(\dfrac{1}{4^k}\right) = 0\,.$

Proof. (i) Suppose $\epsilon > 0$ is given. Then choose an integer $N > \frac{1}{\epsilon}$ (we can do this because of Proposition 13.4), and we have for $n \geq N$

$$\left|\frac{n-1}{n} - 1\right| = \frac{1}{n} \leq \frac{1}{N} < \epsilon\,.$$

(ii) By the Binomial Theorem 5.13, we have for $n \geq 1$,

$$4^n = (1+3)^n = \sum_{r=0}^{n} \binom{n}{r} 3^r = 1 + 3n + 9\binom{n}{2} + \cdots + 3^n > 3n\,.$$

We will use this inequality in a moment. Suppose $\epsilon > 0$ is given. Then choose an integer $N > \frac{1}{3\epsilon}$, and we have for $n \geq N$

$$\left|\frac{1}{4^n} - 0\right| = \frac{1}{4^n} < \frac{1}{3n} \leq \frac{1}{3N} < \epsilon\,.$$ □

Sometimes we are only interested in *the fact that* a sequence converges, rather than what it converges to. So to say (x_k) **converges** means: there exists $L \in \mathbb{R}$ such that (x_k) converges to L. If no such L exists, we say that the sequence **diverges**. So the statement "(x_k) diverges" is the negation of the statement "there exists $L \in \mathbb{R}$ such that (x_k) converges to L"; you should write down this negation in exact terms.

We first prove that the limit of a sequence, if it exists, is unique:

Proposition 13.13. *If (x_k) converges to L and to L' then $L = L'$.*

Project 13.14. *Find the limits of your favorite sequences from calculus. Find sequences that diverge. Prove your assertions.*

Project 13.15. *For which $\alpha \in \mathbb{R}$ does the sequence $\left(\alpha^k\right)_{k=0}^{\infty}$ converge? Prove your assertions.*

Proposition 13.16. *If a sequence converges, then it is bounded.*[1]

[1] People say it this way, but there is a distinction to be made here. A sequence is a function. What we are really saying here is that the image of that function is a bounded subset of $\mathbb{R}$.

Proposition 13.17. *Suppose* $A = \lim_{k\to\infty} a_k$ *and* $B = \lim_{k\to\infty} b_k$*.*

(i) *For any* $c \in \mathbb{R}$, $\displaystyle\lim_{k\to\infty} (c\, a_k) = c\, A$ *.*

(ii) $\displaystyle\lim_{k\to\infty} (a_k + b_k) = A + B$ *.*

(iii) $\displaystyle\lim_{k\to\infty} (a_k b_k) = AB$ *.*

(iv) *If* $A \neq 0$*, then* $\displaystyle\lim_{k\to\infty} \frac{1}{a_k} = \frac{1}{A}$ *.*

Project 13.18. *In calculus you learned about sequences* (x_k) *that "blow up" in the sense that* $\lim_{k\to\infty} x_k = \infty$*; an example is the sequence given by* $x_k = k^2$*. We think of this sequence as converging to infinity; in this sense the limit* $\lim_{k\to\infty} x_k$ *exists (as opposed to, e.g.,* $\lim_{k\to\infty}(-1)^k k^2$*). Come up with a solid mathematical definition for* $\lim_{k\to\infty} x_k = \infty$ *and prove that* $\lim_{k\to\infty} k^2 = \infty$*.*

13.4 Square Roots

As another consequence of Axiom 10.7 we consider square roots.

Theorem 13.19. *There is a positive real number* u *such that* $u^2 = 2$*.*

We will denote this number u by $\sqrt{2}$. Note that $-\sqrt{2}$ is also a solution to the equation $x^2 = 2$, but we reserve the phrase **square root of 2** for the positive solution of this equation.

Proof. Let u be the least upper bound of $A = \{x \in \mathbb{R} : x^2 < 2\}$. Since $1 \in A$, u will certainly be positive. We claim that $u^2 = 2$. To prove this, we will show that both (a) $u^2 > 2$ and (b) $u^2 < 2$ lead to contradictions.

(a) Suppose $u^2 > 2$, and let $\delta = u^2 - 2 > 0$. Now let $h = \frac{\delta}{4u}$; note that $\delta > 0$. Furthermore,

$$u^2 - (u-h)^2 = u^2 - u^2 - h^2 + 2uh = h(2u - h) < h \cdot 2u < \delta\,.$$

Thus

$$2 < (u-h)^2 < u^2,$$

which implies both that $u - h$ is an upper bound for A (because for any positive $x \in A$, we have $x^2 < 2 < (u-h)^2$, and so Proposition 13.7 implies) and that $u - h < u$ (by Proposition 13.7), contradicting the fact that $u = \sup A$.

(b) Suppose $u^2 < 2$; then $\delta = 2 - u^2 > 0$. For $h > 0$

$$(u+h)^2 - u^2 = u^2 + h^2 + 2uh - u^2 = h(2u+h) < h \cdot 3u < \delta\,,$$

if h is chosen so that $h < u$ and $h < \frac{\delta}{3u}$. So taking $h = \min\left(\frac{\delta}{4u}, \frac{u}{2}\right)$ we conclude that

$$u^2 < (u+h)^2 < 2\,.$$

This implies $u + h \in A$, so u is not an upper bound for A, a contradiction. □

As you can imagine, there is nothing special about the number 2 in the above theorem.

Theorem 13.20. *Given any $r \in \mathbb{R}_{>0}$, there is a positive real number u such that $u^2 = r$.*

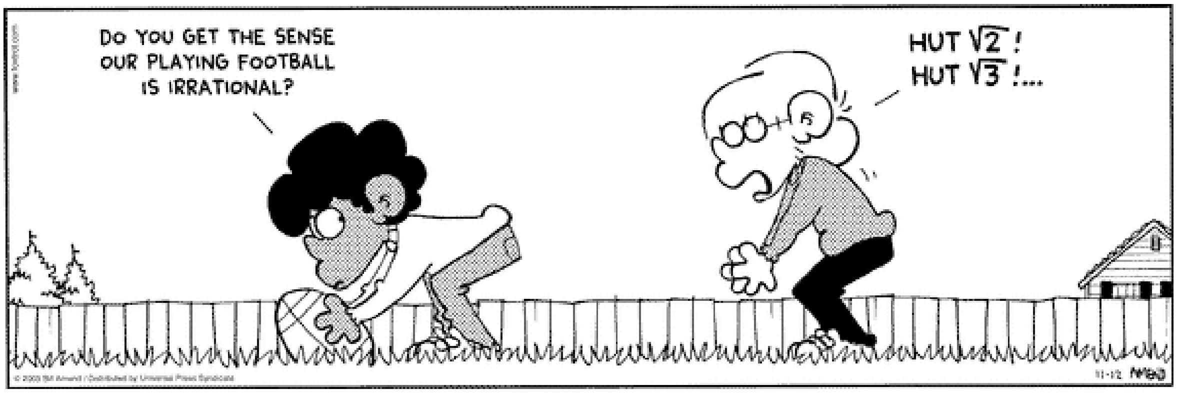
We denote this number u by $\sqrt{r}$ and complement our definition by setting $\sqrt{0} = 0$.

Proposition 13.21. *If $r < 0$ there is no $u \in \mathbb{R}$ such that $u^2 = r$.*

This proposition is the reason why negative real numbers do not have real square roots. In Chapter 17 we will be studying a larger set of numbers than $\mathbb{R}$, called the complex numbers. Negative real numbers do have "square roots" in the complex numbers.

Chapter 14

Irrational Numbers


DO YOU GET THE SENSE OUR PLAYING FOOTBALL IS IRRATIONAL?
HUT $\sqrt{2}$! HUT $\sqrt{3}$!...
11-12

14.1 Irrational Numbers Exist

Recall that a number $x \in \mathbb{R}$ is rational if $x = \frac{m}{n}$, where m and n are integers. We have just defined the real number $\sqrt{2}$. Our first goal in this chapter is to show that $\sqrt{2}$ is not rational. This will prove that $\mathbb{Q}$ is a *proper* subset of $\mathbb{R}$, i.e., $\mathbb{Q} \neq \mathbb{R}$.

Proposition 14.1. *The real number* $\sqrt{2}$ *is irrational.*

Proof. We will give a proof by contradiction. Assume that $\sqrt{2} = \frac{m}{n}$ for some $m, n \in \mathbb{Z}$. Because of Proposition 12.1, we can eliminate any common factors of m and n. Now $2 = \frac{m^2}{n^2}$ implies that we can write

$$\frac{m}{n} = \frac{2n}{m},$$

and since $\frac{m}{n}$ is written in lowest terms, Proposition 12.2 implies that n divides m. But then $\frac{m}{n} = \sqrt{2}$ is an integer, which is a contradiction (e.g., we saw in the proof of Theorem 13.19 that $1 < \sqrt{2} < 2$). □

You're invited to modify this proof to show the following more general result. An integer n is a **perfect square** if $n = m^2$ for some $m \in \mathbb{Z}$.

Proposition 14.2. *If $r \in \mathbb{N}$ is not a perfect square, then $\sqrt{r}$ is irrational.*

Project 14.3. *Here is the outline of an alternative proof for Proposition 14.1:* Again, we assume (hoping to obtain a contradiction) that $\sqrt{2} = \frac{m}{n}$ for some $m, n \in \mathbb{Z}$, and since we may write this fraction in lowest terms, m and n are not both even. Now $m^2 = 2n^2$, so m^2 is even, and Proposition 6.7 implies that m is even. Let's write $m = 2j$ for some integer j; then some quick algebra give that $n^2 = 2j^2$, which means that n^2 is even. We can use Proposition 6.7 once more to deduce that n is even. But then both m and n are even, contrary to the first sentence of this proof. *Compare the two proofs of Proposition 6.7. How do they differ? Are they really different? What are advantages/disadvantages of each?*

We can also define higher roots: Namely, for an integer $n \geq 2$, the n**th root of** $r \in \mathbb{R}_{>0}$ is the positive real number $\sqrt[n]{r}$ that satisfies $\left(\sqrt[n]{r}\right)^n = r$. One can show that these numbers exist, and the following proposition can be proved in an analogous way to our proof of Proposition 14.1.

Proposition 14.4. *The real number $\sqrt[n]{2}$ is irrational.*

Here is our promised result that Axiom 10.7 does not hold in $\mathbb{Q}$:

Project 14.5. *Name a non-empty subset of $\mathbb{Q}$ that is bounded above but has no least upper bound in $\mathbb{Q}$. Justify your claim.*

14.2 Quadratic Equations

Recall that a **quadratic equation** is an equation of the form $ax^2 + bx + c = 0$ where, in our case, $a, b, c \in \mathbb{R}$.[1]

Proposition 14.6. *The equation $x^2 + 1 = 0$ does not have a solution in $\mathbb{R}$.*

Proposition 14.7. *Suppose $a, b, c \in \mathbb{R}$, where a and b are not both zero. Then the equation $ax^2 + bx + c = 0$ has a solution in $\mathbb{R}$ if and only if $b^2 - 4ac \geq 0$.*

In fact, as you well know, $\frac{-b+\sqrt{b^2-4ac}}{2a}$ is a solution (assuming that $a \neq 0$); and $\frac{-b-\sqrt{b^2-4ac}}{2a}$ is also a solution. The number $b^2 - 4ac$ is called the **discriminant** of the equation $ax^2 + bx + c = 0$.

Project 14.8. *Prove that a quadratic equation has at most two solutions. (Why do we say "at most"?)*

Proposition 14.9. *Let $b, c, d, e \in \mathbb{R}$. If $x^2 - bx - c = 0$ has the two solutions s and t, and if we define $a_n = d\, s^n + e\, t^n$, then $a_n = b\, a_{n-1} + c\, a_{n-2}$.*

Setting $b = c = 1$ we get the recurrence relation $a_n = a_{n-1} + a_{.2}$, which (when we set $a_1 = a_2 = 1$) is the defining recurrence relation for the Fibonacci numbers. Thus the quadratic equation $x^2 - x - 1 = 0$ has some connection with the Fibonacci numbers.

Project 14.10. *Prove that the n^{th} Fibonacci number is given by*

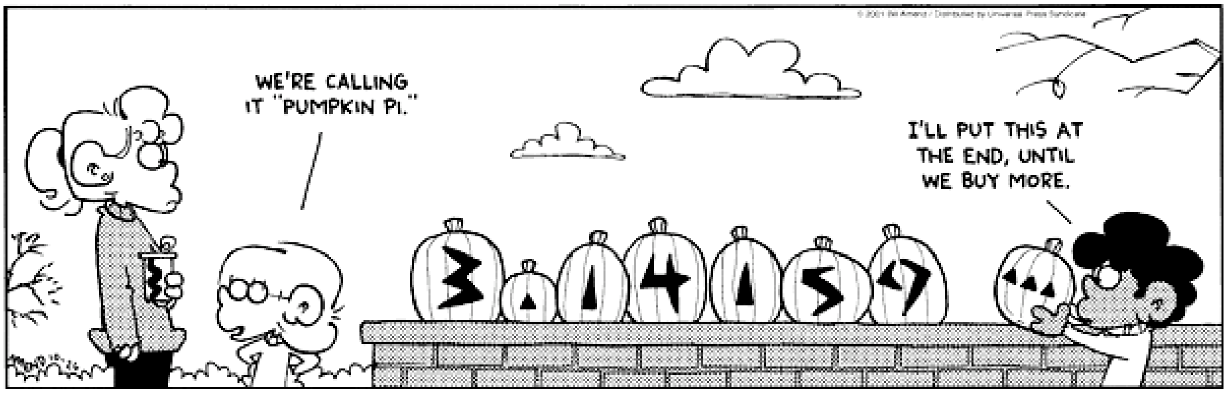
$$f(n) = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right).$$

Compare your proof with the one we have given for Proposition 7.5.

[1] Strictly speaking, we have to demand that $a \neq 0$ for the equation $ax^2 + bx + c = 0$ to be called quadratic. An equation of the form $bx + c = 0$ is called a **linear equation**.

Chapter 15

Decimal Expansions


© 2001 Bill Amend / Distributed by Universal Press Syndicate
WE'RE CALLING IT "PUMPKIN PI."
3.14159
I'LL PUT THIS AT THE END, UNTIL WE BUY MORE.

15.1 Series

We touched upon series briefly in Chapter 5, so the first part of the following definition is just a reminder.

A **series** is a sequence $(a_k)_{k=0}^{\infty}$ whose members are sums of the form $a_k = \sum_{j=0}^{k} b_j$. We call $(b_k)_{k=0}^{\infty}$ the **sequence of terms** of the series. The $a_k = \sum_{j=0}^{k} b_j$ are called the **partial sums** of the series. If $a_k = \sum_{j=0}^{k} b_j$ converges people usually write $\sum_{j=0}^{\infty} b_j$ instead of $\lim_{k\to\infty} \sum_{j=0}^{k} b_j$.

Given two real numbers a and x the series whoes jth term is ax^j is called the **geometric series** whose 0th term is a and whose **common ratio** is x. Consider the case $a = 1$ and $x = \frac{1}{2}$. We saw in Proposition 5.10 that the partial sum that ends with the mth term is

$$\sum_{j=0}^{k} \frac{1}{2^j} = \frac{1 - \left(\frac{1}{2}\right)^{k+1}}{1 - \frac{1}{2}} = 2\left(1 - \left(\frac{1}{2}\right)^{k+1}\right).$$

Since we proved that $\lim_{k\to\infty} \left(\frac{1}{2}\right)^{k+1} = 0$, we see that in the limit $\sum_{j=0}^{\infty} \frac{1}{2^j} = 2$. The following proposition is a generalization.

Proposition 15.1. *The geometric series converges for $|x| < 1$; namely,* $\displaystyle\sum_{j=0}^{\infty} x^j = \frac{1}{1-x}$.

Proof. By Proposition 5.10, we have for $x \neq 1$ and $k \in \mathbb{Z}_{\geq 0}$,

$$\sum_{j=0}^{k} x^j = \frac{1 - x^{k+1}}{1-x},$$

so in light of Proposition 13.17, we only have to show that $\lim_{k\to\infty} x^{k+1} = \lim_{k\to\infty} x^k = 0$ for $|x| < 1$. This limit certainly holds when $x = 0$, so we may assume that $x \neq 0$. To prove $\lim_{k\to\infty} x^k = 0$, suppose $\epsilon > 0$. Choose an integer $N > \log_{|x|} \epsilon$; then for $n \geq N$,

$$|x^n - 0| = |x|^n \leq |x|^N < \epsilon.$$

(Note that the first inequality only holds because $|x| < 1$.) □

Our next goal is to prove the following *comparison test*:

Proposition 15.2. *Suppose $0 \leq a_k \leq b_k$ for all $k \geq 0$. If* $\displaystyle\sum_{j=0}^{\infty} b_j$ *converges then* $\displaystyle\sum_{j=0}^{\infty} a_j$ *converges.*

Before proving this, we discuss a famous principle for sequences: namely, *a monotonic bounded sequence always converges.* We explain:

The sequence $(a_k)_{k=0}^{\infty}$ is **increasing** if $a_{k+1} \geq a_k$ for all $k \geq 0$, and **decreasing** if $a_{k+1} \leq a_k$ for all $k \geq 0$. A sequence is **monotonic** if it is either increasing or decreasing. The sequence $(a_k)_{k=0}^{\infty}$ is **bounded** if there exist $l, u \in \mathbb{R}$ such that $l \leq a_k \leq u$ for all $k \geq 0$.

Theorem 15.3. *Every increasing bounded sequence converges.*

Proof. Suppose the sequence $(a_k)_{k=0}^{\infty}$ is increasing and bounded. Because the set

$$A = \{a_k : k \geq 0\} \subseteq \mathbb{R}$$

is bounded, it has a least upper bound s by Axiom 10.7. We claim that $s = \lim_{k\to\infty} a_k$. To prove this, suppose $\epsilon > 0$ is given. Then $s - \epsilon$ is *not* an upper bound for A (since $s = \sup A$), and so there exists N such that $a_N > s - \epsilon$. But $(a_k)_{k=0}^{\infty}$ is increasing, so $a_n \geq a_N$ for all $n \geq N$. In summary, we have for $n \geq N$

$$s - \epsilon < a_N \leq a_n \leq s < s + \epsilon,$$

and so $|a_n - s| < \epsilon$. We have proved that, given any $\epsilon > 0$, there exists N such that for $n \geq N$, $|a_n - s| < \epsilon$, as claimed. □

Project 15.4. *Prove the analogous statement for* decreasing *bounded sequences. In summary, we then know that every monotonic bounded sequence converges.*

It is important to note, even to remember for the rest of your life, that Theorem 15.3 is an existence theorem: we proved that the sequence converges without finding its limit. We can do this because Axiom 10.7 asserts the existence of real numbers without computing them.

Proposition 15.5. *Suppose the sequence $(a_k)_{k=0}^{\infty}$ is increasing and $L = \lim_{k\to\infty} a_k$. Then $a_k \leq L$ for all $k \geq 0$.*

With this preparation, we are ready to prove the comparison test.

Proof of Proposition 15.2. Suppose $0 \leq a_k \leq b_k$ for all $k \geq 0$, and that $L = \sum_{j=0}^{\infty} b_j$. Since $b_k \geq 0$, the sequence $(B_k)_{k=0}^{\infty}$ of partial sums $B_k = \sum_{j=0}^{k} b_j$ is increasing, and by Proposition 15.5, $B_k \leq L$ for all $k \geq 0$. Let $A_k = \sum_{j=0}^{k} a_j$. Then for all $k \geq 0$

$$a_0 = A_0 \leq A_k = \sum_{j=0}^{k} a_j \leq \sum_{j=0}^{k} b_j = B_k \leq L\,,$$

so the sequence $(A_k)_{k=0}^{\infty}$ of partial sums is bounded. Since $a_k \geq 0$, $(A_k)_{k=0}^{\infty}$ is also increasing, and so by Theorem 15.3, $\sum_{j=0}^{\infty} a_j$ converges. □

15.2 Decimals

A **nonnegative decimal** is a sequence $(m, d_1, d_2, d_3, \dots)$ where $m \geq 0$ is an integer and each d_n is a digit, that is, an integer between 0 and 9. By tradition (as you well know) the notation used for a nonnegative decimal is $m.d_1d_2d_3\cdots$. This nonnegative decimal **represents** the (non-negative) real number $x = m + \sum_{j=1}^{\infty} d_j 10^{-j}$, and we call $(m, d_1, d_2, d_3, \dots)$ a **nonnegative decimal expansion** of x. A decimal expansion of a negative real number x is found by placing a minus sign in front of a decimal expansion of the positive number $-x$.

For all of this to make sense, we must be sure that $\sum_{j=1}^{\infty} d_j 10^{-j}$ converges:

Proposition 15.6. *Let $(d_k)_{k=1}^{\infty}$ be a sequence of digits. Then $\displaystyle\sum_{j=1}^{\infty} d_j 10^{-j}$ converges.*

Proposition 15.7. *Let $(d_k)_{k=1}^{\infty}$ be a sequence of digits and $n \in \mathbb{N}$. Then $\displaystyle\sum_{j=n}^{\infty} d_j 10^{-j} \leq \frac{1}{10^{n-1}}$.*

Proposition 15.6 implies that every decimal expansion represents a real number. Now we will prove the converse:

Theorem 15.8. *Every real number has a decimal representative.*

Proof. We will prove this theorem for *nonnegative* real numbers. The general case follows easily, as we can simply get an expansion of $-x$ if $x < 0$.

Let $x \geq 0$ be given. We will recursively construct a decimal representation $m.d_1d_2\cdots$ of x. Let m be the least integer for which $x < m + 1$. Then $m \leq x$ (otherwise m was not chosen minimally). Next, let d_1 be the least element of $\left\{n \in \mathbb{Z}_{\geq 0} : \ x < m + \frac{n+1}{10}\right\}$. Then $0 \leq d_1 \leq 9$ (otherwise m was not chosen minimally) and $m + \frac{d_1}{10} \leq x$ (otherwise d_1 was not chosen minimally), in summary:

$$m + \frac{d_1}{10} \leq x < m + \frac{d_1 + 1}{10}\,.$$

Now we construct the remaining digits recursively. Assuming that $d_1, d_2, \ldots, d_k$ have been constructed so that

$$m + \sum_{j=1}^{k} d_j 10^{-j} \leq x < m + \sum_{j=1}^{k-1} d_j 10^{-j} + \frac{d_k + 1}{10^k},$$

we let d_{k+1} to be the least element of $\left\{ n \in \mathbb{Z}_{\geq 0} : \ x < m + \sum_{j=1}^{k} d_j 10^{-j} + \frac{n+1}{10^{k+1}} \right\}$. Then

$$m + \sum_{j=1}^{k+1} d_j 10^{-j} \leq x < m + \sum_{j=1}^{k} d_j 10^{-j} + \frac{d_{k+1} + 1}{10^{k+1}},$$

This recursive construction defines the digits and ensures that for $k \in \mathbb{N}$

$$0 \leq x - \left(m + \sum_{j=1}^{k} d_j 10^{-j} \right) < \frac{1}{10^k},$$

which will now allow us to prove that $x = m + \sum_{j=1}^{\infty} d_j 10^{-j}$. Namely, for a given $\epsilon > 0$, we choose $N > \log_{10} \left(\frac{1}{\epsilon} \right)$; then for $n \geq N$,

$$\left| x - \left(m + \sum_{j=1}^{n} d_j 10^{-j} \right) \right| = x - \left(m + \sum_{j=1}^{n} d_j 10^{-j} \right) < \frac{1}{10^n} \leq \frac{1}{10^N} < \epsilon.$$

This means that the partial sums $m + \sum_{j=1}^{k} d_j 10^{-j}$ converge to x as $k \to \infty$, which is what we set out to prove. □

Next, we consider the uniqueness question: can a real number have more than one decimal expansion, and if so how many? We start with the special case of the real number 1.

Proposition 15.9. *Let $m.d_1 d_2 d_3 \cdots$ represent $1 \in \mathbb{R}$. Then either $m = 1$ and every $d_k = 0$ or $m = 0$ and every $d_k = 9$. In other words 1 can be represented by $1.00000 \cdots$ and $0.999999 \cdots$, and by no other decimal.*

Proof. The expansion $1.00000 \cdots = 1 + \sum_{j=1}^{\infty} 0 \cdot 10^{-j}$ certainly represents 1, as does

$$0.999999 \cdots = \sum_{j=1}^{\infty} 9 \cdot 10^{-j} = 9 \sum_{j=1}^{\infty} \left(\frac{1}{10} \right)^j = 9 \left(\frac{1}{1 - \frac{1}{10}} - 1 \right)$$

(by Proposition 15.1). We must show there are no other decimal representatives of 1.

So suppose $m + \sum_{j=1}^{\infty} \frac{d_j}{10^j}$ is a decimal expansion of 1. If $m \geq 2$ or $m \leq -1$ then this expansion differs from 1 by at least 1, so we just have to consider the cases $m = 0$ and $m = 1$.

1. case: $m = 0$. Let N be the lowest index among $n \geq 1$ for which $d_n < 9$. (If all $d_n = 9$ we get the expansion $0.999999 \cdots$.) Then

$$\sum_{j=1}^{N} \frac{d_j}{10^j} = \sum_{j=1}^{N-1} \frac{9}{10^j} + \frac{d_N}{10^N} = 9 \left(\frac{1 - 10^{-N}}{1 - \frac{1}{10}} - 1 \right) + \frac{d_N}{10^N},$$

by Proposition 5.10. The expression on the right-hand side simplifies to

$$\sum_{j=1}^{N} \frac{d_j}{10^j} = 1 - \frac{1}{10^{N-1}} + \frac{d_N}{10^N} = 1 - \frac{10 - d_N}{10^N} \,.$$

But then

$$1 - \sum_{j=1}^{\infty} \frac{d_j}{10^j} = \frac{10 - d_N}{10^N} - \sum_{n=N+1}^{\infty} \frac{d_j}{10^j} \,.$$

Since $d_N < 9$, the first term on the right-hand side is at least $\frac{2}{10^N}$. The second term is bounded above by $\frac{1}{10^N}$, by Proposition 15.7. Hence

$$1 - \sum_{j=1}^{\infty} \frac{d_j}{10^j} \geq \frac{1}{10^N} \,,$$

which implies that $\sum_{j=1}^{\infty} \frac{d_j}{10^j} \neq 1$.

2. case: $m = 1$. Let N be the lowest index among $n \geq 1$ for which $d_n > 0$. (If all $d_n = 0$ we get the expansion $1.000000 \cdots$.) Then

$$1 + \sum_{j=1}^{N} \frac{d_j}{10^j} = 1 + \frac{d_N}{10^N} \,.$$

But then

$$\left(1 + \sum_{j=1}^{\infty} \frac{d_j}{10^j}\right) - 1 = \frac{d_N}{10^N} + \sum_{n=N+1}^{\infty} \frac{d_j}{10^j} \geq \frac{1}{10^N}$$

since $d_N > 0$. This implies that $1 + \sum_{j=1}^{\infty} \frac{d_j}{10^j} \neq 1$. □

The above proof contains all the necessary ingredients for the following more general result.

Theorem 15.10. *Let $m.d_1d_2d_3\cdots$ and $n.e_1e_2e_3\cdots$ represent the same real number.*

(a) *If $m < n$, then $n = m + 1$, every $e_k = 0$ and every $d_k = 9$.*

(b) *If $m = n$ and if there is some j such that $d_j \neq e_j$, let N be the smallest index j such that $d_j \neq e_j$. If $d_N < e_N$ then $e_N = d_N + 1$ and $e_j = 0$ for all $j > N$, and $d_j = 9$ for all $j > N$.*

Thus if the non-negative rational number x is represented by two different decimals then both are repeating. Furthermore, if a real number has two different decimal expansions then one of those expansions has only finitely many nonzero digits. This immediately implies:

Corollary 15.11. *If $r \in \mathbb{R}$ has two different decimal expansions then $r \in \mathbb{Q}$.*

A nonnegative decimal $(m, d_1, d_2, d_3, \dots)$ is *repeating* if there exists $N \in \mathbb{N}$ and $p \in \mathbb{N}$ such that for all $0 \leq n < p$ and for all $k \in \mathbb{N}$

$$d_{N+n+kp} = d_{N+n} \,.$$

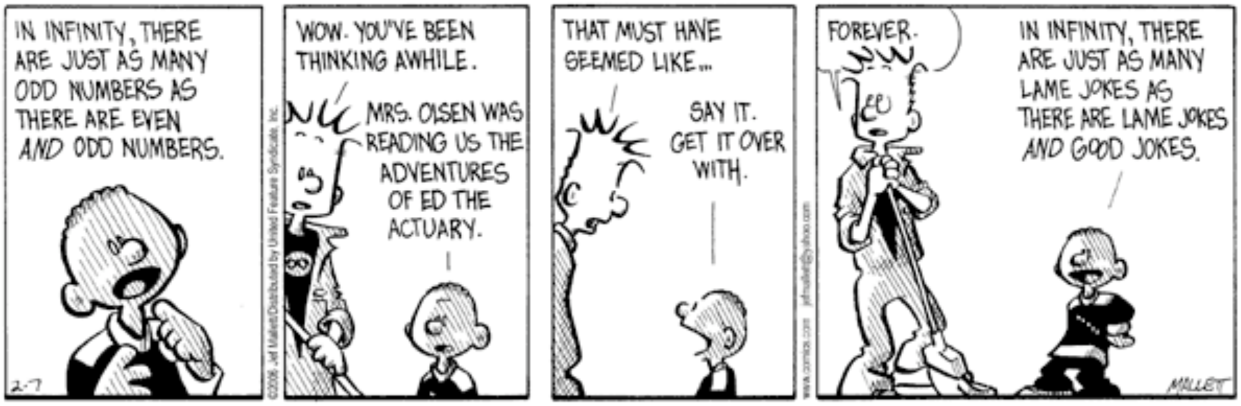
Examples are 5.666... , 0.34712712712712... . It is not hard to prove that a repeating (non-negative) decimal represents a (non-negative) rational number: there is a geometric series hidden here: do you see it? The converse is also true:

Proposition 15.12. *Every non-negative rational number x is represented by a non-negative repeating decimal.*

(*Hint for a possible proof:* Try an example to start: By long division express $\frac{21232444}{9999000}$ as a decimal. This will show you why eventually you must have repetition in the remainder, which forces a repeating decimal. To get a proof that every non-negative rational number has a repeating decimal expansion, use the division algorithm (Theorem 6.5): it expresses the result of long division without the algorithmic or calculational part.)

Chapter 16

Cardinality


IN INFINITY, THERE ARE JUST AS MANY ODD NUMBERS AS THERE ARE EVEN AND ODD NUMBERS.
2-7
WOW. YOU'VE BEEN THINKING AWHILE.
MRS. OLSEN WAS READING US THE ADVENTURES OF ED THE ACTUARY.
THAT MUST HAVE SEEMED LIKE...
SAY IT. GET IT OVER WITH.
FOREVER.
IN INFINITY, THERE ARE JUST AS MANY LAME JOKES AS THERE ARE LAME JOKES AND GOOD JOKES.
MALLET

© UFS, Inc.

The goal of this chapter is to compare the sizes of infinite sets. It is perfectly sensible to say that the sets $\{1, 2, 4\}$ and $\{2, 3, 5\}$ have the same size, but how do we go about comparing $\mathbb{N}$ with $\mathbb{Z}$, with $\mathbb{R}_{>0}$, with $\mathbb{Q}$, or with $\mathbb{R}$?

16.1 Injections, Surjections, and Bijections Revisited

Here is the central definition of what we mean when we say that two sets are "comparable in size": The sets A and B have the same **cardinality** (or **cardinal number**) if there exists a bijection from A to B.

A special case is that of finite sets, the archetype of which is $\{1, \dots, n\}$. Since we will use this set frequently in this chapter, we abbreviate $\{1, \dots, n\}$ by $[n]$.

The set A is **finite** if either $A = \varnothing$ or for some $n \in \mathbb{N}$ there is a bijection from $[n]$ to A. The set A is **countably infinite** if there is a bijection from $\mathbb{N}$ to A. The set A is **countable** if either A is finite or A is countably infinite.

Proposition 16.1. *If $m, n \in \mathbb{N}$ and $m \neq n$ there is no bijection $[m] \to [n]$.*

Thus for finite sets **the number of elements** is well defined: a set A contains n elements if and only if there is a bijection from $[n]$ to A. In that case, we say that A **contains** n **elements**. Then every set having the same cardinal number as A also contains n elements. We say $\varnothing$ contains 0 elements. The following result should appear in your proof of Proposition 16.1.

Corollary 16.2 (Pigeon Hole Principle)**.** *If $m > n$ then a function $[m] \to [n]$ cannot be injective.*

The reason for the name of this corollary is the following interpretation: Corollary 16.2 says that if we label n objects with numbers from 1 to m (where $m > n$) then there exist two objects that have the same label. The Pigeon Hole Principle appears in many different areas in mathematics and beyond. It asserts that, if there are m pigeons at n fishing holes, there are at least two pigeons that have to share a hole, or if there are m people in an elevator and n buttons are pressed, someone is playing a practical joke...

Proposition 16.3. *The non-empty set A is countable if and only if there is a surjection $\mathbb{N} \to A$.*

Proof. Suppose $A \neq \varnothing$ is countable. If A is finite, then there exists $n \in \mathbb{N}$ and a bijection $\phi : [n] \to A$. Let $\psi : \mathbb{N} \to [n]$ be defined by

$$\psi(m) = \begin{cases} m & \text{if } 1 \leq m \leq n, \\ 1 & \text{otherwise.} \end{cases}$$

Then $\phi \circ \psi : \mathbb{N} \to A$ is a surjection.

If A is infinite, then there exists a bijection from $\mathbb{N}$ to A, which is certainly surjective.

Conversely, suppose there exists a surjection $\sigma : \mathbb{N} \to A$. If A is finite, it is countable by definition, and we're done. If A is infinite, we define a bijection $\beta : \mathbb{N} \to A$ recursively through $\beta(1) = \sigma(1)$ and for $n \geq 2$

$$\beta(n) = \sigma(m) \qquad \text{where} \qquad m = \inf \{k \in \mathbb{N} : \sigma(k) \neq \sigma(j) \text{ for all } j < n\}.$$

(The set on the right-hand side is not empty because A is infinite.) □

Proposition 16.4. *A subset of a countable set is countable.*

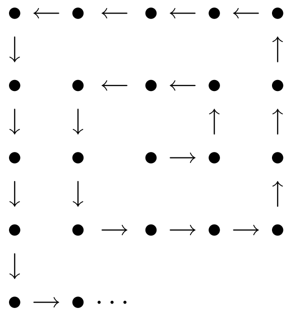
16.2 $\mathbb{Q}$ Is Countable, $\mathbb{R}$ Is Uncountable

The next propositions are counterintuitive at first sight.

Proposition 16.5. *$\mathbb{Z}$ is countable.*

Proposition 16.6. *$\mathbb{Z} \times \mathbb{Z}$ is countable.*

Here is the idea of the proof:



Corollary 16.7. $\mathbb{N} \times \mathbb{N}$ *is countable.*

Corollary 16.8. $\mathbb{Z} \times (\mathbb{Z} - \{0\})$ *is countable.*

Corollary 16.9. $\mathbb{Q}$ *is countable.*

Let's pause here. We know that $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q}$; moreover, each is a proper subset of the next one, i.e., $\mathbb{N} \neq \mathbb{Z}$ and $\mathbb{Z} \neq \mathbb{Q}$. This might make you think that $\mathbb{N}$ is smaller than $\mathbb{Z}$ and that $\mathbb{Z}$ is smaller than $\mathbb{Q}$. But we have just proved that these sets have the same cardinality—i.e., they have the same size. This can be confusing for beginners: if A and B are finite sets and A is a proper subset of B, then A and B have different cardinality by Proposition 16.1. We are seeing here that no such statement holds for infinite sets.

Proposition 16.10. *The countable union of countable sets is countable, i.e., if A_n is a countable set for each $n \in \mathbb{N}$ then $\displaystyle\bigcup_{n=1}^{\infty} A_n$ is countable.*

Theorem 16.11. $\mathbb{R}$ *is not countable.*

Proof. We will prove this by contradiction. Suppose that there is a surjective function $f : \mathbb{N} \to \mathbb{R}$. By Theorem 15.10, every real number has at most two decimal representations. So for each $n \in \mathbb{N}$, $f(n)$ can be written in the form $\pm m^{(n)}.d_1^{(n)} d_2^{(n)} d_3^{(n)} \cdots$. If there is more than one such decimal for $f(n)$, we use the one that has infinitely many nonzero digits.

Now let y be the real number represented by $0.a_1a_2a_3\cdots$, where

$$a_n = \begin{cases} 3 & \text{if } d_n^{(n)} \neq 3, \\ 4 & \text{if } d_n^{(n)} = 3. \end{cases}$$

Then for all $n \in \mathbb{N}$, $y \neq f(n)$ (because the n^{th} decimal places of y and $f(n)$ do not agree). Hence $y \in \mathbb{R}$ is not in the image of f, which contradicts the fact that f is surjective. □

It follows that there is no one-to-one correspondence between the infinite sets $\mathbb{R}$ and $\mathbb{Q}$, i.e., no function $f : \mathbb{Q} \to \mathbb{R}$ that is bijective. In particular, the "inclusion function" $\mathbb{Q} \to \mathbb{R}$ which takes each rational number to itself (regarded as a real number) is not surjective. This gives another proof that there exist irrational real numbers.

So $\mathbb{R}$ and $\mathbb{Q}$ have different cardinality, i.e., different size. This discovery was considered revolutionary in mathematics of the late 19th century. It was not that people had thought the opposite to be true; they just had never seriously considered the idea of infinite sets having different sizes. The foundations of the part of mathematics called analysis had to be completely rethought because of this.

Corollary 16.12. *The set $\mathbb{R} - \mathbb{Q}$ of irrational numbers is uncountable.*

The proof of Theorem 16.11 reveals even more. If shows that the set of decimals

$$\{0.d_1 d_2 d_3 \cdots : \text{ each } d_j = 3 \text{ or } 4\} \subseteq \mathbb{R}$$

is uncountable. In particular, the interval $[0,1] = \{x \in \mathbb{R} : 0 \le x \le 1\}$ is uncountable. This construction can be modified to prove:

Theorem 16.13. *Any interval $[x,y]$, where $x < y$, is uncountable.*

This allows us to complement Theorem 13.5 as follows.

Corollary 16.14. *Let $x, y \in \mathbb{R}$ with $x < y$. Then there is an irrational number z such that $x < z < y$.*

If you can only remember a few things from this book, let the following be one of them: *between any two real numbers lies a rational number and also an irrational number.*

Corollary 16.15. *There is no smallest positive irrational number.*

16.3 Power Sets and an Infinite Hierarchy of Infinities

We write $\operatorname{card} A \le \operatorname{card} B$ if there exists an injection $A \to B$. We write $\operatorname{card} A = \operatorname{card} B$ if A and B have the same cardinal number, i.e., if there exists a bijection $A \to B$. We write $\operatorname{card} A < \operatorname{card} B$ when $\operatorname{card} A \le \operatorname{card} B$ and $\operatorname{card} A \ne \operatorname{card} B$. By Proposition 11.5, $\operatorname{card} A \le \operatorname{card} B$ is equivalent to saying that there exists a surjection $B \to A$.

If A is a set, let $P(A)$ denote the set of all subsets of A, called the **power set** of A.

Example. If $A = \{a,b\}$, $P(A) = \{\varnothing, \{a\}, \{b\}, \{a,b\}\}$.

Our goal in this section is to prove that the power set of A is "bigger" than A.

Theorem 16.16. *For any set A, $\operatorname{card} A < \operatorname{card} P(A)$.*

This theorem is profound. It implies the existence of an infinite hierarchy of infinities. For example, it says that $P(\mathbb{N})$ is not countable, $P(P(\mathbb{N}))$ has larger cardinality than $P(\mathbb{N})$, $P(P(P(\mathbb{N})))$ is yet larger than $P(P(\mathbb{N}))$, etc.

We start with the finite case, for which we can be more precise:

Proposition 16.17. *For any $n \in \mathbb{N}$, $\operatorname{card} P([n]) = \operatorname{card}\,[2^n]$.*

(Before proving this, carefully think through the case $n = 1$.)

The next step on our way to proving Theorem 16.16 is the construction of the **characteristic function** $\chi_B : A \to \{0,1\}$ corresponding to a subset $B \subseteq A$:

$$\chi_B(x) = \begin{cases} 1 & \text{if } x \in B, \\ 0 & \text{otherwise.} \end{cases}$$

This function belongs to the set F_A of all functions of the form $f : A \to \{0, 1\}$:

$$F_A = \{f : \text{ } f \text{ has domain } A \text{ and codomain } \{0, 1\}\} .$$

This set is as large as the power set of A:

Proposition 16.18. *For any set A,* $\operatorname{card} F_A = \operatorname{card} P(A)$.

(*Hint for a possible proof:* Let $\phi : F_A \to P(A)$ be given by $\phi(f) = \{x \in A : f(x) = 1\}$. Show that ϕ is a bijection.)

Proof of Theorem 16.16. The injection $\iota : A \to P(A)$, $\iota(x) = \{x\}$ shows that $\operatorname{card} A \le \operatorname{card} P(A)$; the tricky part is to show that one cannot find a *bijection* $A \to P(A)$. It suffices by Proposition 16.18 to prove that there is no bijection $A \to F_A$. We do so by mimicking the proof of Theorem 16.11.

Again we give a proof by contradiction. Suppose that $\sigma : A \to F_A$ is a bijection. Define $f : A \to \{0, 1\}$ through

$$f(x) = \begin{cases} 0 & \text{if } (\sigma(x))\,(x) = 1\,, \\ 1 & \text{if } (\sigma(x))\,(x) = 0\,. \end{cases}$$

Then for all $x \in A$, $f \neq \sigma(x)$. Hence $f \in F_A$ is not in the image of σ, which contradicts the fact that σ is surjective. □

Cardinality questions are often difficult to answer. For example, it is a famous (and somewhat tricky-to-prove) theorem that $\operatorname{card} A \le \operatorname{card} B$ and $\operatorname{card} B \le \operatorname{card} A$ implies $\operatorname{card} A = \operatorname{card} B$. Even more basic, it is not obvious that the cardinal numbers of any two sets are comparable, but it is indeed true that, if A and B are sets, then either card $A \le$ card B or card $B \le$ card A. We will not include proofs of these theorems here.

16.4 Noncomputable Numbers

Even more remarkable than the noncountability of $\mathbb{R}$ is the existence of *noncomputable* real numbers, in a sense that will become clear shortly.

An **alphabet** A is a finite set whose elements we call **letters**. An **n-letter word** in A is an n-tuple of letters.[1] If we don't specify the length n we just speak of **words** in the alphabet A. If A has k letters then there are k^n distinct n-letter words. We will denote the set of all words in A by

$$W_A = \{w : \text{ for some } n \in \mathbb{N}, \text{ } w \text{ is an } n\text{-letter word in } A\} .$$

For example, A could be the set of all typesetting symbols used in this book (including "space," "period," "left parenthesis," etc.) Then this book is a word in A.

Fix an alphabet A. We call $x \in \mathbb{R}$ **A-computable** if there exists $m \in \mathbb{N}$ and an m-letter word in A which unambiguously describes x.

[1] An **n-tuple** with entries in A is a finite sequence $(x_1, x_2, \dots, x_n)$, where all $x_j \in A$. We can think of $(x_1, x_2, \dots, x_n)$ as an ordered set of these n elements. Note that the same entry may appear more than once in the n-tuple. For example $(2, 2, 3, 2, 4)$ is a 5-tuple.

Example. Let A consist of all symbols used in this book. You learned in calculus that

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots .$$

Thus the sequence of partial sums $\left(\sum_{j=1}^{k} (-1)^{j+1} \frac{1}{2j-1}\right)_{k=1}^{\infty}$ converges to $\frac{\pi}{4}$, and so

$$\sum_{j=1}^{\infty} (-1)^{j+1} \frac{4}{2j-1} = \pi$$

defines π unambiguously. This describes π as a member of W_A, and so π is A-computable.

Example. Again, let A consist of all symbols used in this book. Earlier we saw that $\sqrt{2} = \sup \left\{x \in \mathbb{R} : x^2 < 2\right\}$. This describes $\sqrt{2}$ unambiguously, so $\sqrt{2}$ is A-computable.

Proposition 16.19. *For any alphabet A, the set of words W_A is countable.*

We call $x \in \mathbb{R}$ **computable** if there is an alphabet A such that x is A-computable.

Theorem 16.20. *There exist real numbers that are not computable.*

Proof. It will be useful, for this proof, to think of the set $\mathbb{N}$ as an "infinite alphabet"; but we do not allow infinite words: an n-letter word in $\mathbb{N}$ is, by definition, an n-tuple of letters drawn from $\mathbb{N}$. We will first show that there are real numbers x that are not $\mathbb{N}$-computable.

Suppose all real numbers were $\mathbb{N}$-computable. Then there would be a function $g : W_{\mathbb{N}} \to \mathbb{R}$ which is onto. By Proposition 16.19, $\mathbb{R}$ would be countable, which contradicts Theorem 16.11.

For every alphabet A there is an injection $e : A \to \mathbb{N}$. Let x be a number which is not $\mathbb{N}$-computable. Then x cannot be A-computable, for if x could be described unambiguously using the letters of A, then x could be described unambiguously using the letters of $e(A)$. □

The last theorem is profound. In the earlier history of mathematics, it was thought that there is a chasm separating the "continuous" from the "discrete," or, if you like, $\mathbb{R}$ from $\mathbb{Z}$. Gradually it became clear that all real numbers can be understood in terms of integers via decimals. The above theorem reintroduces the chasm in a more subtle way. While there exists a decimal representation for each real number, the last theorem states the sense in which, for most real numbers, a decimal description cannot actually be written down.

Chapter 17

Complex Numbers

The imaginary number is a fine and wonderful resource of the human spirit, almost an amphibian between being and not being.
Gottfried Leibniz (1646–1716)

One deficiency of the real numbers is that the equation $x^2 = -1$ has no solution $x \in \mathbb{R}$ (Proposition 14.6). In this chapter, we will extend $\mathbb{R}$ to overcome this deficiency; the price that we'll have to pay is that this extension does not have a useful ordering relation.

17.1 Definition and Algebraic Properties

A **complex number** is an ordered pair of real numbers. The set of all complex numbers is denoted by $\mathbb{C} = \{(x, y) : x, y \in \mathbb{R}\}$. $\mathbb{C}$ is equipped with the **addition**

$$(x, y) + (a, b) = (x + a, y + b)$$

and the **multiplication**

$$(x, y) \cdot (a, b) = (xa - yb, xb + ya) .$$

Just as we embedded $\mathbb{Z}$ in $\mathbb{R}$, we embed $\mathbb{R}$ in $\mathbb{C}$ by the injective function $e : \mathbb{R} \to \mathbb{C},\ e(r) = (r, 0)$. Identifying r with $e(r)$ we will write $\mathbb{R} \subseteq \mathbb{C}$ from now on.

Proposition 17.1. *For all $(a, b), (c, d), (e, f) \in \mathbb{C}$:*

(i) $(a, b) + (c, d) = (c, d) + (a, b) .$

(ii) $((a, b) + (c, d)) + (e, f) = (a, b) + ((c, d) + (e, f)) .$

(iii) $(a, b) \cdot ((c, d) + (e, f)) = (a, b) \cdot (c, d) + (a, b) \cdot (e, f) .$

(iv) $(a, b) \cdot (c, d) = (c, d) \cdot (a, b) .$

(v) $((a, b) \cdot (c, d)) \cdot (e, f) = (a, b) \cdot ((c, d) \cdot (e, f)) .$

Proposition 17.2. *There is a complex number $0 \in \mathbb{C}$ such that for all $z \in \mathbb{C}$, $z + 0 = z$.*

Proposition 17.3. *There is a complex number* $1 \in \mathbb{C}$ *with* $1 \neq 0$ *such that for all* $z \in \mathbb{C}$, $z \cdot 1 = z$.

Proposition 17.4. *For all* $z \in \mathbb{C}$, *there exists a complex number denoted* $-z \in \mathbb{C}$, *such that* $z + (-z) = 0$.

Proposition 17.5. *For all* $z \in \mathbb{C} \setminus \{0\}$, *there exists a complex number, denoted* z^{-1}, *such that* $z \cdot z^{-1} = 1$.

The last two propositions allow us to define **subtraction** and **division** of two complex numbers, just like in the real case.

The definition of our multiplication implies the innocent looking statement

$$(0,1) \cdot (0,1) = (-1,0)\,. \tag{17.1}$$

This equation together with the fact that

$$(a,0) \cdot (x,y) = (ax, ay)$$

leads to an alternative notation for complex numbers—the notation which is always used—as we now explain. We can write

$$(x,y) = (x,0) + (0,y) = (x,0) \cdot (1,0) + (y,0) \cdot (0,1)\,.$$

If we think—in the spirit of our remark on the embedding of $\mathbb{R}$ in $\mathbb{C}$—of $(x,0)$ and $(y,0)$ as the real numbers x and y, then we can write any complex number (x,y) as a linear combination of $(1,0)$ and $(0,1)$, with the real coefficients x and y. Now, $(1,0)$ can be thought of as the real number 1. So if we give $(0,1)$ a special name, the traditional choice is i, then the complex number which we have been writing as $z = (x,y)$ can be written as $x \cdot 1 + y \cdot i$, or in short,

$$z = x + iy\,.$$

x is called the **real part** and y the **imaginary part**[1] of the complex number $x + iy$, often denoted as $\mathrm{Re}(x+iy) = x$ and $\mathrm{Im}(x+iy) = y$. We invite you to check that the definitions of our binary operations and the above propositions are coherent with the usual "real" arithmetic rules if we think of complex numbers as given in the form $x + iy$.

The equation (17.1) now reads

$$i^2 = -1\,,$$

so that $i \in \mathbb{C}$ is a solution to the equation $x^2 = -1$.

One can say much more: every polynomial equation has a solution in $\mathbb{C}$. This fact, the *Fundamental Theorem of Algebra*, is too difficult for us to prove here. But you will see a proof if you take a course in complex analysis (which we strongly recommend). While we gained solutions to previously unsolvable equations by extending $\mathbb{R}$ to $\mathbb{C}$, this came for a price:

Project 17.6. *Discuss the sense in which* $\mathbb{C}$ *does not satisfy Axiom 10.6.*

[1] The name has historical origins: people thought of complex numbers as unreal, imagined.

17.2 Geometric Properties

Although we just introduced a new way of writing complex numbers, let's for a moment return to the (x, y)-notation. It suggests that one can think of a complex number as a two-dimensional real vector. When plotting these vectors in the plane $\mathbb{R}^2$, we will call the x-axis the **real axis** and the y-axis the **imaginary axis**. The addition that we defined for complex numbers resembles vector addition. The analogy stops at multiplication: there is no usual multiplication of two vectors which gives another vector.

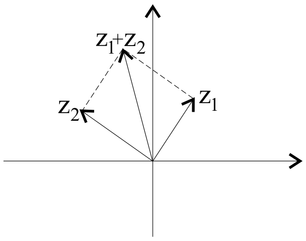


Figure 17.1: Addition of complex numbers.

Any vector in $\mathbb{R}^2$ is defined by its two coordinates. On the other hand, it is also determined by its length and the angle it encloses with, say, the positive real axis; let's define these concepts thoroughly. (Here we need to assume high-school trigonometry.) The **absolute value** (sometimes also called the **modulus**) of $x + iy$ is

$$r = |x + iy| = \sqrt{x^2 + y^2}$$

and an **argument** of $x + iy$ is a number ϕ such that

$$x = r\cos\phi \qquad \text{and} \qquad y = r\sin\phi\,.$$

This means, naturally, that any complex number has many arguments; any two of them differ by a multiple of 2π.

The absolute value of the difference of two vectors has a nice geometric interpretation: it is the *distance* of the (end points of the) two vectors (see Figure 17.2). It is very useful to keep this geometric interpretation in mind when thinking about the absolute value of the difference of two complex numbers.

The first hint that absolute value and argument of a complex number are useful concepts is the fact that they allow us to give a geometric interpretation for the multiplication of two complex numbers.

Proposition 17.7. *Suppose $x_1 + iy_1 \in \mathbb{C}$ has absolute value r_1 and argument ϕ_1, and $x_2 + iy_2 \in \mathbb{C}$ has absolute value r_2 and argument ϕ_2. Then the product $(x_1 + iy_1)(x_2 + iy_2)$ has absolute value r_1r_2 and (one of its) argument $\phi_1 + \phi_2$.*

Geometrically, we are multiplying the lengths of the two vectors representing our two complex numbers, and adding their angles measured with respect to the positive x-axis.[2]

[2] You should convince yourself that there is no problem with the fact that there are many possible arguments for complex numbers, as both cosine and sine are periodic functions with period 2π.

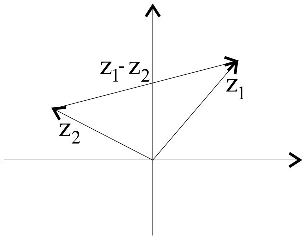


Figure 17.2: Geometry behind the distance of two complex numbers.

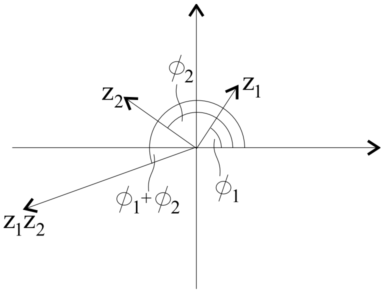


Figure 17.3: Multiplication of complex numbers.

Thus the notation $e^{i\phi} = \cos\phi + i\sin\phi$ can become handy. With this notation, the sentence "The complex number $x + iy$ has absolute value r and argument ϕ" now becomes the equation

$$x + iy = re^{i\phi}.$$

The left-hand side is often called the **rectangular form**, the right-hand side the **polar form** of this complex number.

At this point, this exponential notation is indeed purely a notation. It has an intimate connection to the complex exponential function, which you will see in a course in complex analysis. For now we motivate our use of this notation by the following proposition.

Proposition 17.8. *For any $\phi, \phi_1, \phi_2 \in \mathbb{R}$,*

$$\begin{aligned} e^{i\phi_1}\, e^{i\phi_2} &= e^{i(\phi_1+\phi_2)}. \\ 1/e^{i\phi} &= e^{-i\phi}\,. \\ e^{i(\phi+2\pi)} &= e^{i\phi}. \\ \left|e^{i\phi}\right| &= 1\,. \end{aligned}$$

Proposition 17.9. *For any $z \in \mathbb{C},\ x, y \in \mathbb{R}$,*

(i) $-|z| \le \operatorname{Re} z \le |z|\,.$

(ii) $-|z| \le \operatorname{Im} z \le |z|\,.$

(iii) $|x + iy|^2 = (x + iy)(x - iy)\,.$

The last equation of this proposition is one of many reasons to give the process of passing from $x+iy$ to $x-iy$ a special name: $x-iy$ is called the **(complex) conjugate** of $x+iy$. We denote the conjugate by

$$\overline{x+iy} = x-iy\,.$$

Geometrically, conjugating z means reflecting the vector corresponding to z in the real axis—think of the real axis as a mirror. Here are some basic properties of the conjugate.

Proposition 17.10. *For any $z, z_1, z_2 \in \mathbb{C}$, $\phi \in \mathbb{R}$,*

$$\begin{aligned}
\overline{z_1 \pm z_2} &= \overline{z_1} \pm \overline{z_2}\,. \\
\overline{z_1 \cdot z_2} &= \overline{z_1} \cdot \overline{z_2}\,. \\
\overline{z_1/z_2} &= \overline{z_1}/\overline{z_2}\,. \\
\overline{(\overline{z})} &= z\,. \\
|\overline{z}| &= |z|\,. \\
|z|^2 &= z\overline{z}\,. \\
\operatorname{Re} z &= \tfrac{1}{2}\left(z+\overline{z}\right). \\
\operatorname{Im} z &= \tfrac{1}{2i}\left(z-\overline{z}\right). \\
\overline{e^{i\phi}} &= e^{-i\phi}.
\end{aligned}$$

We finish this chapter with the complex counterpart to Proposition 13.10(iv), the *triangle inequality.*

Proposition 17.11. *For any two complex numbers z_1, z_2,*

$$|z_1+z_2| \le |z_1| + |z_2|\,.$$

Chapter 18

Final Remarks

The greatest reward for a student is not a good grade. It is the willingness of his teacher to listen to him.
Nikolay Konstantinov

We have had several purposes in this course:

1. To teach you the *axiomatic method*. This was described in Chapter 1 but the point may not have been clear at the beginning. Please reread Chapter 1.

2. To teach you to *read mathematics* by forcing you to read your own mathematics: to know the difference between an incorrect argument and a correct argument.

3. To teach you to *do mathematics*: to discover and write down your own proofs, so that what you write down accurately reflects what you discovered and is free of mistakes. As time goes on your style of writing math will improve: watch how the writers of your textbooks write. Develop opinions about good and bad writing.

4. To teach you to *write mathematics* so that it is communicated accurately and clearly to another qualified reader.

5. To teach you *induction*, one of the most fundamental tools.

6. To make you *understand the real numbers* and how the rational numbers are distributed in them.

7. To put the *whole math curriculum* from Sesame Street through calculus in perspective.

Final Project. *Discuss whether the following lines (from the poem* Little Gidding *in "Four Quartets" by T.S. Eliot) are relevant to the course we have just finished:*

We shall not cease from exploration
And the end of all our exploring
Will be to arrive where we started
And know the place for the first time.

Index