To the Student

Mathematics is unique among human intellectual endeavors; it is not art, philosophy, or science, but it shares certain features with each. The example of digital data storage will help convey the nature and uses of mathematics and the flavor of the material covered in this book.

Computers store, manipulate, and transmit data as *bits* or *binary digits*. Physically, bits have been represented and conveyed in a vast array of schemes from historical to modern, including

- Shaking or nodding one's head.
- Dots and dashes in Morse code.
- Holes and "no-holes" in a paper strip: punch cards, ticker tape.
- Magnetic domains: floppy and ZIP disks, PC hard drives.
- Light and dark spots or bands: compact disks, UPC symbols, QR codes.
- Charged and uncharged capacitors: flash memory, RAM.

A mathematician or theoretical computer scientist sees no essential difference between these schemes: The central mathematical "object" is a *pair of contrasting states*. Depending on context, the states might be called (and, in actual practice, *are* called) "zero and one", "false and true", "white and black", "no and yes", "open and closed", "low and high", or "off and on".

Mathematical abstraction extends beyond data, however, encompassing the operations performed on objects.

Binary arithmetic. Think of 0 as representing an arbitrary even integer, and 1 as representing an arbitrary odd integer; namely, identify an integer with its remainder on division by 2. The sum of two odd integers is even ("$1 + 1 = 0$"), the product of an even and an odd

integer is even ("$0 \cdot 1 = 0$"), and so forth. These relationships may be tabulated as

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

Boolean logic. Think of F as representing an arbitrary "false" assertion (such as "$2 + 2 = 5$") and T as representing an arbitrary "true" sentence (such as "$1 + 1 = 2$"). Since, e.g., "$2 + 2 = 5$ or $1 + 1 = 2$, but not both" is true, we write "F xor T = T". ("xor" stands for "exclusive or": one statement is true, but not both.) Since "$2+2 = 5$ and $1+1 = 2$" is false, we write "F and T = F". Tabulating the "truth value" of the statements made by conjoining two statements, according to whether or not each statement is true or false,

xor	F	T
F	F	T
T	T	F

and	F	T
F	F	F
T	F	T

Abstract implementation. The truly mathematical observation is *the entries of these pairs of tables correspond*: Under the correspondence even-False and odd-True, "addition (mod 2)" corresponds to "xor", and "multiplication (mod 2)" corresponds to "and". The tables above are different "implementations" of the same abstract structure, which we might denote by

$\vee$	$\circ$	$\bullet$
$\circ$	$\circ$	$\bullet$
$\bullet$	$\bullet$	$\circ$

$\wedge$	$\circ$	$\bullet$
$\circ$	$\circ$	$\circ$
$\bullet$	$\circ$	$\bullet$

These three pairs of tables exemplify an abstract relationship, known as "isomorphism", between operations on pairs of contrasting states. Any logical consequence that holds for one implementation necessarily holds for other implementations. For example, if we let variables x and y stand for either of two contrasting states, and we denote by x' the state unequal to x, then the identities

$$(x')' = x, \qquad x \vee y = (x \wedge y') \vee (x' \wedge y) = x' \vee y'$$

hold regardless of what values are assigned to x and y, and, more significantly, *no matter which implementation is used*. In this fashion, mathematical structures can be studied and organized with their

extraneous details set aside. More subtle patterns can sometimes be discerned in the abstraction, leading to a deeper understanding of the original structures.

Among of the most remarkable internal features of mathematics is its absoluteness: the perfect and intricate logical meshing of truth even when drawn from widely separated sub-disciplines, the universality of theorems across cultures, the sense among mathematicians that their work describes an objective (if non-physical) reality.

As a language, mathematics is unparalleled in its ability to express features of the natural world, often with astounding accuracy. At the same time, mathematics has no known *a priori* connection to the real world. The objects of mathematics are idealized concepts (such as "a pair of contrasting states"), and do not have physical existence in the same way stars, molecules, or people do. Conversely, stars, molecules, and people are not mathematical objects, though they do possess attributes that can be modeled by mathematical concepts.

This book was written to help you bridge the gap between informal intuition and the more formal language and framework of modern mathematics. Learning mathematics requires active preparation and participation from you, but offers continual rewards, including deepened comprehension of the natural world and the sheer enjoyment of mathematical beauty.

Practice reading actively, with a pencil and scratch paper. When you encounter a new definition, try to construct examples and non-examples before reading further, and ask yourself how you might test an object to see if it satisfies the definition.

Develop the habit of filling in the missing steps of calculations and omitted "standard" steps of proofs. When you first read the statement of a theorem, pause to think about what the theorem claims, and whether or not you *believe* the assertion. Try to sketch out an argument on your own before reading the book's proof.

Situate new general concepts and examples in the context of your existing mathematical knowledge. Pay attention to the overall structure of proofs, not merely to the details. Look for commonalities in arguments, and be sure you are able to use these strategies yourself. Your repertoire of proof techniques and other mathematical idioms will grow steadily.

Work on mathematics outside of class every day, rather than in one or two long "marathon sessions" per week. Don't become discouraged if

new ideas don't immediately "click". Re-read confusing passages after a day or two. Speak with classmates and your instructor for clarification as necessary. At the same time, develop intellectual self-reliance. The more mathematics you have made your own, the easier learning new mathematics becomes.

Above all, cultivate the enjoyment of thinking about new ideas, solving problems, and finding meaningful connections between seemingly disparate concepts. The greatest reward of your mathematical studies will, ideally, be a deeper experience of life itself.

Contents

Part I

The Language of Mathematics

Chapter 1

Logic and Proofs

Mathematics admits no "absolute truth". Instead, most mathematicians work within the axiom system known as Zermelo-Fraenkel with choice, or ZFC for short. ZFC formalizes the concept of a *set*, an abstraction of a collection of objects, called *elements*. For now, the details of ZFC are unimportant. This chapter describes the basic rules of logic. Chapter 2 provides an informal introduction to ZFC.

ZFC is believed to be logically consistent, and the "correctness" of mathematical statements is evaluated according to "provability" and "logical consistency" with respect to ZFC. Theorems proved in ZFC are said colloquially to be "true". Strictly speaking, however, mathematicians do not find metaphysical truths, but instead deduce logical *conclusions* starting from assumptions called *hypotheses*.

1.1 Statements and Negations

A *statement* is a sentence having a *truth value*, T (True) or F (False). Contact with the external world can be made via experience, but in mathematics *true* and *false* may be viewed as undefined terms.

As noted earlier, the basic objects of ZFC are sets, collections of elements. The examples below refer to the set of *integers*, or whole numbers: 0, 1, -1, 2, -2, and so forth.

Example 1.1. -4 is an even integer.

The decimal expansion of π is non-repeating and contains the string '999999'. (True)

$2 + 2 = 5$. (False)

Example 1.2. Sentences that are *not* statements include "n is an even integer" (whose truth value depends on n) "10^{1000} is a large number" ("large" has not been given a mathematical definition), and the self-referential examples, "This sentence is true" (whose truth value must be specified as an axiom) and "This sentence is false" (which cannot be consistently assigned a truth value).

1.2 Negation and Logical Connectives

Conventionally, abstract statements are denoted P and Q.

Not. The *negation* of a statement P is its logical opposite $\neg P$. You may regard the negation as P preceded by the clause "It is not the case that...", but usually a more pleasant wording can be found.

Example 1.3. P: $2+2=4$. $\neg P$: $2+2\neq 4$.

Let P and Q be statements. New statements can be constructed using the "logical connectives" *and*, *or*, and *implies*.

And. The statement "P and Q" has its ordinary meaning: The compound statement is true provided both P and Q are true, and is false otherwise.

Example 1.4. $2+2=4$ and $0<1$. (True)
$2+2=5$ and $0<1$. (False)
$2+2=5$ and $1<0$. (False)

Or. The statement "P or Q" always has the "inclusive" meaning in mathematics: P is true, or Q is true, *or both*.

Example 1.5. $2+2=4$ or $0<1$. (True)
$2+2=5$ or $0<1$. (True)
$2+2=5$ or $1<0$. (False)

Remark 1.6. In colloquial English, "or" is frequently used in the "exclusive" sense. The sentence "You will earn a 70% on the final exam or you will not pass the course" is conventionally interpreted to mean "If you earn a 70% on the final exam, then you will pass the course, and if you do not earn a 70%, then you will not pass."

Mathematicians and computer scientists denote "exclusive or" by "xor" to distinguish it from "or". The statement "P xor Q" means

P is true, or Q is true, *but not both*. When needed, "P xor Q" can be expressed as "(P or Q) and $\neg$(P and Q)". In this book, xor does not appear again.

Implies.A statement of the form "If P then Q", also read "P *implies* Q", is a *logical implication*. P is called the *hypothesis* of the implication, Q the *conclusion*.

By definition, a logical implication "P implies Q" is true provided Q is true whenever P is true. In other words, "P implies Q" is false precisely when P is true and Q is false.

Example 1.7. If $1 \neq 0$, then $1^2 \neq 0$. (True)
If $1 \neq 0$, then $1^2 = 0$. (False)
If $1 = 0$, then $0 = 0$. (True)
If $1 = 0$, then $1^2 = 0$. (True)

Logical implication plays a central role in mathematics. If "P implies Q" is true, we say the implication is *valid*, and view Q as being *deduced* or *derived* from P. The definition of valid implication ensures that by starting with true hypotheses and making valid deductions, we obtain only true conclusions, not falsehoods.

There are two noteworthy and potentially confusing consequences of this convention. First, it is valid (not logically erroneous) to deduce an arbitrary conclusion from a false hypothesis. An implication with false hypothesis is said to be *vacuous*. Humorous examples abound: "If $1 = 0$, then money grows on trees."

In particular, the third and fourth implications of the preceding example are vacuous. Note that in each case, we can give a proof. If $1 = 0$, then subtracting this equation from itself gives $0 = 0$, which proves the third statement. To prove the fourth statement, square both sides, obtaining $1^2 = 0^2 = 0$.

Second, a valid implication need not connect causally related statements. The implication "If $0 = 0$, then 2 is an even integer" is valid because both the hypothesis and conclusion are true, but is effectively a *non sequitur*; the conclusion does not "follow" from the hypothesis in any obvious sense. A valid implication does not, of itself, constitute a proof. In the example at hand, we know the implication is valid only because there exists a separate proof, consisting of implications whose validity can be checked directly.

In these two senses, mathematicians are liberal in deeming an implication to be valid. Again, "validity" is the weakest criterion that

excludes the act of drawing a false conclusion from a true hypothesis.

Remark 1.8. If, in some axiom system, some statement P and its negation $\neg P$ are both true, then *every* statement Q is provable, since either "P implies Q" or "$\neg P$ implies Q" is vacuously true. The pair $\{P, \neg P\}$ is called a *logical contradiction*. An axiom system is *inconsistent* if a contradiction can be derived in it.

Work of K. Gödel in the 1930s showed ZFC cannot be proved consistent without using some other ("more powerful") axiom system whose consistency is unknown. However, if there is a contradiction in ZFC, there is a contradiction in ordinary arithmetic.

Belief in the consistency of ZFC is about as close as mathematics gets to an "article of faith".

In this book, and throughout mathematics in practice, valid deductions do actually link causally related statements. Most implications involve classes of objects, and assert that every object satisfying some condition must also satisfy some other condition.

Negation and Conjunctions

If P and Q are statements, then the statement "P and Q" is false if *at least one* of P and Q is false. If someone assures you two statements are both true, only one has to be false for the assurance to be unfounded. Formally, the compound statements

$$\neg(P \text{ and } Q), \qquad (\neg P) \text{ or } (\neg Q)$$

express the same logical condition.

Analogously, if someone assures you at least one statement of two is true, then both must be false for the assurance to be unfounded. Formally, the compound statements

$$\neg(P \text{ or } Q), \qquad (\neg P) \text{ and } (\neg Q)$$

express the same logical condition.

Together, the two preceding relationships are known as *De Morgan's laws*, after the 19th Century English logician A. De Morgan. Loosely, the conjunctions "and" and "or" are interchanged by negation, perhaps contrary to first impression.

Consequently, the order of negation and a connective matters:

Example 1.9. The integers 1 and 0 are **not both** zero. (True.)

The integers 1 and 0 are **both not** zero. (False.)
The integers 1 and -1 are **both not** zero. (True.)

Remark 1.10. All too frequently, one sees humorous ambiguities of the type "While driving, teens should not use cell phones and obey traffic laws". To avoid confusion, this sentence should be phrased "While driving, teens should obey traffic laws and not use cell phones" (placing the negation where it clearly applies only to one clause) or "While driving, teens should not use cell phones, and should obey traffic laws" (explicitly delimiting the negation).

In formal logic, "$\neg P$ and Q" means "$(\neg P)$ and Q".

1.3 Quantification

To accommodate classes of objects in the framework of statements, we allow statements to contain *variables* standing for elements of a set, so long as each variable is "quantified", accompanied by the phrase "for every" or "there exists". The quantifiers are crucial; pay close attention to them while reading, and *do not omit them when thinking and writing.*

Example 1.11. For every integer n, $n^2 - n$ is an even integer. (True)
For every integer n, $n^2 \geq 0$. (True)
For every integer n, $n^2 = 1$. (False)

The preceding "for every" statements involve *universal quantification.* Each statement encapsulates multiple statements. For example, the first statement of the preceding example encapsulates an infinite collection of statements, one for each integer: $0^2 - 0$ is an even integer; $1^2 - 1$ is an even integer; $(-1)^2 - (-1)$ is an even integer; and so forth.

Example 1.12. There exists an integer n such that $n^2 = 1$. (True)
There exists an integer n such that $n^2 = 2$. (False)
There exists an n such that both n and $n + 1$ are even. (False)

The preceding "there exists" statements involve *existential quantification.* Again, each encapsulates multiple statements. For example, the third expresses that at least one truism is found among the statements: 0 and 1 are both even; 1 and 2 are both even; -1 and 0 are both even; and so forth. The compound statement is false because *every* individual statement is false.

Remark 1.13. The statements of the preceding examples contain only "bound" (i.e., quantified) variables.

Sentences containing "free" or "unbound" variables (such as "n is an even integer" or "$x^2+x-2=0$") are not statements. However, sentences containing unbound variables play the useful role of *conditions* in mathematics, selecting objects (perhaps integers n or real numbers x) for which the resulting statement is true.

Many mathematical theorems take the universally quantified form "For every x satisfying $P(x)$, condition $Q(x)$ is true". For stylistic variety, such statements may be worded as implications involving "arbitrary" values of variables.

Example 1.14. If x is a real number such that $x^2 + x - 2 = 0$, then $x = 1$ or $x = -2$. (True)

If n is an integer, then there exist unique integers q and r such that $n = 4q + r$ and $0 \leq r < 4$. (True)

If a, b, and c are positive integers, then $a^3 + b^3 \neq c^3$. (True)

Quantifiers and Negation

The universal quantifier "for every" may be viewed as an enhancement of the "and" conjunction: "For every integer n, the condition $P(n)$ is true" means that the infinitely many statements $P(0)$, $P(1)$, $P(-1)$, and so forth, are *all* true.

The existential quantifier "there exists" may be viewed similarly as an enhancement of "or": "There exists an integer n such that the condition $P(n)$ is true" means that among the infinitely many statements $P(0)$, $P(1)$, $P(-1)$, ..., *at least one* is true.

Example 1.15. Logical negation "converts" a "for every" statement into a "there exists" statement of negations, and converts a "there exists" statement into a "for every" statement of negations:

P: For every integer n, $n^2 \geq 0$.
$\neg P$: There exists an integer n such that $n^2 < 0$.

P: There exist integers m and n such that $m^2 + n^2 = 8$.
$\neg P$: For all integers m and n, $m^2 + n^2 \neq 8$.

This type of linguistic transformation needs to become second nature. Particularly, a positive assertion regarding a class of objects can be disproved by finding a counterexample, but cannot be proved by finding an example.

Remark 1.16. When the hypothesis of a logical implication contains a variable but no quantifier is explicitly present, the convention is to read "for every". For example, "If $x > 0$ then $x^2 > 0$" should be read "For every real number x, if $x > 0$ then $x^2 > 0$" (assuming the context dictates real numbers as opposed to, say, integers).

If an implicitly-quantified statement is negated, the existential quantifier must be added explicitly: "There exists a real number $x > 0$ such that $x^2 \leq 0$".

To avoid confusion, including your own, include logical quantifiers explicitly. This book makes a special effort to set a good example.

Implications, and Multiple Quantifiers

Among the most subtle conditions in mathematics are those containing multiple quantifiers. Elementary algebra seldom ventures into this territory, but analysis, the mathematics underlying and extending differential and integral calculus, is built upon definitions and theorems of this type. When you encounter multiply-quantified statements, slow down and read several times to ensure you thoroughly understand the dependencies implicit in the ordering.

Example 1.17. For every integer n, there exists an integer M such that $n \leq M$. (True; every integer n is smaller than some other integer M.)

There exists an integer M such that for every integer n, $n \leq M$. (False; there is no largest integer M, i.e., no integer that is greater than every other integer n.)

1.4 Truth Tables and Applications

The logical operators ("not", "and", "or", and "implies") introduced above are neatly summarized by *truth tables*:

P	Q	$\neg P$	P and Q	P or Q	P implies Q
T	T	F	T	T	T
T	F	F	F	T	F
F	T	T	F	T	T
F	F	T	F	F	T

Truth tables furnish a useful tool for studying sentences built of other statements and logical connectives. This section gives a few applications.

Logical Equivalence. Two statements P and Q are *logically equivalent* if each implies the other: P implies Q and Q implies P. For brevity, we may write P iff Q, "iff" being short for "if and only if". A truth table calculation shows P and Q are equivalent precisely when they have the same truth value:

P	Q	P implies Q	Q implies P	P iff Q
T	T	T	T	T
T	F	F	T	F
F	T	T	F	F
F	F	T	T	T

The Converse. The implications "P implies Q" and "Q implies P" are said to be *converse* to each other. The preceding table shows these implications are not equivalent.

The Contrapositive. The implications "P implies Q" and "$\neg Q$ implies $\neg P$" are said to be *contrapositive* to each other. An implication and its contrapositive are logically equivalent:

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

This fact of logic should become second nature to you. Many implications are easier to understand and prove in contrapositive form.

Example 1.18. In each statement, x stands for a real number. Let P be the statement "$x^2 - 1 \neq 0$" and Q be the statement "$x \neq 1$".

The implication P implies Q is true, but may require a few seconds' thought to see.

The converse implication, "If $x \neq 1$, then $x^2 - 1 \neq 0$" is an invalid deduction. The number $x = -1$ is a *counterexample*: It satisfies the converse hypothesis Q, but not the converse conclusion P.

The contrapositive reads, "If $x = 1$, then $x^2 - 1 = 0$." This implication is obviously true, and on general grounds its truth implies the truth of P implies Q.

Example 1.19. In each statement below, n is a positive integer. A positive integer n is said to be *prime* if $n > 1$, and if n has no positive divisors other than 1 and n.

Direct implication: If n is a prime, then $n = 2$ or n is odd. (True.)

Converse: If $n = 2$ or n is odd, then n is a prime. (False: $n = 1$ and $n = 9$ are the two smallest of infinitely many counterexamples.)

Contrapositive: If $n \neq 2$ and n is not odd, then n is not prime. (True. Every such integer has the form $n = 2k$ for some integer $k > 1$.)

One final example, drawn from analysis rather than from algebra, will illustrate the power of the contrapositive.

Example 1.20. In each statement, $x \geq 0$ is a real number, and n is a positive integer.

Direct implication: If $x < 1/n$ for every n, then $x = 0$.

Converse: If $x = 0$, then $x < 1/n$ for every n.

Contrapositive: If $x > 0$, then there exists an n such that $1/n \leq x$.

It turns out that all three statements are true. The second is easily seen, even though the conclusion consists of infinitely many statements: $0 < 1$, $0 < 1/2$, $0 < 1/3$, etc.

The third statement is true, and not difficult to see; informally, $1/k \to 0$ as $k \to \infty$, so if $x > 0$, there is some positive integer n such that $1/n \leq x$.

The direct implication is therefore true, since its contrapositive is true. However, the direct implication exhibits a new phenomenon: The hypothesis consists of infinitely many statements, $x < 1$, $x < 1/2$, $x < 1/3$, etc., but *no finite number of these statements implies the conclusion*. Indeed, if we assume only finitely many inequalities of the form $x < 1/n$, there is a largest denominator, say N, and our collection of inequalities is equivalent to the single inequality $x < 1/N$, which does not imply $x = 0$.

Exercises

Exercise 1.1. In each pair P, Q of conditions, n represents an integer. (i) Give the negations of P and Q, and (ii) Form the implication P implies Q, its converse, and its contrapositive, and determine whether each is true.

(a) P: $n^2 - 4 = 0$. Q: $n = 2$.

(b) P: n is even. Q: n is an integer multiple of 4.

(c) P: n is even. Q: n is the square of an even integer.

Exercise 1.2. Let P and Q be arbitrary statements. Use a truth table to prove that "P implies Q" is logically equivalent to "$\neg P$ or Q".

Exercise 1.3. Let P, Q, and R be arbitrary statements. Use a truth table to prove the following pairs of statements are logically equivalent:

(a) "$\neg(P$ or $Q)$" and "$\neg P$ and $\neg Q$".

(b) "$\neg(P$ and $Q)$" and "$\neg P$ or $\neg Q$".

(c) "$(P$ or $Q)$ and R" and "$(P$ and $R)$ or $(Q$ and $R)$".

(d) "$(P$ and $Q)$ or R" and "$(P$ or $R)$ and $(Q$ or $R)$".

Exercise 1.4. A game-show host presents the contestant with the equation "$a^2 + b^2 = c^2$". The contestant replies, "What is the Pythagorean theorem?"

Why is the contestant's reply logically deficient? Modify it to give a mathematically satisfactory question.

Exercise 1.5. The President, a law-abiding citizen who always tells the truth, has time for one more Yes/No question at a press conference. In an attempt to embarrass the President, a reporter asks, "Have you stopped offering illegal drugs to visiting Heads of State?"

(a) Which answer ("Yes" or "No") is logically truthful?

(b) Suppose the President answers "Yes". Can the public conclude that the President has offered illegal drugs to visiting Heads of State? What if the answer is "No"?

(c) Explain why both answers are embarrassing.

If the President were a Zen Buddhist she might reply "mu" (pronounced "moo"), meaning "Your question is too flawed in its hypotheses to answer meaningfully."

Exercise 1.6. In this question, assume advertisers tell the truth, but selectively. For each statement, discuss uncharitable conditions under which the statement would be true, and explicitly state "the whole truth", the charitable interpretation the advertiser wants the consumer to infer.

(a) Made with all-natural ingredients.

(b) Now with 50% less sugar per serving.

(c) New smaller package, same great price.

(d) Contains real fruit.

Exercise 1.7. Correctly using the phrases "for every", "there exists", and "such that" can be tricky. Explain why each of the following is anomalous, and determine the presumed meaning.

(a) If there exists an x such that $x = y$, then $x^2 - y^2 = 0$.

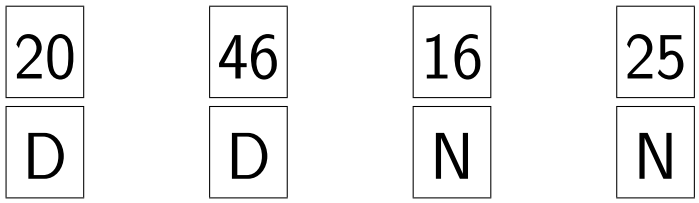
(b) If $\delta > 0$ for every δ such that $\delta > 1$, then $0 < 1 < \delta^2$.

(c) If $y = x^2$ for every $x > 0$, then $y > 0$.

Exercise 1.8. The human brain has evolved to detect "cheating"—behavior violating established rules. These rules may have logical formulations, but the "cheating" interpretation can be remarkably easier to "see".

(a) Each card in a deck is printed with a letter "D" or "N" on one side and a number between 16 and 70 on the other. Your job is to assess whether or not cards satisfy the criterion: "Every 'D' card has a number greater than or equal to 21 printed on the reverse." You are also to separate cards that satisfy this criterion from those that do not.

Write the criterion as an "If..., then..." statement, and determine which of the following cards satisfy the criterion:



(i) (ii) (iii) (iv)

(b) You are shown four cards:

(i) (ii) (iii) (iv)

Which cards must be turned over to determine whether or not they satisfy the criterion of part (a)?

(c) The legal drinking age in a certain state is 21. Your job at a gathering is to ensure that no one under 21 years of age is drinking alcohol, and to report those that are. A group of four people consists of a 20 year old who is drinking, a 46 year old who is drinking, a 16 year old who is not drinking, and a 25 year old who is not drinking. Which of these people is/are violating the law?

After reporting this incident, you find four people at the bar: An 18 year old and a 35 year old with their backs to you, and two people of unknown age, one of whom is drinking. From which people do you need further information to see whether or not they are violating the law?

(d) Explain why the card question is logically equivalent to the drinking question. Which did you find easier to answer correctly?

Chapter 2

An Introduction to Sets

Modern mathematics is built on the concept of a "set", a collection of "elements". These primitive notions will serve in lieu of definitions. This chapter informally introduces the set of complex numbers, connects sets with the basics of logic, and gives advice on constructing and writing mathematical proofs.

2.1 Sets and Set Operations

Example 2.1. The collection of all integers (whole numbers) is a set. Its elements are 0, 1, -1, 2, -2, and so forth. The set of integers is denoted $\mathbf{Z}$, from the German *Zahl* (number). Properties of the integers are developed formally in Chapter 3.

Example 2.2. The collection of "prime numbers", integers p greater than 1 that have no divisors other than 1 and p, is a set. The numbers 2, 5, and $2^{13466917} - 1$ are elements, while 4 and $2^{13466917} = 2 \cdot 2^{13466916}$ are not.

Example 2.3. The set of periodic table entries in the year 1960 has 102 elements. "Hydrogen", "promethium", and "astatine" are elements of this set, while "Massachusetts", "ammonia", and "surprise" are not.

Abstract sets will be denoted with capital letters, such as A or B. Elements are normally denoted with lower case letters, such as a and b. We write "$a \in A$" as shorthand for the statement "a is an element of (the set) A", and "$b \notin A$" for the logical negation "b is not an element of A". For example, $0 \in \mathbf{Z}$, $-7 \in \mathbf{Z}$, and $\frac{1}{2} \notin \mathbf{Z}$.

Definition 2.4. Let A and B be sets. We say A is a *subset* of B, and

write "$A \subseteq B$", if $x \in A$ implies $x \in B$, that is, if every element of A is an element of B. Two sets A and B are *equal* if $A \subseteq B$ and $B \subseteq A$, namely if they have exactly the same elements: $x \in A$ if and only if $x \in B$.

The most basic and explicit way of describing a set is to list its elements. Curly braces are used to denote a list of elements comprising a set. Sets do not "keep track of" what order the elements are listed, or whether their elements are multiply-listed.

Example 2.5. Each of the sets $A = \{-1, 0, 1\}$, $B = \{0, 1, -1\}$, and $C = \{0, 1, 0, -1, 1\}$, contains three elements, and in fact $A = B = C$.

Example 2.6. Let A be a set. For each element a in A, there is a *singleton* set $\{a\}$ contained in A. Take care to distinguish a and $\{a\}$; a is an object, while $\{a\}$ is a "package" that contains exactly one object.

Example 2.7. There exists an *empty set* $\varnothing$ containing *no* elements. For all x, the statement $x \in \varnothing$ is false. In particular, for every set A the logical implication "$x \in \varnothing$ implies $x \in A$" is vacuous (has false hypothesis). Consequently, $\varnothing \subseteq A$ is true for all A.

Remark 2.8. The empty set is unique: If $\varnothing$ and $\varnothing'$ are sets having no elements, then $\varnothing \subseteq \varnothing'$ and $\varnothing' \subseteq \varnothing$ are both true, so $\varnothing = \varnothing'$.

In mathematics, we always restrict attention to sets contained in a fixed set $\mathcal{U}$, called a *universe*. Specific subsets of $\mathcal{U}$ are conveniently described using *set-builder notation*, in which elements are selected according to logical conditions formally known as a *predicates*. The expression $\{x \text{ in } \mathcal{U} : P(x)\}$ is read "the set of all x in $\mathcal{U}$ such that $P(x)$".

Example 2.9. The expression $\{x \text{ in } \mathbf{Z} : x > 0\}$, read "the set of all x in $\mathbf{Z}$ such that $x > 0$", specifies the set $\mathbf{Z}^+$ of *positive integers*.

To personify, if $\mathcal{U}$ is a population whose elements are individuals, then a subset A of $\mathcal{U}$ is a club or organization, and the predicate defining A is a membership card. We screen individuals x for membership in A by checking whether or not x carries the membership card for A, namely whether or not $P(x)$ is true.

Example 2.10. Thanks to *Russell's paradox*, named for the English logician B. Russell, there is no "set $\mathcal{U}$ of all sets". If there were, the set $R = \{x \text{ in } \mathcal{U} : x \notin x\}$ of all sets that are not elements of themselves would have the property that $R \in R$ if and only if $R \notin R$, a contradiction.

Example 2.11. The expression $\{x \text{ in } \mathbf{Z} : x = 2n \text{ for some } n \text{ in } \mathbf{Z}\}$ is the set of *even integers*. We often denote this set $2\mathbf{Z}$, the idea being that the general even integer arises from multiplying some integer by 2.

Similarly, the set of *odd integers* could be expressed as

$$2\mathbf{Z} + 1 = \{x \text{ in } \mathbf{Z} : x = 2n + 1 \text{ for some } n \text{ in } \mathbf{Z}\}.$$

Remark 2.12. For brevity, we sometimes write, e.g., the set of even integers as $\{2n : n \in \mathbf{Z}\}$, read "the set of $2n$ such that n is an element of $\mathbf{Z}$". This way of writing a set is convenient, and the meaning is generally clear, but it isn't technically proper, compare Example 2.10. To define a set formally, first give the universe, then specify the predicate.

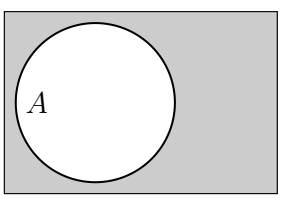
Remark 2.13. The elements of a set may be other sets. For example, the set $A = \{2\mathbf{Z}, 2\mathbf{Z} + 1\}$ has two elements, $2\mathbf{Z}$ and $2\mathbf{Z} + 1$. Note carefully that A is not a subset of $\mathbf{Z}$: The elements of A are not themselves integers, but sets of integers.

Sets and Logic

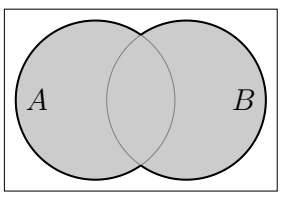
Let $\mathcal{U}$ be a universe, and let A and B be subsets of $\mathcal{U}$. The statements $x \in A$ and $x \in B$ may be viewed as predicates P and Q on elements of $\mathcal{U}$. By definition, the logical implication "$x \in A$ implies $x \in B$" corresponds to the set relation "$A \subseteq B$". Logical negation, disjunction (or), and conjunction (and) similarly have natural interpretations in terms of A and B.

The *complement* of A:	$A^c = \{x \text{ in } \mathcal{U} : x \notin A\}$.
The *union* of A and B:	$A \cup B = \{x \text{ in } \mathcal{U} : x \in A \text{ or } x \in B\}$.
The *intersection* of A and B:	$A \cap B = \{x \text{ in } \mathcal{U} : x \in A \text{ and } x \in B\}$.

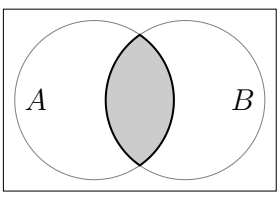
A *Venn diagram* represents subsets of a universe $\mathcal{U}$ pictorially. The universe is depicted as a rectangle, and subsets are disks or, if necessary, more complicated shapes. The complement of A, or the union and intersection of two sets A and B, might be drawn as indicated:



A^c



$A \cup B$



$A \cap B$

Two sets A and B are *disjoint* if $A \cap B = \varnothing$, namely if A and B have no elements in common. A Venn diagram of disjoint sets might be drawn as a pair of non-overlapping disks.

Example 2.14. The sets $2\mathbf{Z}$ and $2\mathbf{Z}+1$ of even and odd integers are disjoint: No integer is both even and odd. The sets $A = 2\mathbf{Z}$ and $B = \mathbf{Z}^+$ are *not* disjoint: For example, 2, 4, and 84 are elements of $A \cap B$, since each is both positive and a multiple of 2.

Definition 2.15. Let A be a set. The *power set* of A, $\mathscr{P}(A)$, is the set of all subsets of A.

Example 2.16. If $A = \{0, 1\}$ has two elements, the power set $\mathscr{P}(A)$ has four elements:

$$\mathscr{P}(A) = \big\{\varnothing, \{0\}, \{1\}, A\big\}.$$

The empty set and A itself are always subsets of A, so a power set is never empty. Indeed, $\mathscr{P}(\varnothing) = \{\varnothing\}$ has a single element.

2.2 Partitions and Mappings

Definition 2.17. Let A be a set, and I a set of indices. A family of subsets $\{A_i\}_{i\in I}$ of A constitutes a *partition* of A if each element of A is an element of *exactly one* of the sets A_i.

In other words, $\{A_i\}_{i\in I}$ is a partition of A if $A_i \cap A_j = \varnothing$ for $i \neq j$ (each pair of sets is disjoint), and A is the union of the sets A_i.

Example 2.18. The sets $A_0 = 2\mathbf{Z}$ and $A_1 = 2\mathbf{Z}+1$ are a partition of $A = \mathbf{Z}$; every integer is either even or odd, and no integer is both. Here the index set is $I = \{0, 1\}$.

The sets $A_0 = 3\mathbf{Z}$, $A_1 = 3\mathbf{Z}+1$, $A_2 = 3\mathbf{Z}+2$ are another partition of $\mathbf{Z}$, since every integer leaves a unique remainder of 0, 1, or 2 upon division by 3:

$\mathbf{Z}$	$\cdots$	-4	-3	-2	-1	0	1	2	3	4	5	6	$\cdots$
A_0	$\cdots$		-3			0			3			6	$\cdots$
A_1	$\cdots$			-2			1			4			$\cdots$
A_2	$\cdots$	-4			-1			2			5		$\cdots$

Example 2.19. We will prove in Chapter 4 (Theorem 4.8) that if $n > 1$ is an integer, there is a partition of $\mathbf{Z}$ into n subsets, $A_k = n\mathbf{Z} + k$ with $k = 0, \ldots, n-1$ an integer. An integer x is an element of A_k if and only if x leaves a remainder of k on division by n.

In Chapter 6, we will write $[k]_n = n\mathbf{Z} + k$, and form a set $\mathbf{Z}_n$ having n elements: $\mathbf{Z}_n = \{[0]_n, [1]_n, \ldots, [n-1]_n\}$. Note that $\mathbf{Z}_n \subseteq \mathscr{P}(\mathbf{Z})$: The *elements* of $\mathbf{Z}_n$ are *subsets* of $\mathbf{Z}$.

Mappings

A "mapping" is a mathematical structure formalizing a particularly useful relationship between elements of two sets.

Definition 2.20. Let A and B be sets. Their *Cartesian product* $A \times B$ is the set of all "ordered pairs" from A and B,

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

Example 2.21. If $A = \{a, b, c\}$ and $B = \{0, 1\}$, then $A \times B$ is the six-element set $\{(a, 0), (b, 0), (c, 0), (a, 1), (b, 1), (c, 1)\}$ in the left-hand diagram in Figure 2.1.

For the same set B, $B \times B = \{(0, 0), (1, 0), (0, 1), (1, 1)\}$.

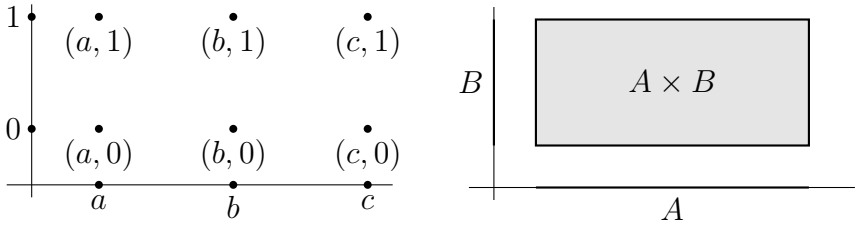


Figure 2.1: Cartesian products.

Example 2.22. If $A = \varnothing$ or $B = \varnothing$, then $A \times B = \varnothing$.

An abstract Cartesian product can be visualized conveniently by depicting the set A on a horizontal axis and the set B on a vertical axis, and taking the set of points lying above or below A *and* to the left or right of B. The right-hand diagram in Figure 2.1 shows the case where A and B are intervals of real numbers.

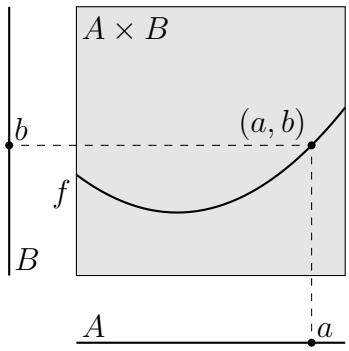
Definition 2.23. A *mapping* $f : A \to B$, read "f from A to B", is a subset f of $A \times B$ satisfying the following condition:

For every a in A, there exists a unique b in B such that $(a, b) \in f$.

The set A is the *domain* of f, and B is the *codomain* of f.

If $(a, b) \in f$, we write $b = f(a)$, and call b the *value* of f at a. We also say a is *mapped to* b by f, or that f *maps* a to b.

Remark 2.24. Conceptually, a mapping $f : A \to B$ associates a unique value b in the codomain to each element a of the domain. If the Cartesian product $A \times B$ is viewed as a rectangle, a mapping is a "graph" in the sense of calculus, namely a subset intersecting each vertical line in the rectangle exactly once. The vertical line at horizontal position a intersects f (a.k.a. the graph of f) at height $b = f(a)$.



Remark 2.25. If B is arbitrary (empty or not), there is a unique mapping $f : \varnothing \to B$, namely the empty set.

If A is non-empty, there exists no mapping $f : A \to \varnothing$. (Why?)

2.3 Advice on Writing Proofs

A mathematical theorem is an idealized contract: *If* certain conditions are met (the hypotheses), *then* other conditions are guaranteed (the conclusion). No theorem has even one exception. To bolster this implicit guarantee of inviolable correctness, a mathematical proof must leave no logical possibility unexamined, no contingency unresolved.

Often, a mathematical theorem (such as an algebraic identity) makes infinitely many assertions, one for each assignment of values to variables. In these circumstances, a proof usually works with variables as symbols; any particular set of values in a theorem statement implicitly expands to a set of values throughout the proof.

Discovering and writing proofs are nearly opposite activities. To discover a proof (and before that, to guess what is true, or *formulate a conjecture*), you must permit yourself to make partial, open-ended, possibly unjustified guesses and see where they lead. Ultimately, you'll

find most of the writing you do in discovering mathematics does not need to be written up; it's just "scaffolding".

Example 2.26. Assume a and b are integers. Prove that

$$(a+b)^2 = (a-b)^2 + 4ab.$$

(Preliminary Work). When proving an identity such as this we have an obvious strategy: Express each side in terms of simpler information and see if the answers agree. Here, it's natural to start by expanding each side:

$$(a+b)^2 = a^2 + 2ab + b^2, \qquad (a-b)^2 + 4ab = (a^2 - 2ab + b^2) + 4ab.$$

These are indeed equal.

(The Written Solution). Assume a and b are integers. Prove that $(a+b)^2 = (a-b)^2 + 4ab$.

Proof: Let a and b be arbitrary integers. We have

$$(a-b)^2 + 4ab = (a^2 - 2ab + b^2) + 4ab = a^2 + 2ab + b^2 = (a+b)^2,$$

as was to be shown.

Remark 2.27. When writing up a formal proof of an algebraic identity Q, the preferred style is to build a chain of equalities from one side to the other. When possible, start with the "more complicated" side and simplify. There is no need to perform only steps that would naturally occur to the reader, however.

Do not write down the desired conclusion Q, then manipulate each side until you have an identity P. At best, this "two-column" argument establishes the converse, Q implies P, which is not equivalent to P implies Q, and does not even imply the truth of Q. See Exercises 2.17 and 2.18 for pitfalls of the "two-column" style of proof.

Example 2.28. Assume a and b are integers. Prove that $a^2 = b^2$ if and only if $a = b$ or $a = -b$.

(Preliminary Work). The condition $a^2 = b^2$ is equivalent to the condition $a^2 - b^2 = 0$. Further, the left-hand side factors as a difference of squares: $a^2 - b^2 = (a-b)(a+b)$. Finally, if a product of integers is zero, then one factor or the other (or both) is zero. Here, we deduce that $a - b = 0$ (i.e., $a = b$) or $a + b = 0$ (i.e., $a = -b$).

We give two write-ups. In the first, each implication ("if" and "only if") is proven separately. This is the default style. In general, the proof of a statement and its converse differ substantially. Separating the two proofs clarifies the argument and helps ensure all the necessary details are addressed.

A second proof can be given because in this particular situation, there is a chain of "if and only if" statements leading from hypothesis to conclusion.

(The Written Solution). Assume a and b are integers. Prove that $a^2 = b^2$ if and only if $a = b$ or $a = -b$.

Proof 1: Let a and b be integers.

If $a^2 = b^2$, then $(a-b)(a+b) = a^2 - b^2 = 0$. Since a product of non-zero integers is non-zero, either $a - b = 0$ or $a + b = 0$. That is, either $a = b$ or $a = -b$.

Conversely, suppose $a = b$ or $a = -b$. If $a = b$, then $a^2 = b^2$ by substitution. If $a = -b$, then $a^2 = (-b)^2 = b^2$. In each case, $a^2 = b^2$.

Proof 2: Let a and b be integers. Since $(a-b)(a+b) = a^2 - b^2$ for all integers a and b, we have

$$\begin{aligned} a^2 = b^2 \quad & \text{if and only if} \quad (a-b)(a+b) = a^2 - b^2 = 0, \\ & \text{if and only if} \quad a - b = 0 \text{ or } a + b = 0, \\ & \text{if and only if} \quad a = b \text{ or } a = -b. \end{aligned}$$

Example 2.29. Prove or disprove: $2\mathbf{Z} + 1 = 2\mathbf{Z} - 1$.

(Preliminary Work). By definition of equality of sets, we are to determine whether each set is a subset of the other. Some initial formalization can be performed mechanically. Give each set a name, write down its definition, and express the question in terms of this framework.

Here, we have two sets of integers,

$$\begin{aligned} A = 2\mathbf{Z} + 1 &= \{x \text{ in } \mathbf{Z} : x = 2u + 1 \text{ for some } u \text{ in } \mathbf{Z}\}, \\ B = 2\mathbf{Z} - 1 &= \{y \text{ in } \mathbf{Z} : y = 2v - 1 \text{ for some } v \text{ in } \mathbf{Z}\}. \end{aligned}$$

We wish to show either that $A \subseteq B$ and $B \subseteq A$ (which by definition means $A = B$ as sets), or that at least one of these inclusions is false.

Next, try to determine intuitively whether or not the statement is false (which can be shown by exhibiting a counterexample, an element of one set that is not an element of the other set) or true. To get an element of $2\mathbf{Z}+1$, add 1 to an even integer: $1 = 0+1$, $3 = 2+1$, $5 = 4+1$,

$-1 = -2 + 1$, and so forth, are elements. Similarly, subtracting 1 from an even integer gives an element of $2\mathbf{Z} - 1$: $-1 = 0 - 1$, $1 = 2 - 1$, $3 = 4 - 1$, $-3 = -2 - 1$, and so forth, are elements.

This evidence doesn't merely suggest the two sets *are* equal, it even points to a strategy of proof: Any integer one greater than an even integer is one less than the next largest even integer. We'll sketch out an informal proof to settle notation and iron out any unforeseen logical wrinkles.

The statement "$A \subseteq B$" may be phrased "if $x \in A$, then $x \in B$". If $x \in A$, then by the definition of A there exists an integer u such that $x = 2u + 1 = 2(u + 1) - 1$. Setting $v = u + 1$ (an integer because u is), we see x has the form $2v - 1$ for some integer v, which by definition means $x \in B$. This shows $A \subseteq B$.

The inclusion $B \subseteq A$ is entirely similar, so at this stage we can write up a formal proof. The considerations above that led to the proof are customarily omitted from the formal write-up. Note, however, that the proof involves choices not easily known ahead of time; the scratch work is important!

(The Written Solution). Show $2\mathbf{Z} + 1 = 2\mathbf{Z} - 1$.

Proof: By definition, $A = \{x \text{ in } \mathbf{Z} : x = 2u + 1 \text{ for some } u \text{ in } \mathbf{Z}\}$ and $B = \{y \text{ in } \mathbf{Z} : y = 2v - 1 \text{ for some } v \text{ in } \mathbf{Z}\}$. Assume $x \in A$. By hypothesis, there exists an integer u such that $x = 2u+1$. Let $v = u+1$, so $u = v - 1$, and note v is an integer. Since

$$x = 2u + 1 = 2(v - 1) + 1 = 2v - 2 + 1 = 2v - 1,$$

$x \in B$. Since x was arbitrary (i.e., x could have been any element of A), we have shown $A \subseteq B$.

Conversely, suppose $y = 2v-1 \in B$ for some integer v. Let $u = v-1$, so that $v = u + 1$. Then

$$y = 2v - 1 = 2(u + 1) - 1 = 2u + 1,$$

so $y \in A$. Since y was arbitrary, we have shown $B \subseteq A$.

Since $A \subseteq B$ and $B \subseteq A$, we have $A = B$.

Writing proofs requires practice. The final result should be a coherent, logical, step-by-step argument starting with the given hypotheses and leading to the conclusion.

Example 2.30. Let A and B be subsets of $\mathcal{U}$. Find the most general conditions on A and B under which $A \cap B = A$.

(Examples). If you're comfortable with sets and operations, go for the frontal assault ("reducing to the definitions", below). Otherwise, proceed by writing out examples on scratch paper or a blackboard. If Venn diagrams are more natural, use those. If concrete sets are easier to think about, use those. At this stage it's all right to let $\mathcal{U} = \mathbf{Z}$, the set of integers, but in the final proof, do not make any assumptions on the nature of $\mathcal{U}$, A, or B.

(Simpler cases). Since the target condition involves two sets, we can reduce to a simpler question by "fixing" one set and letting the other set vary.

If $A = \varnothing$, then $A \cap B = \varnothing \cap B = \varnothing = A$ regardless of B. If $A = \mathcal{U}$, then $A \cap B = \mathcal{U} \cap B = B$, which is not equal to A unless $B = \mathcal{U}$.

These examples show the condition $A \cap B = A$ *can* be true, but is *not always* true. The guiding task is to discover what common aspect these examples possess. If you're still not sure, draw a Venn diagram with a circle representing A, and ask: What condition on B guarantees that $A \subseteq A \cap B$? Draw circles that are disjoint from A, that are contained in A, that partially overlap A, or that contain A. The evidence of this "experiment" should point toward the desired condition.

(Reducing to the definitions). The condition $A \cap B = A$ encapsulates two set inclusions, $A \cap B \subseteq A$ and $A \subseteq A \cap B$. The first inclusion is true for all pairs of sets: If $a \in A \cap B$, then $a \in A$ and $a \in B$, so perforce $a \in A$. Since a is an arbitrary element of $A \cap B$, this argument shows $A \cap B \subseteq A$.

We are therefore seeking the most general conditions under which $A \subseteq A \cap B$, namely, "$a \in A$ implies '$a \in A$ and $a \in B$' ". Clearly, this is equivalent to "$a \in A$ implies $a \in B$," which may be rephrased as $A \subseteq B$. This condition is our supposed conclusion, or conjecture.

As a consistency check, recall $A = \varnothing$ and $A = \mathcal{U} = B$ satisfied the condition. In each case, $A \subseteq B$ holds. If the purported abstract condition is violated by examples, it's definitely wrong.

(Supposed conclusion). As the result of considerations above, we claim that $A \cap B = A$ if and only if $A \subseteq B$. To *prove* this formally, it suffices to establish two logical implications:

$$A \cap B = A \text{ implies } A \subseteq B, \qquad A \subseteq B \text{ implies } A \cap B = A.$$

Here, approximately, is what you'd normally write up:

(The Written Solution). $A \cap B = A$ if and only if $A \subseteq B$.

Proof: ($A \cap B = A$ implies $A \subseteq B$) Assume $A \cap B = A$, namely $A \cap B \subseteq A$ and $A \subseteq A \cap B$. Since the first inclusion holds for all sets, our initial hypothesis is equivalent to $A \subseteq A \cap B$.

Let a be an arbitrary element of A. Since $A \subseteq A \cap B$ by hypothesis, $a \in A \cap B$, so $a \in A$ and $a \in B$. In particular, $a \in B$. We have shown that if $a \in A$, then $a \in B$; this means that $A \subseteq B$, as was to be shown.

($A \subseteq B$ implies $A \cap B = A$) By hypothesis, if $a \in A$, then $a \in B$, so if $a \in A$, then $a \in A$ and $a \in B$. Since a is arbitrary we have $A \subseteq A \cap B$. The reverse inclusion $A \cap B \subseteq A$ holds for all sets A and B. We have shown that if $A \subseteq B$, then $A \cap B = A$. This completes the proof.

Find your own writing style. *Do write accurately and precisely*, but don't be pedantic or excessively wordy. Short, declarative sentences expressing one idea are a good general rule.

Avoid pronouns, especially "it". In the middle of even a simple proof, two or three objects tend to be under consideration, and "it" can often refer to any of them. If you're unable to decide exactly what "it" refers to, you've located something you don't fully understand.

Though it may feel awkward at first, read your solution aloud, either to yourself or someone else. Listening engages different parts of the brain than writing. Lapses of grammar, narrative continuity, and logic are usually more obvious when heard than when read, if only because the more times you re-read your own proof, the more you skim.

Exercises

Exercise 2.1. Let $A = 2\mathbf{Z}$ and $B = 3\mathbf{Z}$.

(a) Find $A \cap B$; that is, determine which integers are in $A \cap B$.

(b) List the elements of $A \cup B$ between -12 and 12.

Exercise 2.2. Prove or disprove:

(a) $3\mathbf{Z} \subseteq 2\mathbf{Z}$. (b) $4\mathbf{Z} \subseteq 2\mathbf{Z}$. (c) $2\mathbf{Z} \subseteq 4\mathbf{Z}$.

Exercise 2.3. Prove or disprove:

(a) $2\mathbf{Z} \cup 3\mathbf{Z} \subseteq 5\mathbf{Z}$. (b) $5\mathbf{Z} \subseteq 2\mathbf{Z} \cup 3\mathbf{Z}$. (c) $8\mathbf{Z} = 2\mathbf{Z} \cap 4\mathbf{Z}$.

Exercise 2.4. Prove or disprove:

(a) $3\mathbf{Z} + 1 \subseteq 2\mathbf{Z}$. (b) $3\mathbf{Z} + 1 \subseteq 4\mathbf{Z}$. (c) $3\mathbf{Z} + 2 = 3\mathbf{Z} - 1$.

Exercise 2.5. Let A be a set and assume $a \in A$. Determine whether each condition is always true, sometimes true, or never true. If sometimes true, give examples of A and/or a for which the condition is true or is false.
(a) $a \in \{a\}$ (b) $a \subseteq A$ (c) $\{a\} \subseteq \varnothing$ (d) $\varnothing \in A$ (e) $\{a\} \in A$

Exercise 2.6. Let A and B be arbitrary subsets of a universe $\mathcal{U}$.

(a) Prove $A \cup B = A$ if and only if $B \subseteq A$.

(b) Prove $A \cap B = B$ if and only if $B \subseteq A$.

Exercise 2.7. Let A and B be subsets of $\mathcal{U}$.

(a) Prove $A \subseteq B$ if and only if $B^c \subseteq A^c$, and illustrate with a Venn diagram.

(b) How is part (a) related to contrapositives?

Exercise 2.8. Let A, B, and C be subsets of a universe $\mathcal{U}$, and let P, Q, and R be the predicates $x \in A$, $x \in B$, and $x \in C$. Use truth tables to establish the indicated identities.

(a) $(A \cup B) \cup C = A \cup (B \cup C)$.

(b) $(A \cap B) \cap C = A \cap (B \cap C)$.

Exercise 2.9. Let A, B and C be subsets of a universe $\mathcal{U}$. As in Exercise 2.8, use truth tables to establish *De Morgan's laws* (a) and (b) and the *distributive laws* (c) and (d).

(a) $(A \cup B)^c = A^c \cap B^c$.

(b) $(A \cap B)^c = A^c \cup B^c$.

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

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

Exercise 2.10. Draw Venn diagrams illustrating each part of the preceding exercise, and compare with Exercise 1.3.

Exercise 2.11. Let A and B be subsets of $\mathcal{U}$. Their *difference* is defined to be $A \setminus B = \{x \text{ in } A : x \notin B\}$.

(a) Prove $A \setminus B = A \cap B^c$, and illustrate with a Venn diagram.

(b) List the elements of $\mathbf{Z} \setminus \mathbf{Z}^+$ between -5 and 5.

(c) List the elements of $2\mathbf{Z} \setminus 3\mathbf{Z}$ between -12 and 12.

(d) List the elements of $3\mathbf{Z} \setminus 2\mathbf{Z}$ between -12 and 12.

Exercise 2.12. Let A and B be subsets of $\mathcal{U}$. Their *symmetric difference* is defined to be $A \triangle B = (A \setminus B) \cup (B \setminus A)$.

(a) Prove $A \triangle B = (A \cup B) \setminus (A \cap B)$ and illustrate with a Venn diagram.

(b) Prove $A \triangle B = \{x \text{ in } \mathcal{U} : x \in A \text{ or } x \in B \text{ but not both}\}$. This condition is called *exclusive or*, denoted "xor".

Exercise 2.13. (a) Let $A = \{a, b, c\}$ be a set with three distinct elements. List the elements of the power set $\mathscr{P}(A)$.

(b) How would your answer to part (a) differ if $A = \{0, 1, 2\}$?

(c) Describe how you could use your answer to part (a) to list the elements of the power set of $A' = \{a, b, c, d\}$. Suggestion: There are two types of subset of A', those having d as an element, and those not having d as an element.

Exercise 2.14. Let A and B be subsets of $\mathcal{U}$.

(a) Suppose $A \subseteq B$. Prove $\mathscr{P}(A) \subseteq \mathscr{P}(B)$ as subsets of $\mathscr{P}(\mathcal{U})$.

(b) Suppose that $\mathscr{P}(A) = \mathscr{P}(B)$ as subsets of $\mathscr{P}(\mathcal{U})$. Prove $A = B$.

Exercise 2.15. Let A be an arbitrary set, and let $f : A \to \mathscr{P}(A)$ be an arbitrary mapping. (Procedurally, f associates some *subset* of A to each *element* of A.) Define

$$X = \{a \text{ in } A : a \notin f(a)\} \subseteq A.$$

Show that if ZFC is consistent, then there exists no x in A such that $f(x) = X$.

Exercise 2.16. The Infinity Gambling Supplies Company has an infinite set of Keno balls, labeled serially with the positive integers. At 11 PM on December 31, balls numbered 1–10 are added to an infinitely large bin, and the ball labeled 1 is removed. At 11:30, balls numbered 11–20 are added, and ball 2 is removed. At 11:45, balls 21–30 are added and ball 3 is removed. The pattern continues in Zeno-like fashion.

At midnight, how many balls are in the bin? Hint: What number(s) are never removed?

Exercise 2.17. Explain in detail what is wrong with this two-column "proof" that $-1 = 1$.

$-1 = 1$	to be shown,
$(-1)^2 = 1^2$	square both sides,
$1 = 1$	true statement.

Therefore $-1 = 1$.

Exercise 2.18. Let a and b denote real numbers, and assume $a = b$.

(a) What is wrong with the following "proof" that $2 = 1$?

$b^2 = ab$	Multiply $b = a$ by b,
$b^2 - a^2 = ab - a^2$	subtract a^2,
$(b + a)(b - a) = a(b - a)$	factor each side,
$(b + a) = a$	cancel common factor,
$2a = a$	$a = b$,
$2 = 1$	cancel common factor.

(b) If the proof is read from bottom to top, is each step valid?

Part II

Discrete Structures

Chapter 3

Natural and Whole Numbers

Counting is surely the oldest mathematical activity discussed in this book, existing in some form even in non-human animals. The act of counting is so basic and instinctive that the profound reasons for its existence may be difficult to see: We live in surroundings containing distinct objects that are similar enough to group conceptually.

Our mathematical survey begins in the familiar territory of counting, but quickly adopts the viewpoint of modern mathematics, with its idealized axioms and formal definitions.

Imagine a shepherd of three thousand years ago. Each morning he lets his sheep out to pasture, and each evening brings them in. How can he be sure he hasn't lost sheep during the day?

One can imagine a scheme: The shepherd gathers a supply of small stones. In the morning, as each sheep passes, he places one of the stones aside. Once all the sheep have left, he puts these stones into a pouch for safe keeping. That evening, he takes a stone from the pouch for each returning sheep. If stones remain in the pouch, he has lost sheep.

The apocryphal shepherd has made a fundamental abstraction about the physical world: There is a meaningful notion of *counting*, or *a number of things*—sheep, or stones, or sunrises, or nicks on a tree branch. By keeping track of small stones (or *calculi* as they will be known in Latin many centuries in the shepherd's future), the shepherd can keep track of his flock.

No matter how many sheep or stones one has, even if the number be like unto the grains of sand on all the beaches of the world, one can imagine having one more.

Two shepherds can combine their flocks, and *calculate* how many

sheep they have between them simply by agglomerating their piles of stones. They quickly discover it does not matter whose stones are appended to whose, the result is the same either way. They have discovered the operation of *addition*, and the *commutative law*.

Two shepherds can compare their flocks to see who has more sheep. No matter whose flocks are compared, one flock is larger than the other, or the flocks are the same size; absolute comparison is always possible. Moreover, if Aleph has more sheep than Beth, who has more sheep than Gimel, then Aleph has a larger flock than Gimel. They have discovered the *ordering* of the natural numbers, and the *transitive law*.

3.1 The Peano Axioms

The *natural numbers* represent possible answers to "how many objects": How many sheep in a flock; how many stones in a pile; how many unicorns over the rainbow; how many characters (typographical symbols) on a page; how many distinct 250-page books with 1500 characters per page, each character chosen from an alphabet of 100 letters, digits, and punctuation marks.

Following modern mathematical custom, we do not say *what natural numbers are*, but instead specify *how natural numbers behave*:

There exists a set $\mathbf{N}$ called the *set of natural numbers*, a notion of *successorship*, and an *initial element* 0 in $\mathbf{N}$, such that:

N1. Every natural number n has a unique natural number $S(n)$ as successor.

N2. For every natural number $n \neq 0$, there exists a unique *predecessor*, a natural number m such that $n = S(m)$. The natural number 0 has no predecessor.

N3. If L is a collection of natural numbers such that 0 is in L, and the successor $S(n)$ is in L for every n in L, than $L = \mathbf{N}$.

For the most part, we view Properties N1–N3 as *axioms*, unquestioned properties whose logical consistency is assured. The rest of this section introduces the basic operations and technical tools in the natural numbers: induction and recursion, comparison, addition and sub-

traction, and multiplication. Section 3.4 contains a sketched construction of the natural numbers, proving that these axioms are as logically consistent as ZFC, and proves the properties stated below.

Definition 3.1. The successor $S(0)$ of 0 is *one*, denoted 1. If k is a natural number, we denote its successor by $S(k) = k + 1$.

Suppose we have a family of statements $P(m)$, one for each natural number m. Let L be the set of natural numbers m such that $P(m)$ is true. To prove that every statement in our family is true, it suffices to establish:

(i) The Base case: $P(0)$ is true, i.e., $0 \in L$;

(ii) The inductive step: For every natural number k, $P(k)$ implies $P(k+1)$, i.e., $k \in L$ implies $k + 1 \in L$.

Property N3 then guarantees $L = \mathbf{N}$, namely that $P(m)$ is true for every natural number m. This foundational proof technique, *mathematical induction*, is systematically developed in Section 3.2.

Our recursive definition of addition formalizes the process of agglomerating heaps of stones. Intuitively, start with natural numbers m and n, and simultaneously replace m with its predecessor and n with its successor "as many times as possible", namely, move stones one at a time from the first pile to the second. This process terminates in finitely many steps, at the end of which the second pile contains a number of stones equal to the sum of m and n.

Definition 3.2. Let m and n be natural numbers. If $m = 0$, define $n + m = n$. Generally, define

$$(*) \qquad n + (m + 1) = (n + m) + 1 \qquad \text{for all } n.$$

That is, if m is a natural number, define $n + S(m) = S(n + m)$ for all n.

Remark 3.3. For each natural number m, let $P(m)$ denote the statement "$n + m$ is defined for all n", and let L be the set of natural numbers m such that $P(m)$ is true. We claim that $L = \mathbf{N}$, namely that $m + n$ is defined for all natural numbers m and n.

The definition $n + 0 = n$ for all n acts as a base case ($0 \in L$). The *recursion relation* (*) gives, for all n, a definition of $n + (m + 1)$ in terms of $n + m$. That is, (*) acts as an inductive step ($m \in L$ implies $m + 1 \in L$). By Property N3, $L = \mathbf{N}$.

Theorem 3.4. *Let k, m, and n denote arbitrary natural numbers.*

(i) *Addition is associative:* $n + (m + k) = (n + m) + k$.

(ii) *Addition is commutative:* $n + m = m + n$.

Remark 3.5. Addition is associative on the set of natural numbers. A sum of three or more natural numbers may therefore be written unambiguously without parentheses: Any two groupings of summands has the same sum. A careful proof could be given now, but for pedagogical reasons is deferred.

Theorem 3.6. *Let k, m, and n denote arbitrary natural numbers. There exists a unique ordering $<$ on the set of natural numbers satisfying:*

(i) $n < S(n)$ *for every* n;

(ii) *If* $k < m$ *and* $m < n$, *then* $k < n$;

(iii) *Exactly one of the following holds:* $m < n$, $m = n$, *or* $n < m$.

Further, $m < n$ if and only if there exists a non-zero natural number k such that $m + k = n$.

Definition 3.7. If $m \leq n$ are natural numbers, their *difference* $n - m$ is the unique natural number satisfying $m + (n - m) = n$.

Theorem 3.8. *If X is a non-empty set of natural numbers, there exists a least element, i.e., a natural number m such that $m \leq x$ for all x in X.*

Definition 3.9. Let m and n be natural numbers. If $m = 0$, define $n \times 0 = 0$. Generally, define

$$n \times (m + 1) = (n \times m) + n \quad \text{for all } n.$$

Theorem 3.10. *Let n, m, and k be natural numbers.*

(i) $n \times (m \times k) = (n \times m) \times k$.

(ii) $n \times m = m \times n$.

(iii) $(n + m) \times k = (n \times k) + (m \times k)$.

(iv) *If* $m \leq n$, *then* $(n - m) \times k = (n \times k) - (m \times k)$.

Remark 3.11. Properties (iii) and (iv) are the *distributive laws*; we say multiplication distributes over addition or subtraction. The corresponding identities with multiplication on the left are also true, since multiplication is commutative.

3.2 Applications of Induction

The natural numbers provide a framework for *recursive definition* and *mathematical induction*. Loosely, a function f on the set of natural numbers is recursive if $f(0)$ is defined explicitly, and for each k, the value $f(k+1)$ is defined in terms of $f(k)$. Addition and multiplication were defined recursively, and their properties established by mathematical induction. This section introduces additional examples, and develops induction as a general proof technique.

Definition 3.12. Let m and n be natural numbers. If $m = 0$, define $n^0 = 1$. If m is an arbitrary natural number, define

$$n^{m+1} = (n^m) \times n \quad \text{for all } n.$$

Remark 3.13. Informally, n^m is the product of m factors of n. Note carefully that we *define* $0^0 = 1$. This definition works "as expected" in the discrete parts of mathematics, and in some parts of calculus, particularly the theory of power series. In the theory of limits of functions of two variables, this definition entails no greater possibility confusion than leaving 0^0 undefined.

Theorem 3.14. *If n, m, and k are arbitrary natural numbers, then:*

(i) $n^{m+k} = (n^m) \times (n^k)$.

(ii) $n^{m \times k} = (n^m)^k$.

Remark 3.15. Unlike addition and multiplication, exponentiation is neither associative nor commutative. The expression n^{m^k} implicitly means $n^{(m^k)}$. Can you see why?

Example 3.16. Expanding the definition for $a = 2$, we have

$$2^3 = 2^2 \cdot 2 = 2^1 \cdot 2 \cdot 2 = 2^0 \cdot 2 \cdot 2 \cdot 2 = 1 \cdot 2 \cdot 2 \cdot 2 = 2 \cdot 2 \cdot 2 = 8.$$

After $2^0 = 1$, the next twenty powers of 2 are:

$2^1 = 2$	$2^6 = 64$	$2^{11} = 2048$	$2^{16} = 65,536$
$2^2 = 4$	$2^7 = 128$	$2^{12} = 4096$	$2^{17} = 131,072$
$2^3 = 8$	$2^8 = 256$	$2^{13} = 8192$	$2^{18} = 262,144$
$2^4 = 16$	$2^9 = 512$	$2^{14} = 16,384$	$2^{19} = 524,288$
$2^5 = 32$	$2^{10} = 1024$	$2^{15} = 32,768$	$2^{20} = 1,048,576.$

Note that $2^{10} = 1024 \approx 1000 = 10^3$ and $2^{20} = 1,048,576 \approx 10^6$.

Definition 3.17. If n is a natural number, the *factorial* of n, denoted $n!$, is defined recursively by

$$0! = 1, \qquad (n+1)! = (n+1) \cdot n! \quad n \geq 0.$$

Example 3.18. Expanding the recursive definition for $n = 3$,

$$3! = 3 \cdot 2! = 3 \cdot 2 \cdot 1! = 3 \cdot 2 \cdot 1 \cdot 0! = 3 \cdot 2 \cdot 1 \cdot 1 = 6.$$

Informally, $m! = m(m-1)(m-2)\cdots 3 \cdot 2 \cdot 1$ is the product of the natural numbers between 1 and m (if $m > 0$). The first several factorials are:

$0! = 1$	$4! = 24$	$8! = 40,320$
$1! = 1$	$5! = 120$	$9! = 362,880$
$2! = 2$	$6! = 720$	$10! = 3,628,800$
$3! = 6$	$7! = 5040$	$11! = 39,916,800.$

Definition 3.19. An ordered list of natural numbers,

$$(b_k)_{k=0}^{\infty} = (b_0, b_1, b_2, \ldots, b_n, \ldots),$$

is called a *sequence*.

The sequence of *partial sums* of the sequence $(b_k)_{k=0}^{\infty}$ is the sequence $(s_n)_{n=0}^{\infty}$ defined recursively by $s_0 = b_0$, and

$$s_{n+1} = s_n + b_{n+1}.$$

Remark 3.20. Informally, $s_n = b_0 + b_1 + b_2 + \cdots + b_n$. For example, repeated application of the recursion rule gives

$$\begin{aligned} s_3 &= s_2 + b_3 = (s_1 + b_2) + b_3 = s_1 + (b_2 + b_3) \\ &= (s_0 + b_1) + (b_2 + b_3) = s_0 + (b_1 + b_2 + b_3) \\ &= b_0 + b_1 + b_2 + b_3. \end{aligned}$$

Partial sums arise often enough to get special notation:

$$s_n = \sum_{k=0}^{n} b_k,$$

read "the sum from $k = 0$ to n of b_k". The $\sum$ sign is Sigma, the Greek letter S, for *sum*.

Inductive Proof

Suppose someone tells you the sum of the first n positive odd numbers is equal to n^2. What basis do you have for believing this claim?

As a start, you might verify a few instances by hand. For example, $1 + 3 + 5 = 9 = 3^2$, so the claim is true when $n = 3$. Perhaps skeptical, you add the first ten odd numbers, or the first twenty, each time verifying the claim. Perhaps you are starting to believe.

Logically, however, testing special cases leaves you no closer to complete certainty. Have you tried adding the first hundred thousand odd numbers? The first billion? Finding a single counterexample would prove the claim false, but no matter how many cases you verify, there remain infinitely many unverified cases.

Mathematical induction allows us to resolve such questions with a finite proof. The idea is to break the statement "For every natural number n, the sum of the first n odd positive numbers is equal to n^2" into an infinite list of statements. Here, we take

$$P(n) \qquad 1 + 3 + 5 + \cdots + (2n - 1) = n^2.$$

To say $P(100)$ is true, for example, means the sum of the first hundred odd numbers is equal to $10,000 = 100^2$. (An "empty" sum is 0, so $P(0)$ reads $0 = 0$.)

The original statement may be rephrased "For every natural number n, $P(n)$ is true." This single statement P encapsulates the infinite list of statements: $P(0)$ is true, $P(1)$ is true, $P(2)$ is true, etc.

In order to establish the truth of P, it suffices to prove $P(0)$ is true (the *base case*), and to prove $P(k)$ implies $P(k + 1)$ for every k in $\mathbf{N}$ (the *inductive step*).

Remark 3.21. It is sometimes convenient to take an index $n_0 > 0$ for the base case. In this event, one must prove $P(n_0)$ is true, and establish that $P(k)$ implies $P(k + 1)$ for $k \geq n_0$. The conclusion is that $P(n)$ is true for all $n \geq n_0$.

To see intuitively why the base case and inductive step are enough, consider the consequences of "$P(0)$ is true, and $P(k)$ implies $P(k + 1)$ for all $k \geq 0$". Taking $k = 0$, the inductive step says $P(0)$ implies $P(1)$. But $P(0)$ is *true* by the base case, so $P(1)$ is also *true* by the inductive step. Now repeat the argument, taking $k = 1$. By the inductive step, $P(1)$ implies $P(2)$, but $P(1)$ is *true*, so $P(2)$ is also true. Continuing in this fashion, $P(3)$ is true, and $P(4)$, and so forth, *ad infinitum*. The

chain of deduction may be represented as a sequence of arrows:

$$\underbrace{P(0)}_{\text{Base case}} \implies P(1) \implies P(2) \implies P(3) \implies \dots$$
$$\implies P(k) \implies P(k+1) \implies \dots$$

Example 3.22. Let's see how induction works in practice, with $P(n)$ as above. As noted earlier, the base case $P(0)$ reads $0 = 0$ because an empty sum is 0. (If this seems suspicious, note that the first odd positive number is equal to 1^2, so $P(1)$ is also true).

Next, assume inductively that $P(k)$ is true for some fixed (but arbitrary) natural number k. The sum of the first $(k+1)$ odd positive numbers is equal to the sum of the first k plus the $(k+1)$th. By hypothesis, the sum of the first k is equal to k^2. We therefore deduce

$$\underbrace{1+3+5+\cdots+(2k-1)}_{=k^2 \text{ by } P(k)}+(2k+1) = k^2 + (2k+1) = (k+1)^2$$

by algebra. This equation says the sum of the first $(k+1)$ odd positive numbers is equal to $(k+1)^2$. By assuming $P(k)$, we proved $P(k+1)$.

To summarize, the base case $P(0)$ is true, and the inductive step, $P(k)$ implies $P(k+1)$, is valid for each natural number k. By mathematical induction, $P(n)$ is true for all $n \geq 0$.

Remark 3.23. Our use of n or k to denote an arbitrary natural number in an inductive proof signifies a subtle but important distinction. In this book, $P(n)$ refers to the general statement of an inductive list, whose truth value is to be established. By contrast, $P(k)$ refers to a general statement that is "inductively true": We assume "for the sake of argument" that $P(k)$ is true for some (particular but arbitrary) k, and try to deduce $P(k+1)$.

Example 3.24. For each natural number n,

$$1 + \sum_{k=0}^{n} 2^k = 2^{n+1}.$$

Call the preceding equation $P(n)$. The base case $P(0)$ reads $1+2^0 = 2^1$, or $1+1=2$, which is true.

Assuming inductively that $P(n)$ is true for some natural number n,

$$\begin{aligned}
1 + \sum_{k=0}^{n+1} 2^k &= \left[1 + \sum_{k=0}^{n} 2^k\right] + 2^{n+1} && \text{Recursion rule for summation,} \\
&= 2^{n+1} + 2^{n+1} && \text{Inductive hypothesis,} \\
&= 2^{n+1} \cdot 2 && \text{Distributivity, } 1 + 1 = 2, \\
&= 2^{(n+1)+1} && \text{Definition of exponentiation.}
\end{aligned}$$

That is, $P(n)$ implies $P(n+1)$ for each natural number n. By induction, $P(n)$ is true for all n.

Example 3.25. Consider the statement, "For all $n \geq 1$, $n^2 + n + 41$ is prime." Checking cases may convince you this statement is true. Taking $n = 2$, 5, 20, and 100 respectively asserts that $2^2 + 2 + 41 = 47$, $5^2 + 5 + 41 = 71$, $20^2 + 20 + 41 = 461$, and $100^2 + 100 + 41 = 10141$ are primes, all true statements.

This example demonstrates the danger of relying merely on checking cases. Note that $n^2 + n + 41 = n(n + 1) + 41$. Can you use this fact to find two (or more) values of n for which $n^2 + n + 41$ is not prime?

Complete mastery of mathematical induction is essential. It is our fundamental technique for proving infinite families of statements when they can be listed in such a way that each statement implies the next.

Example 3.26. The *Tower of Hanoi* puzzle consists of seven disks of decreasing size, stacked on one of three spindles. The object is to move the entire stack to one of the other spindles, moving only one disk at a time, and never placing a larger disk atop a smaller one. The initial configuration is shown in Figure 3.1.

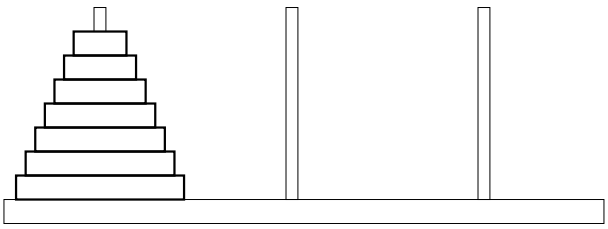


Figure 3.1: The Tower of Hanoi, initial configuration.

How many individual transfers are required to "solve" the Tower of Hanoi? To bring the power of mathematical induction to bear, we

generalize the puzzle, allowing n disks rather than seven. Let $T(n)$ denote the number of individual transfers required to move a stack of n disks subject to the rules above.

Clearly $T(1) = 1$; a single transfer moves a "stack" of one disk. For two disks, a bit of thought shows the task can be done in three transfers and no fewer: $T(2) = 3$. It's worthwhile to experiment with a stack of three or four coins of different sizes before reading further.

The puzzle with $(n+1)$ disks can be solved as follows: Move the top n disks from spindle 1 to spindle 2 (taking $T(n)$ transfers), then move the bottom disk to spindle 3 (one transfer), and finally move the stack of n disks from spindle 2 to spindle 3 (another $T(n)$ transfers). This strategy is clearly optimal, since the bottom disk cannot be transferred until the rest of the stack has been moved away. Tallying the number of transfers, we find $T(n + 1) = 2T(n) + 1$.

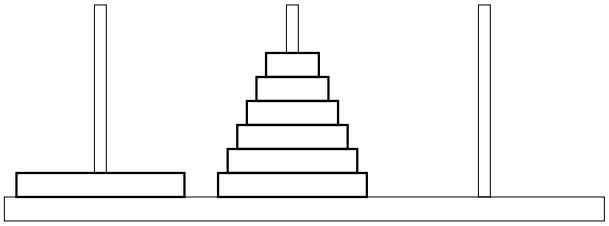


Figure 3.2: The Tower of Hanoi, intermediate configuration.

The original question could now be answered by successively calculating $T(3) = 2T(2) + 1 = 7$, $T(4) = 2T(3) + 1 = 15$, and so forth. Having formulated a general problem, however, we are led to ask "How many transfers are required to move a stack of n disks?" This is no longer a finite issue, since there are infinitely many puzzles, one for each positive number n.

To proceed further, we must *guess a formula for* $T(n)$. The sequence 1, 3, 7, 15 (for towers of one, two, three, and four disks respectively) might lead us to suspect the number of transfers is one less than a power of two: $T(n) = 2^n - 1$. Call this equation $P(n)$. To see if this guess is correct, we will attempt to prove $P(n)$ is true for all n.

The statement $P(1)$ says $T(1) = 1 = 2^1 - 1$, which is true. This establishes the base case. Next, assume inductively that $P(k)$ is true for some k, namely $T(k) = 2^k - 1$. By the hierarchical strategy described earlier,

$$T(k + 1) = 2T(k) + 1 = 2(2^k - 1) + 1 = 2 \cdot 2^k - 2 + 1 = 2^{k+1} - 1.$$

Thus, $P(k)$ implies $P(k+1)$ for each natural number k.

Since the base case is true and the inductive step is valid for each $k \geq 1$, our guess at a formula for $T(k)$ was correct by mathematical induction. As a special case, we find that $2^7 - 1 = 127$ transfers are required to move a stack of seven disks.

Remark 3.27. Both luck and skill are involved in solving this type of problem. Looking at the "data" 1, 3, 7, we might have found other plausible formulas, such as $(n-1)^2 + (n-1) + 1 = n^2 - n + 1$.

The inductive step weeds out incorrect guesses such as this one. If $T(k) = k^2 - k + 1$ for some k, then

$$\begin{aligned} T(k+1) &= (k+1)^2 - (k+1) + 1 = k^2 + k + 1, \\ 2T(k) + 1 &= 2k^2 - 2k + 3, \end{aligned}$$

and these are generally different.

3.3 Counting

If the nouns of mathematics are sets, then the verbs of mathematics are "mappings", usually called "functions" in school mathematics. Chapter 7 contains a formal definition and explores properties of mappings in detail. Here, we introduce mappings informally and procedurally.

Mappings and Subsets

Recall that if A and B are sets, a mapping $f : A \to B$ is a rule that associates to each element a of A an element $b = f(a)$ of B. The set A is called the domain of f, and the set B is the codomain. If $b = f(a)$, we say b is the value of f at a, and we say f maps a to b.

Definition 3.28. The set

$$f(A) = \{b \text{ in } B : b = f(a) \text{ for some } a \text{ in } A\}$$

is the *image* of f.

Remark 3.29. If A is the set of boxes you packed when moving to college and B is a set of labels ("clothes", "books", "kitchenware", etc.), then a mapping $f : A \to B$ is an assignment of a unique label to each box.

The image of f is the set of labels that actually get used.

Note carefully that a mapping assigns *precisely one* label $b = f(a)$ to each box a. It is possible, however, that:

- Two different boxes get the same label. (Maybe you have three boxes of clothes.)
- Some labels are not assigned to any box. (Maybe you have no kitchenware.)

Definition 3.30. A mapping $f : A \to B$ is *injective* if distinct elements of A map to distinct elements of B: If $a_1 \neq a_2$, then $f(a_1) \neq f(a_2)$.

Remark 3.31. Contrapositively, f is injective if $f(a_1) = f(a_2)$ implies $a_1 = a_2$.

If a box-labeling scheme is injective, then the contents of a box are uniquely determined by the label. (You do not have to ask, "Which box of clothes is this?" because there is only one box labeled "clothes".) Generally, the "input" a can be uniquely deduced from the "output" $b = f(a)$.

Definition 3.32. A mapping $f : A \to B$ is *surjective* if every element of B is a value of f: For every b in B, there exists an a in A such that $b = f(a)$.

Remark 3.33. A mapping $f : A \to B$ is surjective if and only if the image is the entire codomain: $B = f(A)$.

If a box-labeling scheme is surjective, then every type of label is used on at least one box.

Definition 3.34. A mapping $f : A \to B$ is *bijective* if f is both injective and surjective: For every b in B, there exists *precisely one* a in A such that $b = f(a)$.

Remark 3.35. A mapping that is injective, surjective, or bijective is often called, respectively, an *injection*, a *surjection*, or a *bijection*.

For the remainder of this section, m and n denote natural numbers, and $\underline{\mathbf{m}}$ and $\underline{\mathbf{n}}$ are sets containing m and n elements. When we need to list elements, we write $\underline{\mathbf{m}} = \{1, 2, \ldots, m\}$, with the understanding that $\underline{\mathbf{m}} = \varnothing$ if $m = 0$.

We give formulas for the number of mappings from $\underline{\mathbf{m}}$ to $\underline{\mathbf{n}}$, the number of *injective* mappings, and the number of distinct images of injective mappings, i.e., the number of m-element subsets of $\underline{\mathbf{n}}$.

Proposition 3.36. *There are precisely n^m mappings from $\underline{\mathbf{m}}$ to $\underline{\mathbf{n}}$.*

Remark 3.37. This result is our first substantial justification for defining $0^0 = 1$. (The laws of exponents are compatible with $0^0 = 0$.)

Proof. If $\underline{\mathbf{m}} = \underline{\mathbf{0}} = \varnothing$, there exists a unique mapping to $\underline{\mathbf{n}}$ if $n \geq 0$. If instead $m > 0$, then each element of $\underline{\mathbf{m}}$ can be sent to any of n distinct values in $\underline{\mathbf{n}}$. Since these choices are independent, the total number of mappings is the product of m factors of n, i.e., n^m. □

Remark 3.38. This formula correctly counts that there exist *no* mappings from $\underline{\mathbf{m}}$ to $\underline{\mathbf{0}}$ if $m \geq 1$. If necessary, re-examine the definition of a mapping to see why the empty set does not define a mapping with non-empty domain and empty target.

Definition 3.39. Let $\mathcal{U}$ be a set of n elements. An *ordering* of $\mathcal{U}$ is a bijection $s : \underline{\mathbf{n}} \to \mathcal{U}$, namely a listing $(s_k)_{k=1}^n$ of the elements of $\mathcal{U}$.

Example 3.40. The set $\mathcal{U} = \{a, b, c\}$ can be ordered in six ways. In "alphabetical" order:

$$(a, b, c), \quad (a, c, b), \quad (b, a, c), \quad (b, c, a), \quad (c, a, b), \quad (c, b, a).$$

Proposition 3.41. *Let $\mathcal{U}$ be a set of n elements. There exist $n!$ distinct orderings of $\mathcal{U}$.*

Remark 3.42. This formula is compatible with the definition $0! = 1$: The unique mapping $f : \varnothing \to \varnothing$ is vacuously bijective. It therefore suffices to prove the proposition for $n \geq 1$.

Proof. If $n \geq 1$, there are n ways to choose s_1, and then $(n-1)$ ways to choose a distinct s_2, then $(n-2)$ ways to choose s_3, and so on. The total number of choices, i.e., the number of ways of ordering $\mathcal{U}$, is therefore $n(n-1)(n-2)\cdots 3 \cdot 2 \cdot 1 = n!$. □

Example 3.43. A 52-card deck of playing cards can be shuffled into

$$52! = 80,658,175,170,943,878,571,660,636,856,403,766,$$
$$975,289,505,440,883,277,824,000,000,000,000,$$

or about $8.065817517 \times 10^{67}$ orderings.

To put the vastness of this number into perspective, the age of the visible universe is roughly 4.4×10^{17} seconds, and the visible universe is estimated to contain roughly 10^{70} atoms, give or take a couple of orders of magnitude. Since a human body contains roughly 10^{27} atoms,

the visible universe contains enough matter to form $10^{70} \div 10^{27} = 10^{43}$ card dealers.* If these dealers shuffled decks once every second (with no rest breaks for 13.7 billion years), they would have performed roughly $4.4 \times 10^{17} \times 10^{43} = 4.4 \times 10^{60}$ shuffles since the big bang, only enough to have seen about one in ten million possible orderings. In our actual universe, where the total number of earthly card shuffles surely does not exceed 10^{15} (a billion dealers each shuffling one million times), the number of shufflings ever seen is a vanishingly small fraction of all possible shufflings.

Definition 3.44. Let $\mathcal{U}$ be a set of n elements. An *ordered m-set* from $\mathcal{U}$ is an injection $f : \underline{\mathbf{m}} \to \mathcal{U}$, i.e., an ordered m-tuple $(a_1, \ldots, a_m)$ whose "terms" are distinct: $a_j \neq a_k$ if $j \neq k$. The image of an ordered m-set is the (*associated*) *unordered m-set* $\{a_1, \ldots, a_m\}$.

Remark 3.45. These terms are not in wide usage, and are introduced primarily for convenience in the remainder of this section.

Remark 3.46. For small values of m, one normally speaks of ordered *pairs* ($m = 2$), *triples* ($m = 3$), *quadruples* ($m = 4$), and so forth. There is no sharp dividing line for switching over to numerical prefixes from Latin, but terms such as *dodecatuple* ($m = 12$) or *vigintuple* ($m = 20$) are sadly under-used. Pronouncing "20-tuple" or "42-tuple" highlights all too clearly the drawbacks of this well-entrenched terminology. In practice, one speaks of "an m-tuple with $m = 20$".

Example 3.47. The 4-element set $\mathcal{U} = \{a, b, c, d\}$ has twelve ordered 2-sets

$$\begin{array}{llllll} (a,b) & (a,c) & (a,d) & (b,c) & (b,d) & (c,d) \\ (b,a) & (c,a) & (d,a) & (c,b) & (d,b) & (d,c). \end{array}$$

Note that each unordered 2-set appears exactly twice. (In what order are these pairs listed? Why does each appear twice? How many ordered and unordered triples are there?)

Proposition 3.48. *Let $\mathcal{U}$ be a set of n elements. If $0 \leq m \leq n$, there exist precisely*

$$n(n-1)(n-2)\cdots(n-m+1) = \frac{n!}{(n-m)!}$$

*Most of the universe consists of hydrogen, while a card dealer is largely made up of elements more than ten times heavier than hydrogen. Further, our estimate puts aside necessary support infrastructure: planets with habitable surface environments, resort cities, and casinos with all-you-can-eat buffets.

distinct ordered m-sets from $\mathcal{U}$.

Proof. Following the idea of Proposition 3.41, there are n ways to choose s_1, and then $(n-1)$ ways to choose a distinct s_2, then $(n-2)$ ways to choose s_3, etc., and $(n-m+1)$ ways to choose s_m. □

Remark 3.49. If $n < m$, the number of ordered m-sets from $\mathcal{U}$ is 0.

We come to the major goal of this subsection: Counting (unordered) m-element subsets of a set of n elements.

Definition 3.50. Let $\mathcal{U}$ be a set of n elements. The *binomial coefficient* $\binom{n}{m}$, read "n choose m", is defined to be the number of distinct subsets of $\mathcal{U}$ having precisely m elements.

Remark 3.51. Each binomial coefficient $\binom{n}{m}$ is a non-negative integer, and $\binom{n}{m} = 0$ unless $0 \leq m \leq n$.

Further, $\binom{n}{m} = \binom{n}{n-m}$: If $\mathcal{U}$ contains n elements, then to each m-element subset A of $\mathcal{U}$ is uniquely associated its complement $\mathcal{U} \setminus A$, having $(n-m)$ elements.

Proposition 3.52. *If m and n are arbitrary integers, then*

$$\binom{n}{m} = \begin{cases} \dfrac{n!}{m!(n-m)!} & \text{if } 0 \leq m \leq n, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. If $m < 0$ or $n < m$, there are no m-sets from $\mathcal{U}$.

Suppose $0 \leq m \leq n$, and let $\mathcal{U}$ be a set of n elements.

By Proposition 3.48, there are precisely $n!/(n-m)!$ ordered m-sets from $\mathcal{U}$. By Proposition 3.41, each unordered m-set from $\mathcal{U}$ is associated to precisely $m!$ ordered m-sets from $\mathcal{U}$. Combining these observations,

$$\frac{n!}{(n-m)!} = m! \cdot \binom{n}{m}, \quad \text{or} \quad \binom{n}{m} = \frac{n!}{m!(n-m)!}. \qquad \square$$

3.4 Construction of the Naturals

This section collects proofs of the properties asserted in the Section 3.1.

Sketch of proof of Theorem 3.1. It suffices to construct natural numbers and "successorship" in terms of sets. Define the empty set $\varnothing$ to be the initial element 0, and for an arbitrary natural number n, define

$$S(n) = n \cup \{n\}.$$

Concretely, $S(0) = \varnothing \cup \{\varnothing\} = \{\varnothing\}$ is a set containing "one element", the empty set itself;

$$S(S(0)) = \{\varnothing, \{\varnothing\}\}$$

is a set containing "two" elements, the empty set, and *the set consisting of the empty set*;

$$S(S(S(0))) = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}$$

is a set containing "three" elements (what are they?), and so forth. Properties N1 and N3 are obvious, while N2 follows (with some work) because for each natural number m, "the set $S(m) \setminus m = \{m\}$ contains precisely one element". □

Remark 3.53. For each natural number n we have, in this construction, $n \subseteq S(n) = n \cup \{n\}$, and $n \neq S(n)$ as sets.

Remark 3.54. If your philosophical tastes run toward "essentialism" (the nature of objects is grounded in *what they are*), it would be natural to conclude that the number zero *is* the empty set, the number one *is* the set $\{\varnothing\}$, and so forth.

In mathematics, this viewpoint is subtly wrong: Mathematical objects are characterized by their *behavior*, not by their *construction*.*

In the language of object-oriented programming, any infinite set together with an "initial element" and a notion of successorship satisfying Properties N1–N3 of Theorem 3.1 is an *implementation* of the natural numbers. The natural numbers themselves are the "abstract essence" characterized by the *public interface* of N1–N3. "Properties of the natural numbers" are precisely consequences of N1–N3.

From our perspective, the construction of the preceding proof shows that N1–N3 are logically consistent so long as ZFC itself is consistent.

Proof of Theorem 3.4. (i). Let $P(k)$ denote the statement

$$n + (m + k) = (n + m) + k \qquad \text{for all natural numbers } n \text{ and } m.$$

*This is not metaphysical hair-splitting: Computer software and digital data files are not specific patterns of bytes, but instructions for causing a system of computer hardware to behave in some specified way. That is, the nature of a piece of software is determined by *how it behaves*, not by its *physical structure as an ordered sequence of bytes*.

You yourself are not a particular collection of atoms, but an entity consisting of relationships divided conceptually into a hierarchy of functional systems, organs, cells, molecules, and atoms. Decades from now, you will retain some identity of who you currently are, but most of the atoms in your body will have been replaced.

By definition of addition, the base case $P(0)$ is true. Suppose $P(k)$ is true for some natural number k. For all natural numbers n and m, we have

$$\begin{aligned} n + \big(m + (k+1)\big) &= n + \big((m+k)+1\big) && (*), \\ &= \big(n + (m+k)\big) + 1 && (*), \\ &= \big((n+m)+k\big) + 1 && P(k), \\ &= (n+m) + (k+1) && (*). \end{aligned}$$

That is, the inductive step is true: $P(k)$ implies $P(k+1)$. By mathematical induction, $P(k)$ is true for all k.

(ii). For each natural number m, let $P(m)$ be the statement

$$n + m = m + n \qquad \text{for all } n.$$

Our goal is to prove $P(m)$ is true for every natural number m. To this end, we first establish $P(0)$ (a necessary part of the conclusion, but not strong enough for inductive purposes) and $P(1)$.

The single statement $P(0)$, which reads $0 + n = n + 0$ for all n, itself constitutes an *infinite family of statements*, one for each natural number n. The base case $n = 0$ is obvious: $0 + 0 = 0 + 0$. Now, if $0 + k = k$ for some natural number k, then

$$\begin{aligned} 0 + (k+1) &= (0+k) + 1 && (*), \\ &= k + 1 && \text{Inductive hypothesis.} \end{aligned}$$

By mathematical induction, $0+n = n = n+0$ for all n. This completes the proof of $P(0)$.

Similarly, $P(1)$ reads $1 + n = n + 1$ for all n. Treating this as an infinite family of statements as above, the base case reads $1+0 = 0+1$; this is the $n = 1$ case of the preceding paragraph. Assuming inductively that $1 + k = k + 1$ for some natural number k, we have

$$\begin{aligned} 1 + (k+1) &= (1+k) + 1 && (*), \\ &= (k+1) + 1 && \text{Inductive hypothesis.} \end{aligned}$$

By mathematical induction, $1+n = n+1$ for all n. This completes the proof of $P(1)$.

We now turn to $P(m)$. Assuming inductively that $P(m)$ is true for some natural number m, we have, for all natural numbers n,

$$\begin{aligned} n + (m + 1) &= (n + m) + 1 && (*), \\ &= (m + n) + 1 && P(m), \\ &= m + (n + 1) && (*), \\ &= m + (1 + n) && P(1), \\ &= (m + 1) + n && \text{Associativity.} \end{aligned}$$

Since $P(0)$ and $P(1)$ are true, and since $P(m)$ implies $P(m + 1)$ for every nonzero natural number m, induction on m guarantees $P(m)$ is true for all m. □

Proof of Theorem 3.6. Let m and n be natural numbers. We say m is *less than* n, and write $m < n$, if in the construction of the natural numbers as sets, $m \subseteq n$ and $m \neq n$.

Property (i), that $n < S(n) = n \cup \{n\}$ for every natural number n, is obvious.

Property (ii) is also obvious: If k is a proper subset of m and m is a proper subset of n, then k is a proper subset of n.

By construction, the natural numbers are strictly nested as sets, so given two natural numbers m and n, either they are equal, or one is a proper subset of the other; this is Property (iii).

Conversely, suppose $<$ satisfies Properties (i)–(iii) in Theorem 3.6, and that m and n are natural numbers such that $m < n$. We cannot have $n \subseteq m$ as sets, since m arises in the chain of successorship starting with n, from which we would deduce the false statement $n \leq m < n$. Thus, $m < n$ implies $m \subseteq n$ (and $m \neq n$) as sets.

This proves that the "proper subset of" relation coincides with the less-than relation, which by definition satisfies (i)–(iii).

If m and n are natural numbers, then $m < n$ if and only if $m \subseteq n$ as sets in the construction, if and only if n arises in the chain of successorship starting with m, if and only if there exists a natural number k such that $m + k = n$. □

For later use, we note the following technical result.

Theorem 3.55. *If n, m, k are natural numbers such that $n+k = m+k$, then $n = m$.*

Proof. Let $P(k)$ be the statement

$$n + k = m + k \text{ implies } n = m \text{ for all natural numbers } m \text{ and } n.$$

The statement $P(0)$ is true by definition of addition. Assume inductively that $P(k)$ is true for some natural number k. For all natural numbers m and n, if $n + (k + 1) = m + (k + 1)$, then

$$(n + k) + 1 = n + (k + 1) = m + (k + 1) = (m + k) + 1.$$

By uniqueness of predecessors, $n+k = m+k$, so $n = m$ by the inductive hypothesis. That is, $P(k)$ implies $P(k+1)$ for every natural number k. By induction, $P(k)$ is true for all k. □

Proof of Theorem 3.8. Assume contrapositively that X is a set of natural numbers having no least element; it suffices to prove X is empty. Let $P(m)$ be the statement

$$\text{If } k \leq m, \text{ then } k \notin X.$$

If $0 \in X$, then X has a least element; thus $0 \notin X$, i.e., $P(0)$ is true.

Now suppose inductively that $P(m)$ is true for some natural number m. If $m+1 \in X$, then $m+1$ would be a least element of X contrary to hypothesis. Thus $m + 1 \notin X$, i.e., $P(m + 1)$ is true.

Mathematical induction implies $P(m)$ is true for all m. Particularly, $m \notin X$ for all m, i.e., X is empty. □

3.5 Construction of the Integers

The size of a flock of sheep can be measured in the natural numbers, and addition corresponds to agglomeration of flocks. Analogously, debts can be accounted and paid by *subtracting*.

Unfortunately, the difference $n - m$ of two natural numbers is a natural number if and only if $m \leq n$. In other words, if m and n are natural numbers, the equation $m + x = n$ is solvable in the natural numbers precisely when $m \leq n$. Consequently, a debt x can be paid only if the amount owed, m, is no large than the amount possessed, n. Or can it?

Aleph: You owe me • • • • • sheep.

Beth: But I possess only • • • sheep.

Aleph: Very well. Give me those, and you only owe me • • sheep.

The set of integers, or whole numbers, is an abstract enlargement of the set of natural numbers in which, for arbitrary integers m and n, the equation $m+x = n$ has a solution x in the integers. A debt corresponds to a "negative" integer, a flock of size less than zero. A mathematically-minded chapter of the Society for Creative Anachronism might stage the preceding discussion:

Aleph: As far as I am concerned, you have -5 sheep.

Beth: But I possess only 3 sheep.

Aleph: Very well. Give me those, and you will have -2 sheep.

Remark 3.56. Negative numbers cannot be interpreted as absolute quantities. A mathematician saw two people go into a house. When three later came out, the mathematician reasoned, "If one person goes in, the house will be empty again."

Despite its whimsy, this joke contain a profound, useful idea: An integer is a *relationship between two natural numbers*. The solution x of $m + x = n$ represents a relationship between m and n. To encode this relationship, we form the *ordered pair* (m, n).

There is a technical hitch: Multiple equations, such as $1 + x = 4$, $6 + x = 9$, $1965 + x = 1968$, etc., all correspond to the same natural number. As integers, we want $(1, 4) = (6, 9) = (1965, 1968)$. These considerations motivate the definition.

Definition 3.57. Let m and n be arbitrary natural numbers. The ordered pair (m, n) is called an *integer representative*. Two representatives (m_1, n_1) and (m_2, n_2) are *equal as integers* if $m_1 + n_2 = m_2 + n_1$.

An *integer* is the "equivalence class" $[m, n]$ of integer representatives that are mutually equal. The set of integers is denoted $\mathbf{Z}$.

Example 3.58. If x is a natural number, the equivalence class $[0, x]$ is the integer representing x. If m is an arbitrary natural number, and if $n = m + x$, then $[0, x] = [m, n]$.

Similarly, the equivalence class $[x, 0] = [n, m]$ represents "negative x", a "debt of size x".

We want to do arithmetic with integers just as we do with natural numbers. Our conceptual identification $[m, n] = n - m$ and the wish

for "familiar laws of arithmetic"

$$(n-m)+(n'-m') = (n+n')-(m+m'),$$
$$(n-m)\cdot(n'-m') = (n\cdot n'+m\cdot m')-(m\cdot n'+m'\cdot n),$$

motivate our rules of addition and multiplication of integers. Note that the definition relies only on operations for natural numbers.

Definition 3.59. Let $N=(m,n)$ and $N'=(m',n')$ be integer representatives. Their *sum* and *product* are the integers defined by

$$(m,n)\oplus(m',n') = [m+m', n+n'],$$
$$(m,n)\odot(m',n') = [(m\cdot m')+(n\cdot n'),(m\cdot n')+(m'\cdot n)].$$

Remark 3.60. We have defined addition and multiplication of *representatives*. Before we can view the formulas in Definition 3.62 as operations on *integers*, we must check that "the result does not depend on the choice of representatives".

Theorem 3.61. *If $N_j=(m_j,n_j)$ and $N_j'=(m_j',n_j')$ are integer representatives for $j=1$, 2, and if $[N_1]=[N_1']$ and $[N_2]=[N_2']$, then*

$$N_1\oplus N_2 = N_1'\oplus N_2', \qquad N_1\odot N_2 = N_1'\odot N_2'.$$

Proof. The idea for addition is to "transform" $N_1\oplus N_2$ into $N_1'\oplus N_2'$ while preserving equality of integers. To accomplish this, we may add or cancel natural numbers in representatives, and use the hypotheses $m_1+n_1'=m_1'+n_1$ and $m_2+n_2'=m_2'+n_2$:

$$\begin{aligned} N_1\oplus N_2 &= [m_1+m_2, n_1+n_2] \\ &= [(m_1+m_1'+m_2+m_2'),(n_1+m_1'+n_2+m_2')] \\ &= [(m_1+m_1'+m_2+m_2'),(n_1'+m_1+n_2'+m_2)] \\ &= [m_1'+m_2', n_1'+n_2'] = N_1'\oplus N_2'. \end{aligned}$$

Multiplication is similar, Exercise 3.9. □

Definition 3.62. If $N=[m,n]$ and $N'=[m',n']$ are integers, we define their *sum* and *product* to be

$$[m,n]\oplus[m',n'] = [m+m', n+n'],$$
$$[m,n]\odot[m',n'] = [(m\cdot m')+(n\cdot n'),(m\cdot n')+(m'\cdot n)].$$

Remark 3.63. If m and n are natural numbers, then

$$[0, m] \oplus [0, n] = [0, m + n], \qquad [0, m] \odot [0, n] = [0, m \cdot n].$$

Loosely, addition and multiplication of natural numbers "work just the same" if we regard natural numbers as integers. For this reason, we may safely use the symbols "+" and "$\cdot$" to denote *integer* addition and multiplication.

Remark 3.64. We do not define exponentiation in the integers. It turns out there is no way to extend the operation on the natural numbers in a way that satisfies the laws of exponents.

Definition 3.65. If $N = [m, n]$ and $N' = [m', n']$ are integers, we say N is *less than or equal to* N', and write $N \leq N'$, if $n + m' \leq n' + m$. We say N is *less than* N', and write $N < N'$, if $n + m' < n' + m$.

Remark 3.66. As with the definitions of addition and multiplication, we must verify that these conditions, which are defined using representatives, do not depend on the choice of representative. Here, the verification is an immediate consequence of the definition of equality.

Theorem 3.67. *Let $(\mathbf{Z}, +, \cdot)$ denote the set of integers equipped with the operations of integer addition and multiplication.*

(i) *Addition is associative: If N, N', and N'' are integers, then*

$$N + (N' + N'') = (N + N') + N''.$$

(ii) *Additive identity element: If $0 = [0, 0]$, then $N + 0 = N$ and $0 + N = N$ for every integer N.*

(iii) *Additive inverses: If $N = [m, n]$, there exists a unique integer N' such that $N + N' = 0$ and $N' + N = 0$.*

(iv) *Addition is commutative: If N and N' are integers, then*

$$N + N' = N' + N.$$

(v) *Multiplication is associative: If N, N', and N'' are integers, then*

$$N \cdot (N' \cdot N'') = (N \cdot N') \cdot N''.$$

(vi) *Multiplicative identity element: If* $1 = [0, 1]$*, then* $N \cdot 1 = N$ *and* $1 \cdot N = N$ *for every integer* N*.*

(vii) *Multiplication distributes over addition: If* N*,* N'*, and* N'' *are integers, then*

$$N \cdot (N' + N'') = (N \cdot N') + (N \cdot N'').$$

(viii) *Multiplication is commutative: If* N *and* N' *are integers, then*

$$N \cdot N' = N' \cdot N.$$

(ix) *Trichotomy: For every integer* N*,* exactly one *of the following is true:* $N < 0$*,* $N = 0$*,* $N > 0$*.*

Proof. Each part reduces to mere computation. For practice, you should provide your own proofs for the omitted parts. Throughout the proof, we write $N = [m, n]$, $N' = [m', n']$, and $N'' = [m'', n'']$.

(Additive inverses). The integer $N' = [n, m]$ is easily checked to satisfy $N + N' = N' + N = 0$. To prove uniqueness, observe that if N' and N'' are additive inverses of N, then

$$N' = N' + 0 = N' + (N + N'') = (N' + N) + N'' = 0 + N'' = N''.$$

(Multiplication is associative). For brevity, multiplication signs are omitted between natural numbers. By direct calculation, freely using associativity, commutativity, and distributivity of operations on natural numbers,

$$\begin{aligned} N \cdot (N' \cdot N'') &= [m, n] \cdot [(m'm'' + n'n''), (m'n'' + m''n')] \\ &= [m(m'm'' + n'n'') + n(m'n'' + m''n'), \\ &\qquad m(m'n'' + m''n') + n(m'm'' + n'n'')] \\ &= [(mm' + nn')m'' + (mn' + nm')n'', \\ &\qquad (mm' + nn')n'' + (mn' + nm')m''] \\ &= (N \cdot N') \cdot N''. \end{aligned}$$

□

Theorem 3.68. *Let* N*,* N'*, and* N'' *be arbitrary integers.*

(i) *If* $N \leq N'$*, then* $N + N'' \leq N' + N''$*.*

(ii) *If* $N \leq N'$ *and* $0 \leq N''$*, then* $NN'' \leq N'N''$.

(iii) *If* $N \leq N'$ *and* $N'' \leq 0$*, then* $N'N'' \leq NN''$.

Remark 3.69. Particularly, the sum of two positive integers is positive, the product of two positive integers is positive, and the product of a positive integer and a negative integer is negative.

"Three times negative five" may be interpreted as the total debt of three shepherds, each having a debt of five sheep.

Attempting to reason along these lines is treacherous, however, as medieval philosophers discovered. What could "negative three shepherds" possibly mean? And what if each had a debt of five sheep? The mind reels. Even in the 21st Century, one can start fruitless arguments in mathematical web forums by asking what multiplying one negative number by another *really means*.

In the end, abstraction lights the way forward. The operation of multiplication has its original meaning when the operands are natural numbers, and the meaning is dictated by conformity to the distributive law one or both operands are negative.

Exercises

Exercise 3.1. *No matter how many sheep or stones one has, even if the number be like unto the grains of sand on all the beaches of the world, one can imagine having one more.*

Discuss the ways in which this assertion is not empirically true. Items to consider include the finite lifetime of a human being, of human culture, of physical conditions capable of sustaining intelligent life; the finite amount of observable matter in the cosmos; the finite speed with which information in the physical universe propagates; the fact that distant galaxies we can currently see will eventually fade to invisibility due to the expansion of space.

Exercise 3.2. Use the indicated strategies to find the sum of the first n positive integers.

(a) Compute a few special cases, formulate a conjecture, and use mathematical induction to prove your formula is correct.

(b) Starting with the formula for the sum of the first n odd positive integers, increment each summand by 1 to get the sum of the first

n even positive integers. Add these sums to get the sum of the first $N = 2n$ integers, and express the result in terms of N.

(c) Observe $1 + 2 + \cdots + (n-1) + n = n + (n-1) + \cdots + 2 + 1$; add these expressions to each other and group the respective first terms, second terms, and so forth. Now solve for the unknown sum.

Exercise 3.3. Use induction to prove that for every positive natural number k,

(a) $\displaystyle\sum_{j=1}^{k} 2j = k(k+1)$.

(b) $\displaystyle\sum_{j=1}^{k} (2j-1) = k^2$.

(c) $\displaystyle\sum_{j=1}^{k} 6j^2 = k(k+1)(2k+1)$.

(d) $\displaystyle\sum_{j=1}^{k} 4j^3 = \big[k(k+1)\big]^2$.

(e) If $r \neq 1$, then $1 + (r-1)\displaystyle\sum_{j=0}^{k} r^j = r^{k+1}$.

Exercise 3.4. Show that if $n \geq 4$, then $2^n < n! < n^n$.

Exercise 3.5. If n, $m \geq 0$ are integers, then $(n+m)! \geq (n+1)^m n!$. The inequality is strict if $m > 1$.

Exercise 3.6. If n is a natural number, define the *double factorial*, $n!!$, by

$$0!! = 1!! = 1, \qquad (n+2)! = (n+2) \cdot n!! \quad \text{for } n \geq 0.$$

(a) Show that

$$(2m)!! = (2m) \cdot (2m-2) \cdot (2m-4) \cdots 6 \cdot 4 \cdot 2,$$
$$(2m-1)!! = (2m-1) \cdot (2m-3) \cdot (2m-5) \cdots 5 \cdot 3 \cdot 1.$$

(a) Without using a calculator, evaluate the double factorials up to $10!!$.

(b) Prove that $(2m)!! = 2^m \cdot m!$ for every m.

(c) Prove that $(2m)!! \cdot (2m-1)!! = (2m)!$ for every m.

(d) Prove that $(2m+1)!! \cdot (2m)!! = (2m+1)!$ for every m.

Exercise 3.7. Prove Theorem 3.10, using only results established up to that point in the text.

Exercise 3.8. Prove Theorem 3.14, using only results established up to that point in the text.

Exercise 3.9. Prove that multiplication of integer representatives is well-defined on integers: If $N_j = (m_j, n_j)$ and $N'_j = (m'_j, n'_j)$ are integer representatives for $j = 1$, 2, and if $[N_1] = [N'_1]$ and $[N_2] = [N'_2]$, then

$$N_1 \odot N_2 = N'_1 \odot N'_2,$$

i.e.,

$$\begin{aligned}(m_1, n_1) \odot (m_2, n_2) &= [(m_1 \cdot m_2) + (n_1 \cdot n_2), (m_1 \cdot n_2) + (m_2 \cdot n_1)] \\ &= [(m'_1 \cdot m'_2) + (n'_1 \cdot n'_2), (m'_1 \cdot n'_2) + (m'_2 \cdot n_1)'] \\ &= (m'_1, n'_1) \odot (m'_2, n'_2).\end{aligned}$$

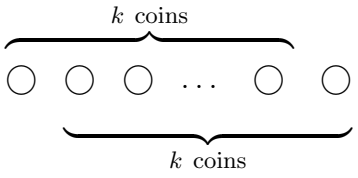
Hints: It suffices to show

$$N_1 \odot N_2 = N'_1 \odot N_2, \qquad N'_1 \odot N_2 = N'_1 \odot N'_2.$$

For the first, add $m'_1 \cdot (m_2 + n_2)$ in the representative of the product, use $m_1 + n'_1 = m'_1 + n_1$ and $m_2 + n'_2 = m'_2 + n_2$, then cancel $m_1 \cdot (m_2 + n_2)$. The second can be handled similarly, or show that $\odot$ is commutative.

Exercise 3.10. Consider the following "proof" that all coins have the same denomination:

Let $P(n)$ be the statement "In a set of n coins, all the coins have the same denomination." Now, $P(1)$ is clearly true (a single coin has a single denomination), so the base case is true. Assume inductively that $P(k)$ is true for some $k > 1$, and divide an arbitrary set of $(k+1)$ coins into two groups as shown at right.



By the inductive hypothesis, the first k coins all have the same denomination, and the last k coins have the same denomination. Since these two sets "overlap" as shown, all the coins have the same denomination. Since $P(k)$ implies $P(k+1)$ for all $k > 1$, $P(n)$ is true for all n, namely, all the coins have the same denomination. Where, exactly, are the logical flaws in this argument?

Chapter 4

Integer Division

In Chapter 3, we constructed the integers from the natural numbers. From now on, we denote integers with lowercase letters as we have done for natural numbers, we view natural numbers as non-negative integers, and we adopt the viewpoint that integers are not a specific implementation, but an arbitrary collection of objects satisfying axioms, Table 4.1, page 58.

4.1 Properties of the Integers

We first show that the identity elements for addition and multiplication, and additive inverses, are uniquely defined. To accomplish this, we use a mathematical idiom you should absorb: Assume two integers satisfy some property, and prove they are equal.

Theorem 4.1. *The additive identity element and the additive inverse of an arbitrary integer a are unique. Precisely:*

(i) *The integer* 0 *is uniquely defined by* A1. *and* A2.

(ii) *The integer* 1 *is uniquely defined by* M1. *and* M2.

(iii) *The integer* $-a$ *is uniquely defined by* A1., A2. *and* A3.

(iv) *For every integer* a, $-(-a) = a$. *In particular,* $-(-1) = 1$.

Proof. (i) Suppose integers 0 and $0'$ satisfy A3. By A3. with $a = 0'$, we have $0 + 0' = 0'$. However, by A3. with $a = 0$, we have $0 + 0' = 0$. Combining, $0 = 0 + 0' = 0'$. The proof of (ii) is entirely analogous, and is left to you.

There exists a set $\mathbf{Z}$, a subset $\mathbf{Z}^+$ of $\mathbf{Z}$, and two operations, $+$ and $\cdot$, satisfying the following axioms.

A1. (Associativity of addition) For all elements a, b, and c in $\mathbf{Z}$, we have $a + (b + c) = (a + b) + c$.

A2. (Additive identity element) There exists an integer 0 such that $a + 0 = 0 + a = a$ for all a in $\mathbf{Z}$.

A3. (Additive inverses) For every a in $\mathbf{Z}$, there exists $-a$ in $\mathbf{Z}$ such that $a + (-a) = (-a) + a = 0$.

A4. (Commutativity of addition) For all a and b in $\mathbf{Z}$, $a+b = b+a$.

M1. (Associativity of multiplication) For all a, b, and c in $\mathbf{Z}$, we have $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.

M2. (Multiplicative identity element) There exists an integer $1 \neq 0$ such that $a \cdot 1 = 1 \cdot a = a$ for all a in $\mathbf{Z}$.

M3. (Commutativity of multiplication) For all a and b in $\mathbf{Z}$, $a \cdot b = b \cdot a$.

M4. (Distributivity of multiplication over addition) For all a, b, and c in $\mathbf{Z}$, $a \cdot (b + c) = a \cdot b + a \cdot c$.

O1. (Law of Trichotomy) If $-\mathbf{Z}^+ = \{b \text{ in } \mathbf{Z} : -b \in \mathbf{Z}^+\}$, then the sets $\mathbf{Z}^+$, $\{0\}$, and $-\mathbf{Z}^+$ are a partition of $\mathbf{Z}$.

O2. (Sum of positive numbers) If a and b are elements of $\mathbf{Z}^+$, then $a + b \in \mathbf{Z}^+$.

O3. (Product of positive numbers) If a and b are elements of $\mathbf{Z}^+$, then $a \cdot b \in \mathbf{Z}^+$.

O4. (Well-ordering) If $A \subseteq \{0\} \cup \mathbf{Z}^+$ is non-empty, then there is a "smallest element" in A, i.e., there exists an a_0 in A such that $a + (-a_0) \in \{0\} \cup \mathbf{Z}^+$ for every a in A.

Table 4.1: Axioms for the integers.

(iii) Assume $a \in \mathbf{Z}$. If $a+b = b+a = 0$ and $a+c = c+a = 0$, then

$$\begin{aligned} b &= b+0 && \text{A2. with } a = b \\ &= b+(a+c) && a+c = 0 \text{ by hypothesis} \\ &= (b+a)+c && \text{A1.} \\ &= 0+c && b+a = 0 \text{ by hypothesis} \\ &= c && \text{A2. with } a = c. \end{aligned}$$

(iv) Part (iii) constitutes a useful principle: Let a and b be integers. To check whether $b = -a$, it suffices to show $a+b = 0$. But the equation $(-a)+a = 0$ may therefore be interpreted as saying the additive inverse of $(-a)$ is a, which is (iv). □

Definition 4.2. Let a and b be integers. The operation of *subtraction* is defined by $a - b = a + (-b)$. The integer $a - b$ is the *difference* of a and b.

Remark 4.3. Subtraction is neither associative nor commutative, and there is no identity element for subtraction.

Next we establish some "obvious" properties of multiplication.

Theorem 4.4. *If a is an arbitrary integer, then*

(i) $a \cdot 0 = 0 \cdot a = 0$.

(ii) $-1 \cdot a = -a$. *In particular,* $(-1) \cdot (-1) = -(-1) = 1$.

Proof. (i) Assume a is an arbitrary integer, and let $b = a \cdot 0$. By A2. with $a = 0$, $0 + 0 = 0$. Multiplying both sides by a and using the distributive law M4. gives

$$b + b = a \cdot 0 + a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 = b.$$

Adding $-b$ to each side,

$$b = 0 + b = \big((-b) + b\big) + b = (-b) + (b + b) = (-b) + b = 0,$$

as claimed.

(ii) By part (ii) of Theorem 4.1, the additive inverse of each integer is unique, so to prove $-1 \cdot a = -a$ it suffices to prove $a + (-1 \cdot a) = 0$. However,

$$a + (-1 \cdot a) = (1 \cdot a) + (-1 \cdot a) = \big(1 + (-1)\big) \cdot a = 0 \cdot a = 0.$$

Be sure you are able to justify each equality using the axioms and/or properties established earlier. □

Theorem 4.5. *Let a, b, c, and d be arbitrary integers.*

(i) *If $a < b$ and $b < c$, then $a < c$.*

(ii) *$0 < a$ if and only if $-a < 0$.*

(iii) *If $a < b$ and $0 < c$, then $ac < bc$.*

(iv) *If $a < b$ and $c < 0$, then $bc < ac$.*

Proof. (i) By definition, $a < b$ means $b + (-a) \in \mathbf{Z}^+$. Similarly, $b < c$ means $c + (-b) \in \mathbf{Z}^+$. By Axiom O2., a sum of elements of $\mathbf{Z}^+$ is an element of $\mathbf{Z}^+$, so

$$c + (-a) = c + (-b + b) + (-a) = \big(c + (-b)\big) + \big(b + (-a)\big) \in \mathbf{Z}^+,$$

proving $a < c$.

(ii) Suppose $0 < a$, i.e., $a \in \mathbf{Z}^+$, and consider the integer $-a$. By the trichotomy property, $a \neq 0$, and exactly one of the following is true: $-a = 0$, $-a \in \mathbf{Z}^+$, or $-a \in -\mathbf{Z}^+$. We will show that the first two of these statements are false.

If $-a = 0$, then $a = a + 0 = a + (-a) = 0$. Contrapositively, $a \neq 0$ implies $-a \neq 0$.

If $-a \in \mathbf{Z}^+$, then $0 = a + (-a) \in \mathbf{Z}^+$ by Axiom O2., since $a \in \mathbf{Z}^+$. But $0 \notin \mathbf{Z}^+$ by the trichotomy property, so $-a \notin \mathbf{Z}^+$.

The third condition, $-a \in -\mathbf{Z}^+$, must therefore hold.

Conversely, suppose $a < 0$, and let $b = -a$. Since $-b = a$ by Theorem 4.1 (iv), $b \in \mathbf{Z}^+$ by definition of $\mathbf{Z}^+$.

(iii) By hypothesis, $b + (-a) \in \mathbf{Z}^+$ and $c = c + (-0) \in \mathbf{Z}^+$. Axiom O3. guarantees their product is an element of $\mathbf{Z}^+$:

$$b \cdot c + (-a) \cdot c = \big(b + (-a)\big) \cdot c \in \mathbf{Z}^+.$$

Since $(-a)\cdot c = (-1\cdot a)\cdot c = -1\cdot(ac) = -ac$, we have $ac < bc$. The proof of (iv) is entirely analogous: Use the same argument with c replaced by $-c$ in $\mathbf{Z}^+$. □

As an application, we prove $a < b$ if and only if $b > a$. Let a and b be arbitrary distinct integers, and let $c = b + (-a)$, so $c \neq 0$. Now, $-c = a + (-b)$ by Theorem 4.1 (iii), since

$$\begin{aligned}
\big(a + (-b)\big) + \big((-a) + b\big) &= a + \big((-b) + (-a)\big) + b && \text{A1.} \\
&= a + \big((-a) + (-b)\big) + b && \text{A4.} \\
&= \big(a + (-a)\big) + \big((-b) + b\big) && \text{A1.} \\
&= 0 + 0 = 0. && \text{A3. and A2.}
\end{aligned}$$

Using part (ii) of the preceding theorem, $a < b$ if and only if $0 < c$, if and only if $0 > -c$, if and only if $b > a$, as was to be shown.

Theorem 4.6. *Let a, b, and c be arbitrary integers.*

(i) *If $a \neq 0$, then $0 < a^2$. In particular, $0 < 1$.*

(ii) *If $ab = 0$, then $a = 0$ or $b = 0$.*

(iii) *If $a \neq 0$ and $ab = ac$, then $b = c$.*

(iv) *If $0 < a$, then $1 \leq a$. In words, 1 is the smallest positive integer.*

Proof. (i) If $a \neq 0$, then by trichotomy and part (ii) of the preceding theorem, either $a \in \mathbf{Z}^+$ or $a \in -\mathbf{Z}^+$. In the first case, $a^2 = a \cdot a \in \mathbf{Z}^+$ by Axiom O3. In the second case, $-a \in \mathbf{Z}^+$, so

$$a^2 = \big((-1) \cdot (-1)\big) \cdot (a \cdot a) = \big((-1) \cdot a\big)^2 = (-a)^2 \in \mathbf{Z}^+.$$

In either case, $a^2 \in \mathbf{Z}^+$, or $0 < a^2$. Since $1 = 1^2$ by M2., $0 < 1$.

(ii) We prove the contrapositive: If $a \neq 0$ and $b \neq 0$, then $ab \neq 0$. By trichotomy, it suffices to consider four cases: $0 < a$ and $0 < b$; $0 < a$ and $b < 0$; $a < 0$ and $0 < b$; $a < 0$ and $b < 0$. The proofs are similar, so the details of one case will convey the idea.

If $0 < a$ and $b < 0$, then $0 < -b$ by part (ii) of Theorem 4.5. Since a product of positive numbers is positive, $0 < a(-b) = -(ab)$. Invoking part (ii) of Theorem 4.5 again, $ab < 0$; in particular, $ab \neq 0$.

(iii) Since $ab = ac$, we have $0 = ab - ac = a \cdot (b - c)$ by the distributive axiom. The preceding part of this theorem implies $a = 0$ or $b - c = 0$. Since $a \neq 0$ by hypothesis, $b - c = 0$, namely $b = c$.

(iv) Since the set $\mathbf{Z}^+ = \mathbf{N} \setminus \{0\}$ of positive integers is non-empty, there exists a smallest positive integer a_0 by the well-ordering axiom. Since 1 is positive by part (i) of this theorem, $a_0 \leq 1$.

By Axiom O3., $0 < a_0^2$. Since a_0 is the smallest positive integer, we have $a_0 \leq a_0^2$, or $0 \leq a_0^2 - a_0 = a_0(a_0 - 1)$. By parts (iii) and (iv) of Theorem 4.5, we have $a_0 - 1 \geq 0$, or $1 \leq a_0$.

Since $a_0 \leq 1$ and $1 \leq a_0$, we have $a_0 = 1$. □

Example 4.7. Suppose a is an integer such that $a^2 = a$. Subtracting and factoring, $a(a - 1) = a^2 - a = 0$. By part (ii), either $a = 0$, or $a - 1 = 0$, i.e., $a = 1$. This conclusion is no surprise, but naive manipulation of the axioms is unlikely to yield as concise a proof.

4.2 The Division Algorithm

Suppose N objects, such as jelly beans or playing cards, are to be divided among n people. Every child learns the algorithm: Put the people into some fixed order (such as counterclockwise around a circle). Following the order cyclically, give one object to each person in succession until none remain. This process "minimizes unfairness" in that either everyone receives the same number q of objects, or else everyone receives at least q objects, but some number r (with $0 < r < n$) receive one extra.

Mathematically, this process is known as the *division algorithm*, Figure 4.1. The numbers q and r are the *quotient* and *remainder* of N on division by n.

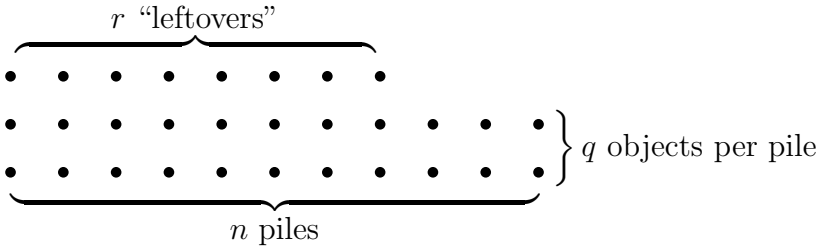


Figure 4.1: Dividing N objects among n piles.

Theorem 4.8. *Assume $N \in \mathbf{Z}$ and $n \in \mathbf{Z}^+$. There exist unique integers q and r, with $0 \leq r < n$, such that $N = nq + r$.*

The proof formalizes the naive algorithm. If $N > 0$, repeatedly subtract n until fewer than n objects remain, and say this occurs after q subtractions. The remainder (i.e., the number of "leftovers") $r = N - nq$ must be between 0 and $n - 1$. If $N \leq 0$, argue similarly, but *add* n repeatedly until, after q additions, the result is positive.

Proof. (Existence). Let N in $\mathbf{Z}$ and $n > 0$ be given. Consider the set S of integers of the form $N - nk$ with k in $\mathbf{Z}$, namely, integers that can be obtained from N by adding or subtracting n repeatedly. Let $S^+ = S \cap \mathbf{N}$ be the set of non-negative integers in S.

The set S^+ is non-empty: If $0 \leq N$, then $N = N - n \cdot 0 \in S^+$. If instead $N < 0$, then since $1 - n \leq 0$, we have $0 \leq (1 - n)N = N - nN$, i.e., $N - nN \in S^+$.

By the well-ordering principle, the non-empty set S^+ has a smallest element, say r. Since $r \in S$, by definition there exists an integer q such that $r = N - nq$, or $N = nq + r$.

Because r is the *smallest* element of S^+, we have $0 \leq r < n$: If instead $n \leq r$ were true, it would follow that $0 \leq r - n < r$, which would mean $r - n$ in S^+ is smaller than r. This completes the proof of the "existence" part of the theorem.

(Uniqueness). We wish to show that N can be written in only one way as $N = nq + r$ with $0 \leq r < n$.

Suppose $N = nq_1 + r_1 = nq_2 + r_2$ with $0 \leq r_1 < n$ and $0 \leq r_2 < n$. We may assume $r_1 \leq r_2$ without loss of generality; if not, swap the names of these numbers. Rearranging,

$$nq_1 - nq_2 = n(q_1 - q_2) = r_2 - r_1.$$

The left-hand side is an integer multiple of n. The right-hand side is non-negative, but no larger than $r_2 < n$. Their common value is therefore a non-negative integer multiple of n that is strictly smaller than n, namely zero. In other words, $r_1 = r_2$ and $q_1 = q_2$. This completes the proof of uniqueness. □

Remark 4.9. The conclusion of the division algorithm may look peculiar when $N < 0$. For example, if $N = -30$ and $n = 11$, the division algorithm gives $-30 = -3 \cdot 11 + 3$, while one might expect it to give $-30 = -2 \cdot 11 - 8$ (cf. Figure 4.1). Both equations are correct, of course, but the condition $0 \leq r < n$ forces us to use $-3 \cdot 11 + 3$ as the unique representation of -30 as a multiple of 11 plus a remainder.

Remark 4.10. In computer programming languages such as C++, bash, and Python, dividing an integer N by a positive integer n performs *integer division* `N/n`, returning the quotient q from Theorem 4.8, while the *modulus operator* `N%n` returns the remainder of N on division by n.

Integer Divisors

Definition 4.11. Let N be an integer. We say N is *even* if there exists an integer q such that $N = 2q$. We say N is *odd* if there exists an integer q such that $N = 2q + 1$.

Remark 4.12. By Theorem 4.8, every integer is either even or odd, and no integer is both.

Definition 4.13. Let a and b be integers. We say a *divides* b, and write $a \mid b$, if there exists an integer q such that $b = aq$. In this situation, we also say a is a *divisor* or a *factor* of b, or that b is a *multiple* of a.

Remark 4.14. Since $a(-q) = (-a)q = -(aq)$, the following are equivalent for all a and b: $a \mid b$, $-a \mid b$, and $a \mid -b$.

Lemma 4.15. *If a, b, and c are integers, and if $a \mid b$ and $b \mid c$, then $a \mid c$.*

Proof. By hypothesis, there exist integers q and r such that $b = aq$ and $c = br$. Substituting, $c = br = (aq)r = a(qr)$, so $a \mid c$. □

Remark 4.16. Clearly, $a \mid 0$ and $-a \mid 0$ are true for every a, by taking $q = 0$. The reverse relation is much more stringent: If $0 \mid a$, then $a = 0$.

Similarly, the statements $1 \mid b$, $-1 \mid b$ are true for every integer b, by taking $q = b$ or $q = -b$ respectively. In words, 1 and -1 divide everything. The converse relation is interesting enough to state formally.

Theorem 4.17. *If $a \in \mathbf{Z}$ and $a \mid 1$, then $a = \pm 1$.*

Proof. We will prove the contrapositive: If $a \neq \pm 1$, then $a \nmid 1$. As noted above, $0 \nmid 1$, and $a \mid 1$ if and only if $-a \mid 1$, so it suffices to consider the case $1 < a$.

If $a \mid 1$, there exists an integer q such that $aq = 1$. However, if $1 < a$, then multiplying by q would give $0 < q < aq = 1$, which is false; there is no integer between 0 and 1. It follows that if $1 < a$, then $a \nmid 1$. □

Subgroups of Integers

Definition 4.18. Let $G \subseteq Z$ be a *non-empty* set of integers. We say G is a *subgroup* of $(\mathbf{Z}, +)$ if

(i) G is *closed under addition*, i.e., if a and b are in G, then $a + b \in G$.

(ii) G is *closed under negation*, i.e., if $a \in G$, then $-a \in G$.

Example 4.19. The set $G = \{0\}$ is a subgroup of $(\mathbf{Z}, +)$, since $0+0 = 0$ and $-0 = 0$.

If $d \neq 0$ is an integer, the set $G = d\mathbf{Z}$ of integer multiples of d is a subgroup of $(\mathbf{Z}, +)$: If a and b are elements of G, then by definition there exist integers m and n such that $a = dm$ and $b = dn$. The distributive law gives $a + b = dm + dn = d(m + n)$, which is in G because $m + n$ is an integer. Similarly, $-a = -(dm) = d(-m)$ is in G since $-m$ is an integer.

The converse of Example 4.19 holds:

Theorem 4.20. *If $G \neq \{0\}$ is a subgroup of $(\mathbf{Z}, +)$, there exists a unique positive integer d such that $G = d\mathbf{Z}$.*

Proof. Let G be a subgroup of $(\mathbf{Z}, +)$, and let $S^+ = G \cap \mathbf{Z}^+$ be the set of positive integers in G. We first show S is non-empty. Since $G \neq \{0\}$, there exists a non-zero element a of G. Condition (ii) in the definition of a subgroup implies $-a \in G$. By trichotomy, either $a > 0$ or $-a > 0$; that is, either $a \in S^+$, or $-a \in S^+$, so S is non-empty.

By well-ordering, the non-empty set $S^+ \subseteq \mathbf{Z}^+$ has a smallest element d. An inductive argument shows that $dq \in G$ for every positive integer q. Condition (ii) in the definition of a subgroup shows $-dq \in G$ for every positive integer q. Finally, $0 = d + (-d) \in G$ by Condition (i) in the definition of a subgroup. That is, $d\mathbf{Z} \subseteq G$. To complete the proof of the theorem, it suffices to show $G \subseteq d\mathbf{Z}$.

Let a be an arbitrary element of G, and use the division algorithm to find the integers q and r such that $0 \leq r < d$ and $a = dq + r$. By the preceding paragraph, $-dq \in G$. Since G is closed under addition, $r = a + (-dq) \in G$. Now, the non-negative integer r must be 0, because $r < d$ and d is the smallest positive element of G. Consequently, $a = dq$, i.e., $a \in d\mathbf{Z}$. Since a was an arbitrary element of G, we have shown $G \subseteq d\mathbf{Z}$. □

4.3 The Greatest Common Divisor

Definition 4.21. Let a and b be integers. An integer c is a *common divisor* of a and b if $c \mid a$ and $c \mid b$.

Remark 4.22. If $a = b = 0$, then every integer is a common divisor of a and b. If at least one of a and b is non-zero, however, there is a common divisor d larger than every other common divisor.

Definition 4.23. Let a and b be integers, not both zero. An integer d is a *greatest common divisor* of a and b if

(i) $d \mid a$ and $d \mid b$ (d is a common divisor of a and b).

(ii) $0 < d$.

(iii) If $c \mid a$ and $c \mid b$, then $c \mid d$ (every common divisor divides d).

Example 4.24. If $a = 12$ and $b = 18$, then the common divisors of a and b are ± 1, ± 2, ± 3, and ± 6. The greatest common divisor is 6, the unique positive divisor into which every other divisor divides, and the largest of the divisors in the ordering of the integers.

Theorem 4.25. *Let a and b be integers, not both zero.*

(i) *There exists a unique greatest common divisor* $\gcd(a, b)$ *of a and b.*

(ii) $\gcd(a, b)$ *is the smallest positive element of the set*

$$\langle a, b\rangle = \{ka + \ell b : k, \ell \in \mathbf{Z}\}$$

of integer linear combinations of a and b.

Proof. (Uniqueness). Suppose d and d' both satisfy conditions (i)–(iii). Since d' divides both a and b, condition (iii) guarantees $d' \,|\, d$. Reversing the roles of d' and d shows $d \,|\, d'$.

Now, two positive integers, each dividing the other, must be equal: By hypothesis, there exist positive integers q_1 and q_2 such that $d = d'q_1$ and $d' = dq_2 = d'(q_1q_2)$, which implies $q_1q_2 = 1$. By Theorem 4.17, $q_1 = 1$, so $d = d'$.

(Existence). The set $G = \langle a, b\rangle$ is easily checked to be a subgroup of $(\mathbf{Z}, +)$ and not equal to $\{0\}$. By Theorem 4.20, G contains a smallest positive element d, and $G = d\mathbf{Z}$ is the set of multiples of d. It remains to show that the integer d satisfies conditions (i)–(iii) of Definition 4.23.

(d is a common divisor of a and b). The integers a and b are elements of G, so each is divisible by d. Property (ii), $0 < d$, is immediate.

(Every common divisor of a and b divides d). If $c \,|\, a$ and $c \,|\, b$, there exist integers q_1 and q_2 such that $a = cq_1$ and $b = cq_2$. Substituting,

$$d = am + bn = cq_1m + cq_2n = c(q_1m + q_2n),$$

which implies $c \,|\, d$. □

Remark 4.26. Changing the sign of a and/or b has no effect on $\langle a, b\rangle$, so $\gcd(-a, b) = \gcd(a, -b) = \gcd(-a, -b) = \gcd(a, b)$. In practice, we may as well assume a and b are both non-negative.

Corollary 4.27. *Let a and b be integers, $b \neq 0$. If $a = bq + r$, then* $\gcd(a, b) = \gcd(b, r)$.

Proof. It suffices to show $\langle a, b\rangle = \langle b, r\rangle$ as sets.

($\langle a, b\rangle \subseteq \langle b, r\rangle$). The general element of $\langle a, b\rangle$ is $ak+b\ell$. If $a = bq+r$, then for arbitrary integers k and ℓ,

$$ak + b\ell = (bq + r)k + b\ell = b(qk + \ell) + rk,$$

which is an element of $\langle b, r\rangle$.

($\langle b, r\rangle \subseteq \langle a, b\rangle$). The general element of $\langle b, r\rangle$ is $bm + rn$. Since $a = bq + r$, we have $r = a - bq$, so if m and n are arbitrary, then

$$bm + rn = bm + (a - bq)n = b(m - qn) + an,$$

which is in $\langle a, b\rangle$.

Since $\langle a, b\rangle = \langle b, r\rangle$, these sets have the same smallest positive element. Theorem 4.25 implies $\gcd(a, b) = \gcd(b, r)$. □

Euclid's Algorithm

The "smallest positive linear combination" characterization of the gcd leads to an efficient algorithm for computing $\gcd(a, b)$: Divide the smaller number into the larger, take the remainder (if non-zero), and repeat using the remainder and smaller divisor, stopping if the division at some stage has remainder zero. This process must terminate after finitely many steps, and the last non-zero remainder is the gcd. In pseudocode, if $0 < b < a$:

```
while (b != 0) // remainder is not zero
  {
    r := remainder(a, b) // a = n*b + r
    a := b
    b := r
  }
return a; // gcd(a, b) = last non-zero remainder
```

Moreover, this algorithm allows us to construct integers m and n such that $\gcd(a, b) = am + bn$. (These integers are not unique!) Let's see how this works in practice before stating a formal theorem.

Example 4.28. Find $d = \gcd(68, 20)$, and write d as a linear combination of 68 and 20.

Repeated long division gives

$$\begin{aligned} 68 &= 3 \cdot 20 + 8, \\ 20 &= 2 \cdot 8 + 4, \\ 8 &= 2 \cdot 4 + 0. \end{aligned}$$

Thus $\gcd(68, 20) = 4$, the last non-zero remainder.

To write 4 in terms of 68 and 20, start with the second-to-last equation just found, and substitute backward up the chain:

$$\begin{aligned} 4 &= 20 - 2 \cdot 8 \\ &= 20 - 2 \cdot (68 - 3 \cdot 20) = 20 - 2 \cdot 68 + 6 \cdot 20 \\ &= 7 \cdot 20 - 2 \cdot 68. \end{aligned}$$

As a check, this reads $4 = 140 - 136$.

Theorem 4.29 (Euclid's algorithm)**.** *Let $0 < b < a$ be integers, and recursively define sequences of quotients and remainders as follows:*

$$\begin{aligned} a &= bq_0 + r_1, && 0 \le r_1 < b \\ b &= r_1 q_1 + r_2, && 0 \le r_2 < r_1 \\ r_1 &= r_2 q_2 + r_3, && 0 \le r_3 < r_2 \\ &\vdots \\ r_k &= r_{k+1} q_{k+1} + r_{k+2}, && 0 \le r_{k+2} < r_{k+1}, \quad \textit{etc.} \end{aligned}$$

If $r_n \neq 0$ and $r_{n+1} = 0$, then $r_n = \gcd(a, b)$.

Proof. Corollary 4.27 says that if $a = bq + r$, then $\gcd(a, b) = \gcd(b, r)$. Applying this relationship to the lines of the algorithm in turn, we have

$$\begin{aligned} \gcd(a, b) = \gcd(b, r_1) &= \gcd(r_1, r_2) \\ &= \quad \vdots \\ &= \gcd(r_{n-1}, r_n) \\ &= \gcd(r_n, r_{n+1}) = \gcd(r_n, 0). \end{aligned}$$

For every positive integer c, $\gcd(c, 0) = c$; thus $\gcd(a, b) = r_n$. □

Exercises

Exercise 4.1. In each part, integers a and b are given. Calculate their gcd, and find integers m and n such that $am + bn = \gcd(a, b)$.

(a) $a = 16$, $b = 10$. (b) $a = 120$, $b = 75$. (c) $a = 121$, $b = -75$.

Exercise 4.2. In each part, a denotes an integer satisfying an equation. Solve for a, giving a proof based on the axioms and theorems in this chapter.

(a) $a^2 = 1$. (b) $a^3 = a$. (c) $a^4 = 1$.

Exercise 4.3. Let a and b be integers. Use the axioms to prove:

(a) $(a - b)(a + b) = a^2 - b^2$.

(b) $(a + b)^2 - (a - b)^2 = 4ab$.

(c) $(a + b)^2 + (a - b)^2 = 2(a^2 + b^2)$.

(d) Suppose a and b are not integers, but entities satisfying all the additive and multiplicative axioms except the commutative law of multiplication. (That is, suppose $ba \neq ab$ in general.) Expand the left-hand side of the identity in each of parts (a)–(c).

Exercise 4.4. An integer N is a *square* if there exists an integer n such that $N = n^2$. Prove that:

(a) Squares of consecutive integers differ by an odd integer.

(b) Every square leaves a remainder of 0 or 1 on division by 4.

(c) If m and n are integers, not both even, then $n^2 \neq 2m^2$.

Exercise 4.5. Suppose n_1 and n_2 are integers that leave respective remainders of r_1 and r_2 on division by 2.

(a) What remainder does $n_1 + n_2$ leave on division by 2?

(b) What remainder does $n_1 n_2$ leave on division by 2?

(c) In each of the preceding parts, make a 2×2 table to list all possible outcomes.

Exercise 4.6. Suppose n_1 and n_2 are integers that leave respective remainders of r_1 and r_2 on division by 5. Describe the remainders left by $n_1 + n_2$ and $n_1 n_2$ on division by 5.

Exercise 4.7. In each part, let a, b, c, and d be integers.

(a) If $a < b$ and $c < d$, then $a + c < b + d$.

(b) If $0 < a < b$ and $0 < c < d$, then $0 < ac < bd$.

Exercise 4.8. Let a and b be integers. Prove there exist integers u and v such that $u + v = a$ and $u - v = b$ if and only if a and b are both even or both odd, and find formulas for u and v in terms of a and b.

Exercise 4.9. Let d_1 and d_2 be integers. Prove that $d_1 \mid d_2$ if and only if $d_2\mathbf{Z} \subseteq d_1\mathbf{Z}$.

Exercise 4.10. Let a and b be non-zero integers, and let $M_{ab} = \langle a \rangle \cap \langle b \rangle$ be the set of common multiples of a and b.

(a) Prove M_{ab} is a subgroup of $(\mathbf{Z}, +)$. The smallest positive element $\operatorname{lcm}(a, b)$ of M_{ab} is called the *least common multiple* of a and b.

(b) Give conditions analogous to those in Definition 4.23, and prove $\operatorname{lcm}(a, b)$ is the unique integer satisfying these conditions.

(c) Prove that $\gcd(a, b) \operatorname{lcm}(a, b) = ab$.
Suggestion: First show that if m and d are positive integers with $ab = md$, then d is a common divisor of a and b if and only if m is a common multiple.

Exercise 4.11. In each part, integers a and b are given. Calculate their lcm.

(a) $a = 16$, $b = 10$. (b) $a = 120$, $b = 75$. (c) $a = 121$, $b = -75$.

Exercise 4.12. The *Fibonacci numbers* are defined recursively by

$$F_1 = F_2 = 1, \qquad F_{n+2} = F_n + F_{n+1}, \quad n \geq 1.$$

(a) Calculate the first twelve Fibonacci numbers.

(b) Show that $F_{n+2}F_{n-1} = F_{n+1}^2 - F_n^2$ for all $n \geq 2$.

(c) Use induction to show $\gcd(F_n, F_{n+1}) = 1$ for every $n \geq 1$.

Chapter 5

Primes

5.1 Primes and Coprimality

Every integer $a > 1$ has at least two positive divisors: 1 and a itself.

Definition 5.1. An integer $p > 1$ is *prime* if its *only* positive divisors are 1 and p. A non-prime integer $n > 1$ is *composite*.

Remark 5.2. The integer 1 is a *unit*, i.e., has a multiplicative inverse. In this chapter, invertibility conveys special status on 1, neither prime nor composite. Our goal is to factor composite integers uniquely into primes. If 1 were prime, uniqueness would be lost; if 1 were composite, existence of a factorization would be lost.

Example 5.3. The primes smaller than 20 are 2, 3, 5, 7, 11, 13, 17, and 19. Of these, only 2 is prime immediately from the definition: The *only* positive integers not exceeding 2 are 1 and 2, so 2 cannot have positive divisors other than 1 and 2!

Remark 5.4. It was known to Euclid around 300 BCE that there are infinitely many primes, a fact we prove below. At this writing, by contrast, it is unknown whether or not there exist infinitely many *twin primes*, pairs of primes differing by 2, such as 3 and 5 or 101 and 103.

In April 2013, Y. Zhang announced the existence of infinitely many pairs of primes differing by no more than 70 million, the first "bounded gap" result. By November 2013, J. Maynard had reduced the bound to 600. By April 2014, the PolyMath project had reduced the bound to 246, the best known bound at this writing (March 2017).

The primes are the "multiplicative building blocks" of the positive integers. This principle culminates in the "Fundamental Theorem of Arithmetic", Theorem 5.19 below. For now we are content to prove a technical result, later carried to its logical conclusion.

Proposition 5.5. *Let $N \geq 2$ be an integer. There exists a prime p such that $p \mid N$.*

Proof. The proof proceeds by mathematical induction on the following statement:

> $P(N)$: For every integer n with $2 \leq n \leq N$, there exists a prime p (depending on n) such that $p \mid n$.

Informally, $P(5)$ says "Each of the integers 2, 3, 4, and 5 has a prime factor".

The statement $P(2)$ is true; $p = 2$ is a divisor of $N = 2$. This establishes the base case.

Assume inductively that $P(k)$ is true for some $k > 1$, namely that *every* integer n with $2 \leq n \leq k$ has a prime factor.

The integer $k+1 > 1$ is either prime or composite. If $k+1$ is prime, then $p = k+1$ is a prime factor; together with $P(k)$, this implies every integer n with $2 \leq n \leq k+1$ has a prime factor, proving $P(k+1)$ in this case.

On the other hand, if $k+1$ is composite, there exist integers n and m, both greater than 1, such that $k+1 = nm$. It follows that $1 < n < k+1$, for if $k+1 \leq n$, then $(k+1)m \leq nm = k+1$, contrary to the inequality $1 < m$. Since $n < k+1$, we have $n \leq k$ by Theorem 4.6 (iv). By $P(k)$, there exists a prime p such that $p \mid n$. Now, $n \mid (k+1)$, so by Lemma 4.15, $p \mid (k+1)$ as well. Thus $k+1$ has a prime divisor, proving $P(k+1)$ in this case.

Since $P(2)$ is true and $P(k)$ implies $P(k+1)$ for each $k \geq 2$, $P(N)$ is true for all $N > 1$ by the principle of mathematical induction. □

Definition 5.6. Integers a and b are *coprime* if $\gcd(a, b) = 1$.

Remark 5.7. By Theorem 4.25, a and b are coprime if and only if there exist integers m and n such that $am + bn = 1$.

Example 5.8. $a = 14 = 2 \cdot 7$ and $b = 15 = 3 \cdot 5$ are coprime.

Example 5.9. $a = 111 = 3 \cdot 37$ and $768 = 2^8 \cdot 3$ are not coprime.

Example 5.10. Among integers between 1 and 11 inclusive, 1, 5, 7, and 11 are coprime to 12.

Proposition 5.11. *If p is prime, then*

(i) $\gcd(a, p) = p$ *if and only if* $p \mid a$.

(ii) $\gcd(a, p) = 1$ *if and only if* $p \nmid a$. *In particular, if* $0 < a < p$, *then* a *is coprime to* p.

Proof. (i) By definition, $\gcd(a, p) \mid a$, so if $\gcd(a, p) = p$, then $p \mid a$.
Conversely, if $p \mid a$, then $\gcd(a, p) = p$ since $p \mid p$ always.

(ii) If p is prime and a is an arbitrary integer, then *a priori* $\gcd(a, p)$ is either p or 1, since those are the only positive divisors of p. The second assertion follows immediately from (i). □

Theorem 5.12. *Suppose* $\gcd(a, b) = 1$. *If* $a \mid bc$, *then* $a \mid c$.

Proof. By hypothesis, there exists an integer q such that $aq = bc$. Since $\gcd(a, b) = 1$, there exist integers m and n such that $am + bn = 1$. Multiplying by c and substituting,

$$c = c(am + bn) = acm + (bc)n = acm + (aq)n = a(cm + qn).$$

But since $cm + qn$ is an integer, $a \mid c$. □

Theorem 5.13 (Euclid's lemma)**.** *Let a and b be integers. If p is prime and $p \mid ab$, then $p \mid a$ or $p \mid b$.*

Proof. If $p \mid a$ there is nothing to prove. Otherwise, $\gcd(a, p) = 1$ by Proposition 5.11, so $p \mid b$ by Theorem 5.12. □

Corollary 5.14. *If $a_1, \ldots, a_n$ are integers, p is prime, and $p \mid a_1 a_2 \ldots a_n$, then there exists an index i such that $p \mid a_i$.*

Proof. The corollary follows by mathematical induction on the number of factors. Euclid's lemma is the base case, for two factors. The details are left as an exercise. □

Remark 5.15. The hypothesis of coprimality cannot be dropped in Theorem 5.12: If $a = 6$, $b = 4$, and $p = 12$, then $p \mid ab$, but $p \nmid a$ and $p \nmid b$.

Of course, each prime factor of 12, namely 2 or 3, *is* a divisor of either a or b (or both), in accordance with Corollary 5.14.

5.2 Prime Factorization

We are aiming for the *Fundamental Theorem of Arithmetic*: Every integer greater than 1 factors "uniquely" into primes.

The word "uniquely" requires explanation here. The integer $n = 60$ factors as $p_1p_2p_3p_4 = 2 \cdot 2 \cdot 3 \cdot 5$, a product of four factors. Since multiplication is commutative, an arbitrary ordering of the same factors gives the same value for the product. Technically, however, $p_1p_2p_3p_4$ and $p_2p_1p_3p_4$ are "distinct products" (since the factors occur in different orders), even though they are *identical* as products of integers (because $p_1 = p_2$).

To avoid this purely linguistic issue, let us agree to organize products of primes so that (i) all occurrences of a given prime are gathered into a single prime power, and (ii) these prime powers are listed with the primes in increasing order.

We say that a product of prime powers satisfying (i) and (ii) is in *standard form*. Symbolically, a product of prime powers

$$N = p_1^{\nu_1} p_2^{\nu_2} \cdots p_m^{\nu_m} = \prod_{i=1}^{m} p_i^{\nu_i}, \quad \nu_i > 0 \text{ for } i = 1, \ldots, m,$$

is in standard form if $p_1 < p_2 < \cdots < p_m$. To say the integer N factors "uniquely" into primes means any two representations of N as products of primes have *identical* standard forms.

Example 5.16. The products $60 = 2^2 \cdot 3 \cdot 5$ and $2352 = 2^4 \cdot 3 \cdot 7^2$ are in standard form, while $2 \cdot 2 \cdot 3 \cdot 5$ (condition (i) unmet) and $3 \cdot 2^2 \cdot 5$ (condition (ii) unmet) are not.

The proof of the Fundamental Theorem is broken into "existence" and "uniqueness", since the required techniques are so different.

Theorem 5.17 (Existence of prime factorizations)**.** *For every integer $N \geq 2$, there exist primes $p_1 < p_2 < \cdots < p_m$ and positive integers $\nu_1, \nu_2, \ldots, \nu_m$ such that*

$$N = p_1^{\nu_1} p_2^{\nu_2} \cdots p_m^{\nu_m} = \prod_{i=1}^{m} p_i^{\nu_i}.$$

Briefly, every integer $N \geq 2$ factors into primes.

Proof. The proof proceeds by mathematical induction on the following statement:

$P(N)$ For every integer n with $2 \leq n \leq N$, n factors into primes.

For example, $P(4)$ asserts "2, 3, and 4 all factor into primes".

The statement $P(2)$, "2 factors into primes", is true because 2 is prime. This establishes the base case.

Assume inductively that $P(k)$ is true for some $k \geq 2$, that is, every integer n with $2 \leq n \leq k$ factors into primes. We wish to show that $(k+1)$ factors into primes, so that every integer n with $2 \leq n \leq k+1$ factors into primes.

By Proposition 5.5, $k+1$ has a prime divisor p. That is, there exists an integer q, $1 \leq q \leq k$, such that $(k+1) = pq$.

If $q = 1$, then $k+1 = p$ is itself prime. Otherwise, we have $2 \leq q \leq k$, so q factors into primes by the inductive hypothesis. Since p is prime, $k+1$ itself factors into primes.

In either case, $P(k)$ implies $P(k+1)$ for arbitrary $k \geq 2$, so by the principle of mathematical induction, $P(N)$ is true for all $N \geq 2$. □

Theorem 5.18 (Uniqueness of prime factorization)**.** *Suppose there exist primes $p_1 < p_2 < \cdots < p_m$ and $q_1 < q_2 < \cdots < q_\ell$, and there exist positive integers $\nu_1, \nu_2, \ldots, \nu_m$ and $\mu_1, \mu_2, \ldots, \mu_\ell$, such that*

$$p_1^{\nu_1} p_2^{\nu_2} \cdots p_m^{\nu_m} = q_1^{\mu_1} q_2^{\mu_2} \cdots q_\ell^{\mu_\ell}.$$

Then $\ell = m$, and for all $i = 1, \ldots, m$, we have $p_i = q_i$ and $\nu_i = \mu_i$.

Proof. The proof is an increasingly-familiar refrain: Proceed by induction on the statement

$P(N)$ For every integer n with $2 \leq n \leq N$, n has a unique prime factorization (in standard form).

A complete proof is left as an exercise, but here is the key step: Assume inductively that every integer n, $2 \leq n \leq k$, factors uniquely into primes, and suppose

$$k+1 = p_1^{\nu_1} p_2^{\nu_2} \cdots p_m^{\nu_m} = q_1^{\mu_1} q_2^{\mu_2} \cdots q_\ell^{\mu_\ell}.$$

Consider the smallest prime factors p_1 and q_1 in the respective products. If $p_1 < q_1$, then $p_1 < q_i$ for all i (the prime factors are listed in increasing order). However, Corollary 5.14 implies $p_1 \,|\, q_i$ for some i, i.e.,

q_i is not prime. Contrapositively, if the q_i are all prime, then $q_1 \leq p_1$. A similar argument proves $p_1 \leq q_1$; thus $p_1 = q_1$.

Writing $k+1 = np_1$ and cancelling $p_1 = q_1$ from the prime factorization of $k+1$ gives

$$p_1^{\nu_1 - 1} p_2^{\nu_2} \cdots p_m^{\nu_m} = q_1^{\mu_1 - 1} q_2^{\mu_2} \cdots q_\ell^{\mu_\ell} = n \leq k.$$

By the inductive hypothesis, these factorizations are identical, so the factorizations of $k+1$ are identical as well. □

Respectively, Theorems 5.17 and 5.18 establish the the existence and uniqueness portions of the following basic result.

Theorem 5.19 (The Fundamental Theorem of Arithmetic)**.** *Let $N \geq 2$ be an integer. There exist primes $p_1 < p_2 < \cdots < p_m$ and positive integers $\nu_1, \nu_2, \ldots, \nu_m$, uniquely defined by N, such that*

$$N = p_1^{\nu_1} p_2^{\nu_2} \cdots p_m^{\nu_m} = \prod_{i=1}^{m} p_i^{\nu_i}.$$

Applications of the Fundamental Theorem

Among the many applications of the Fundamental Theorem of Arithmetic is Euclid's theorem on the infinitude of primes.

Theorem 5.20. *There exist infinitely many primes.*

Proof. Let $S = \{p_1, \ldots, p_n\}$ be an arbitrary finite collection of primes, and let $N = p_1 p_2 \cdots p_n + 1$. By the Fundamental Theorem of Arithmetic, N can be factored into primes. However, no prime factor of N is an element of S, since by construction N leaves a remainder of 1 on division by each element of S. It follows that there exists a prime not in S, which means S is not the set of all primes. Since S was arbitrary, the set of primes is not finite. □

Theorem 5.21. *Let $p_1 < p_2 < \cdots < p_m$ be primes. If*

$$N = p_1^{\nu_1} p_2^{\nu_2} \cdots p_m^{\nu_m} = \prod_{i=1}^{m} p_i^{\nu_i}, \quad \nu_i > 0 \text{ for } i = 1, \ldots, m,$$

then the divisors of N are precisely the integers expressible in the form

$$a = p_1^{\mu_1} p_2^{\mu_2} \cdots p_m^{\mu_m} = \prod_{i=1}^{m} p_i^{\mu_i}, \quad 0 \leq \mu_i \leq \nu_i \text{ for } i = 1, \ldots, m.$$

Corollary 5.22. *With notation as in the theorem, N has exactly*

$$(\nu_1 + 1)(\nu_2 + 1) \cdots (\nu_m + 1)$$

positive divisors.

Example 5.23. The integer $18 = 2 \cdot 3^2$ has exactly $(1+1)(2+1) = 6$ positive divisors. These can be listed by finding all ordered pairs of non-negative exponents (μ_1, μ_2) with $\mu_1 \leq 1$ and $\mu_2 \leq 2$:

(μ_1, μ_2) :	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$	$(0,2)$	$(1,2)$
$2^{\mu_1} \cdot 3^{\mu_2}$:	1	2	3	6	9	18

Example 5.24. $26000 = 2^4 \cdot 5^3 \cdot 13$ has $(4+1)(3+1)(1+1) = 40$ positive divisors.

Example 5.25 (The Sieve of Eratosthenes)**.** To list all primes between 2 and N, it suffices to list the integers from 2 to N, then to perform the following recursive procedure: Circle the first number in the list, 2, then *sieve by* 2: Cross out every other number starting from 2.

Next, circle the first remaining number, 3, and *sieve by* 3: Cross out every third number (some of which have already been crossed out).

Repeat, circling the first remaining number p (which is prime) and sieving by p (crossing out non-trivial multiples of p). When this procedure terminates, all the circled numbers are prime, Exercise 5.11. For instance, the primes smaller than 16 are

$$\textcircled{2}\ \textcircled{3}\ \cancel{4}\ \textcircled{5}\ \cancel{6}\ \textcircled{7}\ \cancel{8}\ \cancel{9}\ \cancel{10}\ \textcircled{11}\ \cancel{12}\ \textcircled{13}\ \cancel{14}\ \cancel{15}$$

Example 5.26. Let $p_1 < p_2 < p_3 < \dots$ be the listing of *all* primes, taken in increasing order. To each sequence $(m_1, m_2, m_3, \dots)$ of non-negative integers with at most finitely many non-zero terms, associate the positive integer

$$N = \prod_{i=1}^{\infty} p_i^{m_i}.$$

The product has only finitely many factors different from 1, so may be viewed as a finite product.

By the Fundamental Theorem of Arithmetic, the mapping f from the set of sequences of non-negative exponents to the positive integers is a *bijection*!

The mapping f satisfies a property reminiscent of the law of exponents: If $\mathbf{m} = (m_1, m_2, \dots)$ and $\mathbf{m}' = (m_1', m_2', \dots)$ are sequences, then

$$f(\mathbf{m} + \mathbf{m}') = f(\mathbf{m}) \cdot f(\mathbf{m}').$$

In principle, one can find the gcd of a and b by finding the associated sequences of exponents, taking the smaller exponent for each prime, and forming the resulting number. Analogously, the lcm is found by taking the larger exponent for each prime. For example,

$$\begin{aligned} 120 &= 2^3 \cdot 3^1 \cdot 5^1 \cdot 7^0 \leftrightarrow (3, 1, 1, 0, \dots) \\ 126 &= 2^1 \cdot 3^2 \cdot 5^0 \cdot 7^1 \leftrightarrow (1, 2, 0, 1, \dots) \\ \gcd(120, 126) &= 2^1 \cdot 3^1 \cdot 5^0 \cdot 7^0 \leftrightarrow (1, 1, 0, 0, \dots) \\ \operatorname{lcm}(120, 126) &= 2^3 \cdot 3^2 \cdot 5^1 \cdot 7^1 \leftrightarrow (3, 2, 1, 1, \dots). \end{aligned}$$

This gives a second proof that $\gcd(a, b) \operatorname{lcm}(a, b) = ab$, see Exercise 4.10.

In practice, using prime factorization to find greatest common divisors is monumentally inefficient, while Theorem 4.29 (Euclid's algorithm) is computationally feasible for any integers small enough to represent conveniently in binary (millions of digits, say).

Exercises

Exercise 5.1. (a) Factor 2754 into primes, and determine the number of positive divisors.

(b) Factor 20400 into primes, and determine the number of positive divisors.

(c) Use the results of parts (a) and (b) to find the prime factorizations of $\gcd(2754, 20400)$ and $\operatorname{lcm}(2754, 20400)$.

Exercise 5.2. Prove Corollary 5.14.

Exercise 5.3. Prove Theorem 5.18.

Exercise 5.4. Pick an arbitrary three-digit number and write it down twice. Divide this integer by 7, then divide by 11, and finally divide by 13. The final quotient is the original number.

For example, starting with 456 yields, successively,

$$456456 \xrightarrow{\div 7} 65208 \xrightarrow{\div 11} 5928 \xrightarrow{\div 13} 456.$$

(a) Explain why this trick works.

(b) Give the details of the analogous trick if a two-digit number is written down three times, as in 424242.

Exercise 5.5. (a) Prove an integer N is divisible by 4 if and only if the number comprising the last two digits of N is divisible by 4.

(b) Prove an integer is divisible by 8 if and only if the number comprising the last three digits is divisible by 8.

Exercise 5.6. Prove an integer is divisible by 9 if and only if the sum of its digits is divisible by 9.
Hint: If $N = a_n \dots a_3 a_2 a_1$ is the decimal representation of N, then

$$N = a_1 + 10\, a_2 + 100\, a_3 + \cdots + 10^{n-1} a_n = \sum_{k=1}^{n} a_k 10^{k-1},$$

while the sum of the digits is

$$S = a_1 + a_2 + a_3 + \cdots + a_n = \sum_{k=1}^{n} a_k.$$

What can you say about $N - S$?

Exercise 5.7. Let $\mathbf{Z}/7\mathbf{Z} = \{0, 1, 2, 3, 4, 5, 6\}$, and define a "multiplication operation" on $\mathbf{Z}/7\mathbf{Z}$ by defining $a \cdot b$ to be the remainder of the ordinary product ab on division by 7. For example, $5{\times}6 = 30 = 4{\times}7{+}2$, so $5 \cdot 6 = 2$.

Make a multiplication table for this operation, and show that for each non-zero a, there exists a b such that $a \cdot b = 1$.

Exercise 5.8. Let p be a prime, $\mathbf{Z}/p\mathbf{Z} = \{0, 1, 2, \dots, p-1\}$, and define a "multiplication operation" on $\mathbf{Z}/p\mathbf{Z}$ by defining $a \cdot b$ to be the remainder of the ordinary product ab on division by p. (Compare the preceding exercise.)

(a) Show that if a is non-zero, there exists a b such that $a \cdot b = 1$.
Hint: $a \cdot b = 1$ if and only if there exists an integer n such that $ab = np + 1$ as integers.

(b) Show that $(\mathbf{Z}/p\mathbf{Z})^{\times}$, the set of non-zero elements of $\mathbf{Z}/p\mathbf{Z}$, is closed under $\cdot$.

Exercise 5.9. Let p be a prime. Prove that $(p-1)!$ leaves a remainder of $p-1$ on division by p.
Hints: In the product $(p-1)!$, use the preceding exercise to pair up elements who product leaves a remainder of 1 on division by p. Handle the case $p = 2$ separately.

Exercise 5.10. Let $N \geq 2$ be an integer. Prove that if N is not prime, then there exists a divisor m such that $m^2 \leq N$.

Exercise 5.11. Referring to the Sieve of Eratosthenes, Example 5.25:

(a) Show the sieving procedure terminates in finitely many steps, and that all circled numbers are prime.

(b) Show that if p is the first uncircled number, and $N < p^2$, then every remaining uncircled number is prime.
Hint: Use the preceding exercise.

(c) Use the Sieve of Eratosthenes to find all primes less than 100. (There are 25 of them, and you may have to sieve fewer times than you first expect.)

Exercise 5.12. Find all primes p such that $p + 2$ and $p + 4$ are also prime.

Exercise 5.13. Let N be an integer greater than 1.

(a) Prove that if N is composite, there exists a divisor k of N such that $k^2 \leq N$.

(b) Prove that if $2 \leq N \leq 120$, and if N is not divisible by any of 2, 3, 5, or 7, then N is prime.

Exercise 5.14. Recall that $10! = 3\,628\,800$ ends with two 0's, while $15! = 1\,307\,674\,368\,000$ ends with three 0's. Without using a computer, determine the number of 0's at the end of 1000! (the factorial of 1000).

Exercise 5.15. Determine (with proof) the number of primes in the sequence 101, 10101, 1010101, ...

Exercise 5.16. Let N be a positive integer. Prove there exist N consecutive composite integers.

Chapter 6

Residue Classes

This chapter constructs and studies a new type of "number system" sharing many properties with the integers, but containing only finitely many elements.

6.1 Congruence $(\mathrm{mod}\, n)$

Fix a positive integer n. For each integer a, the division algorithm (Theorem 4.8) says there exist unique integers q and r such that $a = nq + r$ and $0 \leq r < n$. The concept of even and odd integers generalizes usefully if we sort integers by their remainder r on division by n.

Definition 6.1. Fix a positive integer n. Two integers a and b are *congruent* $(\mathrm{mod}\, n)$, denoted $a \equiv b \pmod{n}$, if a and b leave the same remainder on division by n.

If $a = nq + r$ with $0 \leq r < n$, the set $[a] = n\mathbf{Z} + r$ of integers congruent to a $(\mathrm{mod}\, n)$ is called the *residue class* of a $(\mathrm{mod}\, n)$.

Lemma 6.2. *Integers a and b are congruent $(\mathrm{mod}\, n)$ if and only if $n \mid (b - a)$.*

Proof. Use Theorem 4.8 to write $a = nq_1 + r_1$ and $b = nq_2 + r_2$;

$$b - a = (nq_2 + r_2) - (nq_1 + r_1) = n(q_2 - q_1) + (r_2 - r_1),$$

so $n \mid (b - a)$ if and only if $n \mid (r_2 - r_1)$.

Since $0 \leq r_1, r_2 < n$, we have $-n < -r_1 \leq r_2 - r_1 \leq r_2 < n$. Consequently, $n \mid (r_2 - r_1)$ if and only if $r_2 - r_1 = 0$, if and only if $a \equiv b$ $(\mathrm{mod}\, n)$. □

Remark 6.3. Let n be a positive integer. Every integer is congruent $(\mathrm{mod}\, n)$ to precisely one integer r with $0 \leq r < n$. That is, the residue classes $[0]$, $[1]$, $\ldots$, $[n-1]$ form a partition of $\mathbf{Z}$. The set of residue classes $(\mathrm{mod}\, n)$ is denoted $\mathbf{Z}/n\mathbf{Z}$.

Example 6.4. Equivalence classes mod n and their arithmetic are implicitly familiar. Since there are 12 hours in one turn of a clock, time-keeping works mod 12. At an early age, you learned that six hours after 7 o'clock is 1 o'clock, or that five hours before 3 o'clock is 10 o'clock. For military time, you'd work mod 24 instead, but the idea is the same.

Example 6.5. Days of the week are reckoned mod 7. The labels are not integers, of course, but names: $\{\mathrm{Sunday}, \mathrm{Monday}, \ldots, \mathrm{Saturday}\}$.

Example 6.6. Angular measurements in degrees are made mod 360, because there are 360 degrees in one full turn of a circle. Angles of 270, 630, and -90 degrees represent the same geometric quantity.

Example 6.7. Moving from the sublime to the ridiculous, a cartoon character (usually the cat in a cat-and-mouse conflict) will sometimes acquire amnesia when given a sharp blow on the head. The cure, as everyone knows, is a second blow. The cat's state of mental health (amnesiac or cured) represents the number of cranial blows mod 2 the cat has received. The concept is so simple even young children understand it perfectly.

Equivalence classes mod n may be visualized in at least two useful ways. The first is the "clock" picture of the *set* $\mathbf{Z}/n\mathbf{Z}$ of equivalence

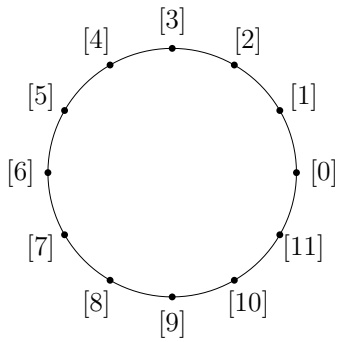


Figure 6.1: The set $\mathbf{Z}/12\mathbf{Z}$ of residue classes mod 12.

classes mod n, which may be drawn as a set of n equally-spaced points on a circle. The case $n = 12$, Figure 6.1 is essentially an ordinary analog clock, though by convention we place [0] at the rightmost position and label counterclockwise, ending with the class $[n-1]$ one space clockwise from [0]. This picture emphasizes the "cyclical" nature of residue classes. Adding 1 corresponds to traveling counterclockwise by one space. Adding n travels one full revolution, returning to the same residue class.

The second picture is the "unwrapping" of the clock onto a number line. For this, choose n distinct symbols, such as $[0], [1], \ldots, [n-1]$, and use these to label integer points on a numbers line.

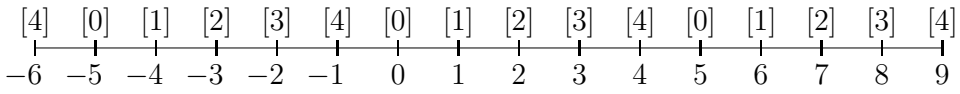


Figure 6.2: Residue classes mod 5 on the number line.

Modular Arithmetic

Timekeeping intuition suggests two residue classes can be added using $[a] + [b] = [a + b]$. This formula hides a subtlety: We might have $[a'] = [a]$ with $a' \neq a$ and $[b'] = [b]$ with $b' \neq b$, and must check that $[a'+b'] = [a+b]$ before assuming the sum of two classes is *well-defined*.

Theorem 6.8. *Fix $n > 0$. Let a, a', b, and b' be integers such that $a \equiv a' \pmod{n}$ and $b \equiv b' \pmod{n}$, and let $[a] = [a']$ and $[b] = [b']$ denote their residue classes mod n. Then*

(i) $a + b \equiv a' + b' \pmod{n}$, *i.e.,* $[a + b] = [a' + b']$.

(ii) $ab \equiv a'b' \pmod{n}$, *i.e.,* $[ab] = [a'b']$.

The theorem is straightforward to prove directly, but breaking the proof into slightly smaller steps clarifies the main idea.

Lemma 6.9. *Let a, a', and c be integers. If $a \equiv a' \pmod{n}$, then $a + c \equiv a' + c \pmod{n}$ and $ac \equiv a'c \pmod{n}$.*

Proof. By definition, $a \equiv a'$ if and only if $n \mid (a' - a)$.

Since $a' - a = (a' + c) - (a + c)$, we have $a \equiv a'$ if and only if $n \mid (a' - a) = (a' + c) - (a + c)$, if and only if $a + c \equiv a' + c$.

Further, if $n \mid (a' - a)$, then $n \mid (a' - a)c = a'c - ac$, i.e., $ac \equiv a'c$. □

Proof of theorem. Let a, a', b, b' be integers such that $a \equiv a' \pmod{n}$ and $b \equiv b' \pmod{n}$. Then

$$\begin{aligned} a + b &\equiv a' + b \pmod{n} && \text{Lemma 6.9 with } c = b \\ &\equiv a' + b' \pmod{n} && \text{Lemma 6.9 with } c = a'. \end{aligned}$$

The proof for products is identical: $ab \equiv a'b \equiv a'b' \pmod{n}$. □

Remark 6.10. If $a \not\equiv 0 \pmod{n}$ and $ab \equiv ac \pmod{n}$, it is *not* valid to deduce $b \equiv c$; the law of cancellation does not generally hold mod n. For example, $2 \cdot 3 \equiv 2 \cdot 0 \pmod{6}$, but even though $2 \not\equiv 0 \pmod{6}$, it is not true that $3 \equiv 0 \pmod{6}$.

Theorem 6.8 allows us to perform arithmetic with residue classes (almost) as if they were integers. Because there are only finitely many residue classes, we can (for small n) represent addition and multiplication $(\mathrm{mod}\, n)$ with a *Cayley table*. List the residue classes $[a]$ in a column on the left, list the residue classes $[b]$ in a row across the top, and place the sum or product in the corresponding table entry.

Example 6.11. The tables for addition and multiplication mod 5 are

+	[0]	[1]	[2]	[3]	[4]
[0]	[0]	[1]	[2]	[3]	[4]
[1]	[1]	[2]	[3]	[4]	[0]
[2]	[2]	[3]	[4]	[0]	[1]
[3]	[3]	[4]	[0]	[1]	[2]
[4]	[4]	[0]	[1]	[2]	[3]

·	[0]	[1]	[2]	[3]	[4]
[0]	[0]	[0]	[0]	[0]	[0]
[1]	[0]	[1]	[2]	[3]	[4]
[2]	[0]	[2]	[4]	[1]	[3]
[3]	[0]	[3]	[1]	[4]	[2]
[4]	[0]	[4]	[3]	[2]	[1]

Each residue class appears exactly once in each row and column of the addition table. This signifies a cancellation law: If $[a] + [c] = [b] + [c]$, then $[a] = [b]$. This cancellation occurs for all n, not just $n = 5$.

The analogous phenomenon occurs with *non-zero* classes mod 5 under multiplication: If $[a][c] = [b][c]$ with $[a]$, $[b]$, and $[c]$ not equal to $[0] \pmod{5}$, then $[a] = [b]$. This phenomenon is not general. We will see shortly that cancellation for multiplication by an arbitrary non-zero class $(\mathrm{mod}\, n)$ occurs if and only if n is prime.

Example 6.12. The tables for addition and multiplication mod 4 are

+	[0]	[1]	[2]	[3]
[0]	[0]	[1]	[2]	[3]
[1]	[1]	[2]	[3]	[0]
[2]	[2]	[3]	[0]	[1]
[3]	[3]	[0]	[1]	[2]

·	[0]	[1]	[2]	[3]
[0]	[0]	[0]	[0]	[0]
[1]	[0]	[1]	[2]	[3]
[2]	[0]	[2]	[0]	[2]
[3]	[0]	[3]	[2]	[1]

Example 6.13. To "solve" the congruence $67x \equiv 54 \pmod 5$ is to find an integer x such that $67x$ leaves the same remainder as 54 on division by 5. With the help of Theorem 6.8, we may reduce 54 and 67 $(\mathrm{mod}\, 5)$, recasting the question in the simpler form $2x \equiv 4 \pmod 5$. By inspection, $x = 2$ is a solution.

Remark 6.14. Generally, when solving $ax \equiv b \pmod n$, the coefficients a and b may be reduced mod n and the resulting congruence is equivalent to the original.

Example 6.15. Consider the problem of simplifying $4^{2000} \pmod{63}$. The naive approach of calculating 4^{2000} (a number of over 1200 digits), then dividing by 63 and taking the remainder, is prohibitively complex. By the second part of Theorem 6.8, however, we may instead compute successive powers of 4, reducing mod 63 each time a running product exceeds 63.

Even this is prohibitive if carried out mechanically, but we can do better still. Indeed, $4^3 = 64 \equiv 1 \pmod{63}$, so $4^{3q} = (4^3)^q \equiv 1^q \equiv 1 \pmod{63}$ for all q. The preferred strategy, therefore, is to divide the original *exponent* 2000 by 3. Writing $2000 = 3 \cdot 666 + 2$, we have

$$4^{2000} = 4^{3\cdot 666+2} = (4^3)^{666} \cdot 4^2 \equiv 4^2 \equiv 16 \pmod{63}.$$

The remainder in question is 16.

Example 6.16. The same idea can be used to compute products mod n. Consider, for example,

$$10! = 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \pmod{11}.$$

Judiciously gathering factors in pairs and reducing mod 11, $10 \equiv -1$, $9 \cdot 6 = 54 \equiv -1$, $8 \cdot 7 = 56 \equiv 1$, $5 \cdot 2 = 10 \equiv -1$, and $4 \cdot 3 = 12 \equiv 1$. Consequently,

$$10! \equiv (-1)(-1)(1)(-1)(1) = -1 \pmod{11}$$

by the second part of the theorem. Compare Exercise 5.9.

6.2 Multiplicative Inverses in $\mathbf{Z}/n\mathbf{Z}$

Fix an integer $n \geq 2$, and let $\mathbf{Z}/n\mathbf{Z}$ be the set of residue classes mod n:

$$\mathbf{Z}/n\mathbf{Z} = \{[0], [1], \ldots, [n-1] = [-1]\}.$$

By Theorem 6.8, there exist well-defined operations + and $\cdot$ on $\mathbf{Z}/n\mathbf{Z}$ satisfying

$$[a] + [b] = [a + b], \qquad [a] \cdot [b] = [a \cdot b]$$

for all integers a and b. These operations are both associative and commutative, since the corresponding integer operations enjoy these properties. For example, if a, b, and c are integers, then

$$\begin{aligned} \big([a] + [b]\big) + [c] &= [a + b] + [c] \\ &= \big[(a + b) + c\big] \\ &= \big[a + (b + c)\big] \\ &= [a] + [b + c] = [a] + \big([b] + [c]\big). \end{aligned}$$

Replacing each "+" by "$\cdot$" proves multiplication is associative. An entirely similar argument shows multiplication distributes over addition. Each operation has an identity element. The residue class [0] is an identity element for addition, while [1] is an identity element for multiplication.

Up to this point, the structures $(\mathbf{Z}/n\mathbf{Z}, +)$ and $(\mathbf{Z}/n\mathbf{Z}, \cdot)$ have completely parallel properties. When we ask about *inverses*, however, the stories quickly diverge. Under addition, matters are trivial: Every element $[a]$ of $\mathbf{Z}/n\mathbf{Z}$ has an additive inverse, $-[a] = [n - a]$. Multiplicative inverses, by contrast, do not always exist.

Definition 6.17. A residue class $[a]$ is *invertible* in $\mathbf{Z}/n\mathbf{Z}$ if there exists a class $[x]$ such that $[a][x] = [1]$. The class $[x]$ is the *inverse* of $[a]$ in $\mathbf{Z}/n\mathbf{Z}$.

Remark 6.18. Applied to residue classes in $\mathbf{Z}/n\mathbf{Z}$, the term "invertible" *always* refers to multiplication. The additive inverse of a residue class is its *negative*.

Example 6.19. In $\mathbf{Z}/7\mathbf{Z}$, $[2][4] = [1]$, so $[2]$ and $[4]$ are invertible in $\mathbf{Z}/7\mathbf{Z}$, and each is the inverse of the other.

Similarly, $[3][5] = [1]$, so $[3]$ and $[5]$ are invertible in $\mathbf{Z}/7\mathbf{Z}$, and each is the inverse of the other.

The class $[6] = [-1]$ is its own inverse.

Example 6.20. The class $[0]$ is *never* invertible in $\mathbf{Z}/n\mathbf{Z}$: For every integer x, $[0][x] = [0] \neq [1]$.

In $\mathbf{Z}/6\mathbf{Z}$, the classes $[2]$, $[3]$, and $[4]$ are not invertible.

Definition 6.21. An invertible class $[a]$ in $\mathbf{Z}/n\mathbf{Z}$ is a *unit* (mod n). The set of units (mod n) is denoted $(\mathbf{Z}/n\mathbf{Z})^\times$.

Proposition 6.22. *Let $n > 1$. If $[a]$ and $[b]$ are invertible classes in $\mathbf{Z}/n\mathbf{Z}$, then $[a]^{-1}$ and $[a][b]$ are invertible in $\mathbf{Z}/n\mathbf{Z}$. That is, $(\mathbf{Z}/n\mathbf{Z})^\times$ is closed under multiplication and under inversion.*

Proof. Let $[a]$ in $\mathbf{Z}/n\mathbf{Z}$ be invertible. The condition $[x] = [a]^{-1}$ says $[a][x] = [1]$. The relationship between an invertible class and its inverse is reciprocal, in the sense that $[x] = [a]^{-1}$ if and only if $[x]^{-1} = [a]$. Particularly, the inverse of $[a]$ is invertible.

Further, if $[b]$ is invertible with $[y] = [b]^{-1}$, then

$$\big([a][b]\big)\big([y][x]\big) = [a]\big([b][y]\big)[x] = [a][1][x] = [a][x] = [1].$$

This means the class $[a][b]$ is invertible in $\mathbf{Z}/n\mathbf{Z}$, and its inverse is the product (in either order) of the inverses of $[a]$ and $[b]$. □

Theorem 6.23. *If a is an integer, the residue class $[a]$ is invertible in $\mathbf{Z}/n\mathbf{Z}$ if and only if* $\gcd(a, n) = 1$.

Proof. By definition, $[a]$ is invertible in $\mathbf{Z}/n\mathbf{Z}$ if and only if there exists an integer x such that $[a][x] = [1]$, namely, $ax \equiv 1 \pmod{n}$. This holds if and only if there exist integers x and y such that $ax + ny = 1$, if and only if $\gcd(a, n) = 1$, see Theorem 4.25. □

Corollary 6.24. *If p is prime and $[a]$ is non-zero in $\mathbf{Z}/p\mathbf{Z}$, then $[a]$ is invertible. That is, $(\mathbf{Z}/p\mathbf{Z})^\times = \mathbf{Z}/p\mathbf{Z} \setminus \{[0]\}$.*

Proof. By Proposition 5.11, if a is an integer, then $\gcd(a, p) = p$ if and only if $p \mid a$, and $\gcd(a, p) = 1$ otherwise. □

Corollary 6.25. *The residue class $[a]$ is invertible in $\mathbf{Z}/n\mathbf{Z}$ if and only if $[-a] = [n - a]$ is invertible in $\mathbf{Z}/n\mathbf{Z}$.*

Proof. By Remark 4.26, $\gcd(-a, n) = \gcd(a, n)$. □

Example 6.26. We have $(\mathbf{Z}/12\mathbf{Z})^\times = \{[1], [5], [7], [11]\}$. Invertible classes occur in pairs of the form $[a]$ and $[-a]$, as guaranteed by Corollary 6.25. Since $[1]^2 = [1]$ and $[5]^2 = [1]$, we have $[7]^2 = [-5]^2 = [1]$ and $[11]^2 = [-1]^2 = [1]$.

Example 6.27. Since $14 = 2 \cdot 7$, we have

$$(\mathbf{Z}/14\mathbf{Z})^\times = \{[1], [3], [5], [9], [11], [13]\} = \{[1], [3], [5], [-5], [-3], [-1]\}.$$

Successive multiplication gives

$$[3]^2 = [9] = [-5], \qquad [3]^4 = [-3] = [11],$$
$$[3]^3 = [-15] = [-1], \qquad [3]^5 = [-9] = [5],$$

and finally $[3]^6 = [3]^0 = [1]$.

A more sensible ordering for the elements of $(\mathbf{Z}/7\mathbf{Z})^\times$ is therefore $\{[3]^0, [3], [3]^2, [3]^3, [3]^4, [3]^5\} = \{[1], [3], [9], [13], [11], [5]\}$. The multiplication table can be worked out from the law of exponents. For example, since $[11] = [3]^4$, we have $[11]^2 = [3]^8 = [3]^6 \cdot [3]^2 = [9]$.

Example 6.28. Since $30 = 2 \cdot 3 \cdot 5$, we have

$$(\mathbf{Z}/30\mathbf{Z})^\times = \{[1], [7], [11], [13], [17], [19], [23], [29]\}.$$

To study multiplication in $((\mathbf{Z}/30\mathbf{Z})^\times, \cdot)$, we can pair off elements and their negatives to avoid multiplying numbers of absolute value greater than $30/2 = 15$, i.e., without computing products larger than $13^2 = 169$.

Direct calculation gives $[7]^2 = [19]$, $[11]^2 = [1]$, and $[13]^2 = [19]$. It follows immediately that

$$[23]^2 = [-7]^2 = [19], \quad [19]^2 = [-11]^2 = [1], \quad [17]^2 = [-13]^2 = [19].$$

Putting these conclusions together, $[7]^4 = [19]^2 = [1]$, $[13]^4 = [1]$, $[17]^4 = [1]$, and $[23]^4 = [1]$.

Each $[x]$ satisfying $[x]^2 = [1]$ is its own inverse. To invert elements satisfying $[x]^4 = [1]$, argue as follows:

$$[7]^{-1} = [7]^3 = [19][7] = [-11][7] = [-17] = [13].$$

The inverses of the other elements may be deduced with no additional effort:

$$[13]^{-1} = [7], \quad [17]^{-1} = [-13]^{-1} = [-7] = [23], \quad [23]^{-1} = [17].$$

There is no ordering of the elements analogous to the preceding example, because no single residue class "generates" the set of units $(\mathrm{mod}\, 30)$. We might instead choose

$$\left\{ \begin{matrix} [1], & [7], & [7]^2, & [7]^3, \\ [11], & [7][11], & [7]^2[11], & [7]^3[11] \end{matrix} \right\} = \left\{ \begin{matrix} [1], & [7], & [19], & [13], \\ [11], & [17], & [29], & [23] \end{matrix} \right\}.$$

For larger n, it may be inconvenient to list the elements of $(\mathbf{Z}/n\mathbf{Z})^\times$. Nonetheless, we can easily test individual residue classes for membership, and can calculate inverses using the Euclidean algorithm.

Example 6.29. Determine whether the following residue classes are invertible in $\mathbf{Z}/105\mathbf{Z}$, and if so find the inverse.

[51]: By the Euclidean algorithm,

$$\begin{aligned} 105 &= 2 \cdot 51 + 3 \\ 51 &= 17 \cdot 3 + 0, \end{aligned}$$

so $\gcd(105, 51) = 3 \neq 1$. Thus [51] is not invertible in $\mathbf{Z}/105\mathbf{Z}$.

[32]: The Euclidean algorithm gives

$$\begin{aligned} 105 &= 3 \cdot 32 + 9 \\ 32 &= 3 \cdot 9 + 5 \\ 9 &= 1 \cdot 5 + 4 \\ 5 &= 1 \cdot 4 + 1. \end{aligned}$$

Thus $\gcd(105, 32) = 1$, so [32] is invertible in $\mathbf{Z}/105\mathbf{Z}$. To find the inverse, write 1 as a linear combination of 105 and 32:

$$\begin{aligned} 1 &= 5 - 4 \\ &= 5 - (9 - 5) = 2 \cdot 5 - 9 \\ &= 2 \cdot (32 - 3 \cdot 9) - 9 = 2 \cdot 32 - 7 \cdot 9 \\ &= 2 \cdot 32 - 7 \cdot (105 - 3 \cdot 32) = 23 \cdot 32 - 7 \cdot 105. \end{aligned}$$

Reducing $(\mathrm{mod}\, 105)$, $[23][32] = [1]$, or $[32]^{-1} = [23]$.

Zero Divisors

As noted earlier, [0] never has a multiplicative inverse in $\mathbf{Z}/n\mathbf{Z}$. To characterize non-invertible elements generally, we introduce the following definition.

Definition 6.30. A residue class $[a]$ in $\mathbf{Z}/n\mathbf{Z}$ is a *zero divisor* if there exists a non-zero class $[b]$ such that $[a][b] = [0]$.

Theorem 6.31. *The class* $[a]$ *in* $\mathbf{Z}/n\mathbf{Z}$ *is a zero divisor if and only if* $\gcd(a, n) > 1$.

Remark 6.32. Combining with Theorem 6.23, each class $[a]$ in $\mathbf{Z}/n\mathbf{Z}$ is either invertible or a zero divisor, but not both.

Proof. (If $[a]$ is a zero divisor, then $\gcd(a, n) > 1$). Contrapositively, suppose $d = \gcd(a, n) = 1$, and that $[a][b] = [0]$ for some $[b]$. We want to show $[b] = [0]$, which will prove $[a]$ is not a zero divisor. By Theorem 6.23, the class $[a]$ is invertible in $\mathbf{Z}/n\mathbf{Z}$, so there exists an $[x]$ with $[x][a] = [1]$. Multiplying $[a][b] = [0]$ by $[x]$ on the left,

$$[0] = [x][0] = [x]\big([a][b]\big) = \big([x][a]\big)[b] = [1][b] = [b].$$

(If $\gcd(a, n) > 1$, then $[a]$ is a zero divisor). Suppose $\gcd(a, n) > 1$. If $a \equiv 0 \pmod{n}$, namely if $[a] = [0]$, there is nothing to prove. Otherwise, we may divide a and n by $d = \gcd(a, n)$ and deduce there exist non-zero integers a' and n' such that $a = da'$ and $n = dn'$. Since $1 < d$, we have $n' < n$, which implies $[n'] \neq [0]$ in $\mathbf{Z}/n\mathbf{Z}$. Since

$$[a][n'] = [(a'd)n'] = [a'(dn')] = [a'n] = [0],$$

$[a]$ is a zero divisor in $\mathbf{Z}/n\mathbf{Z}$. □

Example 6.33. It may help to follow the preceding proof with specific numbers. Let $n = 8$ and $a = 6$. Here, $d = \gcd(6, 8) = 2$, so $a' = a/2 = 3$ and $n' = n/2 = 4$. As expected, $[a][n'] = [6][4] = [0]$ in $\mathbf{Z}/8\mathbf{Z}$, which proves $[a] = [6]$ is a zero divisor.

The dichotomy between units and zero divisors has a pleasant geometric interpretation. View $\mathbf{Z}/n\mathbf{Z}$ as a set of n evenly-spaced points on a circle. To determine whether a residue class $[a]$ is a unit or zero divisor, place a pencil at $[0]$ and count off a spaces at a time around the circle, joining successive values with line segments, Figure 6.3. The corners on this "multiplication diagram" correspond to classes $[k]$ such that $[a][x] = [k]$ has a solution $[x]$ in $\mathbf{Z}/n\mathbf{Z}$.

Algebraically, $ax \equiv k \pmod{n}$ if and only if there exists an integer y with $ax + ny = k$. Geometrically, $ax = k$ is the location reached by counting off a spaces x times. Adding or subtracting an integer multiple of n corresponds to discarding traversals of the entire circle, which have no effect on the location of a corner.

By Theorem 4.25, there exist integers x and y with $ax + ny = k$ if and only if $\gcd(a, n) \mid k$. In other words, consecutive corners on the diagram are separated by $\gcd(a, n)$ spaces in $\mathbf{Z}/n\mathbf{Z}$, and the number of corners in the diagram is $n/\gcd(a, n)$.

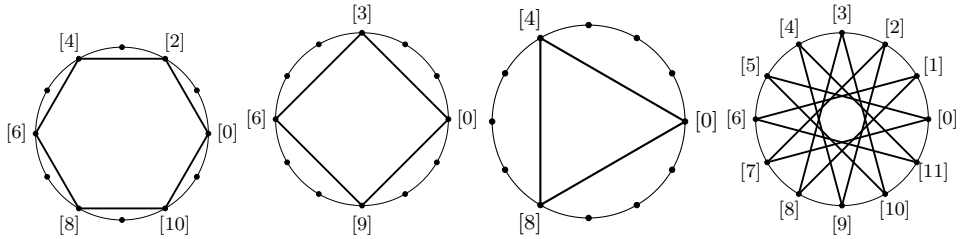


Figure 6.3: Multiplication diagrams for [2], [3], [4], and [5] in $\mathbf{Z}/12\mathbf{Z}$.

The dichotomy implied by Theorems 6.23 and 6.31 is: $[a]$ is a unit in $\mathbf{Z}/n\mathbf{Z}$ if and only if $\gcd(a, n) = 1$, if and only if there are no "gaps" between consecutive corners, if and only if the multiplication diagram touches each element of $\mathbf{Z}/n\mathbf{Z}$.

This mathematics is the basis of the ingenious toy Spirograph, which consists of circular plastic rings of inner circumference n and disk-shaped gears of varying circumference a, Figure 6.4. Both the rings

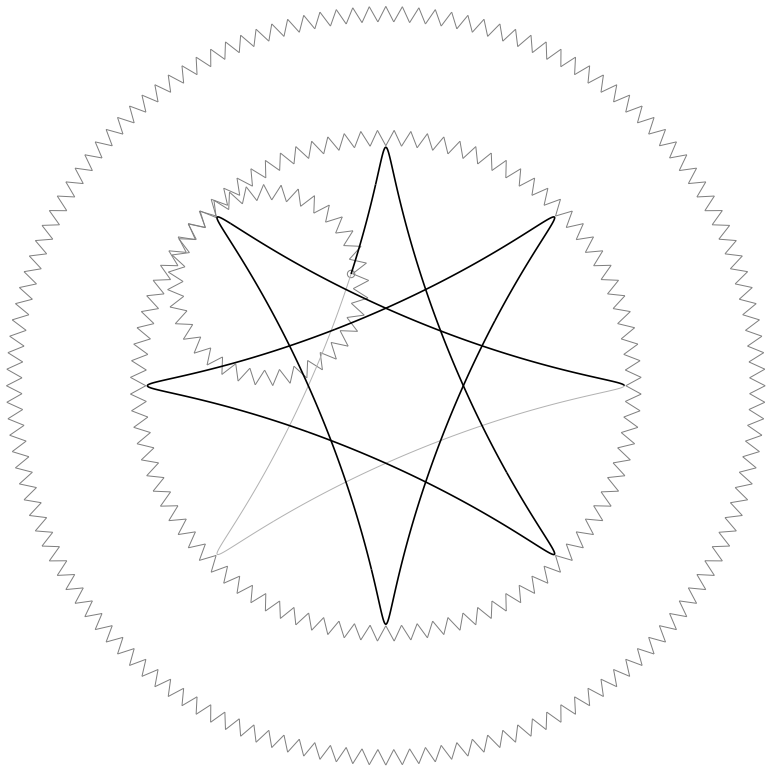


Figure 6.4: Spirograph; a 36-tooth gear in a 96-tooth ring.

and gears have teeth to ensure the gears roll without slipping. The gears have holes through which a pen fits. By tacking the ring to a sheet of paper and rolling a gear around the inside of a ring until the pattern closes, you produce smooth mathematical curves having the overall geometry of the multiplication diagram for $[a]$ in $\mathbf{Z}/n\mathbf{Z}$. See also http://mathcs.holycross.edu/~ahwang/teach/software/Pirouette/index.html

Example 6.34. In Figure 6.4, $n = 96 = 8 \cdot 12$ and $a = 36 = 3 \cdot 12$. Since $\gcd(36, 96) = 12$, there are 12 teeth between consecutive points of the pattern, and the pattern has $96/12 = 8$ points.

Counting by 36 (mod 96), we obtain 0, 36, 72, 12, 48, 84, 24, 60 in succession before returning to 0. These numbers are precisely the multiples of 12 (mod 96) listed in the order they're visited when following the Spirograph pattern.

6.3 Linear Congruences

Example 6.13 introduced a simple linear congruence and solved it by essentially *ad hoc* tricks. We now have the algebraic tools to solve the general linear congruence

$$ax \equiv b \pmod{n}, \tag{6.1}$$

in which a, b, and n are given and x is unknown.

Theorem 6.35. *Let a, n, and b be integers.*

(i) *If $\gcd(a, n) \nmid b$, then (6.1) has no solution in $\mathbf{Z}/n\mathbf{Z}$.*

(ii) *If $\gcd(a, n) = 1$, then (6.1) has a unique solution in $\mathbf{Z}/n\mathbf{Z}$.*

(iii) *If $\gcd(a, n) | b$, then (6.1) has precisely $\gcd(a, n)$ solutions in $\mathbf{Z}/n\mathbf{Z}$.*

Proof. (i) Assume contrapositively that $ax \equiv b \pmod{n}$ has a solution. There exists an integer y such that $ax = b + ny$, or $ax - ny = b$. This expresses b as a linear combination of a and n. By Theorem 4.25, $\gcd(a, n) \mid b$.

(ii) Assume $\gcd(a, n) = 1$. We must prove (6.1) has a solution (existence), and any two solutions are equal (uniqueness).

We *construct* a solution of (6.1). There exist integers s and t such that $as + nt = 1$. Multiply by b to get $a(sb) + n(tb) = b$. Setting

$x = sb$, the preceding equation says $ax \equiv b \pmod{n}$. This establishes existence.

Still assuming $\gcd(a, n) = 1$, if x_1 and x_2 are solutions of (6.1), then $ax_1 \equiv b$ and $ax_2 \equiv b$. Subtracting, $a(x_1 - x_2) \equiv 0 \pmod{n}$, or $n \mid a(x_1 - x_2)$. Since $\gcd(a, n) = 1$, Theorem 5.12 implies $n \mid (x_1 - x_2)$, so $x_1 \equiv x_2 \pmod{n}$. This proves uniqueness of solutions mod n.

(iii) For convenience, write $d = \gcd(a, n)$. By Theorem 4.25, there exist integers s and t such that $as + nt = d$, and there exist integers a' and n' such that $a = a'd$ and $n = n'd$. Further, if $d \mid b$, there exists an integer b' such that $b = b'd$.

Dividing $as + nt = d$ by d gives $a's + n't = 1$, which implies $\gcd(a', n') = 1$. Part (ii) of this theorem guarantees that the congruence $a'x \equiv b' \pmod{n'}$ has a unique solution $x_0 = sb'$. Multiplying $a'x_0 \equiv b' \pmod{n'}$ by d shows x_0 is a solution of (6.1).

It remains to show (6.1) has d solutions. For $i = 0, 1, 2, \ldots, d - 1$, let $x_i = x_0 + in'$. Since $an' = a'dn' = a'n$, each x_i is a solution:

$$ax_i = ax_0 + i(an') = ax_0 + i(a'n) \equiv ax_0 \equiv b \pmod{n}.$$

Moreover, these d numbers are distinct (mod n): If $0 \le j \le i < d$, then $x_i \equiv x_j \pmod{n}$ if and only if

$$n = n'd \mid (x_i - x_j) = (in' - jn') = (i - j)n',$$

if and only if $d|(i-j)$, if and only if $i = j$. We have therefore constructed d distinct solutions of (6.1).

There are no other solutions. If x is an arbitrary solution of $ax \equiv b \pmod{n}$, dividing through by d proves $a'x \equiv b' \pmod{n'}$, so $x \equiv x_0 \pmod{n'}$ by uniqueness. In other words, there exists an integer i such that $x = x_0 + in'$, so x is one of the solutions found above. □

Example 6.36. Solve the congruence $30x \equiv 18 \pmod{216}$.

Here $a = 30$ and $n = 216$, so $d = \gcd(a, n) = 6$. To find x_0, divide through by 6 and solve $5x_0 \equiv 3 \pmod{36}$. Our earlier method using Euclid's algorithm gives $x_0 = 15$. There are $d = 6$ solutions in total, any two differing by a multiple of $n' = 36$: $x_1 = 15 + 36 = 51$, $x_2 = 87$, $x_3 = 123$, $x_4 = 159$, and $x_5 = 195$.

6.4 Fermat's Little Theorem

Theorem 6.37. *Let p be a prime. If $[a]$ is a residue class $(\mathrm{mod}\, p)$, then $[a]^p = [a]$.*

Proof. Fix a prime p, and let $[a]$ be an arbitrary residue class $(\mathrm{mod}\, p)$. The theorem is obvious if $[a] = [0]$.

Assume $[a] \neq [0]$, so that $[a]$ is invertible in $\mathbf{Z}/p\mathbf{Z}$ by Corollary 6.24. If $[x]$ and $[y]$ are arbitrary residue classes $(\mathrm{mod}\, p)$, then $[x] = [y]$ if and only if $[ax] = [ay]$. (The forward direction is obvious; to prove the backward direction, multiply $[ax] = [ay]$ by $[a]^{-1}$.)

The $(p-1)$ residue classes $[a]$, $[2a]$, $\ldots$, $[(p-1)a]$ are all non-zero, and so are a rearrangement of $[1]$, $[2]$, $\ldots$, $[p-1]$. Consequently, the products of the elements in these two sets are equal:

$$\begin{aligned}[a^{p-1}(p-1)!] &= [a] \cdot [2a] \cdot [3a] \cdots [(p-1)a] \\ &= [1] \cdot [2] \cdot [3] \cdots [p-1] = [(p-1)!].\end{aligned}$$

But $[(p-1)!]$ is invertible as a product of invertible elements. Multiplying each side of the preceding equation by $[(p-1)!]^{-1}$ gives $[a]^{p-1} = [1]$. Multiplying each side of *this* by $[a]$ gives $[a]^p = [a]$. □

Corollary 6.38. *Let p be a prime. If a is an integer, $a^p \equiv a \pmod{p}$.*

Proof. Let a be an arbitrary integer, and use the division algorithm to write $a = kp + r$ with $0 \leq r \leq p-1$. Since $a^n \equiv r^n \pmod{p}$ for every positive integer n, and since $r^p \equiv r \pmod{p}$ by Theorem 6.37, we have $a^p \equiv r^p \equiv r \equiv a \pmod{p}$. □

Public-Key Cryptography

The fundamental problem of secure communication is the transmission of messages between two parties, conventionally known as Alice and Bob, in a way that an "eavesdropper" Charlie (i.e., someone with access to the raw content of Alice's and Bob's messages) cannot easily discover the semantic content (i.e., the *meaning*) of Alice's and Bob's messages.

For example, Alice might write Bob a letter on paper, then seal the letter in a strong metal box to which only Bob has the key, and ship the box to Bob. Or, Alice and Bob might agree on a secret code known only to them.

In actual practice on the Internet, the secure communication problem has two other constraints. First, the protocols of the Internet are, by design and intent, known (i.e., available) to everyone. Every text message or email, almost every telephone call, every web page, every electronic purchase travels across a network of machines, each having perfect eavesdropping access to the raw data constituting the

exchange. Second, a typical Internet user communicates with dozens, or thousands, or millions of other entities, each effectively a stranger; creating and distributing a dedicated private code for each pair of users is infeasible.

In 1977, two computer scientists, R. Rivest and A. Shamir, and a mathematician, L. Adelman, described the eponymous RSA *public-key cryptosystem* based on Fermat's little theorem. Each entity who wants to communicate privately creates a *key pair*, one *public* and one *private*.

Alice, for example, creates a key pair and publishes her public key. The public key can be used to *encrypt* arbitrary digital data in a way that is practically impossible to recover *unless one possesses Alice's private key*. Bob uses Alice's public key to encrypt his messages and sends them (in public, eavesdroppable form) over the Internet to Alice. Only Alice knows her private key, so in theory only Alice can decrypt Bob's message.

Digital Data

On contemporary (64-bit) computer hardware, text, audio, images, and video data is stored as sequences of *words*, each word comprising eight *bytes*, each byte having eight *bits* (from **b**inary dig**it**). Abstractly, a binary digit is one of a pair of contrasting states, 0 or 1. An eight-bit byte has $2^8 = 256$ possible values, and is the (hardware-dependent) smallest addressable unit of data. A word is the (again, hardware-dependent) data unit on which a processor naturally operates.

Most Latin-alphabet text is encoded using 7-bit ASCII, the American Standard Code for Information Interchange, which allows $2^7 = 128$ characters to be represented using one byte per character; or some part of Unicode, a 16-bit numbering scheme intended to provide universally compatible representation for all written human languages.

An arbitrary piece of text may be represented as a single string of bytes. For instance, a 400-page book might have 250 English words per page, with an average of five characters per word, for a total length of $500,000$ bytes (500K, or 0.5M).

Audio and image data are also encoded as strings of bytes. The simple encoding of WAV files uses two bytes to store amplitude of an audio signal sampled 44100 times per second (88200 bytes per channel per second, or about 1M per minute of stereo audio). Images are often stored as arrays of pixels, with each byte specifying the intensity of red, green, and blue for one pixel (24-bit color).

The point is, arbitrary data, whether your favorite novel, song, or movie, can be expressed as a single integer (possibly having billions of digits), or as a sequence of integers of bounded size (such as a few hundred to several million integers, each of 1000 digits).

To encrypt such data, we only need to "scramble" the set of thousand-digit integers in a way that can only be unscrambled by someone with the appropriate private key.

Mathematical Models of Encryption

Let X be a finite, non-empty set, whose elements we view as chunks of digital data. An "encryption scheme" is a bijection $f : X \to X$; the corresponding "decryption" is the inverse map $g : X \to X$, defined by $g\big(f(x)\big) = x$ for all x in X.

Example 6.39 (Rotation Ciphers)**.** Represent the letters A, ..., Z of the Roman alphabet as integers 0, ..., 25, or better, as residue classes (mod 26). An encryption scheme is a bijection $f : \mathbf{Z}/26\mathbf{Z} \to \mathbf{Z}/26\mathbf{Z}$.

The map $f(x) = x + 13$ defines the "Rot 13" cipher:

Message:	M	E	E	T	A	T	D	A	W	N
Encoding:	12	4	4	19	0	19	3	0	22	13
Apply f:	25	17	17	6	13	6	16	13	9	0
Cipher Text:	Z	R	R	G	N	G	Q	N	J	A

The *cipher text* "ZRRG NG QNJA" is meaningless to an eavesdropper; only the intended recipient knows to apply Rot 13 and retrieve the secret message.

More generally, if k is an integer, there is a "translation cipher" Rot k implemented by the function $f_k(x) = x + k$. If we place the letters of the alphabet cyclically around the rim of a disk, as on a child's encoder ring, the map f rotates the disk through $k/26$ of a turn. The map $g_k(x) = x - k$ inverts f.

The (additive) rotation ciphers of the preceding example are good childhood fun, but are easy to break. A few surprisingly small enhancements, however, turn the idea into a scheme sufficiently secure for military, diplomatic, and other private communications.

Example 6.40 (Multiplication Ciphers)**.** For simplicity, continue to represent the Roman alphabet by residue classes (mod 26). Pick an integer a coprime to 26 and an integer b, and consider the mapping

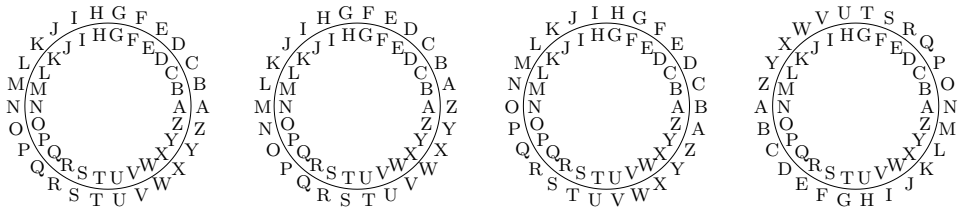


Figure 6.5: Decoder rings: Rot 0, Rot 1, Rot 25, and Rot 13.

$f(x) = ax + b$. The map $g(y) = a^{-1}(y - b)$ inverts f, and therefore decrypts messages encrypted by f.

Taking $a = 3$ and $b = 7$, for example, gives $a^{-1} = 9 \pmod{26}$, so $g(y) = 9y + 15$, since $9(3x + 7) + 15 \equiv 27x + 63 + 15 \equiv x \pmod{26}$. Our secret message encrypts as shown:

Message:	M	E	E	T	A	T	D	A	W	N
Encoding:	12	4	4	19	0	19	3	0	22	13
Apply f:	17	19	19	12	7	12	16	7	21	1
Cipher Text:	R	T	T	M	H	M	Q	H	V	B

Unlike a rotation cipher, which preserves ordering of letters, multiplication by a "scrambles" the ordering of the alphabet, further obfuscating the meaning of the cipher text "RTTM HM QHVB" to an eavesdropper. Or does it?

Any deterministic encryption scheme that merely substitutes letters as in the preceding examples is vulnerable to "attack", to decryption by an unauthorized recipient who has access to a sufficient amount of cipher text.

The number of bijective functions on the Roman alphabet is $26! \approx 4 \times 10^{26}$, too many to try by brute force. On the other hand, letters do not occur with equal frequency in English text; first letters of words are not equally-distributed; certain letters are doubled more frequently than others. For example, E comprises about 12.7% of "ordinary text", T about 9%, A 8.2%, O 7.5%, N 6.75%, R 6%, and so forth. Cipher text from a simple substitution can be trivially analyzed for letter frequencies, for first-letter frequencies, for occurrences of double letters. The eight most common letters account for about two-thirds of ordinary English text. There are only $8! = 40320$ permutations of eight characters, and even if all else fails, checking them all is trivial.

The set of Roman letters, even the 256-element set of bytes, is too small to furnish a safe hiding place for a secret message. One solution is

to divide data into chunks larger than one byte, but still of modest size. To illustrate, a line of text is about 64 characters long. The set of all 64-byte strings is finite, but inconceivably vast, with $256^{64} \approx 1.34 \times 10^{154}$ elements. To say this is larger than the number of elementary particles in the visible universe would be a comical understatement: If each elementary particle in the visible universe were itself a copy of the visible universe, the number of 64-byte strings would be on the order of the number of elementary particles in *all these universes put together*.

Imagine breaking a stream of data into 64-byte chunks. The chance of two chunks being identical is essentially the chance of the same line being deliberately encoded (such as a repeated refrain in song lyrics or a poem, or complete silence in a segment of audio, or a swatch of a single color in an image). The set of 64-byte strings that have ever been or ever will be published is an infinitesimal fraction of the set of *possible* strings. There is no hope of frequency analysis. To say the physical universe will undergo heat death long before patterns emerge in the data is again a comical understatement.

To exploit the vast space of modest-length strings, we only need some computationally-feasible means of "mixing", of mapping strings of some fixed length to other strings of that length, in a way that the inverse mapping is difficult to discover without additional information.

The RSA Algorithm

Definition 6.41. Let p and q be primes, and set $N = (p-1)(q-1)$. An integer e coprime to N is called an *encryption exponent*. Its inverse d in $(\mathbf{Z}/N\mathbf{Z})^\times$ is the corresponding *decryption exponent*.

The pair $m = pq$ and e is the associated *public key*. The RSA *encryption function* is the mapping $f : \mathbf{Z}/m\mathbf{Z} \to \mathbf{Z}/m\mathbf{Z}$ defined by $f(x) = x^e$.

The pair m and d is the associated *private key*. The RSA *decryption function* is the mapping $g : \mathbf{Z}/m\mathbf{Z} \to \mathbf{Z}/m\mathbf{Z}$ defined by $g(y) = y^d$.

Proposition 6.42. *If p and q are primes, $N = (p-1)(q-1)$, e is coprime to N, and $ed \equiv 1 \pmod N$, then $g\big(f(x)\big) = x^{ed} \equiv x \pmod m$ for all integers x.*

Proof. By hypothesis, $ed \equiv 1 \pmod N$, so there exists an integer k such that $ed = 1 + k(p-1)(q-1)$. For all x,

$$g\big(f(x)\big) \equiv (x^e)^d \equiv x^{ed} \equiv x \cdot (x^{(p-1)(q-1)})^k \pmod m.$$

If $[x] = [0]$ in $\mathbf{Z}/m\mathbf{Z}$, we are done. Otherwise, the proof of Fermat's little theorem gives

$$x^{p-1} \equiv 1 \pmod{p}, \qquad x^{q-1} \equiv 1 \pmod{q}.$$

Consequently,

$$x^{ed} = x \cdot (x^{(p-1)(q-1)})^k \equiv x \pmod{p},$$
$$x^{ed} = x \cdot (x^{(p-1)(q-1)})^k \equiv x \pmod{q}.$$

In other words, $p \mid x^{ed} - x$ and $q \mid x^{ed} - x$, which implies $x^{ed} \equiv x \pmod{m}$. □

Remark 6.43. In theory, decrypting an RSA-encoded message entails finding the prime factors p and q when only the product pq is known. While factoring is believed to be difficult, at this writing no mathematical result guarantees that RSA is safe against some other (as-yet undiscovered) attack.

The time required to factor m into pq is roughly proportional to the smaller of p and q. At this writing, a private key is believed to be secure if its prime factors are on the order of 1024 to 4096 bits, about 300 to 1200 digits.

The "prime number theorem" asserts that roughly one in 300 (about one third of one percent) of all 300-digit numbers are prime. Choosing a 300-digit prime therefore amounts to picking a needle from a haystack of some 10^{297} elements. A significant source of potential weakness is a computer's inability to generate random numbers on its own. All modern operating systems use the computer's interactions with the rest of the universe (key presses, mouse motions, arrival of network packets) to maintain an "entropy pool" suitable for generating "truly random" primes. If you have access to a GNU/Linux or MacOSX machine, you can see this in action: Open a terminal, run the command `cat /dev/random`, and watch random bytes scroll out as you move the mouse, or as the kernel communicates with its network router.

Example 6.44. Suppose Alice picks the primes $p = 11$ and $q = 13$, so $m = pq = 143$ and $N = (p-1)(q-1) = 120$. Picking $e = 7$ as encryption exponent (the smallest available choice, since $N = 5!$), Alice finds the decryption exponent d by solving $7d \equiv 1 \pmod{120}$. Since $120 = 7 \cdot 17 + 1$, she has $d \equiv -17 \equiv 103 \pmod{120}$.

Alice's public key is $(m, e) = (143, 7)$. This is shared freely with the world. Anyone with access to the public key (and suitable software)

can encrypt data in a way impossible to recover without the private key.

Alice's private key is $(m, d) = (143, 103)$. This, along with either prime factor p or q, is closely guarded; anyone with the private key can decrypt private messages meant for Alice, and anyone with either prime factor of m can easily calculate the decryption exponent.

To arrange a meeting with Alice, Bob embeds the alphabet into $\mathbf{Z}/143\mathbf{Z}$, letting A through Z correspond to 0 through 25 as before, and mapping the space character to 26. (In a more realistic setting, Bob might use ASCII to encode upper and lower case letters, digits, and punctuation marks in the last seven bits of a byte, i.e., as residue classes modulo $2^7 = 128$, then map these into $\mathbf{Z}/143\mathbf{Z}$.)

The mapping $f(x) = x^e = x^7$ encodes the message as shown:

Message:	M	E	E	T		A	T		D	A	W	N
Encoding:	12	4	4	19	26	0	19	26	3	0	22	13
Apply f:	17	82	82	46	104	0	46	104	42	0	22	117

The sequence of values *is* the cipher text.

To decrypt Bob's message, Alice applies $g(y) = y^d = y^{103}$, then converts the resulting string of residue classes (mod 143) into ordinary characters. At dawn the next day, they rendezvous in the field behind the old elm tree, decide how to assassinate the king, and agree for convenience that next time they'll grab lunch at a cafe.

Exercises

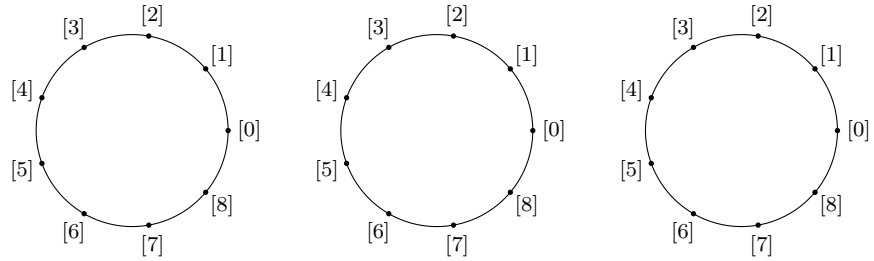
Exercise 6.1. Calculate the following:
(a) $3^{487} \pmod{28}$. (b) $3^{120} \cdot 5^{531} \pmod{26}$. (c) $7^{97} \pmod{11}$.

Exercise 6.2. Solve: $18x \equiv 6 \pmod{24}$. $18x \equiv 6 \pmod{28}$.

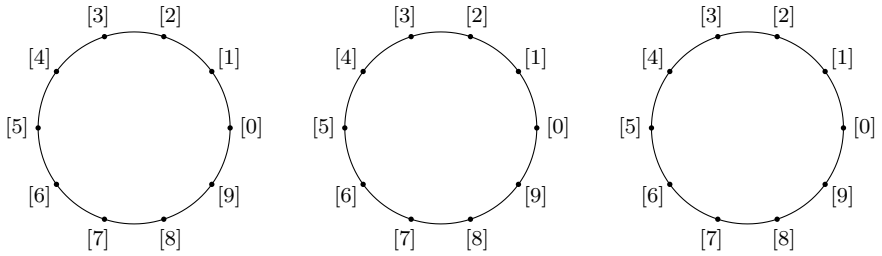
Exercise 6.3. Solve: $150x = 84 \pmod{567}$. $150x = 84 \pmod{210}$.

Exercise 6.4. Let $(\mathbf{Z}/7\mathbf{Z})^\times = \{[1], [2], [3], [4], [5], [6]\}$ be the set of units mod 7. Write out the Cayley table for $((\mathbf{Z}/7\mathbf{Z})^\times, \cdot)$. For each class $[a]$, find the set of powers of $[a]$. Is the set of powers ever all of $(\mathbf{Z}/7\mathbf{Z})^\times$? If so, which element(s) "generate"?

Exercise 6.5. List the elements of $(\mathbf{Z}/9\mathbf{Z})^\times$ and write out the Cayley table for multiplication. For $[a] = [2], [3], [4]$ in $\mathbf{Z}/9\mathbf{Z}$, sketch the multiplication diagram for $[a]$.



Exercise 6.6. List the elements of $(\mathbf{Z}/10\mathbf{Z})^\times$ and write out the Cayley table for multiplication. For $[a] = [2], [3], [4]$ in $\mathbf{Z}/10\mathbf{Z}$, sketch the multiplication diagram for $[a]$.



Exercise 6.7. List the elements of $(\mathbf{Z}/15\mathbf{Z})^\times$, write out the Cayley table for multiplication, and determine the inverse of each element.

Exercise 6.8. List the elements of $(\mathbf{Z}/18\mathbf{Z})^\times$, write out the Cayley table for multiplication, and determine the inverse of each element.

Exercise 6.9. For the residue classes specified, determine whether or not each class is invertible in $\mathbf{Z}/n\mathbf{Z}$, and if so, find the inverse.

(a) In $\mathbf{Z}/48\mathbf{Z}$: $[a] = [17]$, $[a] = [21]$, $[a] = [25]$.

(b) In $\mathbf{Z}/140\mathbf{Z}$: $[a] = [35]$, $[a] = [33]$, $[a] = [81]$.

(c) In $\mathbf{Z}/101\mathbf{Z}$: $[a] = [64]$, $[a] = [100]$.

In the exercises below, characters are encoded using the "visible subset" of ACII from 32 to 126:

	0	1	2	3	4	5	6	7	8	9
30			␣	!	”	#	$	%	&	’
40	(	)	*	+	,	-	.	/	0	1
50	2	3	4	5	6	7	8	9	:	;
60	<	=	>	?		A	B	C	D	E
70	F	G	H	I	J	K	L	M	N	O
80	P	Q	R	S	T	U	V	W	X	Y
90	Z	[	\	]	^	_	‘	a	b	c
100	d	e	f	g	h	i	j	k	l	m
110	n	o	p	q	r	s	t	u	v	w
120	x	y	z	{	|	}	∼			

For example, "Hello, world!" would be encoded

72	101	108	108	111	44	32	119	111	114	108	100	33
H	e	l	l	o	,	␣	w	o	r	l	d	!

Exercise 6.10. Alice's public key is $(m, d) = (161, 7)$.

(a) Calculate Alice's decryption key.

(b) Bob sends the message

52 132 143 143 132 21 152

Decrypt Bob's message.

Exercise 6.11. Bob's public key is $(m, d) = (161, 13)$.

(a) Calculate Bob's decryption key.

(b) Decrypt Alice's message:

127 150 154 81 14 6 76 117 150 65 81 154 6 76 81 156 77 115 56

Exercise 6.12. Pick two 2-digit primes p and q and an integer d coprime to $N = (p-1)(q-1)$. Encrypt a short message, exchange messages and public keys with a classmate, and decrypt each others' messages.

Chapter 7

Mappings and Relations

Recall that if A and B are sets, then a *mapping* $f : A \to B$ is a subset $f \subseteq A \times B$ of the Cartesian product satisfying the condition:

For every a in A, there exists a unique b in B such that $(a, b) \in f$.

As in calculus, a mapping is a rule associating a unique "output" to each "input", but the domain and codomain, the sets of allowable inputs and potential outputs, are an intrinsic part of the definition.

Example 7.1. Consider the familiar squaring function $f(x) = x^2$, where x ranges over the set of real numbers. If we set $y = f(x)$, we might wish to "solve" for x in terms of y. At first glance this is trivial: set $x = \sqrt{y}$. Unfortunately, closer inspection reveals two fatal flaws. First, if $y < 0$, there is no real x satisfying $x^2 = y$. In this context, the square root is *undefined*. Second, if $y > 0$, there exist *two* values of x with $x^2 = y$; the input x is not a function of the output y, so the square root is *not well-defined*. In either event, we have not associated a unique output to each input.

In high school, you learned to avoid complications with square roots by only considering non-negative numbers y, and agreeing that $\sqrt{y}$ always refers to the non-negative square root. Technically you are no longer inverting the function $f(x) = x^2$ with x real, but a *different function defined by the same formula*, for which the allowable inputs and potential outputs have been explicitly restricted.

Example 7.2. Consider longitude (measured in degrees) as a function of position on the earth. Upon circumnavigating the earth to the east, longitude increases by 360°. But this cannot be the whole story; if it

were, each geographic location would have multiple longitudes, any two differing by a whole multiple of 360°.

Instead, when you circumnavigate the globe in an eastward direction, you must cross a line where longitude "jumps down" by 360°. This discontinuity is a mathematical artifact of the impossibility of inverting sine and cosine to recover longitude continuously as a real-valued function of position on the earth.

The earth is approximately spherical and rotates with respect to the distant stars. A *sidereal day*, or 24 hours, is the time required for the earth to rotate 360° with respect to the stars. This *duration* is the same for all points on the earth, but the *starting time* (midnight) depends on one's longitude. By international treaty, the earth's surface is divided into twenty-four *time zones*, each a sector of longitude 15° wide (with substantial allowances for geographical and political boundaries). The times in neighboring zones differ by one hour.

The global discontinuity of longitude has a notable practical consequence: the existence of the International Date Line, an imaginary "cut" along the surface of the earth joining the south and north poles, along which local time "jumps" by 24 hours, affecting global travelers and international stock traders alike.

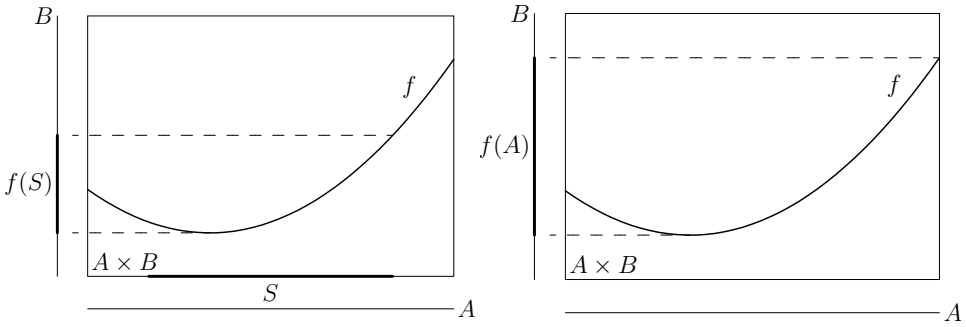
7.1 Images, and Preimages

A variety of mathematical considerations all but require that we study how functions act on sets.

Definition 7.3. Let $f : A \to B$ be a mapping. If $S \subseteq A$, we define the *image* of S under f to be the set

$$f(S) = \{b \text{ in } B : b = f(s) \text{ for some } s \text{ in } S\} \subseteq B.$$

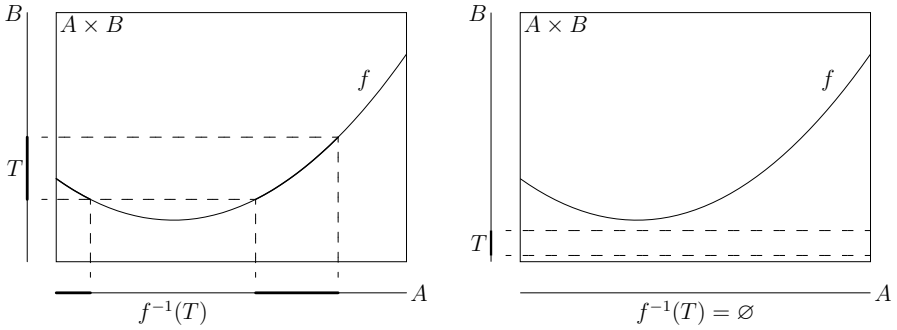
The *image of* f is the set $f(A) \subseteq B$ of all values of f.



Definition 7.4. Let $f : A \to B$ be a mapping. If $T \subseteq B$, we define the *preimage* of T under f to be the set

$$f^{-1}(T) = \{a \text{ in } A : f(a) \in T\} \subseteq A$$

of elements of the domain mapped into T by f.



Remark 7.5. A mapping $f : A \to B$ may be viewed as a "poll" taken of a population A, with responses in the set B. The image under f of a set $S \subseteq A$ is the set of responses from individuals in S. The preimage of a set $T \subseteq B$ is the set of individuals whose responses are in T.

Example 7.6. If A is a non-empty set, we define the *identity mapping* $I_A : A \to A$ by $I_A(a) = a$ for all a in A. Under the identity map, every set is its own image, and its own preimage.

Example 7.7. Let A and B be non-empty sets. For each b in B, there is a *constant mapping* $c_b : A \to B$ defined by $c_b(a) = b$ for all a in A. The image of an arbitrary non-empty subset of A is the singleton $\{b\}$. The preimage of a set T is either the empty set (if $b \notin T$) or the entire domain A (if $b \in T$).

Proposition 7.8. *Let $f : A \to B$ be a mapping, S_1 and S_2 subsets of A, and T_1 and T_2 subsets of B. Then*

(i) $f(S_1 \cup S_2) = f(S_1) \cup f(S_2)$.

(ii) $f(S_1 \cap S_2) \subseteq f(S_1) \cap f(S_2)$.

(iii) $f^{-1}(T_1 \cup T_2) = f^{-1}(T_1) \cup f^{-1}(T_2)$.

(iv) $f^{-1}(T_1 \cap T_2) = f^{-1}(T_1) \cap f^{-1}(T_2)$.

Proof. To prove two sets are equal, we must establish inclusions in both directions. Assume S_1 and S_2 are subsets of A.

(The inclusion $f(S_1 \cup S_2) \subseteq f(S_1) \cup f(S_2)$). If $b \in f(S_1 \cup S_2)$, then by definition there exists an element a in $S_1 \cup S_2$ such that $f(a) = b$. Since either $a \in S_1$ or $a \in S_2$ by definition of the union of sets, either $b \in f(S_1)$ or $b \in f(S_2)$, which means $b \in f(S_1) \cup f(S_2)$. This proves $f(S_1 \cup S_2) \subseteq f(S_1) \cup f(S_2)$.

(The inclusion $f(S_1) \cup f(S_2) \subseteq f(S_1 \cup S_2)$). If $b \in f(S_1) \cup f(S_2)$, there exists an a in $S_1 \subseteq S_1 \cup S_2$ such that $f(a) = b$ or there exists an a in $S_2 \subseteq S_1 \cup S_2$ such that $f(a) = b$. In either case, there exists an a in $S_1 \cup S_2$ such that $f(a) = b$, which means $b \in f(S_1 \cup S_2)$. This proves $f(S_1) \cup f(S_2) \subseteq f(S_1 \cup S_2)$.

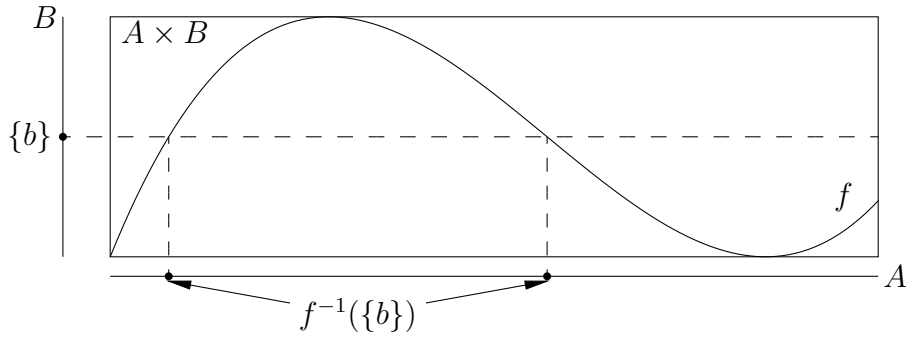
The other parts are entirely similar, and are left to you. □

7.2 Surjectivity and Injectivity

Our inability to invert the map $f : \mathbf{R} \to \mathbf{R}$, $f(x) = x^2$, had two aspects: When we wrote $y = f(x)$, some y were associated with no values of x, and some were associated with multiple values of x.

Definition 7.9. A mapping $f : A \to B$ is *surjective* if for every b in B, there exists an a in A such that $f(a) = b$.

A mapping f is surjective if the preimage $f^{-1}(\{b\})$ is non-empty for each b in B, i.e., if the image of f is the entire codomain, $f(A) = B$. Geometrically, $f : \mathbf{R} \to \mathbf{R}$ is surjective if every horizontal line hits the graph of f at least once.



Definition 7.10. Let $f : A \to B$ be a mapping. Distinct points a_1 and a_2 in A are *identified by* f if $f(a_1) = f(a_2)$, namely if a_1 and a_2 are mapped to the same value by f.

A mapping f is *injective* if no distinct points are identified by f, namely, if $a_1 \neq a_2$, then $f(a_1) \neq f(a_2)$. Contrapositively, $f(a_1) = f(a_2)$ implies $a_1 = a_2$.

A mapping f is injective if every preimage $f^{-1}(\{b\})$ contains at most one element. Geometrically, $f : \mathbf{R} \to \mathbf{R}$ is injective if every horizontal line hits the graph of f at most once.

Remark 7.11. Continuing Remark 7.5, a mapping $f : A \to B$ is surjective if every allowable answer to the poll is given by at least one individual. Similarly, f is injective if no two people give the same response; knowledge of the response uniquely determines the individual who gave that response.

Definition 7.12. A mapping $f : A \to B$ is *bijective* if f is both surjective and injective.

Remark 7.13. If $f : A \to B$ is bijective, then each element a in A corresponds to exactly one element b in B. Procedurally, f "relabels" elements of the set A using elements of B as names.

Remark 7.14. Two finite sets A and B have "the same number of elements" if there exists a bijection $f : A \to B$. That is, bijections formalize the idea of two sets having the same size. This is true even for *infinite* sets.*

Example 7.15. Define $f : \mathbf{Z} \to \mathbf{Z}$ by $f(a) = 1-a$. Prove f is bijective.

(Injectivity). Let a_1 and a_2 be arbitrary integers, and assume that $f(a_1) = f(a_2)$. By the definition of f, $1 - a_1 = 1 - a_2$, so $a_1 = a_2$ by elementary algebra. Since a_1 and a_2 were arbitrary, f is injective.

(Surjectivity). Informally, we wish to solve $b = f(a) = 1 - a$ for a in terms of b. Rearrangement gives $a = 1 - b$.

Formally, let b be an arbitrary integer, and consider the integer $a = 1 - b$. Since $f(a) = f(1 - b) = 1 - (1 - b) = b$, we have shown that for every integer b, there exists an integer a such that $f(a) = b$. This means f is surjective.

*A set A is *infinite* if there exists a bijection from A to some proper subset of A. Intuitively, removing one element does not change the number of elements.

Example 7.16. Let $f : \mathbf{Z} \to \mathbf{Z}$ be defined by $f(a) = 1 - 2a$. Prove f is injective (one-to-one) but not surjective (onto).

(Injectivity). Let a_1 and a_2 be integers, and assume $f(a_1) = f(a_2)$, i.e., $1 - 2a_1 = 1 - 2a_2$. Subtracting the left side from the right gives $0 = 2a_1 - 2a_2 = 2(a_1 - a_2)$. By Theorem 4.6 (ii), $a_1 - a_2 = 0$ as well. Since $f(a_1) = f(a_2)$ implies $a_1 = a_2$, f is injective.

(Non-surjectivity). To show f is not surjective, it suffices to exhibit an integer not in the image of f. Let $b = 0$. The equation $f(a) = b$ becomes $1 - 2a = 0$, or $1 = 2a$. There exists no integer a satisfying this condition, which means 0 is not in the image of f.

Example 7.17. Define $f : \mathbf{Z} \to \mathbf{Z}$ by $f(a) = a^2$. We will show f is neither injective nor surjective.

Since $f(1) = 1 = f(-1)$ but $1 \neq -1$, f is not injective.

Since $f(a) \geq 0$ for all integers a, there exists no integer satisfying $f(a) = -1$, implying f is not onto.

Example 7.18. Define $f : \mathbf{Z}^+ \to \mathbf{Z}$ by

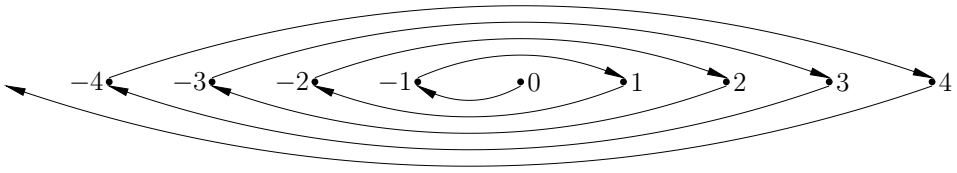
$$f(a) = \begin{cases} \dfrac{a-1}{2} & \text{if } a \text{ is odd,} \\ -\dfrac{a}{2} & \text{if } a \text{ is even.} \end{cases}$$

Prove f is bijective. (Informally, there are just as many positive integers as there are integers!)

(Initial exploration). To understand f intuitively, list its first several values. The inputs (elements of the domain) are 1, 2, 3, 4, ... To find an output, determine whether the input is even or odd, and evaluate the corresponding formula. Thus $f(1) = 0$ (1 is odd), $f(2) = -1$ (2 is even), $f(3) = 1$, $f(4) = -2$, $f(5) = 2$, and so forth:

a	1	2	3	4	5	6	7	8	9
$f(a)$	0	-1	1	-2	2	-3	3	-4	4

In words, f alternately "counts off" one negative, one positive. Using arrows to indicate successive values:



Since the same value is never achieved twice, f is injective. Since every integer value is achieved, f is surjective. We must convert this intuition into a formal proof.

(Injectivity). Let a_1 and a_2 be integers, and assume $f(a_1) = f(a_2)$. Because f is defined "piecewise", it's most convenient to consider three separate cases.

Case 1: a_1 and a_2 both odd. By hypothesis and the definition of f, $(a_1 - 1)/2 = (a_2 - 1)/2$, and elementary algebra implies $a_1 = a_2$.

Case 2: a_1 and a_2 both even. Here, $-a_1/2 = -a_2/2$, and again we find $a_1 = a_2$.

Case 3: a_1 and a_2 have opposite parity (one is odd, one is even). Without loss of generality, we may assume a_1 is odd and a_2 is even. (Otherwise, swap their names.) Since $f(a_2) < 0 \leq f(a_1)$, the hypothesis $f(a_1) = f(a_2)$ is false. Said contrapositively, if $f(a_1) = f(a_2)$, we are not in Case 3.

Since the conclusion $a_1 = a_2$ followed in each case, we have shown f is injective.

(Surjectivity). Let b be an arbitrary integer, and consider two cases:

Case 1: $b < 0$. Let $a = -2b$. Since a is an even integer, we have $f(a) = -a/2 = b$; there exists an a such that $f(a) = b$ provided $b < 0$.

Case 2: $0 \leq b$. Let $a = 1 + 2b$. Since a is odd, $f(a) = (a - 1)/2 = 2b/2 = b$; there exists an a such that $f(a) = b$ provided $0 \leq b$.

Since every integer b is either negative or non-negative, we have handled all possibilities. In each case, there exists an integer a such that $f(a) = b$, so f is onto.

Example 7.19. Let A be an arbitrary set, and let $\mathscr{P}(A)$ be its power set. The following argument of G. Cantor shows there is no surjection $f : A \to \mathscr{P}(A)$.

Let $f : A \to \mathscr{P}(A)$ be an arbitrary mapping. For each a in A, the value $f(a)$ is a *subset* of A, so the statement $a \in f(a)$ is meaningful for each a. Let

$$T = \{a \text{ in } A : a \notin f(a)\}.$$

To prove f is not surjective, it suffices to show $f(t) \neq T$ for all t in A. We will prove that if $f(t) = T$ for some t, then set theory is logically inconsistent. Contrapositively, if set theory is logically consistent then $f(t) \neq T$ for all t in A.

If $f(t) = T$, we may ask which alternative is true: $t \notin T$ or $t \in T$. By the definition of T, if $t \in f(t) = T$, then t fails to satisfy the

criterion for membership in T, so $t \notin T$. However, if $t \notin f(t) = T$, then t satisfies the criterion of membership, so $t \in T$. In summary, the statement $t \in T$ is logically equivalent to its negation $t \notin T$. This completes the proof.

7.3 Composition and Inversion

Definition 7.20. Let $f : A \to B$ and $g : B \to C$ be mappings. Their *composition* is the mapping $g \circ f : A \to C$ defined by

$$(g \circ f)(a) = g\big(f(a)\big) \quad \text{for each } a \text{ in } A.$$

In this situation we say g *is composable with* f.

Remark 7.21. In words, plug the output of f into g; the resulting output is $(g \circ f)(a)$.

When context clearly signifies composition of functions, the operator symbol $\circ$ may be omitted, and the composition $g \circ f$ denoted gf.

Proposition 7.22. *Mapping composition is associative: If $f : A \to B$, $g : B \to C$, and $h : C \to D$ are composable mappings, then $h(gf) = (hg)f$ as mappings from A to D.*

Proof. If a is an arbitrary element of A, then

$$\begin{aligned}\big[h \circ (g \circ f)\big](a) &= h\big[\big(g \circ f\big)(a)\big] = h\big[g\big(f(a)\big)\big] \\ &= \big(h \circ g\big)\big(f(a)\big) = \big[\big(h \circ g\big) \circ f\big](a).\end{aligned}$$ □

Surjectivity and injectivity of mappings f and g are related to whether or not the composition gf is surjective and/or injective. Think of two functions "cooperating", with g acting on the output of f. If f achieves every value in B and g achieves every value in C, then in tandem they achieve every value in C. Similarly, if neither g nor f identifies any pair of distinct points, then gf does not either. Before reading further, you should express these observations formally as logical implications and try to prove them.

Proposition 7.23. *Let $f : A \to B$ and $g : B \to C$ be mappings.*

(i) *If f and g are surjective, then gf is surjective.*

(ii) *If f and g are injective, then gf is injective.*

Proof. (i). Suppose $f : A \to B$ and $g : B \to C$ are surjective. Let c in C be arbitrary. Because g is surjective, there exists a b in B such that $g(b) = c$. Since f is surjective, there exists an a in A such that $f(a) = b$. But $(gf)(a) = g\big(f(a)\big) = g(b) = c$. We have shown that for every c in C, there exists an a in A such that $(gf)(a) = c$, which by definition means gf is surjective.

(ii). Exercise 7.7 (a). □

Conversely, suppose we know gf is surjective, or that gf is injective. What can we deduce about f and g?

In our cooperation metaphor, if gf achieves every value in C, then g itself must as well, since any value not achieved by g is certainly not achieved by gf. Thus, if gf is surjective, then g is surjective.

Similarly, if f identifies some pair of points, then gf identifies that pair as well, since g cannot split asunder what f has joined. Formally, if gf is injective, then f is injective, Exercise 7.7 (b).

Nothing more can be deduced:

Example 7.24. If $A = \{1\}$ and $B = \{-1, 1\}$, then the mapping $f : A \to B$ defined by $f(1) = 1$ is injective but not surjective and the unique mapping $g : B \to A$ is surjective but not injective. The composition $gf : A \to A$ is the identity map (which is bijective), while $fg : B \to B$ is neither injective nor surjective.

Example 7.25. An arbitrary mapping $f : A \to B$ can be "factored" into the composition of an injection followed by a surjection. Define $\gamma_f : A \to A \times B$ and $\pi_2 : A \times B \to B$ by

$$\gamma_f(a) = \big(a, f(a)\big), \qquad \pi_2(a, b) = b.$$

Geometrically, "γ_f lifts a to the graph of f" and "π_2 projects $A \times B$ onto the second factor." Clearly, $f = \pi_2 \circ \gamma_f : A \to B$, γ_f is injective, and π_2 is surjective.

Inversion of Mappings

Definition 7.26. Let A and B be sets. A mapping $f : A \to B$ is *invertible* if there exists a mapping $g : B \to A$ that *inverts* f, i.e., such that $g \circ f$ is the identity map of A and $f \circ g$ is the identity map of B.

Remark 7.27. If $f : A \to B$ is invertible and $g : B \to A$ inverts f, then $(g \circ f)(a) = a$ for all a in A and $(f \circ g)(b) = b$ for all b in B. That is,

(7.1) For all a in A and all b in B, $g(b) = a$ if and only if $f(a) = b$.

Proposition 7.28. *Let A and B be sets, $f : A \to B$ a mapping.*

(i) *f is invertible if and only if f is bijective.*

(ii) *If f is invertible, there exists a unique map inverting f.*

Remark 7.29. Both conclusions hold (with essentially vacuous proof) if either A or B is the empty set. It suffices to assume A, B are non-empty.

Proof. (i). Assume f is invertible, and let g be a mapping that inverts f, i.e., that satisfies $gf = I_A$ and $fg = I_B$. If $f(a_1) = f(a_2)$ for some a_1 and a_2 in A, applying g to both sides gives $a_1 = a_2$, so f is injective. If b is an arbitrary element of b, and if $a = g(b)$, then $f(a) = (fg)(b) = b$, so f is surjective.

Conversely, suppose f is bijective: For each b in B, there exists a unique a in A such that $b = f(a)$. Define $g(b) = a$. This prescription defines a mapping $g : B \to A$ that satisfies (7.1), so f is invertible.

(ii). If g_1, $g_2 : B \to A$ invert f, then

$$g_1 = g_1 \circ I_B = g_1 \circ (f \circ g_2) = (g_1 \circ f) \circ g_2 = I_B \circ g_2 = g_2. \qquad \square$$

A mapping f is invertible if and only if f is injective and surjective. We now consider what happens if each condition holds individually.

Left Inverses

Assume f is one-to-one; not every element of B need be a value of f, but every value (every element of $f(A)$, the image of A under f) is achieved exactly once. We may define $h : f(A) \to A$ by the analog of (7.1): For all b in $f(A)$, $h(b) = a$ if and only if $f(a) = b$.

If we apply f to a in A, then apply h to $b = f(a)$, we find

$$(hf)(a) = h\big(f(a)\big) = h(b) = a \qquad \text{for all } a \text{ in } A.$$

That is, $hf = I_A$, the identity map on A; h is a *left inverse* of f.*

*In general, $fh \neq I_B$, the identity map on B, since (i) h is defined only on the image of f, and (ii) the image of fh, which is a subset of the image of f, may be a *proper* subset of B.

In order to obtain a mapping $g : B \to A$ satisfying $gf = I_A$, we must "enlarge" the domain of h; any convenient "rule" will do. For example, pick an element a_0 in A arbitrarily, and define, for b in B,

$$g(b) = \begin{cases} a & \text{if } b = f(a) \text{ for some } a \text{ in } A \\ a_0 & \text{otherwise} \end{cases}$$

The easy verification that $gf = I_A$ is left to you.

Example 7.30. Define $f : \mathbf{R} \to \mathbf{R}$ by $f(x) = e^x$, see Figure 7.1, left. For each $y > 0$ (namely, for each y in the image of f), we have $y = e^x$ if and only if $x = \ln y$. Define

$$g(y) = \begin{cases} \ln y & \text{if } y > 0, \\ 0 & \text{if } y \leq 0, \end{cases}$$

see Figure 7.1, right. Then $(gf)(x) = g\big(f(x)\big) = x$ for all x; what about $f\big(g(y)\big)$?

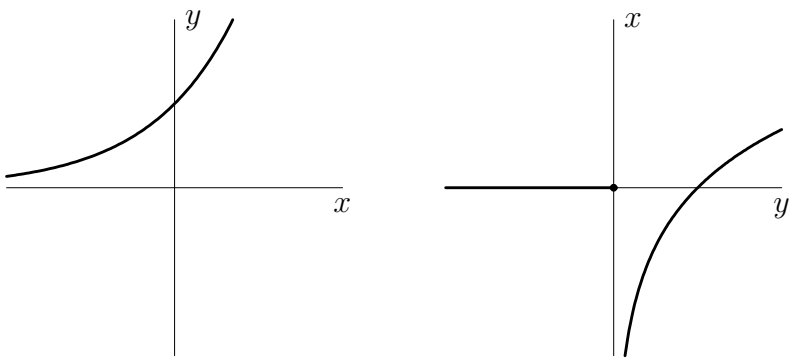


Figure 7.1: A left inverse of $f(x) = e^x$.

Right Inverses

Assume f is onto; every element of B is a value of f, but some values b may be achieved at distinct points: $f(a_1) = f(a_2)$ but $a_1 \neq a_2$. Define $g : B \to A$ by the following prescription: For each b in B, use the Axiom of Choice to pick an a in A such that $f(a) = b$, and define $g(b) = a$.*

*The Axiom of Choice asserts that if $\{S_i\}_{i \in I}$ is a collection of non-empty sets indexed by a set I, it is possible to choose, for each i in I, an element x_i of S_i. For

It is straightforward to check that $fg = I_B$, the identity map on B.* Any particular g defined this way is called a *branch* of f^{-1}.

7.4 Equivalence Relations

Definition 7.31. Let A be a non-empty set. A *relation* on A is a subset $R \subseteq A \times A$. Elements a and b of A are *R-related*, written aRb, if $(a, b) \in R$.

Example 7.32. The *equality* relation $=$ on A is defined by the *diagonal* $R = \Delta = \{(a, a) : a \in A\}$.

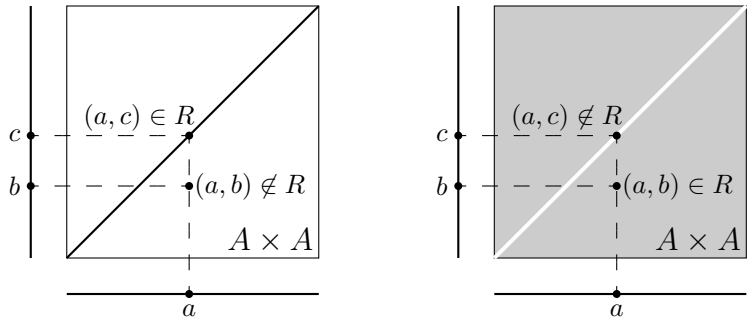


Figure 7.2: Equality and inequality: $a \neq b$ and $a = c$.

Example 7.33. The *inequality* relation $\neq$ is the complement of the equality relation, $A \times A \setminus \Delta = \{(a_1, a_2) : a_1 \neq a_2\}$.

Remark 7.34. Generally, if $R_2 = A \times A \setminus R_1$, the relations R_1 and R_2 are logical opposites: One relation holds for a pair of elements if and only if the other fails for the same pair.

Example 7.35. Let $A = \mathbf{Z}$ be the set of integers. The *less-than* relation is the set $R = \{(n_1, n_2) : n_1 < n_2\}$.

Example 7.36. Again, let $A = \mathbf{Z}$. The *parity* relation on $\mathbf{Z}$ is the set $R = \{(n_1, n_2) : n_2 - n_1 \text{ is even}\}$. Two integers are related if and only if they are both even or both odd.

---

finite or countable index sets I, choice is not very controversial. For "uncountably infinite" index sets, choice is not universally-accepted among mathematicians.

*In general, $gf \neq I_A$, the identity map on A, since if $f(a_1) = b = f(a_2)$ for $a_1 \neq a_2$, we cannot have both $g(b) = a_1$ and $g(b) = a_2$.

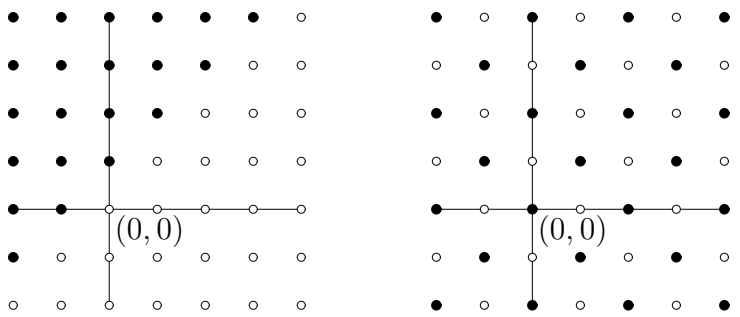


Figure 7.3: Less-than and parity on **Z**.

Example 7.37. Let $f : A \to A$ be a mapping. Viewing f as a subset of $A \times A$ defines the *maps-to-under-f* relation on A: aRb if and only if $f(a) = b$, if and only if a maps to b under f.

Definition 7.38. Let R be a relation on a set A. We say R is

- *reflexive* if aRa for all a in A;
- *symmetric* if, for all a and b in A, aRb implies bRa;
- *transitive* if, for all a, b, and c in A, aRb and bRc imply aRc.

Example 7.39. Though not a formal example, the "friendship" relation may help you assimilate the conditions in the preceding definition. Let A be some set of people, and let aRb mean "b is a friend of a".

R is reflexive if and only if every person is their own friend; R is symmetric if and only if all friendships are mutual; R is transitive if and only if every friend-of-a-friend is a friend.

Definition 7.40. A reflexive, symmetric, and transitive relation is an *equivalence relation.*

If R is an equivalence relation on A and $a \in A$, the *equivalence class* of a is the set

$$[a] = \{x \text{ in } A : aRx\} \subseteq A$$

comprising all elements related to a.

Example 7.41. Equality is an equivalence relation on an arbitrary set: For all a, b, and c, we have $a = a$ (reflexivity), $a = b$ implies $b = a$ (symmetry), and if $a = b$ and $b = c$, then $a = c$ (transitivity).

Inequality is symmetric, but neither reflexive nor transitive.

Less-than is transitive (if $a < b$ and $b < c$, then $a < c$), but neither reflexive nor symmetric.

Example 7.42. Parity is an equivalence relation on the set of integers: For all a, b, and c, $a-a$ is even (reflexivity), if $b-a$ is even (aRb) then $a-b$ is even (bRa), and if $b-a$ and $c-b$ are even (aRb and bRc), then $c-a=(c-b)+(b-a)$ is even (aRc).

Generally, if $n>2$ is an integer then congruence $(\mathrm{mod}\, n)$ is an equivalence relation on the set of integers, Exercise 7.11.

Equivalence Relations and Partitions

Let A be a non-empty set. Recall that a *partition* of A is a collection of non-empty, disjoint subsets whose union is A. Partitions and equivalence relations are two ways of viewing a single mathematical structure: Every equivalence relation gives rise to a partition, every partition gives rise to an equivalence relation, and these associations are inverse to each other.

Proposition 7.43. *Let R be an equivalence relation on A. The equivalence classes of R partition A.*

Proof. Since $a\in[a]$ for each a, every element of A lies in at least one equivalence class. It remains to prove that two arbitrary equivalence classes $[a]$ and $[b]$ are either disjoint or identical. To prove this it suffices to show that if $[a]\cap[b]\neq\varnothing$ (i.e., the classes are not disjoint), then $[a]=[b]$.

Let's first run through the argument using the friendship metaphor. If a and b have a friend in common, then a and b are themselves friends (transitivity). Consequently, every friend of a is a friend of b (transitivity again) and *vice versa*, so a and b have exactly the same set of friends.

Formally, if $[a]\cap[b]\neq\varnothing$, there exists a c in A such that $c\in[a]\cap[b]$. Consequently, aRc and bRc. By symmetry of R, aRc and cRb, and by transitivity aRb. This means $a\in[b]$ and $b\in[a]$.

It is now easy to prove $[a]\subseteq[b]$ and $[b]\subseteq[a]$: If $x\in[a]$, then xRa, and since aRb, transitivity guarantees xRb, meaning $x\in[b]$. Reversing the roles of a and b completes the argument.

We have shown that non-disjoint equivalence classes are identical, so the set of equivalence classes of R is indeed a partition of A. □

Remark 7.44. Conversely, if A is partitioned into subsets $\{A_i\}_{i\in I}$, there is an induced equivalence relation defined by aRb if and only if there exists an index i such that $a\in A_i$ and $b\in A_i$. Informally, aRb if and

only if both elements lie in the same subset of the partition. Be sure to convince yourself that R is an equivalence relation, and that the partition induced by R is the original partition.

Example 7.45. The equivalence classes of the equality relation are the singletons, sets having one element: $[a] = \{a\}$ for each a in A.

Example 7.46. The parity relation on $\mathbf{Z}$ has two equivalence classes: $[0] = 2\mathbf{Z}$ and $[1] = 2\mathbf{Z} + 1$.

Partitions and Prejudice

Our minds organize the external world by categorizing, unconsciously identifying people, objects, or phenomena that share some attribute.

Example 7.47. A physicist, a statistician, and a mathematician saw a flock of 100 sheep, of which one was black. The physicist said, "We can deduce that one in 100 sheep is black." The statistician said, "No, only that *in this sample of* 100 *sheep*, one is black." The mathematician corrected, "No, we can only deduce that one sheep in this sample is black *on one side*."

Often we cope fluently with such hierarchies: a particular mandarin orange, mandarin oranges, oranges, citrus fruit, fruit.... At other times, prejudice deceives us into identifying individuals according to superficial characteristics (such as gender, ethnicity, religion, or scientific field) and incorrectly presuming "all such people are alike".

In mathematics, we can sometimes turn prejudice to good use. Perhaps we don't care which integer we're dealing with, but only if it's even or odd, or if it leaves a remainder of 5 on division by 12. Maybe we're dealing with pairs of points in the plane, but don't care where they're located, only that the second is located one unit to the right of the first. In such cases, an equivalence relation allows us to formalize our prejudice and discard irrelevant information.

Let A be a set, R an equivalence relation on A, and $\{A_i\}_{i \in I}$ the partition of A into equivalence classes. Each "index" i is associated with the non-empty set $A_i \subseteq A$, and the index set I is in bijective correspondence with the set of equivalence classes. We call the set of equivalence classes the *quotient* of A by R, denoted $I = A/R$ and read "A modulo R" (or "A mod R" for short). *Elements* of A/R are *collections* of objects in A. The equivalence relation R is "unable to distinguish" elements of A_i, so when R "looks at" A it "sees" $I = A/R$.

Mappings and Equivalence Classes

Let A be non-empty, R an equivalence relation on A, and $f : A \to B$ a mapping. We will often be interested in trying to define an "induced" map $\bar{f}$ from the quotient set $\bar{A} = A/R$ to B.

Think of the elements of an equivalence class $[a]$ as a clique of friends who are polled by f, the question being "Which element of B do you map to?" If the clique responds unanimously ("We all map to b"), then by fiat $\bar{f}$ maps $[a]$ in $\bar{A}$ to b in B. If *every* clique reaches a unanimous decision, there is a mapping $\bar{f} : A/R \to B$ defined by $\bar{f}\big([a]\big) = f(a)$.

If the responses are mixed for some clique $[a]$, then $\bar{f}$ is undefined; a mapping must be single-valued for every input, but the members of $[a]$ do not decide unanimously where to be mapped by f.

Definition 7.48. Let $f : A \to B$ be a mapping, and R an equivalence relation on A. We say f is *well-defined* modulo R, or f is *constant on equivalence classes* of R, if aRa' implies $f(a) = f(a')$. If f is well-defined modulo R, we define the *induced mapping* $\bar{f} : A/R \to B$ by $\bar{f}\big([a]\big) = f(a)$ for each a in A.

Remark 7.49. If R is an equivalence relation on A, there is a "quotient map" $\Pi : A \to A/R$ defined by $\Pi(a) = [a]$. If $f : A \to B$ is well-defined modulo R and $\bar{f} : A/R \to B$ denotes the induced mapping, then $f = \bar{f} \circ \Pi$. We say "f factors through A/R".

Example 7.50. Let $A = \mathbf{Z}$ be the set of integers, R the parity relation, and $f : \mathbf{Z} \to \{1, -1\}$ the mapping defined by $f(a) = (-1)^a$. Under f, every even integer maps to 1 and every odd integer maps to -1, so f is well-defined modulo parity. Intuitively, to compute $(-1)^a$ for some integer a, we only need to know whether a is even or odd.

The quotient space $A/R = \{[0], [1]\} = \{2\mathbf{Z}, 2\mathbf{Z} + 1\}$ is a set having two elements, and the induced map $\bar{f} : A/R = \{2\mathbf{Z}, 2\mathbf{Z}+1\} \to \{1, -1\}$, defined by

$$\bar{f}\big([0]\big) = (-1)^0 = 1, \qquad \bar{f}\big([1]\big) = (-1)^1 = -1,$$

is bijective.

Example 7.51. Let $A = \mathbf{Z}$, R the parity relation, and $f : \mathbf{Z} \to \mathbf{Z}$ defined by $f(a) = a^2$. The integers 0 and 2 are elements of [0], but $f(0) = 0 \neq 4 = f(2)$. Thus f is not well-defined modulo parity.

This should be no surprise: To compute the square of an integer a, it is not enough to know whether a is even or odd.

Exercises

Exercise 7.1. Define $f : \mathbf{Z} \to \mathbf{Z}$ by $f(a) = a^3$. Determine, with proof, whether or not f is injective or surjective.
Hint: If a and b are integers, then

$$a^2 + ab + b^2 \geq a^2 - |ab| + b^2 \geq a^2 - 2|ab| + b^2 = \big(|a| - |b|\big)^2 \geq 0,$$

and the second inequality is strict unless $ab = 0$.

Exercise 7.2. Let m and b be integers, and define a mapping $f : \mathbf{Z} \to \mathbf{Z}$ by $f(x) = mx + b$.

(a) Prove f is injective if and only if $m \neq 0$.

(b) Find necessary and sufficient conditions on m and b for f to be surjective. If f is bijective, find a formula for the inverse mapping.

Exercise 7.3. Let b and c be integers, and define $f : \mathbf{Z} \to \mathbf{Z}$ by $f(x) = x^2 + bx + c$. Show that f is neither injective nor surjective.

Exercise 7.4. Let $f : A \to B$ be a mapping, and let U and V be subsets of A. Prove the following:

(a) If $U \subseteq V$, then $f(U) \subseteq f(V)$.

(b) $f(A \setminus U) = f(A) \setminus f(U)$.

(c) If f is injective and $f(U) \subseteq f(V)$, then $U \subseteq V$.

Exercise 7.5. Let $f : A \to B$ and $g : B \to C$ be mappings, and assume $S \subseteq A$ and $T \subseteq C$.

(a) Prove $g\big(f(S)\big) = (gf)(S)$.

(b) Prove $f^{-1}\big(g^{-1}(T)\big) = (gf)^{-1}(T)$.

Exercise 7.6. Let $f : A \to B$ be a mapping.

(a) Assume $T \subseteq B$ is arbitrary. Prove $f\big(f^{-1}(T)\big) \subseteq T$, and that equality holds if f is surjective. Give an example of a mapping f and a set T for which the inclusion is proper.

(b) Assume $S \subseteq A$ is arbitrary. Prove $S \subseteq f^{-1}\big(f(S)\big)$, and that equality holds if f is injective. Give an example of a mapping f and a set S for which the inclusion is proper.

Exercise 7.7. (a) Let $f : A \to B$ and $g : B \to C$ be injective. Prove gf is injective. (This is the second assertion of Proposition 7.23.)

(b) Suppose gf is injective. Prove f is injective.
Suggestion: Prove the contrapositive.

Exercise 7.8. Let $f : A \to B$ be a mapping. If $S \subseteq A$, define the *restriction* of f to S to be the mapping $f|_S : S \to B$ defined by $f|_S(a) = f(a)$ for all a in S.

(a) Prove that f is injective if and only if $f|_S$ is injective for *every* subset S of A.

(b) Assume f is a bijection. Prove that if S is a non-empty subset of A, then the restriction $f|_S$ is a bijection from S to $f(S)$ and the restriction $f|_{A\setminus S}$ is a bijection from $A \setminus S$ to $B \setminus f(S)$.

Exercise 7.9. Let m, n, and q be positive integers.

(a) Let A be a set containing m elements, B a set containing n elements, and assume $m > nq$. Prove that if $f : A \to B$ is a mapping, then there exists a b in B such that $f^{-1}(\{b\})$ contains at least $q + 1$ elements. (This result is known as the *Pigeonhole Principle*. If you distribute $m > nq$ pigeons among n holes, then some hole contains more than q pigeons.)
Suggestion: Write B as a union of singleton sets, and use Proposition 7.8. The contrapositive may be more natural to prove.

(b) With the same notation, let $f : A \to B$ be a mapping. Prove that if f is injective, then $m \leq n$, and that if f is surjective, then $m \geq n$. Show by example that both converse statements are false.

(c) With the same notation, assume $m = n$, and let $f : A \to B$ be a mapping. Prove f is injective if and only if f is surjective. (Suggestion: Use part (b) to prove that f is injective if and only if $f(A)$ contains m elements, if and only if f is surjective.)

Exercise 7.10. Let $f : A \to B$ be an arbitrary mapping, and define mappings $\Gamma_f : A \to A \times B$ and $\Pi : A \times B \to B$ by

$$\Gamma_f(a) = \big(a, f(a)\big), \qquad \Pi(a, b) = b.$$

Prove that Γ_f is injective, Π is surjective, and $f = \Pi \circ \Gamma_f$. Illustrate with a sketch. (In words, every mapping factors as an injection followed by a surjection.)

Exercise 7.11. Let $n > 2$ be an integer. Prove that congruence $(\mathrm{mod}\, n)$ is an equivalence relation on the set of integers.

Exercise 7.12. Let A be a non-empty set, and let $R = \varnothing \subseteq A \times A$. Prove R is symmetric and transitive, but not reflexive.

Exercise 7.13. Define a relation R on $\mathbf{Z}$ by aRb if and only if $|a| = |b|$.

(a) Prove R is an equivalence relation.

(b) Let $f : \mathbf{Z} \to \mathbf{Z}$ be defined by $f(a) = a^2$. Is f well-defined mod R?

(c) Let $g : \mathbf{Z} \to \mathbf{Z}$ be defined by $g(a) = 3a$. Is g well-defined mod R?

(d) Prove $f : \mathbf{Z} \to \mathbf{Z}$ is well-defined mod R if and only if f is an even function, see Exercise 9.27.

Exercise 7.14. Define an equivalence relation R on $\mathbf{Z}$ by aRb if and only if $a \equiv b \pmod 4$.

(a) Let $f : \mathbf{Z} \to \mathbf{Z}$ be defined by $f(a) = (-1)^a$. Is f well-defined mod R?

(b) Let $f : \mathbf{Z} \to \mathbf{Z}$ be defined by $f(a) = a^2$. Is f well-defined mod R?

(c) Let $f : \mathbf{Z} \to \mathbf{Z}/4\mathbf{Z}$ be defined by $f(a) = a^2 \pmod 4$. Is f well-defined mod R?

(d) Let $f : \mathbf{Z} \to \mathbf{Z}/4\mathbf{Z}$ be defined by $f(a) = 3a \pmod 4$. Is f well-defined mod R?

Exercise 7.15. Fix an integer $n \geq 2$, and define an equivalence relation R on $\mathbf{Z}$ by aRb if and only if $a \equiv b \pmod n$. Determine whether the given maps $f : \mathbf{Z} \to \mathbf{Z}/n\mathbf{Z}$ are well-defined mod R.

(a) $f(x) = -x$. (b) $f(x) = x^2$. (c) $f(x) = x^3$.

Exercise 7.16. Let $f : A \to A$ be a mapping, and suppose the "maps-to" relation, aRb if and only if $b = f(a)$, is an equivalence relation. What can you say about f?

Exercise 7.17. Let R be an equivalence relation on A. If $f : A \to B$ is a mapping such that $a_1 R a_2$ if and only if $f(a_1) = f(a_2)$, i.e., whose level sets are precisely the equivalence classes of R, prove that the induced mapping $\bar{f} : A/R \to B$ is injective.

Exercise 7.18. Let $\mathscr{M}(\mathbf{R})$ denote the set of all mappings $f : \mathbf{R} \to \mathbf{R}$. Define a binary relation R on $\mathscr{M}(\mathbf{R})$ as follows: If f_1 and f_2 are elements of $\mathscr{M}(\mathbf{R})$, declare that $f_1 R f_2$ if and only if there exists a mapping g in $\mathscr{M}(\mathbf{R})$ such that $f_1 = f_2 \circ g$.

For example, $f_1(x) = x^6$ and $f_2(x) = x^3$ are related, because if $g(x) = x^2$, then $f_1(x) = x^6 = (x^2)^3 = (f_2 \circ g)(x)$.

(a) Show that R is reflexive.

(b) Prove that if $f_1 R f_2$, then $\text{img}(f_1) \subseteq \text{img}(f_2)$. Show (by example) that R is not symmetric.

(c) Is R transitive? (Give a proof or counterexample.)

(d) Suppose $f_1(x) = c$ is constant. Under what condition is $f_1 R f_2$?

Exercise 7.19. Let $\mathscr{M}(\mathbf{R})$ denote the set of all mappings $f : \mathbf{R} \to \mathbf{R}$. If f_1 and f_2 are elements of $\mathscr{M}(\mathbf{R})$, declare that $f_1 R f_2$ if and only if there exists a *bijective* mapping g in $\mathscr{M}(\mathbf{R})$ such that $f_1 = f_2 \circ g$. (Compare preceding question. Intuitively, $f_1 R f_2$ means "f_1 is obtained from f_2 by a change of coordinates".)

For example, $f_1(x) = x^6$ and $f_2(x) = x^2$ are related, because if $g(x) = x^3$ (a bijection), then $f_1(x) = x^6 = (x^3)^2 = (f_2 \circ g)(x)$.

(a) Show that R is an equivalence relation.

(b) Prove that a mapping f in $\mathscr{M}(\mathbf{R})$ is equivalent to the identity mapping $i(x) = x$ if and only if f is bijective.

(c) Are $f_1 = x^3 + x$ and $f_2(x) = x^3 - x$ related?

(d) Are $f_1(x) = x^4$ and $f_2(x) = x^2$ related?

Chapter 8

Binary Operations

We have encountered a variety of "number-like" objects: natural numbers, integers, rational, real and complex numbers, residue classes, matrices.

Ordinary addition of numbers may be viewed as mapping each ordered pair (a, b) of numbers to the number $a + b$. Multiplication of numbers has precisely the same *abstract* description, sending an ordered pair of numbers to a number. Moreover, the properties of these operations share features, including associativity, existence of identity elements, and perhaps commutativity and existence of inverses.

This chapter discusses "binary operations", mathematical functions that accept an ordered pair of objects of some type and return an object of the same type. Algebraic notions, such as associativity and identity elements, make sense in this general setting.

Once we have established a property of general binary operations, such as uniqueness of identity elements, we are assured the property holds automatically each time we encounter a new example, whether it be addition and multiplication of complex numbers, or composition of maps from a set X to itself.

8.1 Definitions

Definition 8.1. Let A be a non-empty set. A *binary operation* on A is a mapping $\mu : A \times A \to A$.

Remark 8.2. Conceptually, a binary operation is a rule for combining two elements of A to obtain an element of A. If a and b are elements of a, we usually write ab instead of $\mu(a, b)$. The expressions $a \cdot b$ or $a * b$

are used to emphasize the operation, especially when more than one binary operation is under consideration.

Example 8.3. The familiar operations of addition, multiplication, and subtraction are binary operations on $\mathbf{Z}$, the set of integers.

Division is *not* a binary operation on $\mathbf{Z}$ since, for example, $1 \div 0$ and $1 \div 2$ do not represent integers.

Example 8.4. Addition, multiplication, and exponentiation are binary operations on the set $\mathbf{N}$ of natural numbers.

Example 8.5. Let X be an arbitrary set, and let $\mathscr{M}(X)$ be the set of all mappings $f : X \to X$. Function composition, $\mu(g, f) = g \circ f$, is a binary operation on $A = \mathscr{M}(X)$.

Example 8.6. Let X be an arbitrary set. The intersection operator defines a binary operation on $A = \mathscr{P}(X)$, the power set of X. Similarly, the union operator defines a binary operation on $\mathscr{P}(X)$.

When A is a finite set of n elements, a binary operation may be represented by a *Cayley table*, an $n \times n$ tabular listing of all products, $\mu(a, b) = ab$ being placed in the "ath row" and "bth column".

Example 8.7. Let $A = \{E, O\}$ be a set with two elements, which we view as representing a general *even* integer (E) and a general *odd* integer (O). The Cayley table

$+$	E	O
E	E	O
O	O	E

expresses the fact that a sum of two even integers or two odd integers is even, while the sum of an even and an odd integer is odd.

Example 8.8. Let $A = \{a, b, c\}$ be a set with three elements. The following Cayley tables define binary operations on A:

μ_1	a	b	c
a	a	c	b
b	b	a	c
c	c	b	a

μ_2	a	b	c
a	a	b	c
b	b	c	a
c	c	a	b

μ_3	a	b	c
a	a	a	a
b	a	b	c
c	a	c	b

We have, e.g., $\mu_1(b, a) = b$, (second row, first column of the first table), while $\mu_1(a, b) = c$, $\mu_2(a, b) = b$, and $\mu_3(a, b) = a$ from the first row, second column of the respective tables.

Example 8.9. Let $B = \{0, 1, 2\}$. The following tables define binary operations on B "isomorphic to" those of Example 8.8:

μ_1	0	1	2
0	0	2	1
1	1	0	2
2	2	1	0

μ_2	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

μ_3	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

The concept of "abstractly identical" operations, formalized in Example 8.10, boils down to *relabeling*.

Example 8.10. Given a binary operation $\cdot$ on a set A and a bijection $\phi : A \to B$, define a binary operation $*$ on B as follows:

$$\phi(a_1) * \phi(a_2) = \phi(a_1 \cdot a_2) \text{ for all } a_1 \text{ and } a_2 \text{ in } A. \tag{8.1}$$

In words, attach each element b in B to its "avatar", the unique element a in A such that $\phi(a) = b$. Since the binary operation $\cdot$ combines pairs of avatars, Equation (8.1) tells us how to combine elements of B.

This relationship can be visualized as a "commutative diagram", Figure 8.1. Starting with a pair (a_1, a_2) in $A \times A$, there are two ways of getting to an element of B: (i) Multiply the elements in A (left edge) and map the product to B by ϕ (bottom edge), or (ii) map the elements individually to B (top edge) then multiply in B (right edge).

The diagram "commutes" because these two mapping compositions yield the same value for all pairs of inputs.

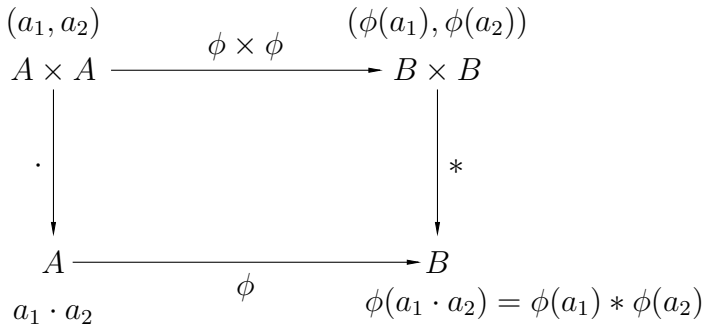


Figure 8.1: Isomorphism depicted as a commutative diagram.

Example 8.11. Consider Examples 8.8 (where $A = \{a, b, c\}$) and 8.9 ($B = \{0, 1, 2\}$). If we define $\phi : A \to B$ by $\phi(a) = 0$, $\phi(b) = 1$, and $\phi(2) = c$, the respective binary operations are related as in Example 8.10. Be sure you understand this example in detail.

8.2 Properties of Binary Operations

Definition 8.12. Suppose μ is a binary operation on A, and $S \subseteq A$ is non-empty. We say S is *closed under* μ if $\mu(s_1, s_2) \in S$ for all s_1 and s_2 in S.

If S is closed under μ, then the "restricted" mapping $\mu : S \times S \to S$ is a binary operation on S.

Example 8.13. Let $\mu : \mathbf{Z} \times \mathbf{Z} \to \mathbf{Z}$ be addition: $\mu(a, b) = a + b$.

The set $S = 2\mathbf{Z}$ of even integers is closed under addition: A sum of even integers is even.

Similarly, the set $S = \mathbf{Z}^+$ of positive integers is closed.

The set $S = \{0, 1\}$ is *not* closed under addition: $s_1 = 1$ and $s_2 = 1$ are elements of S, but $s_1 + s_2 = 2 \notin S$.

The set $S = 2\mathbf{Z} + 1$ of odd integers is not closed in a particularly strong way: A sum of odd integers is *never* odd.

Remark 8.14. Note carefully that the existence of a *single pair* s_1, s_2 in S with $\mu(s_1, s_2)$ not in S is enough to prove S is not closed under μ. Examples do not suffice to show S *is* closed under μ, however.

Associativity

By definition, a binary operation gives rise to "products" involving two elements. In practice, we often wish to combine three or more elements. This gives rise to a potential ambiguity: When we write a product "abc", we might mean either

$$(ab)c = \mu\big(\mu(a, b), c\big) \quad \text{or} \quad a(bc) = \mu\big(a, \mu(b, c)\big).$$

In general these expressions represent different elements of A. For the binary operation μ_1 of Example 8.8, we have

$$(ba)c = bc = c, \quad b(ac) = bb = a.$$

Definition 8.15. A binary operation μ on a set A is *associative* if $a(bc) = (ab)c$ for all a, b, and c in A.

Example 8.16. Addition and multiplication are associative operations on the set of natural numbers, on the set of integers, on the sets of rational numbers, real numbers, and complex numbers.

Example 8.17. Subtraction on $\mathbf{C}$ is *not* associative: If $c \neq 0$, then

$$(a - b) - c \neq a - (b - c) = (a - b) + c.$$

Similarly, division is not associative on the set of non-zero complex numbers.

Example 8.18. Exponentiation defines a binary operation on $\mathbf{N}$, but this operation is not associative. Even in the special case $a = b = c$, we have $a^{(a^a)} = (a^a)^a$ if and only if $a = 1$ or $a = 2$.

Example 8.19. If A is a set, then as noted earlier, mapping composition is a binary operation on $\mathscr{M}(A)$, the set of all mappings $f : A \to A$. This operation is automatically associative by Proposition 7.22.

Example 8.20. Let A be a set. By Exercise 2.8, the operations of intersection and union are associative binary operations on $\mathscr{P}(A)$, the power set of A.

When a binary operation is associative, products of any finite number of factors may be grouped arbitrarily (preserving only the order of the factors) without changing the result. For example,

$$a\big((bc)d\big) = a\big(b(cd)\big) = (ab)(cd) = \big((ab)c\big)d = \dots.$$

Rather than proving directly that any two groupings of a product have the same value, we'll pick one specific grouping, and show that an arbitrary grouping has the same value.

Definition 8.21. Let A be a non-empty set equipped with a binary operation. A product of n elements is *grouped from the right* if pairs of factors are grouped from right to left:

$$a_1\Big(a_2\big(a_3 \dots (a_{n-1}a_n) \dots\big)\Big).$$

Proposition 8.22. *Let A be a non-empty set equipped with an* associative *binary operation. If $a_1, \dots, a_n$ is an ordered n-tuple of elements of A, then every grouping of these n factors has the same value as the grouping from the right.*

In particular, an arbitrary grouping of the factors *taken in order from left to right* has the same value.

Proof. If n is an integer with $n \geq 3$, let $P(n)$ denote the statement:

> Every m-fold product with $3 \leq m \leq n$ can be regrouped from the right without changing the value of the product.

The associative law says every threefold product can be regrouped from the right without changing the value of the product; $P(3)$ is true.

Assume inductively that $P(k)$ is true for some integer $k \geq 3$. Every grouping of a product of $(k+1)$ factors $a_1, \ldots, a_{k+1}$ may be viewed as a product A_1A_2 of two factors, with each of A_1 and A_2 a product of k or fewer factors. By the inductive hypothesis, we may regroup A_1 from the right without changing the value of the product, say $A_1 = a_1A_1'$. By associativity,

$$A_1A_2 = (a_1A_1')A_2 = a_1(A_1'A_2).$$

The product $(A_1'A_2)$ has k factors, and by the inductive hypothesis can be regrouped from the right without changing the value of the product. The previous equation therefore states that A_1A_2 can be regrouped from the right without changing the value of the product. □

Identity Elements

Definition 8.23. Let (A, μ) be a set equipped with a binary operation. An element e in A is an *identity element* for μ if

$$ea = ae = a \quad \text{for all } a \text{ in } A.$$

A binary operation may have no identity element at all. However, there can be at most one, for if e and e' are identity elements for μ, then $e = ee'$ (since e' is an identity element) and $ee' = e'$ (since e is an identity element), so $e = e'$.

Example 8.24. The integer 0 is the identity element for addition on $\mathbf{Z}$. The integer 1 is the identity element for multiplication on $\mathbf{Z}$.

Example 8.25. There is no identity element for subtraction. In other words, there exists no integer e such that $a - e = e - a = a$ for every integer a. (Why not?)

Example 8.26. An identity element can be located at a glance from a Cayley table: The corresponding row and column of the table will contain the same entries as the "index" entries across the top and down the left side. Consider the operations in Example 8.8:

μ_1	a	b	c
a	a	c	b
b	b	a	c
c	c	b	a

μ_2	a	b	c
a	$\boldsymbol{a}$	$\boldsymbol{b}$	$\boldsymbol{c}$
b	$\boldsymbol{b}$	c	a
c	$\boldsymbol{c}$	a	b

μ_3	a	b	c
a	a	$\boldsymbol{a}$	a
b	$\boldsymbol{a}$	$\boldsymbol{b}$	$\boldsymbol{c}$
c	a	$\boldsymbol{c}$	b

μ_1 has no identity element, while a is the identity element for μ_2 and b is the identity element for μ_3

Example 8.27. Let $\mathcal{U}$ be a universe. Since $A \cup \varnothing = \varnothing \cup A = A$ for every subset $A \subseteq \mathcal{U}$, $e = \varnothing$ for the union operator.

Dually, since $A \cap \mathcal{U} = \mathcal{U} \cap A = A$ for all A, $e = \mathcal{U}$ is the identity element for the intersection operator.

Inverse Elements

Definition 8.28. Let (A, μ) be a set equipped with a binary operation, and assume there is an identity element for μ. If $a \in A$, then an element b in A is an *inverse* of a (with respect to μ) if

$$ab = ba = e.$$

Remark 8.29. It makes no sense to ask about inverses unless μ has an identity element. Moreover, even if μ has an identity element, a specific element a in A may or may not have an inverse.

Remark 8.30. If μ is *associative* and has an identity element e, then each element a has at most one inverse, see Exercise 8.10. Briefly, inverses are unique with respect to an associative operation. The inverse of a is normally denoted a^{-1}. A commutative binary operation is customarily denoted $+$, in which case the inverse of a is denoted $-a$.

Example 8.31. In the set $\mathbf{N}$ of natural numbers, addition has identity element 0, but no non-zero natural number has an additive inverse.

Example 8.32. In the set $\mathbf{Z}$ of integers, addition has identity element 0, and every integer a has an additive inverse, $-a$.

Example 8.33. In the set $\mathbf{Z}$ of integers, multiplication has an identity element 1. The only invertible integers are 1 and -1, each being its own inverse.

Remark 8.34. (Non-)existence of inverse elements can be read off a Cayley table. First locate the identity element e; if none exists, nothing further need be done.

To seek the inverse of a specific element a, inspect the ath row of the table, looking for the identity element e. If e is found in the bth column (signifying $ab = e$), check to see whether $ba = e$ as well. If so, $a^{-1} = b$; otherwise a has no inverse.

Example 8.35. The operation μ_1 in Example 8.8 has no identity element, so the concept of inverses makes no sense.

For μ_2, $a^{-1} = a$, $b^{-1} = c$, and $c^{-1} = b$; every element has an inverse.

For μ_3, $b^{-1} = b$ and $c^{-1} = c$, but a has no inverse.

Example 8.36. Let A be a non-empty set, $\mathscr{M}(A)$ the set of mappings from A to A equipped with the operation of function composition.

The identity mapping $I_A : A \to A$ is the identity element for composition. A mapping $f : A \to A$ has an inverse in the sense of binary operations if and only if f is a bijection, if and only if f is invertible in the sense of mappings.

Example 8.37. In Example 8.27, we saw that the union operation on $\mathscr{P}(A)$ has identity element $e = \varnothing$. If $S \subseteq A$, an inverse of S is a set T such that $S \cup T = \varnothing$. No such set exists unless $S = \varnothing$, in which case $T = \varnothing$ satisfies the conditions for an inverse element. In other words, $\varnothing^{-1} = \varnothing$, and no other set has an inverse with respect to the union. See also Exercise 8.11.

Commutativity

Definition 8.38. A binary operation on A is *commutative* if $ab = ba$ for all a and b in A.

Example 8.39. By Axioms A4. and M4., page 58, addition and multiplication are commutative operations on $\mathbf{Z}$: $a + b = b + a$ and $ab = ba$ for all integers a and b.

Example 8.40. Subtraction is not commutative on $\mathbf{Z}$: $1 - 2 \neq 2 - 1$.

Example 8.41. Set union and intersection are commutative on $\mathscr{P}(A)$.

Example 8.42. Function composition is not commutative. For example, if f and $g : \mathbf{Z} \to \mathbf{Z}$ are defined by $f(a) = a + 1$ and $g(a) = a^2$, then $(f \circ g)(a) = a^2 + 1$, while $(g \circ f)(a) = (a + 1)^2 = a^2 + 2a + 1$.

Example 8.43. Many real-life activities do not commute: Putting on your socks and putting on your shoes; removing car keys from the ignition and closing the locked car door; turning off the electricity and repairing the wiring; looking both ways and crossing the street. In each case, the result of one activity has some bearing on the success or failure of the other.

Exercises

Exercise 8.1. Let $A = (\mathbf{Z}/9\mathbf{Z})^\times$ be the set of units $(\mathrm{mod}\, 9)$. Make a Cayley table for A under multiplication, and find the inverse of each element.

Exercise 8.2. Let $A = (\mathbf{Z}/10\mathbf{Z})^\times$ be the set of units $(\mathrm{mod}\, 10)$. Make a Cayley table for A under multiplication, and find the inverse of each element.

Exercise 8.3. Let $A = (\mathbf{Z}/12\mathbf{Z})^\times$ be the set of units $(\mathrm{mod}\, 12)$. Make a Cayley table for A under multiplication, and find the inverse of each element.

Exercise 8.4. Let $A = (\mathbf{Z}/18\mathbf{Z})^\times$ be the set of units $(\mathrm{mod}\, 18)$. Make a Cayley table for A under multiplication, and find the inverse of each element.

Exercise 8.5. Let $A = \mathbf{Z}/6\mathbf{Z}$, and let $+$ denote addition $(\mathrm{mod}\, 6)$.

(a) Make a Cayley table for $(A, +)$. Does $+$ have an identity element? If so, which element(s) are invertible?

(b) Find all proper, non-empty subsets $B \subseteq A$ that are closed under $+$.

Exercise 8.6. Let $A = \mathbf{Z}/7\mathbf{Z}$, and let $\cdot$ denote multiplication $(\mathrm{mod}\, 7)$.

(a) Make a Cayley table for $(A, \cdot)$. Does $\cdot$ have an identity element? If so, which element(s) are invertible?

(b) Find all proper, non-empty subsets $B \subseteq A$ that are closed under $\cdot$.

Exercise 8.7. Let $A = \{0, 2, 4, 6, 8\}$, and define $a * b = ab \pmod{10}$, the remainder of the integer product of division by 10.

(a) Make a Cayley table for $(A, *)$. Does $*$ have an identity element? If so, which element(s) are invertible?

(b) Find all proper, non-empty subsets $B \subseteq A$ that are closed under $*$. For each such subset B, determine whether $*$ restricted to B has an identity element, and if so, determine which element(s) of B are invertible.

Exercise 8.8. Suppose A is a set of n elements and $*$ is a binary operation on A.

(a) How many conditions must be checked to prove $*$ is commutative? (The answer is not n^2.) Describe how a Cayley table for $(A, *)$ can be used to check commutativity.

(b) How many conditions must be checked to prove $*$ is associative?

(c) For each a in A, construct two binary operations

$$\lambda_a(x, y) = (x * a) * y, \qquad \rho_a(x, y) = x * (a * y).$$

Note that the operation $*$ is associative if and only if these operations have the same Cayley table for all a in A. Use this criterion (*Light's test for associativity*) to check whether the following operations are associative:

$\oplus$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

$\odot$	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

Exercise 8.9. Each part refers to the indicated binary operations on the set $A = \{0, 1, 2, 3\}$.

$\ominus$	0	1	2	3
0	0	3	2	1
1	1	0	3	2
2	2	1	0	3
3	3	2	1	0

$\odot$	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

(a) Is $\ominus$ associative? Is $\ominus$ commutative? Does $\ominus$ have an identity element? If so, which elements of A have inverses?

(b) Show $\odot$ is associative. Suggestion: Use Light's test (Exercise 8.8).

(c) Is $\odot$ commutative? Does $\odot$ have an identity element? If so, which elements of A have inverses?

Exercise 8.10. Let (A, μ) be a set equipped with an associative binary operation, assume μ has an identity element e, and assume $a \in A$. Prove that if b and b' are inverses of a, then $b = b'$.

Exercise 8.11. Consider the intersection operation on $\mathscr{P}(A)$. Determine which subsets of A (if any) are invertible, and find the inverse of any invertible set, cf. Example 8.37.

Exercise 8.12. On the set $\mathbf{Z}$ of integers, define a binary operation by $a * b = a + b - 1$.

(a) Prove $*$ is associative and commutative.

(b) Prove $*$ has an identity element.

(c) Prove every integer has an inverse with respect to $*$.

(d) Define $\phi : \mathbf{Z} \to \mathbf{Z}$ by $\phi(a) = a + 1$. Starting with the operation of ordinary addition, use the method of Example 8.10 to define a new binary operation μ on $\mathbf{Z}$. Try to re-do the first three parts of this question by "transferring" a property of addition to the corresponding property of μ.

Exercise 8.13. Let $A = \{0, 1, 2, 3\}$ and $B = \{a, b, c, d\}$. Each part concerns the following binary operation on A:

$\cdot$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

(a) Is $\cdot$ commutative? Does $\cdot$ have an identity element? If so, which elements have inverses?

(b) Define a bijection $\phi : A \to B$ by $\phi(0) = a$, $\phi(1) = b$, $\phi(2) = c$, $\phi(3) = d$. Write out the Cayley table for the induced operation $*$ on B.

(c) Is $*$ commutative? Does $*$ have an identity element? If so, which elements have inverses? How are your answers related to your answers for part (a)?

(d) Show the set $\{a, c\} \subseteq B$ is closed under $*$, and write out the Cayley table for $*$ restricted to $\{a, c\}$.

(e) Find all proper subsets of B that are closed under $*$.

Exercise 8.14. Use Light's test (Exercise 8.8) to determine whether the operation in the Cayley table is associative:

$*$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

It may help to note that $x * x = e$ for all x, and that if x, y, and z are the distinct elements other than e (in any order), then $x * y = z$.

Exercise 8.15. For each integer n, define the mapping $T_n : \mathbf{Z} \to \mathbf{Z}$ by $T_n(x) = x + n$, and let $\mathscr{M}(T) = \{T_n : n \in \mathbf{Z}\}$.

(a) Prove that each T_n is a bijection, and find a formula for the inverse mapping.

(b) Show that $\mathscr{M}(T)$ is closed under composition of functions.

(c) Show that $\mathscr{M}(T)$ contains an identity element, and that $\mathscr{M}(T)$ is closed under inversion.

Exercise 8.16. For each integer n, define the mapping $S_n : \mathbf{Z} \to \mathbf{Z}$ by $S_n(x) = nx$, and let $\mathscr{M}(S) = \{S_n : n \in \mathbf{Z}\}$.

(a) Determine (with proof) which mappings S_n are injective, and which are surjective.

(b) Show that $\mathscr{M}(S)$ is closed under composition of functions.

(c) Show that $\mathscr{M}(S)$ contains an identity element. Which elements of $\mathscr{M}(S)$ are invertible?

Part III

Continuous Structures

Chapter 9

Real and Complex Numbers

In school you were introduced to "real numbers" and the "number line", and possibly to "complex numbers" and the "complex plane". This chapter introduces these number systems via axioms.

9.1 Axioms for the Real Numbers

We start with a "legal contract" for the real number system, a list of axioms that characterizes the real number system, Table 9.1. You should not memorize these axioms, but instead *internalize* them. They fall into three categories:

Algebraic Properties. (A1.–A4., M1.–M4., D.) The algebraic axioms concern the operations of addition and multiplication: What properties each operation has (associativity, commutativity, existence of an identity element and inverses), and how the two operations interact (the distributive law).

Order Properties. (O1.–O3.) These three axioms formalize the idea of one real number being "greater than" or "less than" another, and the fact that every pair of real numbers is "comparable". Geometrically, the order axioms guarantee that the real number system can be visualized as a subset of a line. "Less than" means "lies to the left of", and "greater than" means "lies to the right of".

Completeness. This axiom formalizes the geometric intuition that the real number system "has no gaps", or that "any quantity that can be approximated by real numbers is itself a real number". Geometrically, if A is a non-empty set on the number line, and if *some* point M lies

The *real number system* consists of a non-empty set $\mathbf{R}$, two binary operations, $+$ and $\cdot$, and a subset P of $\mathbf{R}$ (the set of "positive" real numbers) satisfying the following thirteen axioms:

A1. Addition is associative: For all a, b, c in $\mathbf{R}$, $a + (b + c) = (a + b) + c$.

A2. Additive identity: There exists a unique element 0 in $\mathbf{R}$ such that for all a in $\mathbf{R}$, $0 + a = a + 0 = a$.

A3. Additive inverses: For every a in $\mathbf{R}$, there exists a unique $-a$ in $\mathbf{R}$ such that $a + (-a) = (-a) + a = 0$.

A4. Addition is commutative: For all a, b, in $\mathbf{R}$, $a + b = b + a$.

M1. Multiplication is associative: For all a, b, c in $\mathbf{R}$, $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.

M2. Multiplicative identity: There exists a unique element 1 in $\mathbf{R}$, distinct from 0, such that for all a in $\mathbf{R}$, $1 \cdot a = a \cdot 1 = a$.

M3. Multiplicative inverses: For every non-zero a in $\mathbf{R}$, there exists a unique a^{-1} in $\mathbf{R}$ such that $a \cdot a^{-1} = a^{-1} \cdot a = 1$.

M4. Multiplication is commutative: For all a, b, in $\mathbf{R}$, $a \cdot b = b \cdot a$.

D. Multiplication (on the left) distributes over addition: For all a, b, c in $\mathbf{R}$, $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$.

O1. The law of trichotomy: For every real number a, *exactly one* of the following holds: $a \in P$, $-a \in P$, or $a = 0$.

O2. Closure under addition: If a and b are in P, then $a+b$ is in P.

O3. Closure under multiplication: If a and b are in P, then $a \cdot b$ is in P.

C. Completeness: If A is a non-empty subset of $\mathbf{R}$ that is bounded above, then A has a least upper bound in $\mathbf{R}$.

Table 9.1: Axioms for the numbers.

to the right of every point of A, then there exists a *leftmost* point (i.e., smallest number) lying on or to the right of every element of A.

Remark 9.1. The multiplication dot is often omitted: $a \cdot b = ab$.

Remark 9.2. These axioms have minor redundancies built in for convenience. For example, if $a+b=0$, then by the commutative Axiom A4., $b + a = 0$ as well; there is no need to assume both equations in A3. Further, the uniqueness conditions in the axioms for identity elements and inverses can be deduced from associativity.

In fact, the set of real numbers, including the operations of addition and multiplication and the set of positive numbers, can be constructed entirely from the natural numbers. The axioms in Table 9.1 would then be proven as theorems.

Auxiliary Concepts

We define the operations of *subtraction* and *division* in terms of addition and multiplication.

Definition 9.3. If a and b are real numbers, we define their *difference* to be $a - b = a + (-b)$.

If $b \neq 0$, we define their *quotient* to be $a/b = a \cdot b^{-1}$. In particular, $1/b = b^{-1}$.

Remark 9.4. Subtraction and division are neither associative nor commutative, as you should check.

We define the concepts of *positive* and *negative* numbers, and the relations *greater-than* and *less-than*, using Axioms O1.–O3.

Definition 9.5. Let a, b and c be real numbers.

If $c \in P$, we say c is *positive*, or 0 is less than c, and write $0 < c$.

If $b - a \in P$, we say a is *less than* b and write $a < b$. If $a < b$ or $a = b$, we say a is *less than or equal to* b, and write $a \leq b$.

Remark 9.6. If $0 < c$, we also say c is *greater than* 0 and write $c > 0$.

If $a < b$, we also say b is *greater than* a and write $b > a$. If $a \leq b$ we also say b is *greater than or equal to* a and write $b \geq a$.

Whole numbers and fractions constitute some of the most important classes of real numbers.

Definition 9.7. The set $\mathbf{N}$ of *natural numbers* is the smallest subset of $\mathbf{R}$ satisfying the following conditions: (i) $0 \in \mathbf{N}$; (ii) If $a \in \mathbf{N}$, then $a + 1 \in \mathbf{N}$.

A real number a is an *integer* or *whole number* if either a or $-a$ is a natural number. The set of integers is denoted $\mathbf{Z}$, from the German *Zahl*.

A real number a is a *rational number* if there exist integers p and q such that $q > 0$ and $a = p/q$. The set of rational numbers (a.k.a. *ratios* or *quotients*) is denoted $\mathbf{Q}$.

A real number a is *irrational* if a is not a rational number.

Example 9.8. In the set of rational or real numbers equipped with the operation of multiplication, 1 is the identity element, and every non-zero number has an inverse.

Remark 9.9. In mathematics, a *field* is a set F together with two arithmetic operations $+$ and $\cdot$ that satisfy Axioms A1.–A4., M1.–M4., and D. An *ordered field* additionally satisfies Axioms O1.–O3.

The real numbers turn out to be the unique *complete* ordered field. The rational numbers form an ordered field.

The set of integers is not a field; Axiom M3 (page 58) fails.

The set of complex numbers is a field, but not an ordered field. By Corollary 9.13 (ii) below, in an ordered field, if $a \neq 0$, then $a^2 \in P$. However, each of 1 and -1 is the square of some complex number, so the trichotomy axiom does not hold.

Algebraic Properties

Lemma 9.10. *If $a \in \mathbf{R}$, then $-(-a) = a$. If $a \neq 0$, then $(a^{-1})^{-1} = a$, i.e., $1/(1/a) = a$.*

Proof. The equation $a+(-a) = 0$ may be interpreted as saying that the (unique) additive inverse of $-a$ is a itself, i.e., $-(-a) = a$. Similarly, if $a \neq 0$, the equation $a(a^{-1}) = 1$ says $(a^{-1})^{-1} = a$. □

Proposition 9.11. *For all real numbers a and b:*

(i) $0 \cdot a = a \cdot 0 = 0$.

(ii) *If $ab = 0$, then $a = 0$ or $b = 0$.*

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

(iv) $(-a)(-b) = ab$. *Particularly,* $(-1)(-1) = 1$.

Proof. (i) Taking $a = 0$ in A2, we have $0 = 0 + 0$. Now let a denote an arbitrary real number. Multiplying on the right by a and using the distributive law, we have

$$0 + (0 \cdot a) = (0 \cdot a) = (0 + 0) \cdot a = (0 \cdot a) + (0 \cdot a).$$

By the cancellation law, $0 = 0 \cdot a$. By the commutative Axiom M4., $0 = a \cdot 0$ as well.

(ii) Suppose $ab = 0$. If $a = 0$, there is nothing to prove. If $a \neq 0$,

$$b = (a^{-1}a)b = a^{-1}(ab) = a^{-1}(0) = 0.$$

(iii) Multiply $0 = \big(1 + (-1)\big)$ on the right by a:

$$a + (-a) = 0 = 0 \cdot a = \big(1 + (-1)\big) \cdot a = 1 \cdot a + (-1) \cdot a = a + (-1) \cdot a.$$

By cancellation, $-a = (-1) \cdot a$. By commutativity, $-a = a \cdot (-1)$.

(iv) Taking $a = -1$ in (iii) and invoking Lemma 9.10, we have $(-1)(-1) = -(-1) = 1$. Thus

$$(-a)(-b) = \big((-1)a\big)\big((-1)b\big) = (-1)(-1)(ab) = ab. \qquad \square$$

Order Properties

Proposition 9.12. *For all real numbers a, b, and c:*

(i) *If $a < b$ and $b < c$, then $a < c$.*

(ii) *If $a < b$, then $a + c < b + c$.*

(iii) *If $a < b$ and $0 < c$, then $ac < bc$.*

(iv) *If $a < b$ and $c < 0$, then $bc < ac$.*

(v) *If $0 < a < b$, then $0 < 1/b < 1/a$.*

Two additional extensions are useful occasionally:

Corollary 9.13. *Let a, b, c, and d be real numbers.*

(i) *If $a < b$ and $c < d$, then $a + c < b + d$.*

(ii) *If $0 < a < b$ and $0 < c < d$, then $0 < ac < bd$.*

In particular, $0 \leq a \cdot a$, with equality if and only if $a = 0$.

Remark 9.14. In words, (i) says two inequalities can be "added", preserving the sense of the inequality; (ii) expresses a similar guarantee when inequalities are multiplied, provided all quantities involved are positive.

To reduce Proposition 9.12 to Axioms O1.–O3., it is convenient to establish special cases where one comparand is zero:

Lemma 9.15. *For all real numbers a and b,*

(i) *If $0 < a$ and $b < 0$, then $ab < 0$.*

(ii) *If $a < 0$ and $b < 0$, then $0 < ab$.*

(iii) *If $0 < a$, then $0 < 1/a$.*

Proof. (i) By trichotomy, $b < 0$ if and only if $-b \in P$. By Axiom O3. and Proposition 9.11 (iii), $-(ab) = a(-b) \in P$, i.e., $ab < 0$.

(ii) If $-a$ and $-b \in P$, then by Axiom O3. and Proposition 9.11 (iv), $0 < (-a)(-b) = ab$.

(iii) By trichotomy, either $1 \in P$ or $-1 \in P$. Since P is closed under multiplication, either $1 \cdot 1 = 1$ is positive, or $(-1) \cdot (-1) = 1$ is positive. Under either alternative, we conclude 1 is positive, or $0 < 1$.

Suppose $0 < a$. If a^{-1} were negative, then by (i), we would have $1 = a(a^{-1}) < 0$, which we have just seen is false. Contrapositively, if $0 < a$, then $0 < a^{-1}$. □

Proof of Proposition 9.12. (i) By definition, $a < b$ if and only if $b - a \in P$, and $b < c$ if and only if $c - b \in P$. By Axiom O2.,

$$c - a = (c - b) + (b - a) \in P,$$

which is equivalent to $a < c$.

(ii) For all a, b, and c, we have $b - a = (b + c) - (a + c)$. The claim follows immediately.

(iii) If $a < b$ and $0 < c$, then $b - a \in P$ and $c \in P$, so by Axiom O3., their product $(b - a)c = bc - ac$ is in P, which means $ac < bc$.

(iv) Since $a < b$ and $0 < -c$, part (iii) gives

$$0 < (b - a)(-c) = (a - b)c = ac - bc,$$

i.e., $bc < ac$.

(v) By Lemma 9.15 (iii), if $0 < a < b$, then $0 < 1/a$ and $0 < 1/b$. Algebra gives $1/a - 1/b = (b - a)/(ab) > 0$, i.e., $1/b < 1/a$. □

Proof of Corollary 9.13. (i) By Proposition 9.12 (ii), adding a to $c < d$ and adding d to $a < b$ gives $a + c < a + d < b + d$.

(ii) follows similarly from Proposition 9.12 (iii).

For the assertion about real squares, it suffices to note that $0 \cdot 0 = 0$ by Proposition 9.11 (i), while we have just shown that if $a \neq 0$, then $0 < a \cdot a$. □

The "limiting behavior" of α^n as n grows without bound plays a central role in analysis.

Theorem 9.16. *Let $u > 0$ be a real number. We have*

$$1 + nu \leq (1 + u)^n \quad \textit{for all } n \geq 0.$$

Proof. Let $P(n)$ denote the inequality in the theorem. The base case $P(0)$ asserts that $1 \leq 1$, which is true. Assume inductively that $P(k)$ is true for some $k \geq 0$. We have

$$\begin{aligned} 1 + (k+1)u &\leq 1 + (k+1)u + ku^2 && 0 \leq ku^2 \\ &= (1 + ku)(1 + u) && \text{Algebra} \\ &\leq (1 + u)^k(1 + u) && \text{Inductive hypothesis} \\ &= (1 + u)^{k+1} && \text{Definition of exponentiation} \end{aligned}$$

so that $P(k)$ implies $P(k + 1)$. □

Corollary 9.17. *Let $0 < \alpha < 1$. There exists a positive real number u such that $\alpha = 1/(1 + u)$, and*

$$0 < \alpha^n \leq \frac{1}{1 + nu} \quad \textit{for all } n \geq 0.$$

Proof. If $0 < \alpha < 1$, then $1 < 1/\alpha$ by Proposition 9.12 (v), so we may write $1/\alpha = 1 + u$ for some positive real number u. By Theorem 9.16, we have $0 < 1+nu \leq (1/\alpha)^n = 1/(\alpha^n)$ for all $n \geq 0$. Taking reciprocals again establishes the corollary. □

9.2 Complex Numbers

Points in the Cartesian plane may be viewed as numerical entities in a way that extends the real number system.

Definition 9.18. A *complex number* is an expression $\alpha = a + ib$ in which a and b are real numbers and i is a symbol satisfying $i^2 = -1$. The set of complex numbers is the *complex plane* $\mathbf{C}$.

The real numbers a and b are, respectively, the *real part* and *imaginary part* of α. Viewing the real and imaginary parts of a complex number $\alpha = a + bi$ as Cartesian coordinates, we identify α with the point (a, b), Figure 9.1.

A complex number α is *real* if $b = 0$, and *non-real* if $b \neq 0$. The set of complex numbers of the form $\alpha = a + 0 \cdot i = a$ is the *real axis*.

A complex number α is *imaginary* if $a = 0$. The set of all numbers of the form $\alpha = 0 + b \cdot i = bi$ is the *imaginary axis*.

The *conjugate* of α is the complex number $\bar{\alpha} = a - bi$ obtained by reflecting α across the real axis.

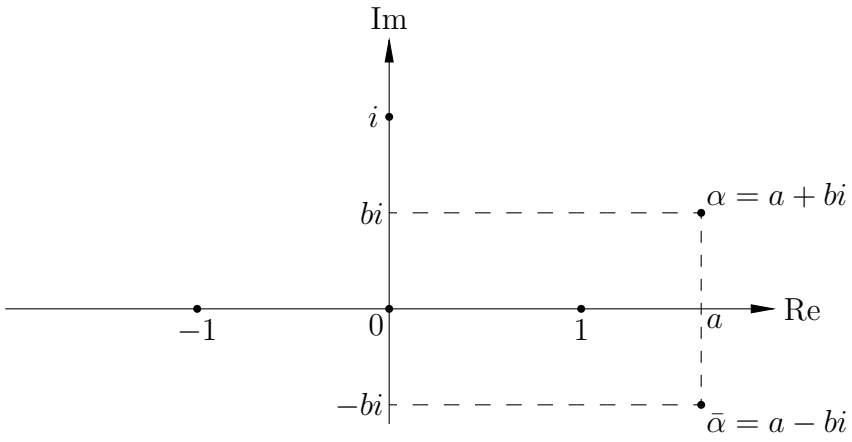


Figure 9.1: The complex plane.

Remark 9.19. Imaginary numbers may seem tainted with suspicion, as if they don't really exist but it's mathematically convenient to pretend they do. This sentiment traces back to the Ancient Greeks, who viewed numbers as lengths, what we now call "real numbers". Indeed, no real number has square equal to -1.

As noted above, hoewever, i has a perfectly concrete existence as the point $(0, 1)$ in the Cartesian plane. Even the mysterious equation $i^2 = -1$ turns out to have a natural interpretation: Multiplication by i corresponds to a counterclockwise quarter-turn of the complex plane about the origin. Performing this operation twice, namely squaring, amounts to a half-turn, which multiplies each complex number by -1.

Complex Addition and Multiplication

From a modern perspective, the complex numbers earn their status as "numbers" by admitting operations of addition, subtraction, multiplication, and division that generalize the familiar algebraic properties of real numbers.

Definition 9.20. Let $\alpha_1 = a_1 + ib_1$ and $\alpha_2 = a_2 + ib_2$ be complex numbers. Their *sum* is defined by the formula

$$\alpha_1 + \alpha_2 = (a_1 + ib_1) + (a_2 + ib_2) = (a_1 + a_2) + i(b_1 + b_2).$$

The formula for subtraction is similar, and left to you to work out. Adding two complex numbers corresponds to the parallelogram law for vector addition in the plane, see Figure 9.2.

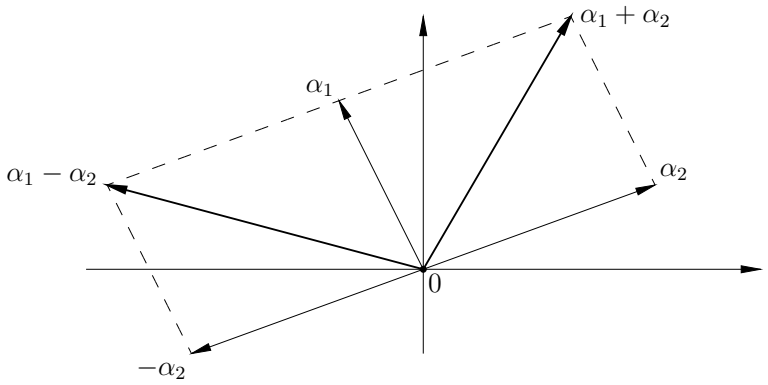


Figure 9.2: Adding and subtracting complex numbers.

Definition 9.21. A set A contained in $\mathbf{C}$ is *closed under addition* if for all α_1 and α_2 in A, the sum $\alpha_1 + \alpha_2$ is in A.

Example 9.22. The set $\{0\}$ is closed under addition, since $0 + 0 = 0$.

Example 9.23. Suppose A is closed under addition and $1 \in A$. Of necessity, $2 = 1 + 1$, $3 = 2 + 1$, $4 = 3 + 1$, and so forth, are in A. That is, A contains the set $\mathbf{Z}^+$ of positive integers. Since the set of positive integers is closed under addition, our hypotheses imply nothing further.

Similarly, if A is closed under addition and $\alpha \neq 0$ is an element of A, the set $\alpha\mathbf{Z}^+$ of positive integer multiples of α is contained in A.

If A is closed under addition, it does not follow that A is "generated" by one element as in the previous examples.

Example 9.24. The set $\mathbf{Z}$ of integers is closed under addition in $\mathbf{C}$, as are the set $\mathbf{Q}$ of rational numbers (ratios of integers) and the set $\mathbf{R}$ of real numbers. None of these sets is obtained by adding a single element to itself repeatedly.

Example 9.25. The set $\mathbf{Z} + i\mathbf{Z} = \{m + in : m, n \in \mathbf{Z}\}$ of *Gaussian integers*, Figure 9.3, is closed under addition: If $\alpha_1 = m_1 + in_1$ and $\alpha_2 = m_2 + in_2$ are Gaussian integers, the addition formula gives $\alpha_1 + \alpha_2 = (m_1 + m_2) + i(n_1 + n_2)$. Since a sum of integers is an integer, the real and imaginary parts of $\alpha_1 + \alpha_2$ are integers. That is, $\alpha_1 + \alpha_2 \in \mathbf{Z} + i\mathbf{Z}$. Since α_1 and α_2 were arbitrary, $\mathbf{Z} + i\mathbf{Z}$ is closed under addition.

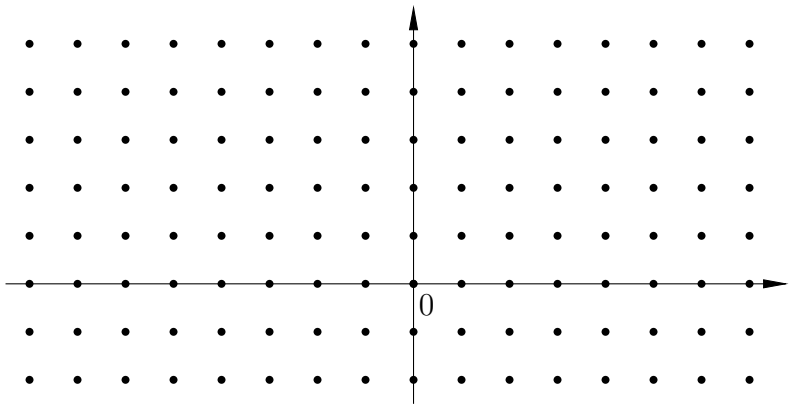


Figure 9.3: The Gaussian integers.

Example 9.26. The *upper half-plane* $\{a + bi : b > 0\}$ is closed under addition.

Example 9.27. The set A of complex numbers that are either real or imaginary, i.e., the union of the real and imaginary axes, is *not* closed under addition. Since "closed under addition" is a "for every" condition, its negation is a "there exists" condition; that is, it suffices to find a *single counterexample*. For instance, $1 \in A$ (since 1 is real) and $i \in A$ (since i is imaginary) but $1 + i \notin A$ (the sum is neither real nor imaginary), so A is not closed under addition.

To define multiplication of complex numbers, we treat i as a symbol distributing over addition of real numbers, commuting with multiplication of real numbers, and satisfying $i^2 = -1$. A short calculation using familiar laws of algebra leads us to

$$\begin{aligned}(a_1 + ib_1)(a_2 + ib_2) &= a_1a_2 + i(a_1b_2 + a_2b_1) + i^2b_1b_2 \\ &= (a_1a_2 - b_1b_2) + i(a_1b_2 + a_2b_1).\end{aligned}$$

Definition 9.28. Let $\alpha_1 = a_1 + ib_1$ and $\alpha_2 = a_2 + ib_2$ be complex numbers. Their *product* is defined by the formula

$$\alpha_1\alpha_2 = (a_1a_2 - b_1b_2) + i(a_1b_2 + a_2b_1).$$

Example 9.29. If $\alpha = a + bi$, then $i\alpha = i(a + bi) = -b + ai$. As expected, the vector $(-b, a)$ is obtained by rotating the vector (a, b) through a quarter turn.

As a consistency check, $i(i\alpha) = i(-b + ai) = -a - bi = -\alpha$.

Example 9.30. If $\alpha = a + bi$, then

$$\alpha\bar{\alpha} = (a + bi)(a - bi) = a^2 - (bi)^2 = a^2 + b^2. \tag{9.1}$$

By the Pythagorean theorem, $\alpha\bar{\alpha} = (\text{distance from } 0 \text{ to } \alpha)^2$.

Complex multiplication is *commutative*: For all complex numbers α_1 and α_2, we have $\alpha_2\alpha_1 = \alpha_1\alpha_2$. We may therefore attempt to define division by declaring $\beta = \alpha_1/\alpha_2$ if and only if $\beta\alpha_2 = \alpha_1 = \alpha_2\beta$.

Remark 9.31. If multiplication were not commutative, the equations $\alpha_1 = \beta\alpha_2$ and $\alpha_1 = \alpha_2\beta$ might well be incompatible conditions for α_1.

To define complex division, let α_1 and α_2 be complex numbers with $\alpha_2 \neq 0$. We wish to write $\alpha_1/\alpha_2 = c_1 + ic_2$, namely, to find formulas for c_1 and c_2 in terms of the real and imaginary parts of the numerator and denominator.

The trick is analogous to rationalizing the denominator in high school algebra: Here we "realify" the denominator, multiplying top and bottom by the conjugate number $\bar{\alpha}_2 = a_2 - ib_2$ and using (9.1):

$$\frac{a_1 + ib_1}{a_2 + ib_2} = \frac{a_1 + ib_1}{a_2 + ib_2} \cdot \frac{a_2 - ib_2}{a_2 - ib_2} = \frac{(a_1a_2 + b_1b_2) + i(-a_1b_2 + a_2b_1)}{a_2^2 + b_2^2}.$$

Example 9.32. To divide $\alpha_1 = 2 - i$ by $\alpha_2 = 4 + 3i$, calculate as follows:

$$\frac{2-i}{4+3i} = \frac{2-i}{4+3i} \cdot \frac{4-3i}{4-3i} = \frac{(8-3)+(-6-4)i}{4^2+3^2}$$
$$= \frac{5-10i}{25} = \frac{1-2i}{5}.$$

In practice, direct calculation is easier than memorizing the formula.

Example 9.33. If $\alpha = a + bi \neq 0$, then

$$\frac{1}{\alpha} = \frac{1}{a+ib} = \frac{a-ib}{a^2+b^2} = \frac{a}{a^2+b^2} - i\,\frac{b}{a^2+b^2}.$$

That is, every non-zero complex number has a reciprocal.

The arithmetic operations on complex numbers satisfy familiar rules of algebra.

Example 9.34. For all *complex* α and β, the *difference of squares* identity holds: $\alpha^2 - \beta^2 = (\alpha + \beta)(\alpha - \beta)$.

Example 9.35. If $\alpha x^2 + \beta x + \gamma = 0$ with α, β, and γ complex and $\alpha \neq 0$, then

$$x = \frac{-\beta \pm \sqrt{\beta^2 - 4\alpha\gamma}}{2\alpha},$$

by the same completing-the-square proof you have seen for real coefficients. There are no "exceptional" cases; every quadratic has exactly two complex solutions, counting multiplicity.

Complex multiplication has a beautiful and useful geometric interpretation, most easily expressed in terms of *polar coordinates*. Recall that every point (a, b) in the plane can be written $(r\cos\theta, r\sin\theta)$ for some radius $r \geq 0$ and some angle θ, measured counterclockwise from the positive x axis and unique up to an added integer multiple of 2π.

Definition 9.36. Let $\alpha = a + bi = r\cos\theta + ir\sin\theta$ be a complex number. The radius r is called the *magnitude* of α, and the polar angle is the *argument* of α. If $-\pi < \theta < \pi$, we say θ is the *principal argument* of α.

Remark 9.37. The magnitude of $\alpha = a + ib$, denoted $|\alpha|$, is given by (9.1):

$$|\alpha| = r = \sqrt{a^2+b^2} = \sqrt{\alpha\bar{\alpha}}.$$

Example 9.38. Since $i = 0 + 1 \cdot i = \cos\frac{\pi}{2} + i\sin\frac{\pi}{2}$, the magnitude of i is 1 and the principal argument of i is $\frac{\pi}{2}$.

Example 9.39. Let θ be a real number. By *Euler's formula* (see appendix), we have $\cos\theta + i\sin\theta = e^{i\theta}$. The magnitude of $e^{i\theta}$ is 1, and the argument is θ.

Generally, $\alpha = |\alpha|(\cos\theta + i\sin\theta) = |\alpha|e^{i\theta}$.

If $e^{i\theta_1} = (\cos\theta_1 + i\sin\theta_1)$ and $e^{i\theta_2} = (\cos\theta_2 + i\sin\theta_2)$ are complex numbers of unit magnitude, the sum formulas for the cosine and sine functions allow us to write their product as

$$\begin{aligned} e^{i\theta_1} \cdot e^{i\theta_2} &= (\cos\theta_1 + i\sin\theta_1)(\cos\theta_2 + i\sin\theta_2) \\ &= (\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2) + i(\cos\theta_1\sin\theta_2 + \cos\theta_2\sin\theta_1) \\ &= \cos(\theta_1+\theta_2) + i\sin(\theta_1+\theta_2) = e^{i(\theta_1+\theta_2)}. \end{aligned}$$

That is, *the law of exponents holds for imaginary exponents.* Since every complex number has polar form $\alpha = |\alpha|\, e^{i\theta}$, complex multiplication satisfies

$$\alpha_1\alpha_2 = \big(|\alpha_1|\, e^{i\theta_1}\big)\big(|\alpha_2|\, e^{i\theta_2}\big) = \big(|\alpha_1|\,|\alpha_2|\big)\, e^{i(\theta_1+\theta_2)}.$$

Geometrically, we multiply two complex numbers by multiplying their magnitudes and adding their arguments (polar angles).

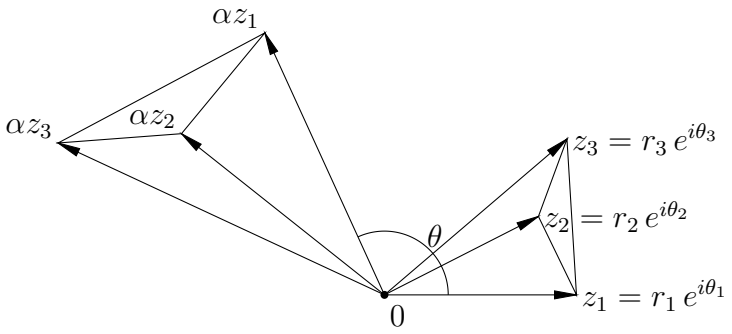


Figure 9.4: Complex multiplication by $\alpha = |\alpha|e^{i\theta}$.

Example 9.40. Since $i = \cos\frac{\pi}{2} + i\sin\frac{\pi}{2} = e^{i\frac{\pi}{2}}$, we have

$$i\alpha = i|\alpha|\, e^{i\theta} = |\alpha|\, e^{i(\theta+\frac{\pi}{2})};$$

again we see that multiplication by i rotates the plane about the origin by a quarter turn counterclockwise.

Definition 9.41. A set A contained in $\mathbf{C}$ is *closed under multiplication* if, for all α_1 and α_2 in A, the product $\alpha_1 \cdot \alpha_2$ is an element of A.

Example 9.42. The set of complex numbers of magnitude 1 is the *unit circle*

$$U(1) = \{z \text{ in } \mathbf{C} : |z| = 1\} = \{z \text{ in } \mathbf{C} : z = e^{i\theta} \text{ for some real } \theta\}.$$

The set $U(1)$ is closed under multiplication: If $|\alpha_1| = 1$ and $|\alpha_2| = 1$, i.e., α_1, $\alpha_2 \in U(1)$, then $|\alpha_1\alpha_2| = |\alpha_1||\alpha_2| = 1$, so $\alpha_1\alpha_2 \in U(1)$.

Example 9.43. The *finite* subsets $\{1\}$ and $\{-1, 1\}$ of $U(1)$ are also closed under multiplication. More generally, for each positive integer n there exists a subset U_n of $U(1)$ that contains exactly n elements and is closed under multiplication:

$$\begin{aligned} U_n &= \{1 = e^0,\ e^{i\,2\pi/n},\ e^{i\,4\pi/n},\ \ldots,\ e^{i\,2\pi(n-1)/n}\} \\ &= \{e^{i\,2\pi k/n} : k = 0, \ldots, n-1\}. \end{aligned}$$

Figure 9.5 depictes the cases $n = 4$ and $n = 6$.

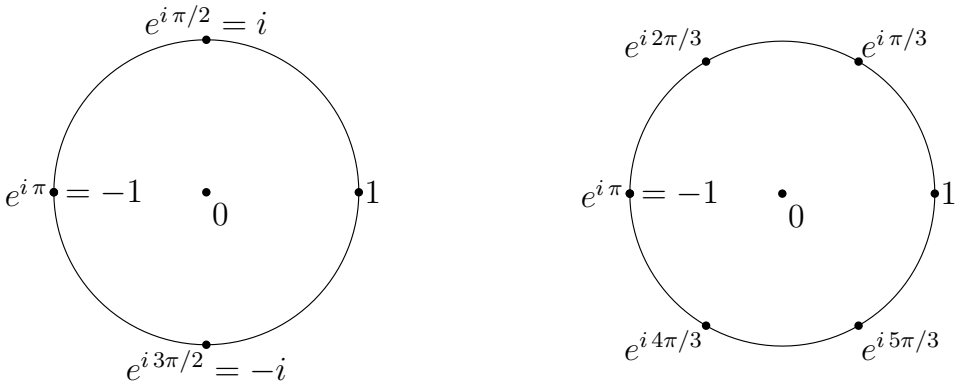


Figure 9.5: The unit circle, $U(1)$, and two finite subsets, U_4 and U_6, that are closed under multiplication.

The elements of U_n are precisely the complex numbers $\zeta = re^{i\theta}$ satisfying the equation $\zeta^n = 1$, namely the so-called *nth roots of unity.* To see why, note that $1 = \zeta^n = r^n e^{in\theta}$ precisely when $r = 1$ and $n\theta$ is an integer multiple of 2π. Assuming without loss of generality that $0 \le \theta < 2\pi$, we have $0 \le n\theta < 2n\pi$, so that $n\theta = 0,\ 2\pi,\ 4\pi,\ \ldots,\ 2(n-1)\pi$, or $n\theta = 2k\pi$ for some integer k with $0 \le k < n$.

To see that the set of nth roots of unity is closed under multiplication, note that if $\zeta_1^n = 1$ and $\zeta_2^n = 1$, then $(\zeta_1\zeta_2)^n = \zeta_1^n\,\zeta_2^n = 1$, which means $\zeta_1\zeta_2$ is an nth root of unity.

9.3 Integer Powers

The Laws of Exponents

Theorem 9.44 (The Law of Exponents)**.** *If α and β are non-zero complex numbers, then*

$$\left.\begin{array}{ll} \text{(i)} & (\alpha\beta)^n = (\alpha^n)(\beta^n), \\ \text{(ii)} & \alpha^{m+n} = \alpha^m \cdot \alpha^n, \\ \text{(iii)} & \alpha^{nm} = (\alpha^n)^m, \end{array}\right\} \quad \textit{for all integers } m \textit{ and } n.$$

In particular, $\alpha^{-n} = (\alpha^{-1})^n$ for all $\alpha \neq 0$ and all integers n.

Remark 9.45. Conceptually, these results amount to counting the number of factors in a product. For instance, α^{m+n} represents a product of $(m+n)$ factors all equal to α; such a product can be separated into a product of m factors and a product of n factors, i.e., into $\alpha^m \cdot \alpha^n$.

Proof. (Proof of (i)). For each natural number n, consider the statement

$$P(n): \qquad (\alpha\beta)^n = (\alpha^n)(\beta^n).$$

The base case $P(0)$ reduces to $1 = 1$, which is true. Assume inductively that $P(k)$ is true for some $k \geq 0$. We have

$$\begin{aligned} (\alpha\beta)^{k+1} &= (\alpha\beta)^k(\alpha\beta) && \text{Definition of exponentiation,} \\ &= (\alpha^k\beta^k)(\alpha\beta) && \text{Inductive hypothesis,} \\ &= \alpha^k(\beta^k \cdot \alpha)\beta && \text{Associativity,} \\ &= \alpha^k(\alpha \cdot \beta^k)\beta && \text{Commutativity,} \\ &= (\alpha^k \cdot \alpha)(\beta^k \cdot \beta) && \text{Associativity,} \\ &= (\alpha^{k+1})(\beta^{k+1}) && \text{Definition of exponentiation.} \end{aligned}$$

Since $P(0)$ is true and $P(k)$ implies $P(k+1)$ for all $k \geq 0$, the statement $P(n)$ is true for all $n \geq 0$ by the principle of mathematical induction. If $n < 0$, replace α and β by their multiplicative inverses.

Taking $\beta = \alpha^{-1}$, we have

$$(\alpha^n)(\alpha^{-1})^n = 1^n = 1 = (\alpha^n)(\alpha^n)^{-1} = (\alpha^n)(\alpha^{-n});$$

by cancellation, $(\alpha^{-1})^n = \alpha^{-n}$ for all n.

(Proof of (ii)). We first assume m and n are non-negative integers. For each n in $\mathbf{N}$, consider the statement

$$P(n) \qquad \alpha^{m+n} = \alpha^m \cdot \alpha^n \quad \text{for all } m \text{ in } \mathbf{N}.$$

This *single statement* may be viewed as an infinite family of statements, one for each natural number m, with n a fixed natural number.

The base case $P(0)$ asserts $\alpha^{m+0} = \alpha^m \cdot \alpha^0$ for all m, which is true since $m + 0 = m$ for all m and $\alpha^0 = 1$. Next, assume inductively that $P(k)$ is true for some k, namely that

$$\alpha^{m+k} = \alpha^m \cdot \alpha^k \quad \text{for all } m \text{ in } \mathbf{N}.$$

For all m, we have

$$\begin{aligned}
\alpha^{m+k+1} &= \alpha^{m+k} \cdot \alpha && \text{Definition of exponentiation} \\
&= (\alpha^m \cdot \alpha^k) \cdot \alpha && \text{Inductive hypothesis} \\
&= \alpha^m \cdot (\alpha^k \cdot \alpha) && \text{Associativity} \\
&= \alpha^m \cdot \alpha^{k+1} && \text{Definition of exponentiation}
\end{aligned}$$

which establishes the inductive step. By the principle of mathematical induction, $\alpha^{m+n} = \alpha^m \cdot \alpha^n$ for all non-negative m and n.

If m and n are both non-positive, conclusion (ii) of the theorem now follows immediately by replacing α with α^{-1}.

It remains to check the case where one exponent is positive, the other negative. Without loss of generality, say $-m$ and n are positive.

Suppose first that $0 \leq m + n$. Since $-m$ is positive, the preceding argument shows

$$\alpha^n = \alpha^{(m+n)+(-m)} = \alpha^{m+n} \cdot \alpha^{-m}.$$

Multiplying both sides by α^m gives $\alpha^{m+n} = \alpha^m \cdot \alpha^n$, as claimed.

If instead $m + n < 0$, then $0 < -(m + n)$, and

$$\alpha^{-m} = \alpha^{-(m+n)+n} = \alpha^{-(m+n)} \cdot \alpha^n;$$

rearranging establishes the asserted claim.

(Proof of (iii)). For m and n non-negative, this follows by induction on the statement

$$P(n) \qquad \alpha^{nm} = (\alpha^n)^m \quad \text{for all } m \text{ in } \mathbf{N}.$$

To establish the inductive step, note that

$$\alpha^{n(m+1)} = \alpha^{nm+n} = \alpha^{nm} \cdot \alpha^n = (\alpha^n)^m \cdot \alpha^n = (\alpha^n)^{m+1}$$

by (ii) and the definition of exponentiation. As a fringe benefit, we find that

$$(\alpha^m)^n = \alpha^{mn} = \alpha^{nm} = (\alpha^n)^m \quad \text{for all } \alpha \neq 0, \text{ all } m \text{ and } n \text{ in } \mathbf{N}.$$

If $m < 0$ or $n < 0$, replace α by α^{-1} and use $(\alpha^{-1})^n = \alpha^{-n}$. □

The Binomial Theorem

The identity $(\alpha+\beta)^2 = \alpha^2+2\alpha\beta+\beta^2$ is doubtless familiar. The *binomial theorem* generalizes to arbitrary positive integer powers $(\alpha + \beta)^n$.

Theorem 9.46 (Binomial Theorem)**.** *If α and β are complex numbers and n is a non-negative integer,*

$$\begin{aligned}(\alpha + \beta)^n &= \sum_{k=0}^{n} \binom{n}{k} \alpha^{n-k}\beta^k \\ &= \tbinom{n}{0}\alpha^n + \tbinom{n}{1}\alpha^{n-1}\beta + \tbinom{n}{2}\alpha^{n-2}\beta^2 + \cdots + \tbinom{n}{n}\beta^n.\end{aligned}$$

Proof. Conceptually, the n-fold product

$$(\alpha + \beta)^n = (\alpha + \beta)(\alpha + \beta)\cdots(\alpha + \beta)$$

is expanded by the following procedure:

- Pick an arbitrary integer k with $0 \leq k \leq n$;
- Distribute k check marks among the n copies of $(\alpha + \beta)$;
- If a copy of $(\alpha + \beta)$ is unchecked, choose α from that copy; otherwise choose β. Multiply the resulting n factors to get $\alpha^{n-k}\beta^k$.
- Sum over all k and all ways of distributing k check marks.

By the first and third points, the expanded product has the form

$$(\alpha+\beta)^n = \_\_\alpha^n + \_\_\alpha^{n-1}\beta + \_\_\alpha^{n-2}\beta^2 + \cdots + \_\_\alpha\beta^{n-1} + \_\_\beta^n$$

for some coefficients. By the second and fourth points, the coefficient of $\alpha^{n-k}\beta^k$ is $\binom{n}{k}$, the number of distinct ways of distributing k check marks among n parenthesized binomials. This completes the proof. □

Pascal's Triangle

The binomial coefficients for any particular exponent n can be found in the $(n+1)$th row of a recursive diagram known as "Pascal's triangle".

$$
\begin{array}{ccccccccccc}
 & 0 & & 0 & & 0 & & \mathbf{1} & & 0 & & 0 & & 0 & & 0 & \dots \\
0 & & 0 & & 0 & & \mathbf{1} & & \mathbf{1} & & 0 & & 0 & & 0 & \dots & \\
 & 0 & & 0 & & \mathbf{1} & & \mathbf{2} & & \mathbf{1} & & 0 & & 0 & & 0 & \dots \\
0 & & 0 & & \mathbf{1} & & \mathbf{3} & & \mathbf{3} & & \mathbf{1} & & 0 & & 0 & \dots & \\
 & 0 & & \mathbf{1} & & \mathbf{4} & & \mathbf{6} & & \mathbf{4} & & \mathbf{1} & & 0 & & 0 & \dots \\
0 & & \mathbf{1} & & \mathbf{5} & & \mathbf{10} & & \mathbf{10} & & \mathbf{5} & & \mathbf{1} & & 0 & \dots & \\
 & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots &
\end{array}
$$

	$k=-1$	$k=0$	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$\cdots$
$n=0$	0	**1**	0	0	0	0	0	$\cdots$
$n=1$	0	**1**	**1**	0	0	0	0	$\cdots$
$n=2$	0	**1**	**2**	**1**	0	0	0	$\cdots$
$n=3$	0	**1**	**3**	**3**	**1**	0	0	$\cdots$
$n=4$	0	**1**	**4**	**6**	**4**	**1**	0	$\cdots$
$n=5$	0	**1**	**5**	**10**	**10**	**5**	**1**	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	

Table 9.2: Pascal's triangle, "classic" format (top) and tabular.

Imagine expanding successive powers of $(\alpha+\beta)$ recursively, in as lazy a manner as possible. Because

$$(\alpha+\beta)^{n+1} = (\alpha+\beta)^n(\alpha+\beta) = (\alpha+\beta)^n\alpha + (\alpha+\beta)^n\beta,$$

knowledge of $(\alpha+\beta)^n$ can be "recycled" in calculating $(\alpha+\beta)^{n+1}$.

Starting from $n=2$ and $(\alpha+\beta)^2 = \alpha^2 + 2\alpha\beta + \beta^2$, we find

$$
\begin{aligned}
(\alpha+\beta)^3 &= (\alpha^2+2\alpha\beta+\beta^2)\alpha + (\alpha^2+2\alpha\beta+\beta^2)\beta \\
&= \alpha^3 + 2\alpha^2\beta + \alpha\beta^2 \\
&\qquad + \alpha^2\beta + 2\alpha\beta^2 + \beta^3 \\
&= \alpha^3 + 3\alpha^2\beta + 3\alpha\beta^2 + \beta^3.
\end{aligned}
$$

Recycling this formula gives

$$\begin{aligned}(\alpha+\beta)^4 &= (\alpha^3+3\alpha^2\beta+3\alpha\beta^2+\beta^3)\alpha + (\alpha^3+3\alpha^2\beta+3\alpha\beta^2+\beta^3)\beta \\ &= \alpha^4+3\alpha^3\beta+3\alpha^2\beta^2+\ \alpha\beta^3 \\ &\qquad +\ \alpha^3\beta\ +3\alpha^2\beta^2+3\alpha\beta^3+\beta^4 \\ &= \alpha^4+4\alpha^3\beta+6\alpha^2\beta^2+4\alpha\beta^3+\beta^4.\end{aligned}$$

To be sure you understand, check in detail that

$$(\alpha+\beta)^5 = \alpha^5+5\alpha^4\beta+10\alpha^3\beta^2+10\alpha^2\beta^3+5\alpha\beta^4+\beta^5.$$

A useful pattern is emerging: If the coefficients for $(\alpha+\beta)^n$ are laid out in a row, "duplicate" the row underneath itself, shift the second copy one entry to the right, and add in columns to get the coefficients for $(\alpha+\beta)^{n+1}$. Alternatively, "pad" the coefficients on either side with an infinite row of 0's. Then, to get the entries in the "next" row, add each entry to its neighbor on the left. The result is *Pascal's triangle*.

Table 9.2 shows two versions of Pascal's triangle. On top is the "classic" format, in which each entry below the first row is the sum of its "parents", the two nearest entries in the preceding row.

In the tabular formatting of Pascal's triangle, the entries are arranged so that n indexes the rows and k indexes the columns.

Example 9.47. The binomial theorem can be used with specific numbers. For example,

$$\begin{aligned}11^3 = (10+1)^3 &= 10^3 \ +3\cdot 10^2\cdot 1+3\cdot 10\cdot 1^2+1^3 \\ &= 1000+300 \qquad +30 \qquad +1 = 1331.\end{aligned}$$

The digits constitute the fourth row of Pascal's triangle.

9.4 Real and Complex Mappings

Example 9.48. Define $f_1 : \mathbf{R} \to [0,\infty)$ by $f_1(x) = x^2$. This mapping is surjective (every non-negative real y can be written as $x^2 = f_1(x)$ for at least one real x), but not injective (since $f_1(-1) = 1 = f_1(1)$, but $-1 \neq 1$).

Example 9.49. Define $f_2 : (0,\infty) \to \mathbf{R}$ by $f_2(x) = x^2$. This mapping is not surjective (there is no real x such that $x^2 = f_2(x) = -1$), but *is*

injective. To establish injectivity, suppose $a_1^2 = f_2(a_1) = f_2(a_2) = a_2^2$. Subtracting and factoring, we find $0 = a_2^2 - a_1^2 = (a_2 - a_1)(a_2 + a_1)$, which implies $a_1 = a_2$ or $a_1 = -a_2$. The latter is impossible since a_1 and a_2 are positive by hypothesis.

We have shown that if $f_2(a_1) = f_2(a_2)$, then $a_1 = a_2$. Since a_1 and a_2 were arbitrary, f_2 is injective.

Note carefully that the mappings f_1 and f_2 in these examples are defined by the same formula, but have distinct domains and/or codomains.

Example 9.50. Let $A = \{-1, 0, 1\} \subseteq \mathbf{C}$. The mapping $f : A \to A$ defined by $f(z) = z^2$ is neither injective nor surjective. As is easily checked, $f(z) = -1$ has no solution (so f is not surjective) while $f(-1) = 1 = f(1)$ (so f is not injective).

The mapping $g : A \to A$ defined by $g(z) = z^3$ is bijective. In fact, $g(z) = z$ for all z in A.

Example 9.51. Let $\zeta = e^{2\pi i/3} = \frac{1}{2}(-1 + i\sqrt{3})$, and consider the set $A = \{1, \zeta, \zeta^2\} \subseteq \mathbf{C}^\times$, the set of non-zero complex numbers. Since ζ is a cube root of unity, $(\zeta^2)^2 = \zeta^4 = \zeta$ and $(\zeta^2)^3 = 1$.

The mapping $f : A \to A$ defined by $f(z) = z^2$ is bijective: $1 = f(1)$, $\zeta = f(\zeta^2)$, and $\zeta^2 = f(\zeta)$.

The mapping $g : A \to A$ defined by $g(z) = z^3$ is neither injective nor surjective. Indeed, $f(z) = 1$ for every z in A.

Example 9.52. Let $f : [-1, 1] \to \mathbf{R}$ and $g : \mathbf{R} \to [-1, 1]$ be defined by $f(x) = \arcsin x$, $g(x) = \sin x$. The mapping f is injective but not surjective (why?), g is surjective but not injective (why?), while $gf : [-1, 1] \to [-1, 1]$ is the identity map (which is bijective), and $fg : \mathbf{R} \to \mathbf{R}$ is neither injective nor surjective.

Example 9.53. Define $f : \mathbf{R} \to [0, \infty)$ by $f(x) = x^2$, see Figure 9.6, left. For each $y > 0$, there exist two real x such that $f(x) = y$, namely $x = \pm\sqrt{y}$. In particular, there are two "obvious" branches of f^{-1}, defined by $g_\pm(y) = \pm\sqrt{y}$ for $y \geq 0$, see Figure 9.6, right. (There are infinitely many other choices, though all are discontinuous.) For any such choice, $(fg)(y) = f\big(g(y)\big) = y$ for all $y \geq 0$. What about $g\big(f(x)\big)$?

Example 9.54. The squaring mapping $f : \mathbf{C} \to \mathbf{C}$ sends $z = x + yi$ to $z^2 = (x^2 - y^2) + 2xyi$. In polar coordinates, the mapping takes the form $f(re^{i\theta}) = r^2 e^{2i\theta}$.

Pairs of points z and $-z$ are identified by f. Conversely, if $z_1^2 = z_2^2$, then $0 = z_2^2 - z_1^2 = (z_2 - z_1)(z_2 + z_1)$, so either $z_2 = z_1$ or $z_2 = -z_1$.

The open half-plane $H = \{z \text{ in } \mathbf{C} : \operatorname{Re} z > 0\}$ contains no pair $\{z, -z\}$, so f restricted to H is injective. On H, the polar angle function satisfies $|\theta| < \pi/2$, so the image of H under f is contained in the "slit plane" where $|\theta| < \pi$, namely, $\mathbf{C} \setminus (-\infty, 0]$.

There is a branch of inverse $g_+ : \mathbf{C} \setminus (-\infty, 0] \to H$ that sends each complex number $\rho e^{i\phi}$ with $|\phi| < \pi$ to $\sqrt{\rho} e^{i\phi/2}$.

There is also a branch of inverse $g_- : \mathbf{C} \setminus (-\infty, 0] \to -H$ sending each complex number $\rho e^{i\phi}$ with $|\phi| < \pi$ to $-\sqrt{\rho} e^{i\phi/2} = \sqrt{\rho} e^{i(\phi/2+\pi)}$.

Each branch is undefined on the negative real axis, and neither can be defined continuously on the negative real axis. Instead, the values of g_+ along the upper edge of the negative real axis "patch together with" the values of g_- along the lower edge of the negative real axis, and *vice versa*.

Example 9.55. Two real numbers θ_1 and θ_2 determine the same longitude on the earth if and only if their difference is a multiple of one full turn, say 360°. To formalize this in the language of quotients, let $A = \mathbf{R}$ be the set of real numbers (a.k.a. the number line), and define the relation R by $\theta_1 R \theta_2$ if and only if $\theta_2 - \theta_1$ is an integer multiple of 360. By an argument entirely similar to that given for the parity relation in Example 7.41, R is an equivalence relation.

The set of equivalence classes is indexed by the half-open interval $[0, 360)$, since every angle is equivalent mod R to a unique number between 0 and 360 (excluding 360, which is equivalent to 0). We call this set the "space of angles".

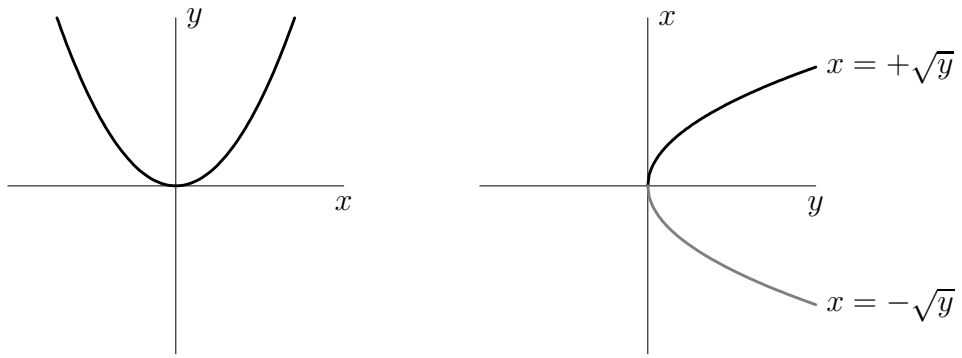


Figure 9.6: Right inverses of $f(x) = x^2$.

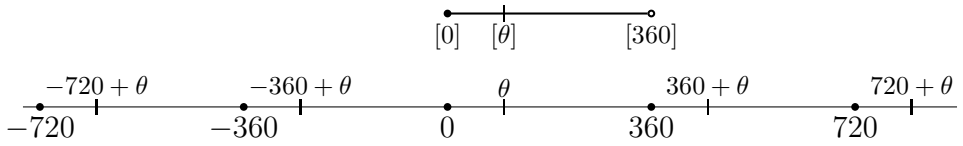
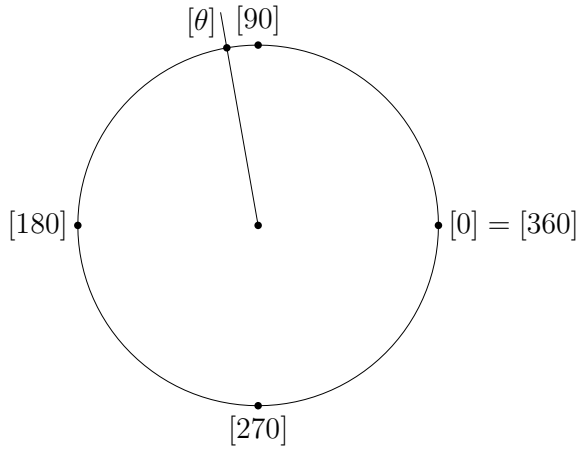


Figure 9.7: The number line, and the space of angles.

Example 9.56. Let $A = \mathbf{R}$ be the set of real numbers, R the "longitude" relation, and define $f : \mathbf{R} \to \mathbf{R}^2$ by $f(t) = (\cos t, \sin t)$, the standard trigonometric parametrization of the circle, with trig functions in "degrees mode".



If $\theta_2 - \theta_1$ is an integer multiple of 360, then $\cos\theta_1 = \cos\theta_2$ and $\sin\theta_1 = \sin\theta_2$, so $f(\theta_1) = f(\theta_2)$. Consequently, there is an induced mapping from the space of angles to the unit circle in the plane. In words, f factors through locations on the earth.

Since $\cos\theta_1 = \cos\theta_2$ and $\sin\theta_1 = \sin\theta_2$ if *and only if* $\theta_2 - \theta_1$ is an integer multiple of 360, the mapping $\bar{f}$ is bijective, so *the space of angles may be regarded as the unit circle.* Geometrically, each equivalence class $[\theta]$ corresponds to a unique point of the unit circle.

Exercises

Exercise 9.1. Use the binomial theorem to calculate $9^3 = (10-1)^3$, $11^4 = (10+1)^4$, and $12^3 = (10+2)^3$. Do not use a calculator.

Exercise 9.2. Without using a calculator, express each of the fractions $\frac{1}{7}$, $\frac{2}{7}$, $\frac{3}{7}$, $\frac{4}{7}$, $\frac{5}{7}$, and $\frac{6}{7}$ as repeating decimals.

Suggestion: You'll need to do long division at least once, but can reduce the amount of computation by exploiting patterns.

Exercise 9.3. Let r be a complex number, and let

$$S_n(r) = \sum_{k=0}^{n} r^k = 1 + r + r^2 + \cdots + r^n.$$

(a) Use induction to prove $1 + rS_n(r) = S_{n+1}(r)$ for all $n \geq 0$.

(b) Use part (a) and the identity $S_{n+1}(r) = S_n(r) + r^{n+1}$ to prove $(1 - r)S_n(r) = 1 - r^{n+1}$ for all $n \geq 0$.

(c) Find a closed expression (the *finite geometric series formula*) for $S_n(r)$. (Handle the cases $r = 1$ and $r \neq 1$ separately.)

(d) Calculate (and simplify): $\sum_{k=0}^{n} 9 \cdot \left(\frac{1}{10}\right)^k$, $\sum_{k=0}^{n} (-1)^k$, $\sum_{k=0}^{100} \frac{1}{2} \cdot \left(\frac{1}{4}\right)^k$.

Exercise 9.4. Express $1.\overline{12}$, $0.24\overline{9}$, and $0.041\overline{6}$ as reduced fractions.

Exercise 9.5. Let x and y be complex numbers. Prove that

$$xy = \tfrac{1}{2}\big((x + y)^2 - x^2 - y^2\big).$$

In words, multiplication is uniquely determined by squaring.

Exercise 9.6. Let x, y, u, and v be complex. Prove that $2u = x + y$ and $2v = x - y$ if and only if $x = u + v$ and $y = u - v$.

Exercise 9.7. Let x and y be real. Prove that $x^2 = y^2$ if and only if $x = y$ or $x = -y$. (Use axioms and results from the text. Do not use square roots, whose existence has not yet been established.)

Exercise 9.8. Let x and y be real numbers. Prove that

(a) $2|xy| \leq x^2 + y^2$, and the inequality is strict unless $|x| = |y|$.

(b) $0 \leq x^2 + xy + y^2$, and the inequality is strict unless $|x| = |y| = 0$.

Exercise 9.9. Let x_0 and $r > 0$ be real. Prove that if x is real, then

(a) $|x - x_0| < r$ if and only if $x_0 - r < x < x_0 + r$.

(b) $0 < |x - x_0| < r$ if and only if $x_0 - r < x < x_0$ or $x_0 < x < x_0 + r$.

Exercise 9.10. Alice and Bob have two sandwiches. Because Alice isn't very hungry, they will split one sandwich evenly, and Bob will eat the other.

(a) In order to distribute the food most nearly equally, should they split the smaller or the larger sandwich?

(b) Formulate a precise mathematical model for the sandwiches, and use your model to give a proof of your assertion in part (a).

Exercise 9.11. Let x and y be real numbers, and assume $x < y$.

(a) Prove that $x < \frac{1}{2}(x+y) < y$ and $x < \frac{1}{3}(2x+y) < \frac{1}{3}(x+2y) < y$.

(b) More generally, show that if $0 < t < 1$, then $x < (1-t)x + ty < y$. (The expression $(1-t)x + ty$ is called a *convex linear combination* of x and y.)

(c) If $s < t$ are real, then $(1-s)x + sy < (1-t)x + ty$.

Exercise 9.12. If $a \in \mathbf{R}$, define the *absolute value* of a by

$$|a| = \begin{cases} a & \text{if } 0 \le a \\ -a & \text{if } a < 0 \end{cases}$$

For arbitrary real numbers a and b, prove:

(a) $|-a| = |a|$.

(b) $|ab| = |a| \cdot |b|$.

(c) $-|a| \le a \le |a|$.

(d) $|b| \le a$ if and only if $-a \le b \le a$.

(e) $|a+b| \le |a| + |b|$. (The *triangle inequality.*)
Hint: Apply part (c) to a and b separately, add the inequalities, and use part (d).

(f) $\big||a| - |b|\big| \le |a-b|$. (The *reverse triangle inequality.*)
Hint: First apply the triangle inequality to $a = (a-b) + b$ and to $b = a + (b-a)$.

Exercise 9.13. Let a and b be real numbers. Define the *minimum* and *maximum* of a and b by

$$\min(a,b) = \begin{cases} a & \text{if } a \le b, \\ b & \text{if } b < a, \end{cases} \qquad \max(a,b) = \begin{cases} b & \text{if } a \le b, \\ a & \text{if } b < a. \end{cases}$$

(a) Prove $a + b = \min(a, b) + \max(a, b)$.
Give two proofs: One involving verbal explanation, and one based on manipulation of formulas.

(b) Prove $|a - b| = \max(a, b) - \min(a, b)$.
Suggestion: Again, give two proofs. For the algebraic proof, consider two cases, $a \leq b$ and $b < a$.

(c) Use parts (a) and (b) to find algebraic formulas for $\min(a, b)$ and $\max(a, b)$.

Exercise 9.14. If $A = \{a_1, a_2, \ldots, a_n\}$ is a *finite* set of real numbers, there exist unique elements $\max A$ and $\min A$ in A such that

$$\min A \leq x \leq \max A \quad \text{for all } x \text{ in } A.$$

Suggestion: Use induction on the number of elements of A to prove existence.

Exercise 9.15. Let a, b, and c be real numbers, and assume $a > 0$. By completing the square, show that $ax^2 + bx + c \geq (4ac - b^2)/(4a)$ for all real x, with equality if and only if $x = -b/(2a)$.

Exercise 9.16. Let $n > 0$ be an odd integer. Prove that if $x < y$ are real numbers, then $x^n < y^n$.

Exercise 9.17. Give a formal proof by mathematical induction that if x_1, x_2, $\ldots$, x_n are real numbers, then

$$|x_1 + x_2 + \cdots + x_n| \leq |x_1| + |x_2| + \cdots + |x_n|.$$

Exercise 9.18. Let $u > 0$ be real and n a non-negative integer.

(a) Use induction to prove $1 + nu + \frac{n(n-1)}{2}\, u^2 \leq (1 + u)^n$.

(b) Use the binomial theorem to prove $1 + nu + \frac{n(n-1)}{2}\, u^2 \leq (1 + u)^n$.

Exercise 9.19. With the help of the binomial theorem, expand:

(a) $(a + b)^3$, $(a - b)^3$, and $\frac{1}{2}\big[(a + b)^3 \pm (a - b)^3\big]$.

(b) $(a + b)^4$, $(a - b)^4$, and $\frac{1}{2}\big[(a + b)^4 \pm (a - b)^4\big]$.

(c) $(a + b)^6$. (Hint: Extend Pascal's triangle.)

Exercise 9.20. Suppose $i^2 = -1$. Use the binomial theorem to expand:

(a) $(x+iy)^2$. (b) $(x+iy)^3$. (c) $(x+iy)^4$.

In each part, separate the real and imaginary parts.

Exercise 9.21. Use the binomial theorem to establish the identities:

(a) $\displaystyle\sum_{k=0}^{n} \binom{n}{k} = 2^n$ for $n \geq 0$. (b) $\displaystyle\sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0$ for $n \geq 1$.

Exercise 9.22. Let α, β, and γ be complex numbers, and $n \geq 0$ an integer. State and prove a "trinomial theorem" for $(\alpha+\beta+\gamma)^n$.

Exercise 9.23. Let $A = \mathbf{Z}$, and define a relation R by aRb if and only if $b-a$ is an integer multiple of 4.

(a) Prove R is an equivalence relation, and find the equivalence classes of R, and describe the quotient $\mathbf{Z}/R$.

(b) Let $f : \mathbf{Z} \to \mathbf{C}$ be defined by $f(a) = (-1)^a$. Prove f is well-defined mod R. Is $\bar{f}$ injective?

(c) Let $g : \mathbf{Z} \to \mathbf{C}$ be defined by $g(a) = i^a$. Is g well-defined mod R? If so, is $\bar{g}$ injective?

Exercise 9.24. Define $f : \mathbf{C} \to \mathbf{C}$ by $f(z) = z^2$.

(a) By writing $z = x + iy$ with x and y real, calculate the real and imaginary parts of $f(z)$.

(b) By writing $z = re^{i\theta}$ with $r \geq 0$ and θ real, re-calculate $f(z)$, and use your result to describe the geometric action of the mapping f.

(c) Find the preimages of the singletons $\{1\}$, $\{-1\}$, $\{i\}$, and $\{\rho e^{i\phi}\}$.

Exercise 9.25. Repeat the preceding question for $f : \mathbf{C} \to \mathbf{C}$ defined by $f(z) = z^3$. In part (c), do you notice any "geometric pattern"?

Exercise 9.26. Let $n > 1$ be an integer, and define $f : \mathbf{C} \to \mathbf{C}$ by $f(z) = z^n$. By writing $z = re^{i\theta}$, describe the geometric action of f, and find the preimage of $\{\rho e^{i\phi}\}$. If $\rho > 0$, how many points are in the preimage, and how are these points situated geometrically in $\mathbf{C}$?

Exercise 9.27. Let $g : \mathbf{R} \to \mathbf{R}$ be a real-valued function of one real variable. We say g is *even* if $g(-x) = g(x)$ for all x in $\mathbf{R}$, and that g is *odd* if $g(-x) = -g(x)$ for all x in $\mathbf{R}$. (Analogous formulas define the notions of "even" and "odd" functions whose domain and/or codomain is $\mathbf{Z}$ or any other set in which negatives are defined.)

(a) Find all functions that are *both* even and odd.

(b) Let $f : \mathbf{R} \to \mathbf{R}$ be an arbitrary function. Show the functions

$$f_{\text{even}}(x) = \tfrac{1}{2}\big[f(x) + f(-x)\big], \qquad f_{\text{odd}}(x) = \tfrac{1}{2}\big[f(x) - f(-x)\big]$$

are even and odd, respectively.

(c) Suppose there exist an even function E and an odd function O such that $f(x) = E(x) + O(x)$ for all real x. Find formulas for E and O. Hint: Compute $f(-x)$.

(d) Prove every function $f : \mathbf{R} \to \mathbf{R}$ can be written *uniquely* as the sum of an even function and an odd function. These functions are called the *even part* and *odd part* of f.

(e) Find the even and odd parts of $f(x) = x^3 - 2x^2 + x + 1$, $g(x) = e^x$, and $h(x) = \cos x$.

Exercise 9.28. Let $f : A \to B$ be a mapping, and define a relation on A by $a_1 R a_2$ if and only if $f(a_1) = f(a_2)$.

(a) Prove R is an equivalence relation, and the equivalence classes of R are preimages of singletons, namely *level sets* of f: $f^{-1}(\{b\})$ for some b in B.

(b) Let $f : \mathbf{R} \to \mathbf{R}$ be defined by $f(x) = x^2$. Describe the equivalence classes of f.

(c) Let $f : \mathbf{R}^2 \to \mathbf{R}$ be defined by $f(x, y) = x^2 + y^2$. Describe the equivalence classes of f.

Exercise 9.29. Let $f : A \to B$ be a mapping. A mapping $g : A \to B$ is said to be *constant on the level sets of* f if $f(a_1) = f(a_2)$ implies $g(a_1) = g(a_2)$. (See preceding question.)

(a) Define $f : \mathbf{R}^2 \to \mathbf{R}$ by $f(x, y) = x^2 + y^2$. Which of the following are constant on the level sets of f?

$$g_1(x, y) = \big(1 - \sqrt{x^2 + y^2}\big)^2, \qquad g_2(x, y) = x^2 - y^2, \qquad g_3(x, y) = 1.$$

(b) For a mapping $f : A \to B$, prove the following are equivalent:

(i) g is constant on the level sets of f.

(ii) There exists a mapping $\phi : B \to B$ such that $g = \phi \circ f$. (In this situation we say "g is a function of f".)

Exercise 9.30. For each non-zero real number a, define the mapping $S_a : \mathbf{R} \to \mathbf{R}$ by $S_a(x) = ax$, and let $\mathscr{M}(S) = \{S_a : a \in \mathbf{R}^\times\}$.

(a) Prove that each mapping S_a is a bijection, and find a formula for the inverse mapping.

(b) Show that $\mathscr{M}(S)$ is closed under composition of functions.

(c) Show that $\mathscr{M}(S)$ contains an identity element, and that $\mathscr{M}(S)$ is closed under inversion.

Exercise 9.31. For each non-zero real number a and each real number b, define the mapping $T_{a,b} : \mathbf{R} \to \mathbf{R}$ by $T_{a,b}(x) = ax + b$, and let $\mathscr{M}(T) = \{T_{a,b} : a \in \mathbf{R}^\times, b \in \mathbf{R}\}$.

(a) Prove that each mapping $T_{a,b}$ is a bijection, and find a formula for the inverse mapping.

(b) Show that $\mathscr{M}(T)$ is closed under composition of functions.

(c) Show that $\mathscr{M}(T)$ contains an identity element, and that $\mathscr{M}(T)$ is closed under inversion.

Exercise 9.32. Let A be a non-empty set. The set $\mathbf{R}^A$ of real-valued functions on A consists of *all mappings* $f : A \to \mathbf{R}$. Define binary operations of addition and multiplication on $\mathbf{R}^A$ by

$$(f + g)(a) = f(a) + g(a), \quad (f \cdot g)(a) = f(a) \cdot g(a), \text{ for all } a \text{ in } A.$$

(a) Prove that $+$ is an associative and commutative binary operation on $\mathbf{R}^A$, there is an identity element for $+$, and every element of $\mathbf{R}^A$ has an additive inverse.

(b) Prove that $\cdot$ is an associative and commutative binary operation on $\mathbf{R}^A$. Does $\cdot$ have an identity element? If so, which elements of $\mathbf{R}^A$ have a multiplicative inverse?

Exercise 9.33. This question concerns a set of mappings from the plane $\mathbf{R}^2$ to itself. Let e be the identity map, α the counterclockwise quarter-turn about the origin, α^2 the half-turn about the origin (i.e., α performed twice), and α^3 the clockwise quarter turn about the origin (α performed three times). Let $A = \{e, \alpha, \alpha^2, \alpha^3\}$.

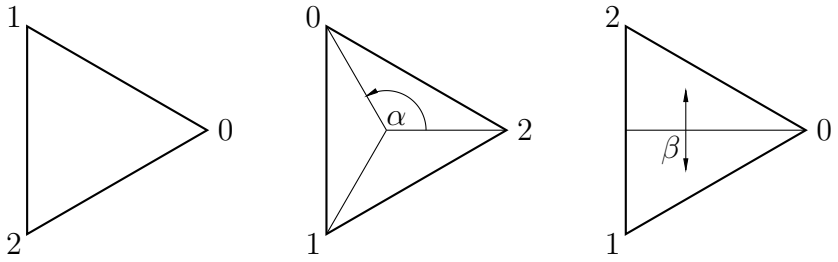
(a) Find formulas for each element of A, and check that your formulas are geometrically sensible. (For example, $e(x, y) = (x, y)$.)

(b) Show A is closed under mapping composition, and write out the Cayley table.

(c) Is composition commutative on A? Does composition have an identity element? If so, which elements of A have inverses?

(d) Explain how your work in this exercise shows the binary operations of Exercise 8.13 are associative.

Exercise 9.34. Consider an equilateral triangle with vertices labeled $0, 1, 2$ (below, left). Let α be a counterclockwise rotation by one-third of a turn about the center (middle), and β the reflection about the horizontal axis (right).



(a) Sketch the six possible configurations of vertex labels on the triangle. (Three of them are shown.)

(b) For each of the following compositions of maps, determine which of your sketches gives the corresponding vertex configuration: α^2, α^3, β^2, $\alpha\beta$, $\alpha^2\beta$, $\beta\alpha$, $\beta\alpha^2$. (Note: Function composition is read right to left. For example, $\alpha\beta$ means "first apply β, then α".)

(c) Let $A = \{e, \alpha, \alpha^2, \beta, \alpha\beta, \alpha^2\beta\}$. Write out the Cayley table for mapping composition on A.
Hints: Find a formula $\beta\alpha = \alpha^k\beta^\ell$ for suitable exponents k and ℓ. To simplify an arbitrary product of αs and βs, use this formula to move all the factors of α to the left.

(d) Is composition commutative on A? Does composition have an identity element? If so, which elements of A have inverses?

(e) Find all proper, non-empty subsets of A that are closed under composition.
Hints: There are five of them. If $A' \subseteq A$ is closed under composition, and if $\alpha \in A'$ and $\beta \in A'$, then $A' = A$. (Why?)

The next two questions concern composition of certain functions. For notational brevity, write f^2 instead of ff, f^3 instead of fff, etc.

Exercise 9.35. In each part, $f(x) = \dfrac{1+x}{1-x}$ for $x \neq -1, 0, 1$.

(a) Compute the compositions f^2, f^3, f^4 and f^5. On the basis of your findings, what are f^{10} and f^{100}?

(b) Let S denote the set of distinct functions found in part (a). List the elements of S, show S is closed under composition, and make a Cayley table for S.

(c) Show composition is a commutative operation on S, there is an identity element, and every element of S has an inverse in S.

Exercise 9.36. In each part, $f(x) = \dfrac{1}{1-x}$ and $g(x) = \dfrac{1}{x}$ for $x \neq 0, 1$.

(a) Compute the compositions f^2, f^3, and f^4. On the basis of your findings, what are f^{10} and f^{2010}?

(b) Compute the composition g^2. Based on this finding, what is g^{1729}?

(c) Compute gf, gf^2 and fg, f^2g.

(d) Let S denote the set of distinct functions found in parts (a)–(c). List the elements of S (there are six), and show each element of S has the form $f^k g^\ell$ for some integers k and ℓ. Show the set S is closed under composition, and make a Cayley table for S.
Hints: Show gf can be written in the stated form. Then argue that in any composition of f and g (in arbitrary order), the "factors" of f can be gathered on the left.

(e) Let $h(x)$ be an arbitrary rational function obtainable by repeatedly composing f and/or g in arbitrary order. Use part (d) to show $h \in S$. (If you can see how, set up a formal argument using mathematical induction.)

(f) Is composition a commutative operation on S? Is there an identity element? Does every element of S have an inverse in S?

Exercise 9.37. The *hyperbolic functions* cosh and sinh are defined by

$$\cosh x = \tfrac{1}{2}(e^x + e^{-x}), \quad \sinh x = \tfrac{1}{2}(e^x - e^{-x}), \qquad x \text{ real.}$$

(a) Show that $\cosh^2 - \sinh^2 = 1$. Carefully sketch the graphs of cosh and sinh on a single set of axes. Suggestion: First calculate $\cosh \pm \sinh$.

(b) Show that for all real x,

$$\cosh(2x) = \cosh^2 x + \sinh^2 x, \qquad \sinh(2x) = 2\cosh x \sinh x.$$

(c) Show that $\cosh' = \sinh$ and $\sinh' = \cosh$.

(d) The *hyperbolic tangent* and *hyperbolic secant* functions are

$$\tanh = \frac{\sinh}{\cosh}, \qquad \operatorname{sech} = \frac{1}{\cosh^2}.$$

Carefully sketch their graphs on a single set of axes, show that $\tanh^2 = 1 + \operatorname{sech}^2$, and find formulas for $\tanh'$ and sech'.

(e) Find an algebraic formula for the inverse function $\tanh^{-1}$. Hint: Solve $y = \tanh x$ for x by cross-multiplying and rearranging.

(f) Find algebraic formulas for $\sinh^{-1}$, and for two branches of $\cosh^{-1}$. Use algebra to show the branches of $\cosh^{-1}$ differ by a sign.
Hint: Solve (e.g.) $y = \sinh x$ for x by multiplying through by e^x and rearranging to get a quadratic in e^x; then use the quadratic formula.

Exercise 9.38. Let x and y be arbitrary real numbers. Show that

$$\begin{aligned}
\cosh(x+y) &= \cosh x \cosh y + \sinh x \sinh y, \\
\sinh(x+y) &= \sinh x \cosh y + \cosh x \sinh y, \\
\tanh(x+y) &= \frac{\tanh x + \tanh y}{1 + \tanh x \tanh y}.
\end{aligned}$$

Exercise 9.39. Let ϕ be a real number, and recall Euler's formula

$$e^{i\phi} = \cos\phi + i \sin\phi.$$

(a) Express $e^{-i\phi}$ in terms of $\cos\phi$ and $\sin\phi$.

(b) Show that

$$\cos\phi = \frac{e^{i\phi} + e^{-i\phi}}{2}, \qquad \sin\phi = \frac{e^{i\phi} - e^{-i\phi}}{2i}.$$

(c) Show that for all real ϕ,

$$\cosh(i\phi) = \cos\phi, \qquad \sinh(i\phi) = i\sin\phi.$$

(The hyperbolic functions are defined in Exercise 9.37.)

Exercise 9.40. Let $n \geq 2$ be an integer, and consider the complex number $\zeta = e^{2\pi i/n}$, an nth root of unity, see Example 9.43. Prove that

$$1 + \zeta + \zeta^2 + \cdots + \zeta^{n-1} = \sum_{k=0}^{n-1} \zeta^k = 0$$

in two ways:

(a) Using the geometric series formula from Exercise 9.3.

(b) Calling the unknown sum S, and multiplying by ζ.

Exercise 9.41. Let $n \geq 2$ be an integer.

(a) Show that the polynomial $z^n - 1$ factors as

$$z^n - 1 = (z-1)(z-\zeta)(z-\zeta^2)\ldots(z-\zeta^{n-1}) = \prod_{k=0}^{n-1}(z-\zeta^k).$$

Hint: Each side has the same roots and the same leading coefficient.

(b) Use part (a) and the geometric series formula to show that

$$\sum_{j=0}^{n-1} z^j = 1 + z + z^2 + \cdots + z^{n-1} = \prod_{k=1}^{n-1}(z-\zeta^k).$$

(c) By setting $z = 1$ in part (b), prove that

$$n = \prod_{k=1}^{n-1}(1-\zeta^k) = \prod_{k=1}^{n-1}|1-\zeta^k|.$$

[This identity has a beautiful geometric interpretation: Inscribe a regular n-gon in the unit circle. Fix a vertex, and consider the $(n-1)$ chords joining that vertex to each of the other vertices. The product of the lengths of these chords is n, the number of sides of the polygon.]

Chapter 10

Completeness and Topology

Calculus is built on the idea of a "limit". Informally, "$f(x)$ approaches L as x approaches a" means "we can make $f(x)$ as close to L as we like by taking x sufficiently close to a".

The problem is, two distinct real or complex numbers $f(x)$ and L are separated by a "gap" of fixed, finite size. Eventually, mathematicians came to formulate the definition of a limit in terms of *sets of numbers*. This chapter introduces "prototypical" sets, intervals and disks, and lays groundwork for the study of limits in Chapter 11.

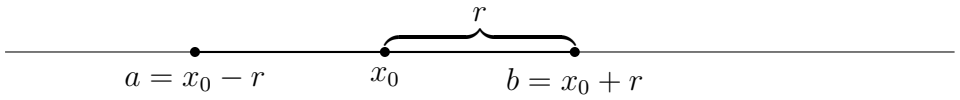
10.1 Sets of Real Numbers

Definition 10.1. Let a and b be real numbers with $a < b$. The sets

$$(a, b) = \{x \text{ in } \mathbf{R} : a < x < b\}$$
$$[a, b] = \{x \text{ in } \mathbf{R} : a \leq x \leq b\}$$

are called, respectively, the *open interval* and the *closed interval* with endpoints a and b.

For either of these intervals, the *center* is the midpoint $x_0 = \frac{1}{2}(b+a)$, the *length* is $b - a = 2r$, and half the length is the *radius*, $r = \frac{1}{2}(b - a)$.



Definition 10.2. Let x_0 be a real number, and r a positive real number. When we wish to emphasize the center and radius of an interval, we

use the notation

$$B_r(x_0) = (x_0 - r, x_0 + r),$$

and call this set the (*open*) *interval of radius* r *about* x_0.

Remark 10.3. By Exercise 9.9, if $x_0 = \frac{1}{2}(b + a)$ and $r = \frac{1}{2}|b - a|$, then

$$(a, b) = \{x \text{ in } \mathbf{R} : |x - x_0| < r\}, \qquad [a, b] = \{x \text{ in } \mathbf{R} : |x - x_0| \leq r\}.$$

Definition 10.4. Let a be a real number. We define *unbounded* open and closed intervals by

$$(-\infty, a) = \{x \text{ in } \mathbf{R} : x < a\} \qquad (a, \infty) = \{x \text{ in } \mathbf{R} : a < x\},$$
$$(-\infty, a] = \{x \text{ in } \mathbf{R} : x \leq a\} \qquad [a, \infty) = \{x \text{ in } \mathbf{R} : a \leq x\}.$$

Remark 10.5. The symbols $-\infty$ and ∞ do not denote real numbers, but are merely place-holders for an omitted inequality.

Example 10.6. If $a < b$, then $(a, b) = (-\infty, b) \cap (a, \infty)$, etc.

Operations on Sets

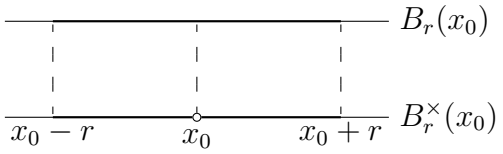
In Chapter 2, we encountered the complement of a set, the difference of two sets, and the union or intersection or an arbitrary family of sets. Starting from closed and open intervals, these operations can be used to construct interesting sets of real numbers, or to express properties of the real number system.

Example 10.7. For all real a, we have $(-\infty, a) = \mathbf{R} \setminus [a, \infty)$, etc.

Definition 10.8. Let x_0 be a real number, and let $r > 0$. We define the *deleted interval* $B_r^\times(x_0)$ of radius r about x_0 to be the set

$$\begin{aligned} B_r(x_0) \setminus \{x_0\} &= (x_0 - r, x_0) \cup (x_0, x_0 + r) \\ &= \{x \text{ in } \mathbf{R} : 0 < |x - x_0| < r\}. \end{aligned}$$

That is, $B_r^\times(x_0)$ is the interval $B_r(x_0)$ from which the center x_0 has been removed.



Example 10.9. Let a be a real number. The one-element set $\{a\}$ is called a *singleton*. Every set can be written as the union of its singleton subsets:

$$A = \bigcup_{a \in A} \{a\}.$$

Definition 10.10. Assume $A \subseteq \mathbf{R}$, and let c be a real number. The set

$$c + A = \{c + a : a \in A\}$$

is called the *translation* of A by c. The set

$$cA = \{ca : a \in A\}$$

is called the *scaling* of A by c. In particular, $-A = \{-a : a \in A\}$ is the reflection of A across the origin.

Remark 10.11. If $c = 0$, then $c + A = A$ and $cA = \{0\}$. In practice, we almost always assume $c \neq 0$.

Definition 10.12. If A and B are non-empty sets of real numbers, we define their *sum* and *product* to be

$$A + B = \{a + b : a \in A \text{ and } b \in B\} = \bigcup_{a \in A} a + B,$$
$$AB = \{ab : a \in A \text{ and } b \in B\} = \bigcup_{a \in A} aB.$$

Example 10.13. The sum of two integers is an integer, and the sum of two rational numbers is a rational number: $\mathbf{Z}+\mathbf{Z} = \mathbf{Z}$, and $\mathbf{Q}+\mathbf{Q} = \mathbf{Q}$.

Since every real number x may be written uniquely as the sum of an integer n and a real number r with $0 \leq r < 1$, as have $\mathbf{Z} + [0, 1) = \mathbf{R}$.

Example 10.14. Let $q \geq 1$ be an integer. The set

$$\tfrac{1}{q}\mathbf{Z} = \{p/q \text{ in } \mathbf{R} : p \in \mathbf{Z}\}$$

consists of all rational numbers that can be represented as a (possibly improper and/or non-reduced) fraction whose denominator is *precisely* q.

The elements of $\frac{1}{q}\mathbf{Z}$ are spaced regularly along the number line, with adjacent elements separated by a distance of $1/q$, Figure 10.1.

If q and q' are positive, then $\frac{1}{q}\mathbf{Z} \subseteq \frac{1}{q'}\mathbf{Z}$ if and only if q divides q'.

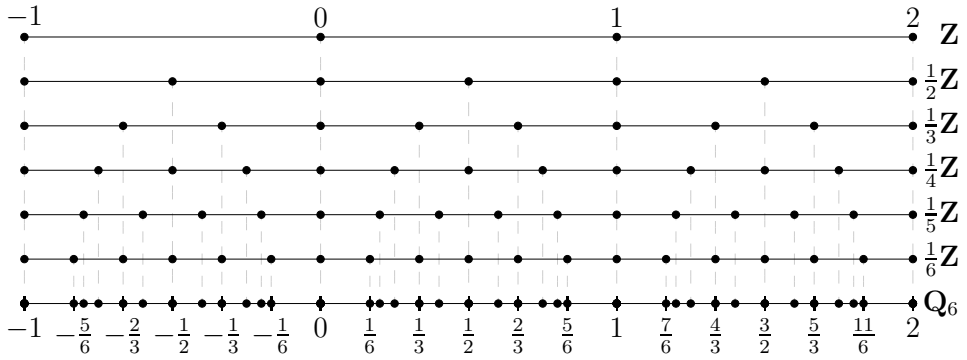


Figure 10.1: The sets $\frac{1}{q}\mathbf{Z}$ and $\mathbf{Q}_N$.

Example 10.15. Let $N \geq 1$ be an integer. The set

$$\mathbf{Q}_N = \bigcup_{q=1}^{N} \tfrac{1}{q}\mathbf{Z}$$

consists of all rational numbers that can be represented as a (possibly improper and/or non-reduced) fraction whose denominator is *no larger than* N.

The sets $\mathbf{Q}_N$ are *nested outward*: $\mathbf{Q}_N \subseteq \mathbf{Q}_{N+1} = \mathbf{Q}_N \cup \frac{1}{N+1}\mathbf{Z}$.

Example 10.16. With the notation of Examples 10.14 and 10.15, the set of rational numbers can be expressed in two ways as an infinite union:

$$\mathbf{Q} = \bigcup_{q=1}^{\infty} \tfrac{1}{q}\mathbf{Z} = \bigcup_{N=1}^{\infty} \mathbf{Q}_N.$$

Example 10.17 (The Cantor ternary set)**.** Let $[a, b]$ be an arbitrary closed interval. The weighted averages

$$a = \tfrac{1}{3}(3a) < \tfrac{1}{3}(2a + b) < \tfrac{1}{3}(a + 2b) < \tfrac{1}{3}(3b) = b$$

subdivide the interval into three pieces of equal length. The result of "removing the (open) middle third" of $[a, b]$ is the set

$$[a, b]^{\smile} = [a, b] \setminus \big(\tfrac{1}{3}(2a + b), \tfrac{1}{3}(a + 2b)\big) = \big[a, \tfrac{1}{3}(2a + b)\big] \cup \big[\tfrac{1}{3}(a + 2b), b\big].$$

The closed intervals $\big[a, \frac{1}{3}(2a + b)\big]$ and $\big[\frac{1}{3}(a + 2b), b\big]$ will be called the *components* of $[a, b]^{\smile}$ in this context.

Recursively construct a family of sets as follows. Let $K_0 = [0, 1]$ be the closed unit interval. Then let K_1 be the result of removing the middle third of K_0, let K_2 be the result of removing the middle third of each component of K_1, and generally, let K_{n+1} be the result of removing the middle third of each component of K_n.

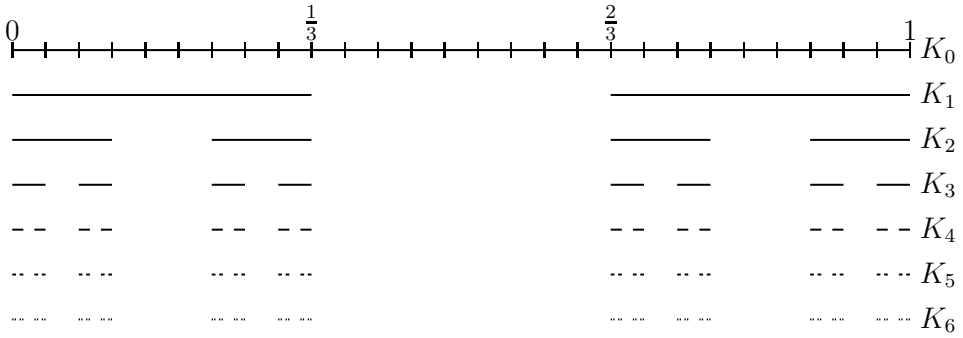


Figure 10.2: Approximations to the Cantor ternary set.

Mathematical induction shows the set K_n is a union of 2^n components, each of length 3^{-n}. The sets K_n are *nested inward*: $K_n \supset K_{n+1}$ for each $n \geq 0$.

The intersection

$$K = \bigcap_{n=0}^{\infty} K_n$$

is called the *Cantor ternary set*, or simply "the Cantor set". Each endpoint of K_n is an element of K_{n+1}, and consequently "survives" in the intersection. The Cantor set therefore contains the union of the endpoints of the sets K_n. Since K_n has 2^{n+1} endpoints, K has infinitely many elements. (K also contains "non-endpoint" elements.)

The Cantor set is *self-similar*. Precisely, $K = \frac{1}{3}K \cup \frac{1}{3}(K + 2)$; the Cantor set is a union of two disjoint subsets, each a scaled copy one-third the size of the entire set.

10.2 Upper and Lower Bounds

Definition 10.18. Let A be a set of real numbers. A real number U is an *upper bound* of A if $x \leq U$ for every x in A. If there exists an upper bound of A, we say A is *bounded above* (in $\mathbf{R}$).

A real number L is a *lower bound* of A if $L \leq x$ for every x in A. If there exists a lower bound of A, we say A is *bounded below* (in $\mathbf{R}$).

The set A is *bounded* if A is bounded above and bounded below.

Example 10.19. Let $a < b$ be real numbers. The interval $[a, b]$ is bounded. The left-hand endpoint a is a lower bound, and the right-hand endpoint b is an upper bound.

Every subset of a bounded set is obviously bounded. For example, the open interval (a, b) and the Cantor set $K \subseteq [0, 1]$ are bounded.

Example 10.20. The set $\mathbf{N}$ of natural numbers is bounded below; 0 is a lower bound. As we will see presently (Theorem 10.44), $\mathbf{N}$ is not bounded above in $\mathbf{R}$. In other words, for every real number x, there exists a natural number n such that $x < n$.

The set of integers is not bounded above or below in $\mathbf{R}$, nor is any set that contains $\mathbf{Z}$; thus $\mathbf{Q}$ and $\mathbf{R}$ are not bounded above or below.

Remark 10.21. If $A \subseteq \mathbf{R}$, then the following are equivalent:

(i) L is a lower bound of A and U is an upper bound of A.

(ii) $A \subseteq [L, U]$.

Remark 10.22. If U is an upper bound of A and if $U < U'$, then U' is an upper bound of A by transitivity of inequality. Similarly, if L is a lower bound of A and if $L' < L$, then L' is also a lower bound of A.

Proposition 10.23. *A subset A of $\mathbf{R}$ is bounded if and only if there exists a positive real number M such that $|x| \leq M$ for all x in A.*

Proof. If there exists an $M > 0$ such that $|x| \leq M$ for all x in A, then $-M \leq x \leq M$ for all x in A, i.e., $L = -M$ and $U = M$ are lower and upper bounds of A, respectively.

Conversely, if $L \leq x \leq U$ for all x in A, take $M = 1 + \max\big(|L|, |U|\big)$. For all x in A, we have $-M < |L| \leq L \leq x \leq U \leq |U| < M$. □

Remark 10.24. In practice, we often choose "symmetrical" upper and lower bounds for a bounded set.

Maxima and Minima

Definition 10.25. Let A be a non-empty set of real numbers. A real number M in A is called the *maximum* of A or the *largest element* of A if $x \leq M$ for every x in A.

A real number m in A is called the *minimum* of A or the *smallest element* of A if $m \leq x$ for every x in A.

Example 10.26. A finite set of real numbers $A = \{a_1, a_2, \ldots, a_n\}$ has a maximum and a minimum, see Exercise 9.14.

Example 10.27. A closed interval $[a, b]$ has a smallest element a and a largest element b.

Example 10.28. Let $a < b$ be real numbers. The open interval (a, b) is bounded below by a and bounded above by b, but contains no smallest or largest element: If $x \in (a, b)$, i.e., if $a < x < b$, then by Exercise 9.11,

$$a < \tfrac{1}{2}(a + x) < x < \tfrac{1}{2}(x + b) < b.$$

In words, $\frac{1}{2}(a+x)$ in (a, b) is smaller than x (so x is not the minimum), and $\frac{1}{2}(x + b)$ in (a, b) is larger.

Remark 10.29. We often write $M = \max A$ and $m = \min A$. If a set A *has* a largest element $\max A$, then A is bounded above by $\max A$. Contrapositively, if A is not bounded above, then A has no largest element. Analogous remarks hold for lower bounds and smallest elements.

Note carefully that a bounded, non-empty set A *need not have* a maximum element or minimum element, see Example 10.28.

Suprema and Infima

Definition 10.30. Let A be a set of real numbers that is bounded above. A real number β is called a *least upper bound* or *supremum* of A if

(i) $x \leq \beta$ for all x in A, i.e., β is an upper bound of A.

(ii) For every upper bound U of A, we have $\beta \leq U$.

Lemma 10.31. *If $A \subseteq \mathbf{R}$, and β and β' are suprema of A, then $\beta = \beta'$.*

Proof. By hypothesis, β' is an upper bound of A, so $\beta \leq \beta'$ by condition (ii). Reversing roles, $\beta' \leq \beta$. □

Remark 10.32. We are therefore justified in writing $\sup A$ to denote the least upper bound of A; the symbol $\sup A$ may signify no number at all, but it never signifies more than one.

If A has a largest element, then $\sup A = \max A$. Otherwise, $\sup A$ is the leftmost number lying to the right of every element of A.

Remark 10.33. The completeness axiom says that if A is a *non-empty* set of real numbers that is bounded above, then $\sup A \in \mathbf{R}$.

Proposition 10.34. *Let A be a set of real numbers having a least upper bound β. Then β is the unique number satisfying the conditions*

(i) *If $x \in A$, then $x \leq \beta$.*

(ii)$'$ *For every $\varepsilon > 0$, there exists an x in A such that $\beta - \varepsilon < x$.*

Proof. Conceptually, the conditions (ii) and (ii)$'$ are contrapositives.

In other words, β is the least upper bound of A if and only if every upper bound U of A satisfies $\beta \leq U$. The contrapositive says that for every $U < \beta$, U *fails* to be an upper bound of A. That is, for every $\varepsilon > 0$, there exists an x in A such that $U = \beta - \varepsilon < x$. □

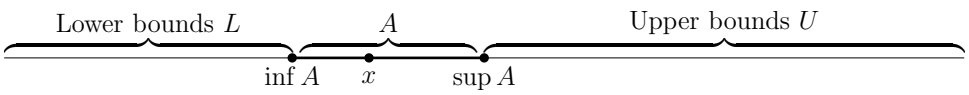
Everything said above for upper bounds has a corresponding concept or statement for lower bounds.

Definition 10.35. A number α is called a *greatest lower bound* or *infimum* of A if

(i) α is a lower bound of A.

(ii) If L is a lower bound of A, then $L \leq \alpha$.

Remark 10.36. Geometrically, an infimum of A lies to the left of A, but is the rightmost such point. Infima are unique (if they exist), so we are justified in writing $\inf A$. For practice, write out an alternative characterization of infima corresponding to Proposition 10.34.



Extended Real Numbers

Definition 10.37. Let $+\infty$ and $-\infty$ denote objects that are not real numbers, and extend the ordering on the real numbers by declaring that $-\infty < x < +\infty$ for every real number x. The set $\overline{\mathbf{R}} = \mathbf{R} \cup \{-\infty, +\infty\}$ is called the *extended real number system*.

Remark 10.38. As yet we make no attempt to define addition or multiplication involving $+\infty$ or $-\infty$.

Definition 10.39. If $A \subseteq \mathbf{R}$ is arbitrary, we write $\sup A = +\infty$ if and only if A is not bounded above, and write $\inf A = -\infty$ if and only if A is not bounded below.

Remark 10.40. *Every* set A of real numbers has an infimum and a supremum in the extended real number system.

If $A = \varnothing$, then $\sup A = -\infty$ (because every extended real number is an upper bound) and $\inf A = +\infty$. The empty set is the only set whose supremum is smaller than its infimum. If $a \in A$, then

$$-\infty \leq \inf A \leq a \leq \sup A \leq +\infty.$$

Proposition 10.41. *Let A and B be non-empty sets of real numbers. If $A \subseteq B$, then* $\inf B \leq \inf A \leq \sup A \leq \sup B$.

Proof. Every lower bound of B is *a fortiori* a lower bound of A. In particular, $\inf B \leq \inf A$. Similarly, $\sup A \leq \sup B$. □

10.3 More About Suprema

Scaling or translating a bounded set of real numbers affects the supremum in a natural way.

Proposition 10.42. *Let A be a bounded set of real numbers, and let k be a real number.*

(i) $\sup(k + A) = k + \sup A$.

(ii) *If $k \geq 0$, then* $\sup(kA) = k \sup A$.

(iii) *If $k > 0$, then* $\sup(-kA) = -k \inf A$.

Example 10.43. If $A = [a, b]$ is an interval, then $k + A = [k + a, k + b]$. If $k \geq 0$, then $kA = [ka, kb]$ and $-kA = [-kb, -ka]$. In each case, the suprema can be read off by inspection. This is a useful way of remembering the conclusion of Proposition 10.42.

Proof. The strategy in each part is to show that the right-hand side satisfies the two conditions for a supremum in Definition 10.30.

(i). Since $k + A = \{x' \text{ in } \mathbf{R} : x' = k + x \text{ for some } x \text{ in } A\}$, a real number U is an upper bound of A if and only if $x \leq U$ for all x in A,

$$\begin{aligned} &\text{if and only if } x' = k + x \leq k + U \text{ for all } x' \text{ in } k + A, \\ &\text{if and only if } k + U \text{ is an upper bound of } k + A. \end{aligned}$$

But $\sup A$ is an upper bound of A, so $k + \sup A$ is an upper bound of $k + A$. By Definition 10.30 (ii), $\sup(k + A) \leq k + \sup A$.

Conversely, $\sup(k+A)$ is an upper bound of $k+A$, so $\sup(k+A)-k$ is an upper bound of A. By Definition 10.30 (ii), $\sup A \leq \sup(k+A)-k$, or $k + \sup A \leq \sup(k + A)$.

(ii). If $k = 0$, then $kA = \{0\}$ and the claim is obvious. Assume, therefore, that $0 < k$. The proof in this case is nothing but a mechanical modification of the preceding argument. The main idea is that multiplication or division by k preserves the sense of an inequality since $0 < k$, so a real number U is an upper bound of A if and only if kU is an upper bound of kA. Just as in the preceding argument, $k \sup A$ is an upper bound of kA, so $\sup(kA) \leq k \sup A$, and $\sup(kA)/k$ is an upper bound of A, so $k \sup A \leq \sup(kA)$.

(iii). Since $-k < 0$, multiplication or division by $-k$ *reverses* inequalities. That is, L is a *lower* bound of A if and only if $U' = -kL$ is an *upper* bound of $-kA$. Particularly, $-k \inf A$ is an upper bound of $-kA$, so $\sup(-kA) \leq -k \inf A$, and $\sup(-kA)/(-k)$ is a lower bound of A, so $\sup(-kA)/(-k) \leq \inf A$, or $-k \inf A \leq \sup(-kA)$. □

The Archimedean Property

Theorem 10.44 (The Archimedean property)**.** *For every number x, there exists a non-negative integer n such that $x < n$.*

Remark 10.45. In words, the set $\mathbf{N}$ of natural numbers is not bounded above in $\mathbf{R}$. (Note that $\mathbf{N}$ *is* bounded above in $\overline{\mathbf{R}}$; $+\infty$ is an upper bound.)

Proof. Suppose to the contrary that there exists a real number x such that $n \leq x$ for every natural number n. This means, by definition, that $\mathbf{N}$ is bounded above in $\mathbf{R}$. By the completeness axiom, there exists a least upper bound ω in $\mathbf{R}$.

By Proposition 10.34 with $\varepsilon = 1$, the real number $\omega - 1 < \omega$ is not an upper bound of $\mathbf{N}$, so there exists a natural number n_0 such that $\omega - 1 < n_0$. But we would then have $\omega < n_0 + 1$, and $n_0 + 1 \in \mathbf{N}$ by the definition of the natural numbers, contradicting the fact that $\omega = \sup \mathbf{N}$. Contrapositively, if the axioms for the real numbers are logically consistent, then $\mathbf{N}$ is not bounded above in $\mathbf{R}$. □

Corollary 10.46 (The generalized Archimedean Property)**.** *If M is an arbitrary real number and $\varepsilon > 0$ is an arbitrary positive real number, there exists a positive integer n such that $M < n\varepsilon$.*

Proof. Let $x = M/\varepsilon$. By the Archimedean property, there exists a positive integer n such that $M/\varepsilon < n$. Multiplying through by the positive number ε gives $M < n\varepsilon$. □

Remark 10.47. Metaphorically, a journey of 1000 miles (M) can be accomplished one step (ε) at a time, no matter how small the steps are.

Corollary 10.48. *If ε is a real number and $0 < \varepsilon$, there exists a positive integer n such that $1/n < \varepsilon$. Contrapositively, if $\varepsilon \leq 1/n$ for every positive integer n, then $\varepsilon \leq 0$.*

Proof. Set $x = 1/\varepsilon$, and note that $0 < x$. By the Archimedean property, there exists an integer n such that $0 < x < n$. Proposition 9.12 (v) implies that $1/n < \varepsilon$. □

Remark 10.49. This corollary asserts, informally, that there are no infinitesimal real numbers, i.e., no positive numbers that are smaller than every reciprocal of a positive integer.

Example 10.50. The Archimedean property has noteworthy consequences for families of sets. For each positive integer n, define the open intervals $A_n = (0, n)$ and $B_n = (0, \frac{1}{n})$. The Archimedean property implies

$$\bigcup_{n=1}^{\infty} A_n = \{x \text{ in } \mathbf{R} : x > 0\}, \qquad \bigcap_{n=1}^{\infty} B_n = \varnothing.$$

In particular, a union of bounded intervals may be unbounded, and an intersection of nested, non-empty sets may be empty.

Density of the Rational Numbers

Definition 10.51. A set A of real numbers is *dense* (in $\mathbf{R}$) if for every real number x and every $\varepsilon > 0$, there exists a number x_0 in A such that $|x - x_0| < \varepsilon$.

Remark 10.52. If A is dense in $\mathbf{R}$, then every real number can be approximated arbitrarily closely by elements of A.

Theorem 10.53. *Let $x < y$ be distinct real numbers. There exists a non-zero rational number $r = m/n$ such that $x < r < y$.*

Remark 10.54. In words, the set $\mathbf{Q}$ of rational numbers is dense in $\mathbf{R}$: If $x \in \mathbf{R}$ and if $\varepsilon > 0$, there exists a rational number $r = p/q$ such that $x < r < y = x + \varepsilon$.

Proof. Assume first that $0 \leq x < y$. The real number $\varepsilon = y - x$ is positive because $x < y$. By Corollary 10.48, there exists a positive integer n such that $1/n < \varepsilon$. It suffices to prove that some integer multiple $r = m/n$ lies between x and y. Geometrically this is plausible: By taking steps of size $1/n$ across the interval between x and y, we must step in the interior at least once.

Formally, consider the set $S = \{p \text{ in } \mathbf{Z}^+ : y \leq p/n\}$. By the generalized Archimedean property, there is a positive integer p such that $y < p/n$; that is, the set S is non-empty. Now, a non-empty set of positive integers has a smallest element, say $m + 1$. By definition of the set S, we have $m/n < y \leq (m + 1)/n$. To complete the proof, it suffices to prove $x < m/n$. But

$$y - \frac{m}{n} \leq \frac{m+1}{n} - \frac{m}{n} = \frac{1}{n} < y - x.$$

Rearranging the inequality gives $x < m/n$. We have proven that if $0 \leq x < y$, then there exists a non-zero rational number $r = m/n$ such that $x < r < y$.

If instead $x < y \leq 0$, then $0 \leq -y < -x$. By the preceding argument, there exists a rational number $r \neq 0$ such that $-y < r < -x$, so the non-zero rational number $-r$ satisfies $x < -r < y$.

Finally, if neither alternative holds, then $x < 0 < y$, and by the first part of the proof there is a rational number $x < 0 < r < y$. □

Remark 10.55. Theorem 10.53 says that between any two real numbers can be found a rational number. Moreover, between any two real numbers can be found an irrational number. For example, if $x < y$, then

by Theorem 10.53 there exists a non-zero rational number r such that $x/\sqrt{2} < r < y/\sqrt{2}$, i.e., such that $x < r\sqrt{2} < y$. But $r\sqrt{2}$ is irrational.

It is tempting to conclude that rational and irrational numbers "alternate" along the number line, as if by painting rational numbers red and irrational numbers blue, the number line would consist of alternating red and blue points. Unfortunately, this picture is utterly incorrect. Distinct real numbers are never adjacent to each other, but instead are endpoints of an interval containing infinitely many rational and irrational numbers.

10.4 Sets of Complex Numbers

Let $\alpha = a + bi$ be a complex number. Recall that the *absolute value* of α,

$$|\alpha| = \sqrt{\alpha\overline{\alpha}} = \sqrt{a^2 + b^2},$$

represents the distance from 0 to α according to the Pythagorean theorem. If α is real, this reduces to the definition for real numbers, since $\sqrt{x^2} = |x|$ for all real x.

The complex numbers are not ordered, so the definition of an interval does not have an obvious generalization. In practice, two types of sets, disks and rectangles, generally play the role of intervals.

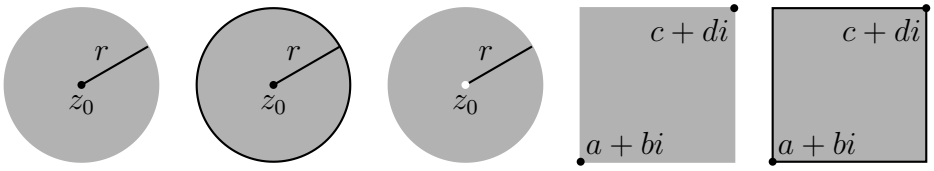
Definition 10.56. Let z_0 be a complex number. For each positive real number r, the sets

$$B_r(z_0) = \{z \text{ in } \mathbf{C} : |z - z_0| < r\},$$
$$\overline{B_r(z_0)} = \{z \text{ in } \mathbf{C} : |z - z_0| \leq r\},$$

are called the *open disk* and the the *closed disk* of *center* z_0 and *radius* r.

The *deleted open disk* is the set

$$B_r^{\times}(z_0) = \{z \text{ in } \mathbf{C} : 0 < |z - z_0| < r\}.$$



Definition 10.57. Let $a < b$ and $c < d$ be real numbers. The sets

$$(a,b) \times i(c,d) = \{z \text{ in } \mathbf{C} : a < \operatorname{Re} z < b, c < \operatorname{Im} z < d\},$$
$$[a,b] \times i[c,d] = \{z \text{ in } \mathbf{C} : a \le \operatorname{Re} z \le b, c \le \operatorname{Im} z \le d\},$$

are called an *open rectangle* and a *closed rectangle*.

The operations of scaling and translation of a set, and of the sum or product of two sets, are defined exactly as for sets of real numbers.

Example 10.58. If $B = B_1(0)$ is the open unit disk, the set

$$\bigcup_{n=-\infty}^{\infty} n + B$$

is the result of placing a copy of U at each integer, while

$$\bigcup_{n=0}^{\infty} \left[(2 - 3 \cdot 2^{-n}) + 2^{-n} B\right]$$

is a sequence of shrinking disks, each tangent to its neighbors.

Example 10.59. The *upper half-plane* $H = \{z \text{ in } \mathbf{C} : \operatorname{Im}(z) > 0\}$ contains an open disk about each of its points: A complex number $z_0 = x_0 + iy_0$ is in H if and only if $y_0 > 0$, in which case $B_{y_0}(z_0) \subseteq H$.

The Triangle Inequalities

If the lengths of two sides of a planar triangle are known, the third side cannot be longer than the sum of the lengths, and cannot be shorter than the difference of the lengths. These inequalities, the *triangle inequality* and *reverse triangle inequality*, are ubiquitous in analysis.

Theorem 10.60. *If α and β are complex numbers, then*

(i) $|\alpha + \beta| \le |\alpha| + |\beta|$.

(ii) $|\alpha - \beta| \ge \big||\alpha| - |\beta|\big|$.

Remark 10.61. Since $|-\beta| = |\beta|$ for all β, the signs on the left-hand sides are arbitrary; that is,

$$\big||\alpha| - |\beta|\big| \le |\alpha \pm \beta| \le |\alpha| + |\beta| \quad \text{for all complex } \alpha \text{ and } \beta.$$

Proof. (i). The complex numbers $\alpha\overline{\beta}$ and $\overline{\alpha}\beta = \beta\overline{\alpha}$ are conjugates, so their sum is twice the real part of either. Since $|\operatorname{Re}(z)| \leq |z|$ for all z,

$$\alpha\overline{\beta} + \beta\overline{\alpha} = 2\operatorname{Re}(\alpha\overline{\beta}) \leq 2|\alpha\overline{\beta}| = 2|\alpha|\,|\beta|.$$

By definition,

$$\begin{aligned}|\alpha+\beta|^2 &= (\alpha+\beta)(\overline{\alpha+\beta}) = (\alpha+\beta)(\overline{\alpha}+\overline{\beta}) = \alpha\overline{\alpha} + (\alpha\overline{\beta}+\beta\overline{\alpha}) + \beta\overline{\beta} \\ &\leq |\alpha|^2 + 2|\alpha|\,|\beta| + |\beta|^2 = \big(|\alpha|+|\beta|\big)^2.\end{aligned}$$

Since the terms of each side of the desired inequality are non-negative, $|\alpha+\beta| \leq |\alpha| + |\beta|$.

(ii). Apply (i) to the identity $\alpha = \beta + (\alpha - \beta)$, obtaining

$$|\alpha| \leq |\beta| + |\alpha-\beta|, \quad \text{or} \quad |\alpha| - |\beta| \leq |\alpha-\beta|.$$

Reversing the roles of α and β shows $|\beta| - |\alpha| \leq |\beta-\alpha|$, and therefore $|\alpha| - |\beta| \geq -|\beta-\alpha| = -|\alpha-\beta|$. Combining,

$$-|\alpha-\beta| \leq |\alpha| - |\beta| \leq |\alpha-\beta|.$$

By Exercise 9.12, $\big||\alpha| - |\beta|\big| \leq |\alpha-\beta|$. □

Definition 10.62. Let α and β be complex numbers. The *distance* between α and β is $|\alpha-\beta| = |\beta-\alpha|$.

Corollary 10.63. *If α, β and γ are complex numbers, then*

$$\big||\alpha-\beta| - |\gamma-\beta|\big| \leq |\alpha-\gamma| \leq |\alpha-\beta| + |\beta-\gamma|.$$

Proof. Apply Theorem 10.60 to the identity

$$(\alpha-\beta) - (\gamma-\beta) = (\alpha-\gamma) = (\alpha-\beta) + (\beta-\gamma).$$ □

Exercises

Exercise 10.1. Prove that:

(a) $3 + 5(-4, 2) = (-17, 23)$.

(b) $x_0 + r(-1, 1) = (x_0 - r, x_0 + r)$ if $r > 0$ is real.

Exercise 10.2. Suppose $2 \leq x \leq 4$ and $3 \leq y \leq 7$. Based solely on this information, find the best possible lower and upper bounds on:

(a) $x + y$. (b) $x - y$. (c) xy. (d) x/y.

Exercise 10.3. Suppose $-2 \leq x \leq 4$ and $3 \leq y \leq 7$. Based solely on this information, find the best possible lower and upper bounds on:

(a) $x + y$. (b) $x - y$. (c) xy. (d) x/y.

Exercise 10.4. Suppose $-2 \leq x \leq 4$ and $3 \leq |y| \leq 7$. Based solely on this information, find the best possible lower and upper bounds on:

(a) $x + y$. (b) $x - y$. (c) xy. (d) x/y.

Exercise 10.5. Alphaton lies between 7 and 8 miles from the town of Origin, and Betaville lies between 2 and 5 miles from Origin.

Based only on this information, how close could Alphaton and Betaville be? How far apart could Alphaton and Betaville be?

Exercise 10.6. Suppose $0 < a < b$ and $0 < c < d$, and let R be the *closed rectangle* $[a, b] \times [c, d]$.

(a) If $(x_0, y_0) \in R$, does there always exist an (x, y) in R such that $x_0^2 + y_0^2 < x^2 + y^2$?

(b) Find the supremum of $x^2 + y^2$ for (x, y) in R.

Exercise 10.7. Suppose $0 < a < b$ and $0 < c < d$, and let R be the *open rectangle* $(a, b) \times (c, d)$.

(a) Show that if $(x_0, y_0) \in R$, there exists an (x, y) in R such that $x_0^2 + y_0^2 < x^2 + y^2$.

(b) Find the supremum of $x^2 + y^2$ for (x, y) in R.

Exercise 10.8. Let $a < b$ be real numbers. Find real numbers $r > 0$ and x_0 such that:

(a) $[a, b] = x_0 + r[0, 1]$. (b) $[a, b] = x_0 + r[-1, 1]$.

Exercise 10.9. In each part, k denotes an integer, and (for example) $\{\frac{4k-1}{k} : k \geq 1\}$ is shorthand for the set

$$\{x \text{ in } \mathbf{R} : x = \tfrac{4k-1}{k} \text{ for some integer } k \geq 1\}.$$

Determine whether each set is bounded above and/or below, and if so, find the supremum and/or infimum with proof.

(a) $\{\frac{(-1)^k}{k}\}$. (c) $\{\frac{4k-1}{k}\}$. (e) $\{(-1)^k + k\}$.

(b) $\{(-1)^k k\}$. (d) $\{(-1)^k - \frac{1}{k}\}$. (f) $\{\frac{4k-1}{2k-5}\}$.

Exercise 10.10. In each part, let $A = \{x \text{ in } \mathbf{R} : x^2 < 2\}$.

(a) Prove that A is bounded. (Suggestion: Show that 2 is an upper bound. Do not use results that have not been established in this book, such as existence of square roots.)

(b) Prove that if x is an upper bound of A, then $x' = \frac{1}{2}(x + \frac{2}{x})$ is an upper bound, and $x' < x$.

Exercise 10.11. Let A be a bounded, non-empty set of real numbers, and let p be a polynomial function. Prove that the image $p(A)$ is bounded. If you like, use the following outline:

(a) There exists a real number $M > 0$ such that $A \subseteq [-M, M]$.

(b) If a is real and $k \geq 0$ is an integer, the function $f(x) = ax^k$ is bounded on $[-M, M]$.

(c) A sum of bounded functions on $[-M, M]$ is bounded on $[-M, M]$.

Exercise 10.12. For $n > 1$, let $A_n = [\frac{1}{n}, \frac{n}{n+1}]$. Express the union $\bigcup_n A_n$ as an interval, with proof.

Exercise 10.13. In each part, a family of sets $\{A_n : n \geq 1\}$ is given. Determine (with proof) whether the union of the A_n is all of $(0, \infty)$.

(a) $A_n = (0, n)$. (b) $A_n = [-\frac{1}{n}, n]$. (c) $A_n = (n-1, n]$.

Exercise 10.14. If $A \subseteq \mathbf{R}$ is non-empty and if x is real, define the *distance* from x to A to be the infimum of $|x - y|$ taken over y in A.

(a) If $x \notin A$, does it follow that $d(x, A) > 0$?

(b) Let $a < b$ be real, $A = [a, b]$. If $x \notin A$, prove $d(x, A) > 0$.

(c) Let K be the Cantor set. Prove that if $x \notin K$, then $d(x, K) > 0$.

Exercise 10.15. For each positive real number r, let $O_r = B_r^{\times}(0)$ be the deleted neighborhood of radius r about 0. Determine $\bigcap_r O_r$ with justification.

Exercise 10.16. Let A be a bounded set of real numbers, and let k be a real number. Formulate and prove a version of Proposition 10.42 for infima.

Exercise 10.17. For each positive integer n, let $I_n = [a_n, b_n]$ be a *closed*, bounded interval of real numbers. Prove that if the intervals are nested inward, i.e., if $I_{n+1} \subseteq I_n$, then the infinite intersection $\bigcap_n I_n$ is non-empty. (Compare Example 10.50, in which the sets B_n are not closed.)
Hint: Consider the sets $A = \{a_n\}$ and $B = \{b_n\}$. Show that A is bounded above, B is bounded below, and $\sup A \leq \inf B$; conclude that $[\sup A, \inf B]$ is contained in $\bigcap_n I_n$.

Exercise 10.18. Let O be a non-empty open set of real numbers. This exercise outlines a proof that O is a union of open intervals.

(a) Suppose $x \in O$, and put

$$a_x = \inf\{t : (t, x) \subseteq O\}, \qquad b_x = \sup\{t : (x, t) \subseteq O\}.$$

Prove that $a_x < x < b_x$ (note strict inequalities).

(b) In the notation of part (a), prove that $(a_x, b_x) \subseteq O$, and that no strictly larger open interval is contained in O. (An interval of the form (a_x, b_x) is called a *component* of O.)

(c) Prove that O is the union over x in O of the components (a_x, b_x).

Exercise 10.19. Let $O = \{x \text{ in } \mathbf{R} : x \neq 0, x \neq \pm 1/n \text{ for all } n \geq 1\}$. Sketch the set O, show O is open, and find the components of O.

Chapter 11

Sequences and Convergence

The placid depiction of the real number system as a line belies complexity utterly beyond human comprehension. In a precise technical sense, there are fewer algorithms (finite-state Turing machines and initial conditions) than there are real numbers. Consequently, *most real numbers cannot be output by any computer program.*

Instead, we approximate numbers by rational numbers (ratios of integers) or by "mildly irrational" numbers. One common scheme is to form running totals from a list of numbers, and to take "the limit as the number of summands grows without bound", as in

$$\sum_{k=0}^{\infty} \frac{1}{k!} = 1 + 1 + \tfrac{1}{2!} + \tfrac{1}{3!} + \tfrac{1}{4!} + \tfrac{1}{5!} + \dots,$$

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} = 1 - \tfrac{1}{3} + \tfrac{1}{5} - \tfrac{1}{7} + \tfrac{1}{9} - \tfrac{1}{11} + \dots.$$

11.1 Sequences

Definition 11.1. Let k_0 be an integer, $k_0 + \mathbf{N} = \{k \text{ in } \mathbf{Z} : k \geq k_0\}$ the set of integers greater than or equal to k_0.

A *real sequence* $(a_k)_{k=k_0}^{\infty}$ is a mapping $\mathbf{a} : k_0 + \mathbf{N} \to \mathbf{R}$.

Remark 11.2. Unpacking the definition of a mapping, a sequence is a collection of ordered pairs (k, a_k) in which $k \geq k_0$ is an integer and a_k is a number. We call a_k the *kth term* of the sequence.

The first coordinate k imposes an ordering on the terms a_k. The ordering of the terms is crucial; a sequence must be carefully distinguished from its image, i.e., from its *set of terms* $\{a_k : k \geq k_0\} \subseteq \mathbf{R}$.

We often assume $k_0 = 0$ or 1, and write (a_k) for brevity when the initial index is unimportant.

Example 11.3. The formulas $a_k = 1/(k+1)$ and $b_k = (-1)^k$ define real sequences. The respective sets of terms are $A = \{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots\}$, and the finite set $B = \{1, -1\}$.

Example 11.4. For each real number x, the formula $a_k = x^k$ defines a real sequence.

Example 11.5. The recursive specification

$$a_0 = 2, \quad a_{k+1} = \frac{1}{2}\left(a_k + \frac{2}{a_k}\right)$$

defines a sequence. The next three terms are $a_1 = \frac{3}{2}$, $a_2 = \frac{17}{12}$, $a_3 = \frac{577}{408}$.

Example 11.6. The recursive rule $a_{k+1} = a_k + \frac{1}{k+1}$, with $a_0 = 0$, defines a sequence whose nth term ($n \geq 1$) may be written

$$a_n = 1 + \tfrac{1}{2} + \tfrac{1}{3} + \cdots + \tfrac{1}{n} = \sum_{k=1}^{n} \frac{1}{k}.$$

11.2 Convergence

In this section, we define formally what it means for a sequence to have a limit. Though not difficult to state, the definition is easily the most complicated logical criterion in this book. It takes practice, concerted effort, and time to develop technical intuition for whether or not a specific sequence has or does not have a limit.

Definition 11.7. Let (a_k) be a real sequence, and let $a_\infty \in \mathbf{R}$. We say (a_k) *converges* to a_∞, and write $(a_k) \to a_\infty$, if the following condition holds:

> For every $\varepsilon > 0$, there exists a natural number N such that if $k \geq N$, then $|a_k - a_\infty| < \varepsilon$.

Remark 11.8. In words, the terms a_k can be made "as close as we like" to a_∞ (i.e., within an error of $\varepsilon > 0$, no matter how small ε is chosen) by "going sufficiently far out into the sequence" (i.e., considering only terms a_k with $k \geq N$). Loosely, "the a_k eventually get arbitrarily close to a_∞". Note carefully that ε is chosen *before* N. Loosely, we choose what is meant by "arbitrarily close" *before* we decide what qualifies as "sufficiently far out".

Proposition 11.9. *If (a_k) is a real sequence that converges to a_∞ and to a'_∞, then $a_\infty = a'_\infty$.*

Proof. Let $\varepsilon > 0$ be arbitrary. Since $(a_k) \to a_\infty$, there exists an index N_1 such that if $k \geq N_1$, then $|a_k - a_\infty| < \varepsilon/2$. Similarly, since $(a_k) \to a'_\infty$, there exists an index N'_1 such that if $k \geq N'_1$, then $|a_k - a'_\infty| < \varepsilon/2$.

Let $N = \max(N_1, N'_1)$. Since $N \geq N_1$ and $N \geq N'_1$, the triangle inequality implies

$$\begin{aligned} |a'_\infty - a_\infty| &= |(a'_\infty - a_N) + (a_N - a_\infty)| \\ &\leq |a'_\infty - a_N| + |a_N - a_\infty| < (\varepsilon/2) + (\varepsilon/2) = \varepsilon. \end{aligned}$$

But $\varepsilon > 0$ was arbitrary, which means $a'_\infty - a_\infty = 0$. □

Remark 11.10. Proposition 11.9 says that a real sequence converges to at most one number. If $(a_k) \to a_\infty$, we call a_∞ the *limit* of (a_k), and write $a_\infty = \lim_{k\to\infty} a_k$.

Remark 11.11. Informally, we say "a_k *approaches* a_∞ as $k \to \infty$". Indeed, the notation a_∞ is meant to suggest "setting $k = \infty$ in the limit". Logically, however, this is not what convergence means. The terms a_k of a sequence (a_k) are merely individual real numbers, which do not "approach" anything. It is the *sequence* (i.e., the *ordered list of terms*) that converges (or fails to converge) to a limit.

The ε-N Game, and Examples

Convergence of a sequence may be understood as an adversarial game. A sequence (a_k) and a putative limit a_∞ are specified in advance. The first player chooses a positive "tolerance" ε, which defines a target, the interval $B_\varepsilon(a_\infty) = (a_\infty - \varepsilon, a_\infty + \varepsilon)$.

The second player now tries to "hit the target", i.e., to ensure that $|a_k - a_\infty| < \varepsilon$, solely by taking k to be sufficiently large. A "successful response" to the first player's "challenge" is a positive integer N such that if $k \geq N$, then $|a_k - a_\infty| < \varepsilon$.

If Player N is able to respond successfully to a particular ε, they "win the round". Otherwise Player ε wins the round.

To say $(a_k) \to a_\infty$ means that Player N *has a winning strategy against a perfect opponent*: No matter how "skillful" Player ε is (i.e., no matter how small $\varepsilon > 0$ is chosen), Player N can always issue a successful response.

Success in analysis is largely a matter of learning to play the ε-N game and its variants. When you read proofs, you may notice that seemingly magical choices of N are made. To make these choices, the author imagined an arbitrary $\varepsilon > 0$ was given, and formulated a strategy for choosing a "winning" index N. The proof itself is merely a written demonstration that Player N wins.

Remark 11.12. Yet another interpretation of convergence of a sequence has the procedural appeal of a solitaire game. Imagine an infinite deck of playing cards in which the kth card has the term a_k of the sequence written on its face. The cards are placed in a row in order, face up. Now an $\varepsilon > 0$ is specified, and for each k, if $|a_k - a_\infty| \geq \varepsilon$, the kth card is turned face down. At issue is whether all the cards past some point remain face up, i.e., whether *only finitely many cards must be turned*. If this is the case *no matter how* $\varepsilon > 0$ *is specified*, we say $(a_k) \to a_\infty$.

Example 11.13. Let c be real. The sequence $a_k = c$ is called a *constant sequence*. A constant sequence obviously converges to $a_\infty = c$: For every $\varepsilon > 0$ and for every k, we have $|a_k - a_\infty| = |c - c| = 0 < \varepsilon$. That is, Player N *cannot lose against a perfect opponent* when playing with a constant sequence!

Example 11.14. The sequence $a_k = 1/(k+1)$ converges to 0. Before giving a proof, we'll play a few rounds of the ε-N game.

If $\varepsilon = 100$, Player N *cannot lose*, and in particular may take $N = 0$: Indeed, $|a_k - 0| = 1/(k+1) \leq 1 < \varepsilon$ regardless of k. Note carefully that this fact by itself does not prove $(a_k) \to 0$; Player N must be able to win against an *arbitrary* $\varepsilon > 0$.

If $\varepsilon = 0.01 = 1/100$, player N may take $N = 100$: If $k \geq 100$, then $|a_k - 0| = 1/(k+1) \leq 1/101 < 1/100 = \varepsilon$.

If $\varepsilon = 1/\sqrt{200}$ (assuming such a real number exists), Player N's goal is to find an N such that $1/(N+1) < \varepsilon = 1/\sqrt{200}$, or after rearranging, $200 < (N+1)^2$. The "best" (i.e., smallest) choice, $N = 14$, is easy to find in this example, but there is no harm in taking, for example, $N = 200$, which is surely sufficient.

To show $(a_k) \to 0$, it suffices to construct a winning strategy for Player N. Let $\varepsilon > 0$ be arbitrary. By Corollary 10.48 of the Archimedean property, there exists a positive integer N such that $1/N < \varepsilon$. This N is our response. If $k \geq N$, then $|a_k - 0| = 1/(k+1) < 1/N < \varepsilon$.

Remark 11.15. If an index N "wins" against some challenge $\varepsilon > 0$, then *every larger integer* $N' \geq N$ *also wins*, because $k \geq N'$ implies $k \geq N$. It is not necessary (or desirable) to pick the smallest winning N.

Correspondingly, making ε smaller makes the target smaller, which makes the condition $|a_k - a_\infty| < \varepsilon$ "harder to meet", and generally forces N to be larger.

Remark 11.16. The standard idiom for picking a *single* N satisfying *finitely many conditions* is to pick multiple Ns, each satisfying one condition, then let our response be the largest of our choices. This was done already in the proof of Proposition 11.9.

Example 11.17. The sequence $a_k = (-1)^k$ has terms that are alternately 1 and -1; precisely, the "even" terms $a_{2\ell}$ are all 1 and the "odd" terms $a_{2\ell+1}$ are all -1. We will show that (a_k) has no limit. That is, for every real number a_∞, the statement "$(a_k) \to a_\infty$" is false. To prove this, we fix a putative limit a_∞ arbitrarily, then take the side of Player ε and look for a winning strategy:

> No matter what N is, $k = 2N \geq N$ is even, and $k = 2N+1 \geq N$ is odd. We are therefore assured that $|a_k - a_\infty|$ takes both values $|1 - a_\infty|$ and $|-1 - a_\infty| = |1 + a_\infty|$ for some $k \geq N$. In order to win, Player N must make *both* of these quantities smaller than ε. But in that event, the triangle inequality would imply $2 \leq |1 - a_\infty| + |1 + a_\infty| < \varepsilon + \varepsilon = 2\varepsilon$. If this inequality is not satisfied, Player N loses.

Having reasoned thusly, Player ε chooses any ε with $0 < \varepsilon \leq 1$. For definiteness, take $\varepsilon = 1$. By the preceding reasoning, there does not exist a positive integer N such that if $k \geq N$, then $|a_k - a_\infty| < \varepsilon = 1$; if such an N existed, we would have

$$2 \leq |1 - a_\infty| + |1 + a_\infty| = |a_{2N} - a_\infty| + |a_{2N+1} + a_\infty| < \varepsilon + \varepsilon = 2,$$

which is false. Since "$(a_k) \to a_\infty$" is false for every real number a_∞, the sequence $(a_k) = \big((-1)^k\big)$ has no limit.

Theorem 11.18. *Let x be a real number satisfying $-1 < x < 1$. The sequence $(a_k) = (x^k)$ converges to 0.*

Proof. Set $a_\infty = 0$. By Corollary 9.17, if we write $|x| = 1/(1+u)$, then $|x^k| \leq 1/(1+ku)$ for all $k \geq 0$.

Fix $\varepsilon > 0$ arbitrarily. By the generalized Archimedean property, there exists a natural number N such that $1/(Nu) < \varepsilon$. If $k \geq N$, then

$$|x^k - a_\infty| = |x^k| \leq \frac{1}{1+ku} < \frac{1}{ku} \leq \frac{1}{Nu} < \varepsilon.$$

By definition, this means $(x^k) \to a_\infty = 0$. □

Example 11.19. For $x = 1/2$ or $x = -4/5$, say, the conclusion of Theorem 11.18 is not intuitively surprising. However, for a number such as $x = 0.99999999999999999 = 1 - 10^{-20}$, a "fairly large" exponent n may be needed to make the power x^n "small".

Remark 11.20. The sequences in Examples 11.5 and 11.6 cannot be handled so naively from the definition. We turn next to theoretical criteria that help resolve these examples.

11.3 Convergence Criteria

Definition 11.21. A real sequence (a_k) is *bounded* if there exists a real number M such that $|a_k| \leq M$ for all k.

We say (a_k) is *bounded above* if its set of terms is bounded above, i.e., if there exists a real number M such that $a_k \leq M$ for all k. Any particular M is called an *upper bound* for the sequence.

Similarly, (a_k) is *bounded below* if there exists a real number m such that $m \leq a_k$ for all k.

Remark 11.22. Geometrically, a real sequence is bounded if there exists an interval (of finite radius) centered at 0 that contains every term.

A real sequence is bounded above if every term lies to the left of some fixed real number, and is bounded below if every term lies to the right of some (other) number.

Proposition 11.23. *If (a_k) is a convergent sequence with limit a_∞, then*

(i) *(a_k) is bounded.*

(ii) *The sequence $\big(|a_k|\big)$ converges to $|a_\infty|$.*

Proof. By hypothesis, for every $\varepsilon > 0$, there exists an N such that if $k \geq N$, then $|a_k - a_\infty| < \varepsilon$.

(i). Taking $\varepsilon = 1$, there exists an N such that $|a_k - a_\infty| < 1$ for $k \geq N$. Let $M = \max(|a_0|, |a_1|, \ldots, |a_{N-1}|, |a_\infty| + 1)$. It suffices to show $|a_k| \leq M$ for all k.

If $0 \leq k < N$, then $|a_k| \leq M$ by construction. If $k \geq N$, the triangle inequality implies

$$|a_k| \leq |a_\infty| + |a_k - a_\infty| < |a_\infty| + 1 \leq M.$$

In summary, $|a_k| \leq M$ for all k.

(ii). Let $\varepsilon > 0$, and choose N such that if $k \geq N$, then $|a_k - a_\infty| < \varepsilon$. By the reverse triangle inequality, $\big||a_k| - |a_\infty|\big| \leq |a_k - a_\infty| < \varepsilon$ if $k \geq N$. Since $\varepsilon > 0$ was arbitrary, $\big(|a_k|\big) \to |a_\infty|$. □

Monotone Sequences

Definition 11.24. Let (a_k) be a *real* sequence.

If $a_k \leq a_{k+1}$ for all k, namely if $a_k \leq a_{k'}$ whenever $k \leq k'$, we say (a_k) is *non-decreasing*.

If $a_{k+1} \leq a_k$ for all k, namely if $a_{k'} \leq a_k$ whenever $k \leq k'$, we say (a_k) is *non-increasing*.

A sequence that is either non-decreasing or non-increasing is said to be *monotone*.

A slightly more general condition is useful in practice, and does not require much extra work to accommodate.

Definition 11.25. A real sequence (a_k) is *eventually non-decreasing* if there exists an index N such that $a_k \leq a_{k+1}$ for all $k \geq N$.

Remark 11.26. By induction, if $N \leq k \leq k'$ then $a_N \leq a_k \leq a_{k'}$.

Example 11.27. The real sequence defined by $a_k = k^2/(k^2 - 500)$ is eventually non-decreasing, since

$$\frac{k^2}{k^2 - 500} = \frac{k^2 - 500 + 500}{k^2 - 500} = 1 - \frac{500}{k^2 - 500},$$

and if $k^2 > 500$, i.e., if $k \geq 23 = N$, the denominator increases with k.

Example 11.28. The real sequences defined by $a_k = (-1)^k k$ or by $a_k = (-1)^k/k$ are *not* eventually non-decreasing.

For practice, give definitions of *eventually non-increasing* and *eventually monotone* real sequences, and modify of the statement and proof of Theorem 11.29 below to handle these types of sequence.

Theorem 11.29. *If (a_k) is an eventually non-decreasing sequence, then*

(i) *(a_k) is bounded below.*

(ii) *(a_k) converges if and only if it is bounded above.*

Proof. By hypothesis, there exists an integer N such that $a_k \leq a_{k'}$ whenever $N \leq k \leq k'$. Particularly, $a_N \leq a_k$ if $N \leq k$.

(i). The real number $m = \min(a_0, a_1, \ldots, a_N)$ is a lower bound for (a_k): By construction, $m \leq a_k$ for $0 \leq k \leq N$. On the other hand, if $N \leq k$, then $m \leq a_N \leq a_k$ as noted above.

(ii). (Convergent implies bounded above). By Proposition 11.23, a convergent real sequence is bounded, hence bounded above.

(Bounded above implies convergent). Suppose there exists a real number M such that $a_k \leq M$ for all k. The set $A = \{a_k : N \leq k\}$ is non-empty and bounded above by M, so $a_\infty = \sup A$ exists by the completeness axiom. It suffices to prove $(a_k) \to a_\infty$.

Let $\varepsilon > 0$ be arbitrary. The number $a_\infty - \varepsilon < a_\infty$ is not an upper bound of A, so there exists an integer $N_0 \geq N$ such that $a_\infty - \varepsilon < a_{N_0}$. If $N_0 \leq k$, then $a_{N_0} \leq a_k \leq \sup A$, so

$$a_\infty - \varepsilon < a_{N_0} \leq a_k \leq a_\infty < a_\infty + \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, we have shown $(a_k) \to a_\infty$. □

Example 11.30. The sequence $a_k = 1/(k+1)$ is non-increasing, and is bounded below since every term is positive. Consequently, (a_k) converges to $a_\infty = \inf\{a_k\}$. Of course, we have already seen that this sequence converges to 0; here we have deduced a less specific conclusion using a more general theorem.

Example 11.31. Consider the sequence (a_k) in Example 11.5. By mathematical induction, each term is positive and satisfies $a_k^2 - 2 > 0$. Further, the sequence is non-increasing since

$$a_k - a_{k+1} = a_k - \frac{1}{2}\left(a_k + \frac{2}{a_k}\right) = \frac{1}{2}\left(a_k - \frac{2}{a_k}\right) = \frac{a_k^2 - 2}{2a_k} > 0.$$

Consequently, $\lim a_k$ exists. Note that we have shown that this sequence converges without explicitly exhibiting the limit.

11.4 Algebraic Properties of Limits

Theorem 11.32. *Let (a_k) and (b_k) be real sequences converging to a_∞ and b_∞ respectively.*

(i) *$(a_k + b_k)$ converges to $a_\infty + b_\infty$.*

(ii) *$(a_k b_k)$ converges to $a_\infty b_\infty$.*

(iii) *If $b_k \neq 0$ for all k and if $b_\infty \neq 0$, then (a_k/b_k) converges to a_∞/b_∞.*

Proof. (i). Let $\varepsilon > 0$ be arbitrary. By hypothesis, there exists an integer N_1 such that if $k \geq N_1$, then $|a_k - a_\infty| < \varepsilon/2$. Similarly, there exists an integer N_2 such that if $k \geq N_2$, then $|b_k - b_\infty| < \varepsilon/2$. Let $N = \max(N_1, N_2)$.

If $k \geq N$, the triangle inequality implies

$$|a_k + b_k - (a_\infty + b_\infty) \leq |a_k - a_\infty| + |b_k - b_\infty| < (\varepsilon/2) + (\varepsilon/2) = \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, $(a_k + b_k) \to a_\infty + b_\infty$.

(ii). Our goal is to make $|a_k b_k - a_\infty b_\infty|$ small, knowing only that we can make $|a_k - a_\infty|$ and $|b_k - b_\infty|$ small. The following algebra idiom is worth remembering:

$$\begin{aligned} a_k b_k - a_\infty b_\infty &= a_k b_k - a_\infty b_k + a_\infty b_k - a_\infty b_\infty \\ &= (a_k - a_\infty) b_k + a_\infty (b_k - b_\infty). \end{aligned}$$

It suffices to make the absolute value of each summand small.

By Proposition 11.23, there exist positive real numbers L and M such that $1 + |a_k| \leq L$ and $1 + |b_k| \leq M$ for all k. Let ε be an arbitrary real number with $0 < \varepsilon < 1$, and put $r = \varepsilon/2M < \varepsilon$ and $s = \varepsilon/2L$.

Since $(a_k) \to a_\infty$, there exists an integer N_1 such that if $k \geq N_1$, then $|a_k - a_\infty| < r$, and since $(b_k) \to b_\infty$, there exists an integer N_2 such that if $k \geq N_2$, then $|b_k - b_\infty| < s$.

Let $N = \max(N_1, N_2)$, so that $|a_\infty| < |a_N| + \varepsilon < L$. If $k \geq N$, the preceding algebra idiom gives

$$\begin{aligned} |a_k b_k - a_\infty b_\infty| &\leq |a_k - a_\infty|\,|b_k| + |a_\infty|\,|b_k - b_\infty| \\ &< \frac{\varepsilon}{2M} \cdot M + \frac{\varepsilon}{2L} \cdot L < \varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, $(a_k b_k) \to a_\infty b_\infty$.

(iii). We first prove that $(1/b_k) \to 1/b_\infty$. Fix $\varepsilon > 0$ arbitrarily, and put $r = \frac{1}{2}\min(|b_\infty|,\ \varepsilon \cdot |b_\infty|^2)$. Note that $r > 0$ because $|b_\infty| > 0$, and $2r/|b_\infty|^2 \leq \varepsilon$ by construction.

Because $(b_k) \to b_\infty$, there exists an integer N such that if $k \geq N$, then $|b_k - b_\infty| < r$. Consequently, $|b_k| > \frac{1}{2}|b_\infty|$, and

$$\left|\frac{1}{b_k} - \frac{1}{b_\infty}\right| = \frac{|b_\infty - b_k|}{|b_k|\,|b_\infty|} < \frac{2r}{|b_\infty|^2} \leq \varepsilon.$$

(iii) now follows immediately from (ii) by writing $\dfrac{a_k}{b_k} = a_k \cdot \dfrac{1}{b_k}$. □

Example 11.33. Let m be a positive integer. The sequence $(a_k)_{k=1}^{\infty}$ defined by $a_k = k^{-m} = 1/k^m$ converges to 0. When $m = 1$, this assertion is proven in Example 11.14. For larger m, Theorem 11.32 (ii) establishes the inductive step.

Example 11.34. In practice, Theorem 11.32 is often invoked without formally defining a sequence. The theorem grants license to move a limit into or out of an arithmetic expression:

$$\lim_{k\to\infty} \frac{k-1}{k+1} = \lim_{k\to\infty} \frac{1-(1/k)}{1+(1/k)} = \frac{1-\lim_k(1/k)}{1+\lim_k(1/k)} = \frac{1}{1} = 1;$$

$$\lim_{k\to\infty} \frac{2k}{k^2+1} = \lim_{k\to\infty} \frac{(2k)/k^2}{1+(1/k^2)} = \frac{2\lim_k(1/k)}{1+\lim_k(1/k^2)} = \frac{0}{1} = 0;$$

$$\lim_{k\to\infty} \left(\frac{k+1}{k}\right)^m = \left(\lim_{k\to\infty} \frac{k+1}{k}\right)^m = (1)^m = 1; \quad \text{etc.}$$

In the third example, the integer exponent $m \geq 0$ is arbitrary, but "fixed", i.e., independent of k. Bringing the limit inside the parentheses implicitly involves an inductive argument together with Theorem 11.32.

Note carefully that Theorem 11.32 does *not* imply

$$\lim_{k\to\infty} \left(\frac{k+1}{k}\right)^k = 1.$$

In fact, this limit is not 1, but e, an irrational number that pervades pure and applied mathematics.

Example 11.35. Consider once again the convergent sequence (a_k) in Example 11.5. Because the limit exists and is positive, we can evaluate the limit by letting $k \to \infty$ on each side of the recursion relation:

$$a_{k+1} = \frac{1}{2}\left(a_k + \frac{2}{a_k}\right) \to a_\infty = \frac{1}{2}\left(a_\infty + \frac{2}{a_\infty}\right).$$

Algebra gives $a_\infty^2 = 2$. That is, 2 has a real square root, $\sqrt{2}$.

Limits and Inequalities

Proposition 11.36. *Let (a_k) and (b_k) be convergent real sequences, with respective limits a_∞ and b_∞.*

(i) *If $0 \leq a_k$ for all but finitely many k, then $0 \leq a_\infty$.*

(ii) *If $a_k \leq b_k$ for all but finitely many k, then $a_\infty \leq b_\infty$.*

(iii) *If $c \leq a_k \leq d$ for all but finitely many k, then $c \leq a_\infty \leq d$.*

Remark 11.37. In words, "non-strict inequality is preserved in the limit". Note that *strict* inequality is not generally preserved in the limit: We have $1/k > 0$ for all $k \geq 1$, but $\lim_k 1/k = 0$.

Proof. (i) We prove the contrapositive: If $a_\infty < 0$, then $a_k < 0$ for infinitely many k.

Put $\varepsilon = -a_\infty/2$, so $\varepsilon > 0$ by hypothesis. Since $(a_k) \to a_\infty$, there exists an N such that if $k \geq N$, then $|a_k - a_\infty| < \varepsilon$. But $|a_k - a_\infty| < \varepsilon$ if and only if $a_\infty/2 < a_k - a_\infty < -a_\infty/2$, and this implies $a_k < a_\infty/2 < 0$. That is, if $k \geq N$, then $a_k < 0$.

(ii) Define $c_k = b_k - a_k$. By hypothesis, $0 \leq c_k$ for all but finitely many k. By Theorem 11.32, (c_k) converges to $b_\infty - a_\infty$. Part (i) implies $0 \leq b_\infty - a_\infty$, i.e., $a_\infty \leq b_\infty$.

(iii) This follows immediately from (ii). □

Divergence to Infinity

The definition of sequential convergence makes no sense if $a_\infty = \infty$ or $a_\infty = -\infty$. The symbols ∞ and $-\infty$ are not real numbers, so formal inequalities such as $|a_k - \infty| < \varepsilon$ have no meaning.

Nonetheless, it is useful to be able to study sequences that "approach" ∞ or $-\infty$. Such sequences diverge (i.e., do not have a real limit), but they still enjoy some special properties of convergent sequences.

Definition 11.38. Let (a_k) be a real sequence. We say (a_k) *diverges to* ∞, denoted $(a_k) \to \infty$, if the following condition holds:

> For every real number M, there exists a natural number N such that if $k \geq N$, then $a_k > M$.

We say (a_k) *diverges to* $-\infty$, denoted $(a_k) \to -\infty$, if:

> For every real number M, there exists a natural number N such that if $k \geq N$, then $a_k < M$.

Example 11.39. The sequence $a_k = k$ diverges to infinity: By the Archimedean property, if $M \in \mathbf{R}$, there exists a natural number N such that $N > M$. If $k \geq N$, then $a_k = k \geq N > M$.

Example 11.40. Fix $x > 1$. The sequence $a_k = x^k$ (see Example 11.4) diverges to infinity. To see this, note that $x = 1 + u$ for some $u > 0$. By Theorem 9.16, $1 + ku \leq x^k = a_k$ for all $k \geq 0$. Fix $M \in \mathbf{R}$, and use the generalized Archimedean property to choose a natural number N such that $M < Nu$. If $k \geq N$, then

$$M < Nu < 1 + Nu \leq 1 + ku \leq x^k = a_k.$$

Since M was an arbitrary real number, $(x^k) \to \infty$.

Remark 11.41. If (a_k) is a real sequence that is eventually non-decreasing, then either (a_k) is bounded above (hence convergent to a finite limit), or not bounded above (hence divergent to ∞).

Analogous remarks hold for a real sequence that is eventually non-increasing. Consequently, an eventually monotone sequence *always* has an extended real limit.

Theorem 11.42. *Assume* $(a_k) \to \infty$ *and* $(b_k) \to b_\infty$ *with* b_∞ *real. Then*

(i) $(a_k + b_k) \to \infty$.

(ii) *If* $b_\infty > 0$, *then* $(a_k b_k) \to \infty$. *If* $b_\infty < 0$, *then* $(a_k b_k) \to -\infty$.

(iii) *If* $a_k \neq 0$ *for all* k, *then* $(b_k/a_k) \to 0$.

Proof. (i) Since a convergent sequence is bounded (Proposition 11.23) and $(b_k) \to b_\infty$, there exists a real number $B > 0$ such that $|b_k| \leq B$ for all k. Let M be arbitrary. Since $(a_k) \to \infty$, there exists an N such that if $k \geq N$, then $a_k > M + B$, which implies

$$a_k + b_k \geq a_k - |b_k| > (M + B) - B = M.$$

Since M was arbitrary, $(a_k + b_k) \to \infty$.

(ii) Assume $b_\infty > 0$. Taking $\varepsilon = b_\infty/2 > 0$, there exists an N_1 such that if $k \geq N_1$, then $|b_k - b_\infty| < \varepsilon = b_\infty/2$. Rearranging gives $b_\infty/2 < b_k$ for $k \geq N_1$.

Fix M arbitrarily. Since $(a_k) \to \infty$, there exists an $N \geq N_1$ such that if $k \geq N$, then $a_k > 2M/b_\infty$. But this implies

$$a_k b_k > (2M/b_\infty) \cdot (b_\infty/2) = M,$$

and since M was arbitrary, $(a_k b_k) \to \infty$. To handle the case $b_\infty < 0$, multiply appropriately by -1 in the preceding proof.

(iii) Suppose $a_k \neq 0$ for all k, so the quotient sequence (b_k/a_k) is defined. As in (i), let $B > 0$ be a bound for $|b_k|$. Fix $\varepsilon > 0$ arbitrarily. Since $(a_k) \to \infty$, there is an N such that if $k \geq N$, then $a_k > B/\varepsilon$, which implies $b_k/a_k < B(\varepsilon/B) = \varepsilon$. Since $\varepsilon > 0$ was arbitrary, $(b_k/a_k) \to 0$. □

Remark 11.43. Theorem 11.42 may be interpreted as assigning values to certain arithmetic expressions containing infinity: If $L > 0$ is real, then

$$\infty \pm L = \infty, \qquad \infty \cdot (\pm L) = \pm\infty, \qquad \pm L/\infty = 0.$$

Easy modifications of the preceding arguments establish that

$$\infty + \infty = \infty, \qquad -\infty - \infty = -\infty, \qquad \pm\infty \cdot \infty = \pm\infty.$$

However, ∞ and $-\infty$ are not real numbers, and the preceding "equations" must be understood as theorems about limits. By contrast, the following expressions are *undefined*, in the sense that their value depends on the sequences used to approximate them:

$$\infty - \infty, \qquad 0 \cdot (\pm\infty), \qquad \infty/\infty, \qquad 0/0.$$

Consequently, care is required when manipulating algebraic expressions involving ∞. For example, it is true (in the "limited" sense above) that $1 + \infty = \infty$, but not legitimate to subtract ∞, "deducing" that $1 = 0$.

11.5 Subsequences

Definition 11.44. An *index sequence* is a strictly increasing sequence ν of natural numbers, i.e., $\nu(k) \in \mathbf{N}$ and $\nu(k) < \nu(k+1)$ for all k.

Two useful properties are easily verified by mathematical induction.

Lemma 11.45. *Let ν be an index sequence.*

(i) *If $k < k'$, then $\nu(k) < \nu(k')$.*

(ii) *$k \leq \nu(k)$ for all k, and if the inequality is strict for some k_0, then the inequality is strict for all $k \geq k_0$.*

Definition 11.46. Let $\mathbf{a}$ be a real sequence and ν an index sequence. The sequence $\mathbf{b}$ defined by $b_k = a_{\nu(k)}$ is called a *subsequence* of $\mathbf{a}$.

Remark 11.47. In words, a subsequence of (a_n) is a sequence obtained by selecting an infinite number of terms $a_{\nu(1)}$, $a_{\nu(2)}$, …, $a_{\nu(k)}$, …, *in their original ordering*, i.e., subject to $\nu(1) < \nu(2) < \cdots < \nu(k) < \ldots$.

Example 11.48. The subsequence (a_{2k}), for which we take $\nu(k) = 2k$, consists of the *even terms* of (a_n). The subsequence (a_{2k+1}), taking $\nu(k) = 2k + 1$, consists of the *odd terms* of (a_n).

Example 11.49. Let $(a_n)_{n=0}^{\infty}$ be a real sequence. For each N in $\mathbf{N}$, the subsequence $(a_n)_{n=N}^{\infty} = (a_{N+k})_{k=0}^{\infty}$ is a *tail* of (a_n), obtained by discarding the terms a_0, …, a_{N-1}. Here $\nu(k) = N + k$.

Remark 11.50. Convergence of a sequence is determined solely by convergence of an arbitrary tail; intuitively, prepending or omitting finitely many terms cannot change the convergence or divergence of a sequence.

Remark 11.51. If some tail of a sequence has a property X, we say the original sequence is "eventually X". Terms such as "eventually non-decreasing" have already been introduced. Similarly, we might say a sequence is "eventually positive", "eventually no larger than 1 in absolute value", or "eventually constant".

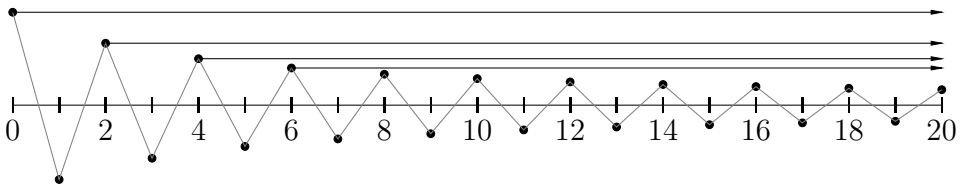
Conditions such as "eventually bounded" or "eventually convergent" are syntactically "legal", but carry no meaning, because a tail cannot be bounded, unbounded, convergent, or divergent unless the original sequence already possesses the same property.

Proposition 11.52. *Let (a_n) be a real sequence that converges to a_∞. If $(b_k) = (a_{\nu(k)})$ is an arbitrary subsequence, then $(b_k) \to a_\infty$.*

Proof. Fix $\varepsilon > 0$ arbitrarily, and pick an index N such that if $n \geq N$, then $|a_n - a_\infty| < \varepsilon$. If $k \geq N$, then $\nu(k) \geq N$ by Lemma 11.45, so $|b_k - a_\infty| = |a_{\nu(k)} - a_\infty| < \varepsilon$. □

Definition 11.53. Let **a** be a real sequence. An index n is a *peak* of **a** if $a_m \leq a_n$ for all $m > n$.

Remark 11.54. Intuitively, a peak is a location n from which, standing at height a_n and looking to the right, you can see all the way to infinity.



Example 11.55. If $a_n = (-1)^n 4/(4+n)$, then every even natural number is a peak. (Arrows indicate unobstructed lines of sight.)

Example 11.56. A sequence **a** is non-increasing if and only if every natural number is a peak of **a**.

If **a** is non-decreasing, then **a** has no peaks.

If **a** is unbounded above, then **a** has no peaks.

Theorem 11.57. *Every real sequence has a monotone subsequence.*

Proof. We consider two cases: The sequence **a** has infinitely many peaks, or only finitely many. In each case, we construct a monotone subsequence recursively.

(Infinitely many peaks). Let $\nu(1)$ be a peak. Now let $m \geq 1$, and assume inductively that there exist peaks $\nu(1) < \nu(2) < \cdots < \nu(m)$. Since **a** has infinitely many peaks, there exists a peak $\nu(m+1) > \nu(m)$. The subsequence $(a_{\nu(k)})$ is non-increasing: By definition of a peak, $a_{\nu(k+1)} \leq a_{\nu(k)}$ for all k.

(Finitely many peaks). Since there are only finitely many peaks, there exists an integer $\nu(1)$ that is greater than every peak. Now let $m \geq 1$, and assume inductively that there exist indices $\nu(1) < \nu(2) < \cdots < \nu(m)$ such that $a_{\nu(1)} < a_{\nu(2)} < \cdots < a_{\nu(m)}$. Since $\nu(m)$ is not a peak, there exists an index $\nu(m+1) > \nu(m)$ such that $a_{\nu(m)} < a_{\nu(m+1)}$. This completes the recursive step.

The subsequence $(a_{\nu(k)})$ is increasing by construction. □

Corollary 11.58 (The Bolzano-Weierstrass theorem)**.** *Every bounded real sequence has a convergent subsequence.*

Proof. Let **a** be a bounded sequence. By the theorem, there exists a monotone subsequence $(a_{\nu(k)})$. But a bounded, monotone sequence converges by Theorem 11.29. □

11.6 Cauchy Sequences

In order to prove from the definition that a real sequence (a_k) converges, the correct "candidate" limit a_∞ must be known. In practice, unfortunately, the limit is not known. The "Cauchy criterion" reformulates the definition of convergence in a way that refers only to the terms of the sequence itself.

Definition 11.59. A real sequence **a** is a *Cauchy sequence* if the following condition holds:

> For every $\varepsilon > 0$, there exists an index N such that if k and $k' \geq N$, then $|a_{k'} - a_k| < \varepsilon$.

Remark 11.60. Exchanging k and k' has no effect on $|a_{k'} - a_k|$, so (as dictated by convenience) we may assume without loss of generality that $k' > k$, or that $k' < k$.

Instead of specifying that $k' > k$, we often write $k' = k + m$ with m a positive integer. The "Cauchy predicate" becomes "If $k \geq N$ and $m > 0$, then $|a_{k+m} - a_k| < \varepsilon$".

Theorem 11.61. *Let (a_k) be a real sequence. Then (a_k) converges to some real number a_∞ if and only if (a_k) is a Cauchy sequence.*

Proof. (Convergent implies Cauchy). Assume $(a_k) \to a_\infty$, and fix $\varepsilon > 0$ arbitrarily. There exists a natural number N such that if $k \geq N$, then $|a_k - a_\infty| < \varepsilon/2$. Consequently, if k and $k' \geq N$, then

$$|a_{k'} - a_k| \leq |a_{k'} - a_\infty| + |a_\infty - a_k| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, (a_k) is Cauchy.

(Cauchy implies convergent). The proof proceeds as follows. We first prove that every Cauchy sequence is bounded. By the Bolzano-Weierstrass theorem, a Cauchy sequence has a convergent subsequence. Finally, we prove that a Cauchy sequence having a convergent subsequence is itself convergent to the same limit.

Taking $\varepsilon = 1$ in the Cauchy criterion, there is an index N such that if k and $k' \geq N$, then $|a_{k'} - a_k| < 1$. In particular, $|a_k - a_N| < 1$ for all $k \geq N$. Let $M = \max(|a_0|, |a_1|, \ldots, |a_{N-1}|, |a_N| + 1)$. Just as in the proof of Proposition 11.23 (ii), it follows that $|a_k| \leq M$ for all k.

Since the Cauchy sequence (a_k) is bounded, the Bolzano-Weierstrass theorem guarantees there is a subsequence $(a_{\nu(k)})$ converging to some real number a_∞. To complete the proof, it suffices to show $(a_k) \to a_\infty$.

Fix $\varepsilon > 0$ arbitrarily, and pick an index N_1 such that if $k \geq N_1$, then $|a_{\nu(k)} - a_\infty| < \varepsilon/2$. Now use the Cauchy criterion to pick $N \geq N_1$ such that if k and $k' \geq N$, then $|a_k - a_{k'}| < \varepsilon/2$.

Since $\nu(N) \geq N$, we have $|a_{\nu(N)} - a_\infty| < \varepsilon/2$, and $|a_k - a_{\nu(N)}| < \varepsilon/2$ for all $k \geq N$. By the triangle inequality, if $k \geq N$, then

$$|a_k - a_\infty| \leq |a_k - a_{\nu(N)}| + |a_{\nu(N)} - a_\infty| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, $(a_k) \to a_\infty$. □

11.7 Infinite Series

Definition 11.62. Let $(a_k)_{k=0}^\infty$ be a real sequence. We define the sequence (s_n) of *partial sums* as follows:

$$s_n = \sum_{k=0}^{n} a_k = a_0 + a_1 + a_2 + \cdots + a_n.$$

Precisely, the partial sums are defined recursively:

$$s_0 = a_0, \qquad s_{n+1} = s_n + a_{n+1} \quad \text{for } n \geq 0.$$

The sequence (a_k) is *summable* if the sequence of partial sums converges to a finite limit s. In this event, we write

$$s = \lim_{n\to\infty} s_n = \lim_{n\to\infty} \sum_{k=0}^{n} a_k = \sum_{k=0}^{\infty} a_k.$$

The expression on the right is called the *infinite series* with *terms* or *summands* a_k, and is said to *converge* to s.

If the sequence of partial sums does not converge, we say the infinite series $\sum_k a_k$ *diverges*.

Remark 11.63. Think of a real sequence as an infinite list of credits (non-negative terms) and debits (negative terms). The partial sums are the "running totals" of the terms *taken in a specified order*. The sum of a series, if it exists, is the net value in the limit, when "all the terms have been added" in their specified order.

Remark 11.64. Since $s_{n+1} - s_n = a_{n+1}$, if $a_k \geq 0$ for all k, the sequence of partial sums is non-decreasing. More generally, if only finitely many a_k are negative, the sequence of partial sums is eventually non-decreasing. Similar remarks hold if at most finitely many summands are positive.

A few general properties are useful to record before we turn to examples.

Theorem 11.65. *Let (a_k) and (b_k) be summable real sequences, and assume $c \in \mathbf{R}$. The sequences $(a_k + b_k)$ and (ca_k) are summable, and*

$$\sum_{k=0}^{\infty}(a_k + b_k) = \sum_{k=0}^{\infty} a_k + \sum_{k=0}^{\infty} b_k, \qquad \sum_{k=0}^{\infty}(ca_k) = c\sum_{k=0}^{\infty} a_k.$$

Proof. If (s_n) and (t_n) denote the respective sequences of partial sums of (a_k) and (b_k), then $(s_n + t_n)$ and (cs_n) are the respective partial sums of $(a_k + b_k)$ and (ca_k). The theorem follows immediately from Theorem 11.32. □

Proposition 11.66. *Let (a_k) be a real sequence. For all natural numbers N and m, we have:*

$$s_{N+m} - s_N = \sum_{k=N+1}^{N+m} a_k, \quad \textit{and therefore} \quad |s_{N+m} - s_N| \leq \sum_{k=N+1}^{N+m} |a_k|.$$

Theorem 11.67. *A real sequence (a_k) is summable if and only if the sequence of partial sums is Cauchy. In particular, if (a_k) is summable, then $(a_k) \to 0$.*

Proof. The first assertion is immediate from Theorem 11.61. For the second, fix $\varepsilon > 0$, and choose an index N such that if k and $k' \geq N$, then $|s_{k'} - s_k| < \varepsilon$. In particular, if $k \geq N$, then $|a_{k+1}| = |s_{k+1} - s_k| < \varepsilon$. This means $(a_k) \to 0$. □

Example 11.68. Let $a \neq 0$ and r be real numbers. The infinite series

$$\sum_{k=0}^{\infty} ar^k = a + ar + ar^2 + ar^3 + \dots$$

is called the *geometric series* with first term a and ratio r.

If $|r| \geq 1$, then $|ar^k| = |a|\,|r|^k$ does not converge to 0, see Example 11.40, so the geometric series diverges.

Suppose $-1 < r < 1$. The key to analyzing the geometric series is the algebraic observation that multiplying a partial sum by r "shifts"

the summands:

$$s_n = \sum_{k=0}^{n} ar^k = a+ar + ar^2 + ar^3 + \cdots + ar^n,$$
$$rs_n = \sum_{k=1}^{n+1} ar^k = \quad ar + ar^2 + ar^3 + \cdots + ar^n + ar^{n+1}.$$

Subtracting the second line from the first, $(1-r)s_n = a - ar^{n+1}$. Since $r \neq 1$, this equation can be solved for s_n:

$$s_n = \sum_{k=0}^{n} ar^k = a + ar + ar^2 + ar^3 + \cdots + ar^n = \frac{a(1-r^{n+1})}{1-r},$$

the *finite geometric series formula*. (This formula is correct for all real r other than 1.)

By Example 11.4, $(r^n) \to 0$ since $-1 < r < 1$. By Theorem 11.32,

$$\sum_{k=0}^{\infty} ar^k = \lim_{n\to\infty} \sum_{k=0}^{n} ar^k = \lim_{n\to\infty} \frac{a(1-r^{n+1})}{1-r} = \frac{a(1-\lim r^{n+1})}{1-r} = \frac{a}{1-r}.$$

Example 11.69. The so-called *harmonic series*

$$\sum_{k=1}^{\infty} \frac{1}{k} = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} + \ldots$$

introduced in Example 11.6 has terms that decrease to 0. Nonetheless, the harmonic series diverges. To prove this, it suffices to show the partial sums are unbounded.

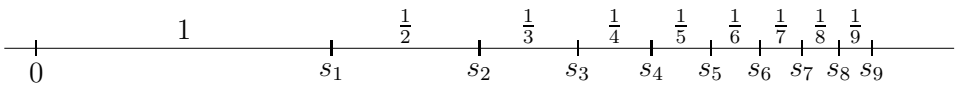
Each of the first two terms is at least 1/2. The next two terms, 1/3 and 1/4, are each no smaller than 1/4, so their sum is greater than $2 \cdot 1/4 = 1/2$.

The next four terms, 1/5, 1/6, 1/7, and 1/8, are each at least 1/8, so their sum is greater than $4 \cdot 1/8 = 1/2$.

Similarly, the next eight terms, 1/9, 1/10, ..., 1/16, sum to at least $8 \cdot 1/16 = 1/2$, the sixteen terms after that sum to at least 1/2, and so on *ad infinitum*.

In more detail, if $m \geq 1$ is an integer, then

$$\sum_{k=2^{m-1}+1}^{2^m} \frac{1}{k} \geq \sum_{k=2^{m-1}+1}^{2^m} \frac{1}{2^m} = \frac{2^{m-1}}{2^m} = \frac{1}{2}.$$



Consequently,

$$\sum_{k=1}^{2^n} \frac{1}{k} = 1 + \sum_{m=1}^{n} \sum_{k=2^{m-1}+1}^{2^m} \frac{1}{k} \geq 1 + \sum_{m=1}^{n} \frac{1}{2} = 1 + \frac{n}{2},$$

which is unbounded by the Archimedean property.

Tests for Summability

No single, easily-applied test determines whether a general infinite series is convergent or divergent. Instead, we develop techniques to handle special classes of series.

Lemma 11.70. *Let (a_k) and (b_k) be real sequences, and let (s_n) and (t_n) denote the sequences of partial sums of (a_k) and (b_k) respectively. If (b_k) is summable, and if there exists a natural number N_0 such that*

$$|s_{m'} - s_m| \leq |t_{m'} - t_m| \quad \textit{whenever } m' > m \geq N_0,$$

then (a_k) is summable.

Proof. If (b_k) is summable, i.e., if the sequence (t_n) of partial sums converges, then (t_n) is Cauchy. Thus, for every $\varepsilon > 0$, there exists an index $N \geq N_0$ such that if $m' > m \geq N$, then $|t_{m'} - t_m| < \varepsilon$. By the hypotheses of the lemma, if $m' > m \geq N$, then $|s_{m'} - s_m| < \varepsilon$ as well. This implies the sequence (s_n) is Cauchy, hence convergent by Theorem 11.61. □

Theorem 11.71 (The comparison test)**.** *Let (a_k) and (b_k) be* nonnegative *real sequences, and assume $a_k \leq b_k$ for all but at most finitely many k. If (b_k) is summable, then (a_k) is summable. Contrapositively, if (a_k) is not summable, then (b_k) is not summable.*

Proof. Let (s_n) and (t_n) denote the sequences of partial sums of (a_k) and (b_k) respectively. By hypothesis, there exists an index N_0 such that $a_k \leq b_k$ for all $k \geq N_0$. If $n \geq N_0$ and $m > 0$, then

$$|s_{n+m} - s_n| = \sum_{k=n+1}^{n+m} a_k \leq \sum_{k=n+1}^{n+m} b_k = |t_{n+m} - t_n|.$$

The theorem follows from Lemma 11.70. □

Absolute Summability

Theorem 11.72. *Let (a_k) be a real sequence. If the sequence $(|a_k|)$ of absolute values is summable, then (a_k) is summable, and*

$$\left|\sum_{k=0}^{\infty} a_k\right| \leq \sum_{k=0}^{\infty} |a_k|.$$

Proof. Let (s_n) and (t_n) denote the sequences of partial sums of (a_k) and $(|a_k|)$ respectively. By the triangle inequality,

$$|s_{n+m} - s_n| = \left|\sum_{k=n+1}^{n+m} a_k\right| \leq \sum_{k=n+1}^{n+m} |a_k| = |t_{n+m} - t_n|$$

for all natural numbers n and m. If (t_n) converges, then (s_n) converges by Lemma 11.70. Since

$$\left|\sum_{k=0}^{n} a_k\right| \leq \sum_{k=0}^{n} |a_k| \quad \text{for all } n \geq 0$$

the inequality in the theorem holds by Proposition 11.36. □

Definition 11.73. A real sequence (a_k) is *absolutely summable* if the sequence $(|a_k|)$ is summable.

If (a_k) is summable but $(|a_k|)$ is not, then (a_k) is *conditionally summable*.

Remark 11.74. Alternatively, an infinite series $\sum_k a_k$ is *absolutely convergent* if $\sum_k |a_k|$ is convergent, and is *conditionally convergent* if $\sum_k a_k$ converges but $\sum_k |a_k|$ diverges.

Definition 11.75. Let (a_k) be a real sequence. The sequences (a_k^+) and (a_k^-) defined by

$$a_k^+ = \max(a_k, 0) = \frac{|a_k| + a_k}{2}, \qquad a_k^- = -\min(a_k, 0) = \frac{|a_k| - a_k}{2},$$

are called the sequence of *positive terms* of (a_k) and the sequence of *negative terms* of (a_k), respectively.

Example 11.76. If $a_k = (-1/2)^k$ for $k \geq 0$, then

$k =$	0	1	2	3	4	5	...
$a_k =$	1	$-1/2$	$1/4$	$-1/8$	$1/16$	$-1/32$	...
$\lvert a_k \rvert =$	1	$1/2$	$1/4$	$1/8$	$1/16$	$1/32$	...
$a_k^+ =$	1	0	$1/4$	0	$1/16$	0	...
$a_k^- =$	0	$1/2$	0	$1/8$	0	$1/32$	...

Remark 11.77. Each sequence $(a_k^\pm)$ is non-negative, so each is summable if and only if its sequence of partial sums is bounded. Further,

$$a_k = a_k^+ - a_k^-, \qquad |a_k| = a_k^+ + a_k^-,$$

and $0 \leq a_k^\pm \leq |a_k|$ for all k.

Proposition 11.78. *Let (a_k) be a real sequence.*

(i) *(a_k) is absolutely summable if and only if both sequences (a_k^+) and (a_k^-) are summable.*

(ii) *If (a_k) is conditionally summable, then both sequences (a_k^+) and (a_k^-) are non-summable.*

Proof. (i) If (a_k) is absolutely summable, i.e., if $(|a_k|)$ is summable, then each sequence $(a_k^\pm)$ is summable by comparison with $(|a_k|)$.

Conversely, if the sequences (a_k^+) and (a_k^-) are both summable, then $(|a_k|) = (a_k^+ + a_k^-)$ is summable by Theorem 11.65.

(ii) By hypothesis, (a_k) is summable but $(|a_k|)$ is not. If either of $(a_k^\pm)$ were summable, then $(|a_k|)$ would be as well by Theorem 11.65, since $|a_k| = 2a_k^\pm \mp a_k$. Contrapositively, $(a_k^\pm)$ is not summable. □

The Ratio Test

Theorem 11.79 (The ratio test)**.** *Let (a_k) be a real sequence. If the limiting ratio*

$$\rho = \lim_{k\to\infty} \left| \frac{a_{k+1}}{a_k} \right|$$

exists, and if $\rho < 1$, then $\sum_k a_k$ is absolutely convergent.

If $\rho > 1$, then $\sum_k a_k$ diverges.

Remark 11.80. If $\rho = 1$, the ratio test is inconclusive: $\sum_k a_k$ may converge absolutely, converge conditionally, or diverge.

Proof. Let $r = (1+\rho)/2$, so that $\rho < r < 1$, and put $\varepsilon = r - \rho$, see diagram. Since $|a_{k+1}/a_k| \to \rho$, there exists an index N such that if $k \geq N$, then

$$\left|\frac{a_{k+1}}{a_k}\right| < \rho + \varepsilon = r < 1.$$

Rearranging, $|a_{k+1}| < |a_k| r$ for all $k \geq N$. By induction on m,

$$|a_{N+m}| < |a_N| r^m \quad \text{for all } m > 0.$$

Consequently, we have

$$\sum_{k=N+1}^{\infty} |a_k| = \sum_{m=1}^{\infty} |a_{N+m}| \leq |a_N| \sum_{m=1}^{\infty} r^m.$$

This upper bound is a convergent geometric series, so $\sum_k |a_k|$ converges by comparison.

The assertion for $\rho > 1$ is left to you, Exercise 11.8. □

Corollary 11.81. *Let r be a real number with $|r| < 1$. For each positive integer m, the series $\sum_k k^m r^k$ is absolutely convergent. Consequently, the sequence $(k^m r^k)$ converges to 0, and in particular is bounded.*

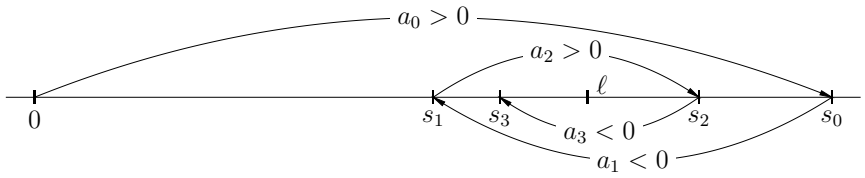
Proof. Setting $a_k = k^m r^k$, we have

$$\lim_{k\to\infty} \left|\frac{(k+1)^m r^{k+1}}{k^m r^k}\right| = \lim_{k\to\infty} \big(1 + (1/k)\big)^m |r| = |r| < 1,$$

so $\sum_k k^m r^k$ is absolutely convergent by the ratio test. The remaining assertions are immediate. □

Alternating Series

Definition 11.82. If (a_k) is a sequence of positive terms, an infinite series $\pm\sum_k (-1)^k a_k$ is said to be *alternating*.



Theorem 11.83 (The alternating series test)**.** *If* (a_k) *is a* non-increasing *positive real sequence, and if* $(a_k) \to 0$*, then the alternating series* $\sum_k (-1)^k a_k$ *converges, and*

$$\left| \sum_{k=n+1}^{\infty} (-1)^k a_k \right| \leq |a_{n+1}|.$$

Proof. Assume for definiteness that the first summand, say a_0, is positive, and let

$$s_n = \sum_{k=0}^{n} (-1)^k a_k.$$

If n is an arbitrary even index, then $s_{n+2} = s_n - a_{n+1} + a_{n+2} \leq s_n$ since $a_{n+2} \leq a_{n+1}$. With the same n, $s_{n+3} = s_{n+1} + a_{n+2} - a_{n+3} \geq s_{n+1}$. Writing $n = 2m$ and combining these inequalities,

$$s_{2m+1} \leq s_{2m+3} \leq s_{2m+2} \leq s_{2m}$$

for all $m \geq 0$. Consequently, the even partial sums form a non-increasing sequence that is bounded below by every odd partial sum, and is therefore convergent to a real number $\ell^+ = \inf_m s_{2m}$. Similarly, the odd partial sums form a non-decreasing sequence that is bounded above by every even partial sum, and is therefore convergent to a real number $\ell^- = \sup_m s_{2m+1}$. Since $(a_k) \to 0$ and

$$\ell^+ - \ell^- \leq s_{2m} - s_{2m+1} = a_{2m+1}$$

for all $m \geq 0$, we have $\ell^+ = \ell^-$, i.e., the sequence of partial sums converges to some real number ℓ, and

$$\left| \sum_{k=n+1}^{\infty} (-1)^k a_k \right| = |\ell - s_n| \leq |a_{n+1}|. \qquad \square$$

Exercises

Exercise 11.1. Consider the real sequence defined by $a_k = k^{(-1)^k}$. Show that (a_k) has a subsequence converging to 0 and a subsequence diverging to ∞.

Exercise 11.2. Construct a sequence $\mathbf{a}$ of positive real numbers such that $\mathbf{a} \to 0$ but $\mathbf{a}$ is not eventually non-decreasing.

Exercise 11.3. Let $\ell_1, \ell_2, \ldots, \ell_m$ be arbitrary distinct real numbers. Prove there exists a real sequence $\mathbf{a}$ such that for each $i = 1, \ldots, m$, some subsequence of $\mathbf{a}$ converges to ℓ_i.

Exercise 11.4. Let (a_k) be a sequence of *positive* real numbers, and let $b_k = 1/a_k$. Prove that if $(a_k) \to 0$, then $(b_k) \to \infty$.

Exercise 11.5. Let (a_k) be an integer sequence, i.e., a_k is an integer for all k. Prove that (a_k) converges if and only if (a_k) is eventually constant.

Exercise 11.6. Let (a_k) and (b_k) be real sequences. Prove that if (b_k) is bounded and $(a_k) \to 0$, then $(a_k b_k) \to 0$.

Exercise 11.7. (The squeeze theorem) Suppose (a_k), (b_k), and (c_k) are real sequences, and assume there exists an N_0 such that if $k \geq N_0$, then $a_k \leq c_k \leq b_k$. Prove that if (a_k) and (b_k) converge to the same limit L, then $\lim_k c_k$ exists is equal to L.

Exercise 11.8. Prove the second part of Theorem 11.79: Let (a_k) be a real sequence. If $\rho = \lim\limits_{k\to\infty} |a_{k+1}/a_k| > 1$, then $\sum_k a_k$ diverges.

Exercise 11.9. Use the ratio test to determine whether the following converge:

(a) $\displaystyle\sum_{k=0}^{\infty} \frac{10^k}{k!}$; (b) $\displaystyle\sum_{k=0}^{\infty} \frac{k!}{k^k}$; (c) $\displaystyle\sum_{k=0}^{\infty} \frac{k^k}{10^k\, k!}$; (d) $\displaystyle\sum_{k=0}^{\infty} \frac{(k!)^2}{(2k)!}$.

Exercise 11.10. Determine the set of real x for which the following converge:

(a) $\displaystyle\sum_{k=0}^{\infty} \frac{x^k}{k!}$; (b) $\displaystyle\sum_{k=0}^{\infty} \frac{x^k\, k!}{k^k}$; (c) $\displaystyle\sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!}$; (d) $\displaystyle\sum_{k=1}^{\infty} \frac{x^k}{k}$.

Exercise 11.11. Determine the set of real x for which the following converge:

(a) $\displaystyle\sum_{k=0}^{\infty} \frac{(2x-3)^k}{k}$; (b) $\displaystyle\sum_{k=0}^{\infty} \frac{(2x-3)^k}{k^2}$; (c) $\displaystyle\sum_{k=0}^{\infty} k(2x-3)^k$.

Exercise 11.12. Let (a_k) be an absolutely summable sequence. Prove the sequence (a_k^2) is absolutely summable.

Exercise 11.13. If (a_k) and (b_k) are real sequences, we may define a new sequence (c_k) by "shuffling" the two given sequences, i.e., $c_{2k} = a_k$ and $c_{2k+1} = b_k$ for $k \geq 0$.

Write out the first six terms of (c_k). Prove that (c_k) converges if and only if (a_k) and (b_k) converge to the same limit.

Exercise 11.14. Let (a_k) be a real sequence and a_∞ a real number. Consider the following conditions:

(i) For every $\varepsilon > 0$, there exists an N such that if $k \geq N$ then $|a_k - a_\infty| < \varepsilon$.

(ii) There exists an N such that for every $\varepsilon > 0$, if $k \geq N$ then $|a_k - a_\infty| < \varepsilon$.

Are these conditions logically equivalent? If so, give a proof. If not, find a "familiar" condition equivalent to (ii), and give an example of a sequence satisfying (ii) but not (i).

Exercise 11.15. Let (a_k) be a real sequence and a_∞ a real number. Consider the following conditions:

(i) There exists an $\varepsilon > 0$ such that for every integer $N \geq 0$, if $k \geq N$ then $|a_k - a_\infty| < \varepsilon$.

(ii) For every integer $N \geq 0$, there exists an $\varepsilon > 0$ such that if $k \geq N$ then $|a_k - a_\infty| < \varepsilon$.

Are these conditions logically equivalent? If so, give a proof. If not, find a "familiar" condition equivalent to (ii), and give an example of a sequence satisfying (ii) but not (i).

Exercise 11.16. Let $\mathbf{a}$ be a sequence of *positive* real numbers, and assume $\mathbf{a} \to 0$.

(a) Prove that for every natural number m, there exists a natural number $n > m$ such that $a_n < a_m$.

(b) Prove that **a** has a strictly decreasing subsequence. Suggestion: Use part (a) to construct a decreasing subsequence inductively.

Exercise 11.17. Let **a** be a real sequence that is not bounded above. Prove there exists a strictly increasing subsequence.

Exercise 11.18. Let **a** be a sequence that is not eventually constant.

(a) Prove that if **a** is non-increasing, then **a** has a strictly decreasing subsequence.

(b) Use the proof of Theorem 11.57 to prove **a** has a strictly monotone subsequence.

Exercise 11.19. Let **a** be a non-increasing sequence of positive real numbers, and assume $\mathbf{a} \to 0$. Prove there exists a positive, non-increasing sequence **b** such that $\mathbf{b} \to 0$ but $(b_k/a_k) \to \infty$.

Exercise 11.20 (Convergence in the mean)**.** Let $(a_k)_{k=1}^{\infty}$ be a sequence of real numbers, and define the *sequence of means* by

$$s_n = \frac{1}{n}\sum_{k=1}^{n} a_k = \frac{a_1 + a_2 + \cdots + a_n}{n}.$$

(a) Prove that if $(a_k) \to a_\infty$, then $(s_n) \to a_\infty$.

(b) Show that if $a_k = (-1)^k$, then the sequence of means converges.

Exercise 11.21. Let $(a_k)_{k=0}^{\infty}$ be a *summable* sequence. If $\nu : \mathbf{N} \to \mathbf{N}$ is a bijection, the infinite series with terms $b_k = a_{\nu(k)}$ is called a *rearrangement* of $\sum_k a_k$.

Prove that if (a_k) is *absolutely* summable, then every rearrangement is summable, and has the same sum.
Hint: Fix $\varepsilon > 0$ arbitrarily, and choose a positive integer N such that

$$\sum_{k=N+1}^{\infty} |a_k| < \varepsilon/2.$$

If $b_k = a_{\nu(k)}$ is a rearrangement, choose N' greater than $\max\{\nu(k) : 0 \le k \le N\}$, and show that if $n \ge N'$, then

$$\left| \sum_{k=1}^{n} b_k - \sum_{k=1}^{\infty} a_k \right| < \varepsilon.$$

Exercise 11.22. Let $(a_k)_{k=1}^{\infty}$ be a "digit sequence", i.e., a sequence all of whose terms are integers between 0 and 9 inclusive.

(a) Prove that the series

$$\sum_{k=1}^{\infty} \frac{a_k}{10^k} = \frac{a_1}{10} + \frac{a_2}{100} + \frac{a_3}{1000} + \cdots = 0.a_1a_2a_3\ldots$$

converges to a real limit between 0 and 1.
Suggestion: The geometric series formula may be useful in obtaining an upper bound.

(b) Prove that the digit sequences $(5, 0, 0, \ldots)$ and $(4, 9, 9, 9, \ldots)$ give rise to the same sum in part (a).

(c) Under what conditions do two distinct digit sequences define the same sum in part (a)?
Suggestions: Let (a_k) and (b_k) be distinct digit sequences defining the same real number, and assume (i) $a_k = b_k$ for $1 \leq k < n$, and (ii) $a_n < b_n$. How much smaller than b_n can a_n be? What can you say about all subsequent digits of each sequence?

Exercise 11.23. Let $(a_k)_{k=1}^{\infty}$ be a sequence all of whose terms are either 0 or 1, and $(b_k)_{k=1}^{\infty}$ be a sequence all of whose terms are either 0 or 2.

(a) Prove that the series

$$\sum_{k=1}^{\infty} \frac{a_k}{2^k} = \frac{a_1}{2} + \frac{a_2}{2^2} + \frac{a_3}{2^3} + \ldots$$

converges to an element of the unit interval $[0, 1]$. Conversely, show that every element of $[0, 1]$ can be expressed in this form.

(b) Prove that the series

$$x = \sum_{k=1}^{\infty} \frac{b_k}{3^k} = \frac{b_1}{3} + \frac{b_2}{3^2} + \frac{b_3}{3^3} + \ldots$$

converges to an element of the Cantor set K. Conversely, show that every element of K can be expressed in this form *in exactly one way.*
Suggestion: Consider how successive "digits" of x are related to the location of x with respect to the approximating sets K_n.

(c) Construct a surjection from K to $[0, 1]$.

Index