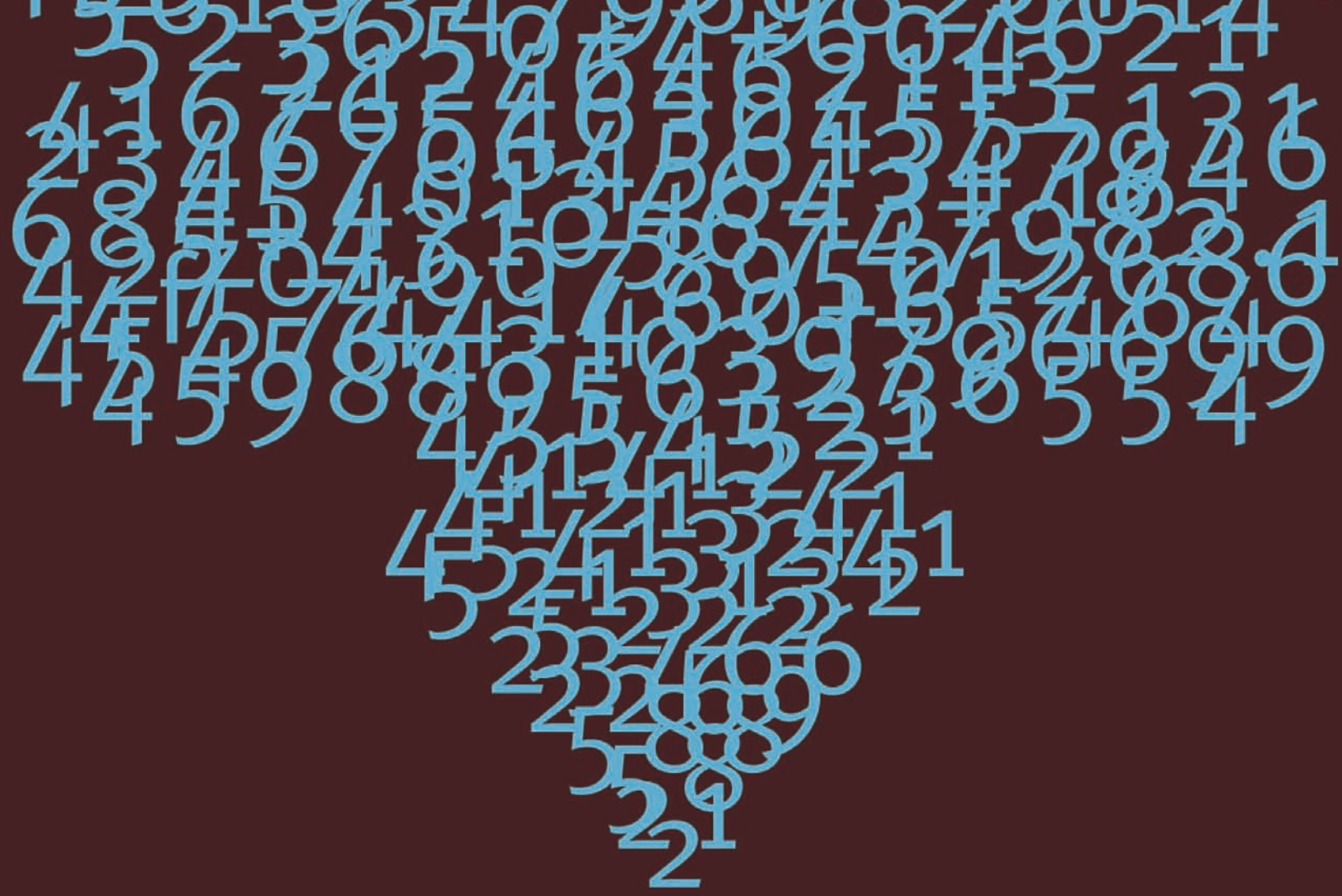
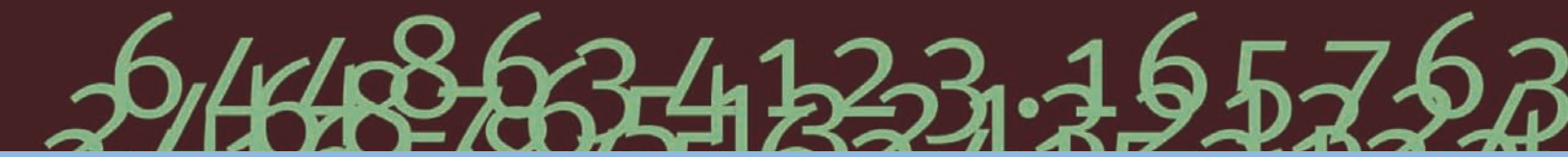


Pearson New International Edition

Elementary Number Theory

Kenneth H. Rosen

Sixth Edition



ALWAYS LEARNING™

PEARSON®

Pearson New International Edition

Elementary Number Theory

Kenneth H. Rosen
Sixth Edition

PEARSON®

Pearson Education Limited
Edinburgh Gate
Harlow
Essex CM20 2JE
England and Associated Companies throughout the world

Visit us on the World Wide Web at: www.pearsoned.co.uk

ISBN 10: 1-292-03954-X
ISBN 13: 978-1-292-03954-1

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Printed in the United States of America

Table of Contents

What Is Number Theory?

There is a buzz about number theory: Thousands of people work on communal number theory problems over the Internet . . . the solution of a famous problem in number theory is reported on the PBS television series NOVA . . . people study number theory to understand systems for making messages secret . . . What is this subject, and why are so many people interested in it today?

Number theory is the branch of mathematics that studies the properties of, and the relationships between, particular types of numbers. Of the sets of numbers studied in number theory, the most important is the set of positive integers. More specifically, the *primes,* those positive integers with no positive proper factors other than 1, are of special importance. A key result of number theory shows that the primes are the multiplicative building blocks of the positive integers. This result, called the *fundamental theorem of arithmetic,* tells us that every positive integer can be uniquely written as the product of primes in nondecreasing order. Interest in prime numbers goes back at least 2500 years, to the studies of ancient Greek mathematicians. Perhaps the first question about primes that comes to mind is whether there are infinitely many. In *The Elements,* the ancient Greek mathematician Euclid provided a proof, that there are infinitely many primes. This proof is considered to be one of the most beautiful proofs in all of mathematics. Interest in primes was rekindled in the seventeenth and eighteenth centuries, when mathematicians such as Pierre de Fermat and Leonhard Euler proved many important results and conjectured approaches for generating primes. The study of primes progressed substantially in the nineteenth century; results included the infinitude of primes in arithmetic progressions, and sharp estimates for the number of primes not exceeding a positive number x. The last 100 years has seen the development of many powerful techniques for the study of primes, but even with these powerful techniques, many questions remain unresolved. An example of a notorious unsolved question is whether there are infinitely many twin primes, which are pairs of primes that differ by 2. New results will certainly follow in the coming decades, as researchers continue working on the many open questions involving primes.

The development of modern number theory was made possible by the German mathematician Carl Friedrich Gauss, one of the greatest mathematicians in history, who in the early nineteenth century developed the language of *congruences*. We say that two integers a and b are congruent modulo m, where m is a positive integer, if m divides $a - b$. This language makes it easy to work with divisibility relationships in much the same way that we work with equations. Gauss developed many important concepts in number theory; for example, he proved one of its most subtle and beautiful results, the *law of quadratic reciprocity.* This law relates whether a prime p is a perfect square modulo

a second prime q to whether q is a perfect square modulo p. Gauss developed many different proofs of this law, some of which have led to whole new areas of number theory.

Distinguishing primes from composite integers is a key problem of number theory. Work on this problem has produced an arsenal of *primality tests*. The simplest primality test is simply to check whether a positive integer is divisible by each prime not exceeding its square root. Unfortunately, this test is too inefficient to use for extremely large positive integers. Many different approaches have been used to determine whether an integer is prime. For example, in the nineteenth century, Pierre de Fermat showed that p divides $2^p - 2$ whenever p is prime. Some mathematicians thought that the converse also was true (that is, that if n divides $2^n - 2$, then n must be prime). However, it is not; by the early nineteenth century, composite integers n, such as 341, were known for which n divides $2^n - 2$. Such integers are called *pseudoprimes*. Though pseudoprimes exist, primality tests based on the fact that most composite integers are not pseudoprimes are now used to quickly find extremely large integers which are are extremely likely to be primes. However, they cannot be used to prove that an integer is prime. Finding an efficient method to prove that an integer is prime was an open question for hundreds of years. In a surprise to the mathematical community, this question was solved in 2002 by three Indian computer scientists, Manindra Agrawal, Neeraj Kayal, and Nitin Saxena. Their algorithms can prove that an integer n is prime in polynomial time (in terms of the number of digits of n).

Factoring a positive integer into primes is another central problem in number theory. The factorization of a positive integer can be found using trial division, but this method is extremely time-consuming. Fermat, Euler, and many other mathematicians devised imaginative factorization algorithms, which have been extended in the past 30 years into a wide array of factoring methods. Using the best-known techniques, we can easily find primes with hundreds or even thousands of digits; factoring integers with the same number of digits, however, is beyond our most powerful computers.

The dichotomy between the time required to find large integers which are almost certainly prime and the time required to factor large integers is the basis of an extremely important secrecy system, the *RSA cryptosystem*. The RSA system is a public key cryptosystem, a security system in which each person has a public key and an associated private key. Messages can be encrypted by anyone using another person's public key, but these messages can be decrypted only by the owner of the private key. Concepts from number theory are essential to understanding the basic workings of the RSA cryptosystem, as well as many other parts of modern cryptography. The overwhelming importance of number theory in cryptography contradicts the earlier belief, held by many mathematicians, that number theory was unimportant for real-world applications. It is ironic that some famous mathematicians, such as G. H. Hardy, took pride in the notion that number theory would never be applied in the way that it is today.

The search for integer solutions of equations is another important part of number theory. An equation with the added proviso that only integer solutions are sought is called *diophantine,* after the ancient Greek mathematician Diophantus. Many different types of diophantine equations have been studied, but the most famous is the *Fermat equation* $x^n + y^n = z^n$. *Fermat's last theorem* states that if n is an integer greater than 2, this

equation has no solutions in integers x, y, and z, where $xyz \neq 0$. Fermat conjectured in the seventeenth century that this theorem was true, and mathematicians (and others) searched for proofs for more than three centuries, but it was not until 1995 that the first proof was given by Andrew Wiles.

As Wiles's proof shows, number theory is not a static subject! New discoveries continue steadily to be made, and researchers frequently establish significant theoretical results. The fantastic power available when today's computers are linked over the Internet yields a rapid pace of new computational discoveries in number theory. Everyone can participate in this quest; for instance, you can join the quest for the new *Mersenne primes,* primes of the form $2^p - 1$, where p itself is prime. In August 2008, the first prime with more than 10 million decimal digits was found: the Mersenne prime $2^{43,112,609} - 1$. This discovery qualified for a $100,000 prize from the Electronic Frontier Foundation. A concerted effort is under way to find a prime with more than 100 million digits, with a $150,000 prize offered. After learning about some of the topics covered in this text, you may decide to join the hunt yourself, putting your idle computing resources to good use.

What is elementary number theory? You may wonder why the word "elementary" is part of the title of this book. This book considers only that part of number theory called *elementary number theory,* which is the part not dependent on advanced mathematics, such as the theory of complex variables, abstract algebra, or algebraic geometry. Students who plan to continue the study of mathematics will learn about more advanced areas of number theory, such as analytic number theory (which takes advantage of the theory of complex variables) and algebraic number theory (which uses concepts from abstract algebra to prove interesting results about algebraic number fields).

Some words of advice. As you embark on your study, keep in mind that number theory is a classical subject with results dating back thousands of years, yet is also the most modern of subjects, with new discoveries being made at a rapid pace. It is pure mathematics with the greatest intellectual appeal, yet it is also applied mathematics, with crucial applications to cryptography and other aspects of computer science and electrical engineering. I hope that you find the many facets of number theory as captivating as aficionados who have preceded you, many of whom retained an interest in number theory long after their school days were over.

Experimentation and exploration play a key role in the study of number theory. The results in this book were found by mathematicians who often examined large amounts of numerical evidence, looking for patterns and making conjectures. They worked diligently to prove their conjectures; some of these were proved and became theorems, others were rejected when counterexamples were found, and still others remain unresolved. As you study number theory, I recommend that you examine many examples, look for patterns, and formulate your own conjectures. You can examine small examples by hand, much as the founders of number theory did, but unlike these pioneers, you can also take advantage of today's vast computing power and computational engines. Working through examples, either by hand or with the aid of computers, will help you to learn the subject—and you may even find some new results of your own!

1 The Integers

In the most general sense, number theory deals with the properties of different sets of numbers. In this chapter, we will discuss some particularly important sets of numbers, including the integers, the rational numbers, and the algebraic numbers. We will briefly introduce the notion of approximating real numbers by rational numbers. We will also introduce the concept of a sequence, and particular sequences of integers, including some figurate numbers studied in ancient Greece. A common problem is the identification of a particular integer sequence from its initial terms; we will briefly discuss how to attack such problems.

Using the concept of a sequence, we will define countable sets and show that the set of rational numbers is countable. We will also introduce notations for sums and products, and establish some useful summation formulas.

One of the most important proof techniques in number theory (and in much of mathematics) is mathematical induction. We will discuss the two forms of mathematical induction, illustrate how they can be used to prove various results, and explain why mathematical induction is a valid proof technique.

Continuing, we will introduce the intriguing sequence of Fibonacci numbers, and describe the original problem from which they arose. We will establish some identities and inequalities involving the Fibonacci numbers, using mathematical induction for some of our proofs.

The final section of this chapter deals with a fundamental notion in number theory, that of divisibility. We will establish some of the basic properties of division of integers, including the "division algorithm." We will show how the quotient and remainder of a division of one integer by another can be expressed using values of the greatest integer function (we will describe a few of the many useful properties of this function, as well).

1.1 Numbers and Sequences

In this section, we introduce basic material that will be used throughout the text. In particular, we cover the important sets of numbers studied in number theory, the concept of integer sequences, and summations and products.

Numbers

To begin, we will introduce several different types of numbers. The *integers* are the numbers in the set

$$\{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}.$$

The integers play center stage in the study of number theory. One property of the positive integers deserves special mention.

The Well-Ordering Property Every nonempty set of positive integers has a least element.

The well-ordering property may seem obvious, but it is the basic principle that allows us to prove many results about sets of integers, as we will see in Section 1.3.

The well-ordering property can be taken as one of the axioms defining the set of positive integers or it may be derived from a set of axioms in which it is not included. (See Appendix A for axioms for the set of integers.) We say that the set of positive integers is *well ordered.* However, the set of all integers (positive, negative, and zero) is not well ordered, as there are sets of integers without a smallest element, such as the set of negative integers, the set of even integers less than 100, and the set of all integers itself.

Another important class of numbers in the study of number theory is the set of numbers that can be written as a ratio of integers.

Definition. The real number r is *rational* if there are integers p and q, with $q \neq 0$, such that $r = p/q$. If r is not rational, it is said to be *irrational.*

Example 1.1. The numbers $-22/7$, $0 = 0/1$, $2/17$, and $1111/41$ are rational numbers. ◀

Note that every integer n is a rational number, because $n = n/1$. Examples of irrational numbers are $\sqrt{2}$, π, and e. We can use the well-ordering property of the set of positive integers to show that $\sqrt{2}$ is irrational. The proof that we provide, although quite clever, is not the simplest proof that $\sqrt{2}$ is irrational. You may prefer the proof that we will give in Chapter 4, which depends on concepts developed in that chapter. (The proof that e is irrational is left as Exercise 44. We refer the reader to [HaWr08] for a proof that π is irrational. It is not easy.)

Theorem 1.1. $\sqrt{2}$ is irrational.

Proof. Suppose that $\sqrt{2}$ were rational. Then there would exist positive integers a and b such that $\sqrt{2} = a/b$. Consequently, the set $S = \{k\sqrt{2} \mid k \text{ and } k\sqrt{2} \text{ are positive integers}\}$ is a nonempty set of positive integers (it is nonempty because $a = b\sqrt{2}$ is a member of S). Therefore, by the well-ordering property, S has a smallest element, say, $s = t\sqrt{2}$.

We have $s\sqrt{2} - s = s\sqrt{2} - t\sqrt{2} = (s-t)\sqrt{2}$. Because $s\sqrt{2} = 2t$ and s are both integers, $s\sqrt{2} - s = s\sqrt{2} - t\sqrt{2} = (s-t)\sqrt{2}$ must also be an integer. Furthermore, it is positive, because $s\sqrt{2} - s = s(\sqrt{2} - 1)$ and $\sqrt{2} > 1$. It is less than s, because $\sqrt{2} < 2$ so that $\sqrt{2} - 1 < 1$. This contradicts the choice of s as the smallest positive integer in S. It follows that $\sqrt{2}$ is irrational. ■

The sets of integers, positive integers, rational numbers, and real numbers are traditionally denoted by $\mathbf{Z}$, $\mathbf{Z}^+$, $\mathbf{Q}$, and $\mathbf{R}$, respectively. Also, we write $x \in S$ to indicate that x belongs to the set S. Such notation will be used occasionally in this book.

We briefly mention several other types of numbers here, though we do not return to them until Chapter 12.

Definition. A number α is *algebraic* if it is the root of a polynomial with integer coefficients; that is, α is algebraic if there exist integers $a_0, a_1, \ldots, a_n$ such that $a_n\alpha^n + a_{n-1}\alpha^{n-1} + \cdots + a_0 = 0$. The number α is called *transcendental* if it is not algebraic.

Example 1.2. The irrational number $\sqrt{2}$ is algebraic, because it is a root of the polynomial $x^2 - 2$. ◀

Note that every rational number is algebraic. This follows from the fact that the number a/b, where a and b are integers and $b \neq 0$, is the root of $bx - a$. In Chapter 12, we will give an example of a transcendental number. The numbers e and π are also transcendental, but the proofs of these facts (which can be found in [HaWr08]) are beyond the scope of this book.

The Greatest Integer Function

In number theory, a special notation is used for the largest integer that is less than or equal to a particular real number.

Definition. The *greatest integer* in a real number x, denoted by $[x]$, is the largest integer less than or equal to x. That is, $[x]$ is the integer satisfying

$$[x] \leq x < [x] + 1.$$

Example 1.3. We have $[5/2] = 2$, $[-5/2] = -3$, $[\pi] = 3$, $[-2] = -2$, and $[0] = 0$. ◀

Remark. The greatest integer function is also known as the *floor function*. Instead of using the notation $[x]$ for this function, computer scientists usually use the notation $\lfloor x \rfloor$. The *ceiling function* is a related function often used by computer scientists. The ceiling function of a real number x, denoted by $\lceil x \rceil$, is the smallest integer greater than or equal to x. For example, $\lceil 5/2 \rceil = 3$ and $\lceil -5/2 \rceil = -2$.

The greatest integer function arises in many contexts. Besides being important in number theory, as we will see throughout this book, it plays an important role in the analysis of algorithms, a branch of computer science. The following example establishes

a useful property of this function. Additional properties of the greatest integer function are found in the exercises at the end of this section and in [GrKnPa94].

Example 1.4. Show that if n is an integer, then $[x+n]=[x]+n$ whenever x is a real number. To show that this property holds, let $[x]=m$, so that m is an integer. This implies that $m \le x < m+1$. We can add n to this inequality to obtain $m+n \le x+n < m+n+1$. This shows that $m+n=[x]+n$ is the greatest integer less than or equal to $x+n$. Hence, $[x+n]=[x]+n$. ◀

Definition. The *fractional part* of a real number x, denoted by $\{x\}$, is the difference between x and the largest integer less than or equal to x, namely, $[x]$. That is, $\{x\} = x - [x]$.

Because $[x] \le x < [x]+1$, it follows that $0 \le \{x\} = x - [x] < 1$ for every real number x. The greatest integer in x is also called the *integral part* of x because $x = [x] + \{x\}$.

Example 1.5. We have $\{5/4\} = 5/4 - [5/4] = 5/4 - 1 = 1/4$ and $\{-2/3\} = -2/3 - [-2/3] = -2/3 - (-1) = 1/3$. ◀

Diophantine Approximation

We know that the distance of a real number to the integer closest to it is at most $1/2$. But can we show that one of the first k multiples of a real number must be much closer to an integer? An important part of number theory called *diophantine approximation* studies questions such as this. In particular, it concentrates on questions that involve the approximation of real numbers by rational numbers. (The adjective *diophantine* comes from the Greek mathematician Diophantus, whose biography can be found in Section 13.1.)

Here we will show that among the first n multiples of a real number α, there must be at least one at a distance less than $1/n$ from the integer nearest it. The proof will depend on the famous *pigeonhole principle,* introduced by the German mathematician Dirichlet.[1] Informally, this principle tells us if we have more objects than boxes, when these objects are placed in the boxes, at least two must end up in the same box. Although this seems like a particularly simple idea, it turns out to be extremely useful in number theory and combinatorics. We now state and prove this important fact, which is known as the pigeonhole principle, because if you have more pigeons than roosts, two pigeons must end up in the same roost.

Theorem 1.2. ***The Pigeonhole Principle.*** If $k+1$ or more objects are placed into k boxes, then at least one box contains two or more of the objects.

[1] Instead of calling Theorem 1.2 the pigeonhole principle, Dirichlet called it the *Schubfachprinzip* in German, which translates to the *drawer principle* in English. A biography of Dirichlet can be found in Section 3.1.

Proof. If none of the k boxes contains more than one object, then the total number of objects would be at most k. This contradiction shows that one of the boxes contains at least two or more of the objects. ■

We now state and prove the approximation theorem, which guarantees that one of the first n multiples of a real number must be within $1/n$ of an integer. The proof we give illustrates the utility of the pigeonhole principle. (See [Ro07] for more applications of the pigeonhole principle.) (Note that in the proof we make use of the *absolute value function.* Recall that $|x|$, the absolute value of x, equals x if $x \geq 0$ and $-x$ if $x < 0$. Also recall that $|x - y|$ gives the distance between x and y.)

Theorem 1.3. ***Dirichlet's Approximation Theorem.*** If α is a real number and n is a positive integer, then there exist integers a and b with $1 \leq a \leq n$ such that $|a\alpha - b| < 1/n$.

Proof. Consider the $n + 1$ numbers $0, \{\alpha\}, \{2\alpha\}, \ldots, \{n\alpha\}$. These $n + 1$ numbers are the fractional parts of the numbers $j\alpha$, $j = 0, 1, \ldots, n$, so that $0 \leq \{j\alpha\} < 1$ for $j = 0, 1, \ldots, n$. Each of these $n + 1$ numbers lies in one of the n disjoint intervals $0 \leq x < 1/n$, $1/n \leq x < 2/n, \ldots, (j-1)/n \leq x < j/n, \ldots, (n-1)/n \leq x < 1$. Because there are $n + 1$ numbers under consideration, but only n intervals, the pigeonhole principle tells us that at least two of these numbers lie in the same interval. Because each of these intervals has length $1/n$ and does not include its right endpoint, we know that the distance between two numbers that lie in the same interval is less than $1/n$. It follows that there exist integers j and k with $0 \leq j < k \leq n$ such that $|\{k\alpha\} - \{j\alpha\}| < 1/n$. We will now show that when $a = k - j$, the product $a\alpha$ is within $1/n$ of an integer, namely, the integer $b = [k\alpha] - [j\alpha]$. To see this, note that

$$\begin{aligned} |a\alpha - b| &= |(k-j)\alpha - ([k\alpha] - [j\alpha])| \\ &= |(k\alpha - [k\alpha]) - (j\alpha - [j\alpha])| \\ &= |\{k\alpha\} - \{j\alpha\}| < 1/n. \end{aligned}$$

Furthermore, note that because $0 \leq j < k \leq n$, we have $1 \leq a = k - j \leq n$. Consequently, we have found integers a and b with $1 \leq a \leq n$ and $|a\alpha - b| < 1/n$, as desired. ■

Example 1.6. Suppose that $\alpha = \sqrt{2}$ and $n = 6$. We find that $1 \cdot \sqrt{2} \approx 1.414$, $2 \cdot \sqrt{2} \approx 2.828$, $3 \cdot \sqrt{2} \approx 4.243$, $4 \cdot \sqrt{2} \approx 5.657$, $5 \cdot \sqrt{2} \approx 7.071$, and $6 \cdot \sqrt{2} \approx 8.485$. Among these numbers $5 \cdot \sqrt{2}$ has the smallest fractional part. We see that $|5 \cdot \sqrt{2} - 7| \approx |7.071 - 7| = 0.071 \leq 1/6$. It follows that when $\alpha = \sqrt{2}$ and $n = 6$, we can take $a = 5$ and $b = 7$ to make $|a\alpha - b| < 1/n$. ◄

Our proof of Theorem 1.3 follows Dirichlet's original 1834 proof. Proving a stronger version of Theorem 1.3 with $1/(n + 1)$ replacing $1/n$ in the approximation is not difficult (see Exercise 32). Furthermore, in Exercise 34 we show how to use the Dirichlet approximation theorem to show that, given an irrational number α, there are infinitely many different rational numbers p/q such that $|\alpha - p/q| < 1/q^2$, an important result in the theory of diophantine approximation. We will return to this topic in Chapter 12.

Sequences

A *sequence* $\{a_n\}$ is a list of numbers $a_1, a_2, a_3, \ldots$. We will consider many particular integer sequences in our study of number theory. We introduce several useful sequences in the following examples.

Example 1.7. The sequence $\{a_n\}$, where $a_n = n^2$, begins with the terms 1, 4, 9, 16, 25, 36, 49, 64, ... This is the sequence of the squares of integers. The sequence $\{b_n\}$, where $b_n = 2^n$, begins with the terms 2, 4, 8, 16, 32, 64, 128, 256, ... This is the sequence of powers of 2. The sequence $\{c_n\}$, where $c_n = 0$ if n is odd and $c_n = 1$ if n is even, begins with the terms 0, 1, 0, 1, 0, 1, 0, 1, ... ◄

There are many sequences in which each successive term is obtained from the previous term by multiplying by a common factor. For example, each term in the sequence of powers of 2 is 2 times the previous term. This leads to the following definition.

Definition. A *geometric progression* is a sequence of the form $a, ar, ar^2, ar^3, \ldots, ar^k, \ldots$, where a, the *initial term*, and r, the *common ratio,* are real numbers.

Example 1.8. The sequence $\{a_n\}$, where $a_n = 3 \cdot 5^n$, $n = 0, 1, 2, \ldots$, is a geometric sequence with initial term 3 and common ratio 5. (Note that we have started the sequence with the term a_0. We can start the index of the terms of a sequence with 0 or any other integer that we choose.) ◄

A common problem in number theory is finding a formula or rule for constructing the terms of a sequence, even when only a few terms are known (such as trying to find a formula for the nth triangular number $1 + 2 + 3 + \cdots + n$). Even though the initial terms of a sequence do not determine the sequence, knowing the first few terms can lead to a conjecture for a formula or rule for the terms. Consider the following examples.

Example 1.9. Conjecture a formula for a_n, where the first eight terms of $\{a_n\}$ are 4, 11, 18, 25, 32, 39, 46, 53. We note that each term, starting with the second, is obtained by adding 7 to the previous term. Consequently, the nth term could be the initial term plus $7(n - 1)$. A reasonable conjecture is that $a_n = 4 + 7(n - 1) = 7n - 3$. ◄

The sequence proposed in Example 1.9 is an *arithmetic progression,* that is, a sequence of the form $a, a + d, a + 2d, \ldots, a + nd, \ldots$. The particular sequence in Example 1.9 has $a = 4$ and $d = 7$.

Example 1.10. Conjecture a formula for a_n, where the first eight terms of the sequence $\{a_n\}$ are 5, 11, 29, 83, 245, 731, 2189, 6563. We note that each term is approximately 3 times the previous term, suggesting a formula for a_n in terms of 3^n. The integers 3^n for $n = 1, 2, 3, \ldots$ are 3, 9, 27, 81, 243, 729, 2187, 6561. Looking at these two sequences together, we find that the formula $a_n = 3^n + 2$ produces these terms. ◄

Example 1.11. Conjecture a formula for a_n, where the first ten terms of the sequence $\{a_n\}$ are 1, 1, 2, 3, 5, 8, 13, 21, 34, 55. After examining this sequence from different perspectives, we notice that each term of this sequence, after the first two terms, is the sum of the two preceding terms. That is, we see that $a_n = a_{n-1} + a_{n-2}$ for $3 \leq n \leq 10$. This is an example of a recursive definition of a sequence, discussed in Section 1.3. The terms listed in this example are the initial terms of the Fibonacci sequence, which is discussed in Section 1.4. ◀

Integer sequences arise in many contexts in number theory. Among the sequences we will study are the Fibonacci numbers, the prime numbers (covered in Chapter 3), and the perfect numbers (introduced in Section 7.3). Integer sequences appear in an amazing range of subjects besides number theory. Neil Sloane has amassed a fantastically diverse collection of more than 170,000 integer sequences (as of early 2010) in his *On-Line Encyclopedia of Integer Sequences*. This collection is available on the Web. (Note that in early 2010, the OEIS Foundation took over maintenance of this collection.) (The book [SlPl95] is an earlier printed version containing only a small percentage of the current contents of the encyclopdia.) This site provides a program for finding sequences that match initial terms provided as input. You may find this a valuable resource as you continue your study of number theory (as well as other subjects).

We now define what it means for a set to be countable, and show that a set is countable if and only if its elements can be listed as the terms of a sequence.

Definition. A set is *countable* if it is finite or it is infinite and there exists a one-to-one correspondence between the set of positive integers and the set. A set that is not countable is called *uncountable*.

An infinite set is countable if and only if its elements can be listed as the terms of a sequence indexed by the set of positive integers. To see this, simply note that a one-to-one correspondence f from the set of positive integers to a set S is exactly the same as a listing of the elements of the set in a sequence $a_1, a_2, \ldots, a_n, \ldots$, where $a_i = f(i)$.

Example 1.12. The set of integers is countable, because the integers can be listed starting with 0, followed by 1 and -1, followed by 2 and -2, and so on. This produces the sequence $0, 1, -1, 2, -2, 3, -3, \ldots$, where $a_1 = 0$, $a_{2n} = n$, and $a_{2n+1} = -n$ for $n = 1, 2, \ldots$. ◀

Is the set of rational numbers countable? At first glance, it may seem unlikely that there would be a one-to-one correspondence between the set of positive integers and the set of all rational numbers. However, there is such a correspondence, as the following theorem shows.

Theorem 1.4. The set of rational numbers is countable.

Proof. We can list the rational numbers as the terms of a sequence, as follows. First, we arrange all the rational numbers in a two-dimensional array, as shown in Figure 1.1. We put all fractions with a denominator of 1 in the first row. We arrange these by placing the fraction with a particular numerator in the position this numerator occupies in the list of

all integers given in Example 1.12. Next, we list all fractions on successive diagonals, following the order shown in Figure 1.1. Finally, we delete from the list all fractions that represent rational numbers that have already been listed. (For example, we do not list 2/2, because we have already listed 1/1.)

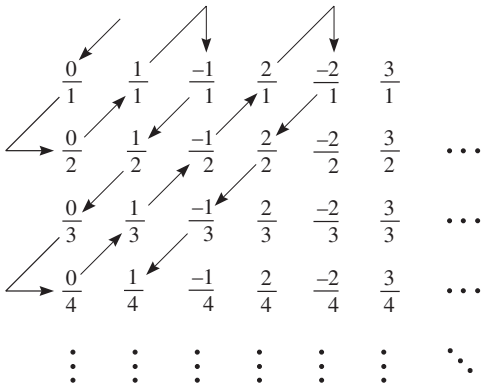


Figure 1.1 *Listing the rational numbers.*

The initial terms of the sequence are $0/1 = 0$, $1/1 = 1$, $-1/1 = -1$, $1/2$, $1/3$, $-1/2$, $2/1 = 2$, $-2/1 = -2$, $-1/3$, $1/4$, and so on.) We leave it to the reader to fill in the details, to see that this procedure lists all rational numbers as the terms of a sequence. ■

We have shown that the set of rational numbers is countable, but we have not given an example of an uncountable set. Such an example is provided by the set of real numbers, as shown in Exercise 45.

1.1 Exercises

1. Determine whether each of the following sets is well ordered. Either give a proof using the well-ordering property of the set of positive integers, or give an example of a subset of the set that has no smallest element.

a) the set of integers greater than 3

b) the set of even positive integers

c) the set of positive rational numbers

d) the set of positive rational numbers that can be written in the form $a/2$, where a is a positive integer

e) the set of nonnegative rational numbers

➢ **2.** Show that if a and b are positive integers, then there is a smallest positive integer of the form $a - bk$, $k \in \mathbf{Z}$.

3. Prove that both the sum and the product of two rational numbers are rational.

4. Prove or disprove each of the following statements.

a) The sum of a rational and an irrational number is irrational.

b) The sum of two irrational numbers is irrational.

c) The product of a rational number and an irrational number is irrational.
d) The product of two irrational numbers is irrational.

* **5.** Use the well-ordering property to show that $\sqrt{3}$ is irrational.

6. Show that every nonempty set of negative integers has a greatest element.

7. Find the following values of the greatest integer function.
a) $[1/4]$ c) $[22/7]$ e) $[[1/2]+[1/2]]$
b) $[-3/4]$ d) $[-2]$ f) $[-3+[-1/2]]$

8. Find the following values of the greatest integer function.
a) $[-1/4]$ c) $[5/4]$ e) $[[3/2]+[-3/2]]$
b) $[-22/7]$ d) $[[1/2]]$ f) $[3-[1/2]]$

9. Find the fractional part of each of these numbers:
a) $8/5$ b) $1/7$ c) $-11/4$ d) 7

10. Find the fractional part of each of these numbers:
a) $-8/5$ b) $22/7$ c) -1 d) $-1/3$

11. What is the value of $[x]+[-x]$ where x is a real number?

12. Show that $[x]+[x+1/2]=[2x]$ whenever x is a real number.

13. Show that $[x+y]\geq[x]+[y]$ for all real numbers x and y.

14. Show that $[2x]+[2y]\geq[x]+[y]+[x+y]$ whenever x and y are real numbers.

15. Show that if x and y are positive real numbers, then $[xy]\geq[x][y]$. What is the situation when both x and y are negative? When one of x and y is negative and the other positive?

16. Show that $-[-x]$ is the least integer greater than or equal to x when x is a real number.

17. Show that $[x+1/2]$ is the integer nearest to x (when there are two integers equidistant from x, it is the larger of the two).

18. Show that if m and n are integers, then $[(x+n)/m]=[([x]+n)/m]$ whenever x is a real number.

* **19.** Show that $\left[\sqrt{[x]}\right]=\left[\sqrt{x}\right]$ whenever x is a nonnegative real number.

* **20.** Show that if m is a positive integer, then

$$[mx]=[x]+[x+(1/m)]+[x+(2/m)]+\cdots+[x+(m-1)/m]$$

whenever x is a real number.

21. Conjecture a formula for the nth term of $\{a_n\}$ if the first ten terms of this sequence are as follows.
a) 3, 11, 19, 27, 35, 43, 51, 59, 67, 75
b) 5, 7, 11, 19, 35, 67, 131, 259, 515, 1027
c) 1, 0, 0, 1, 0, 0, 0, 0, 1, 0
d) 1, 3, 4, 7, 11, 18, 29, 47, 76, 123

22. Conjecture a formula for the nth term of $\{a_n\}$ if the first ten terms of this sequence are as follows.
a) 2, 6, 18, 54, 162, 486, 1458, 4374, 13122, 39366
b) 1, 1, 0, 1, 1, 0, 1, 1, 0, 1

c) 1, 2, 3, 5, 7, 10, 13, 17, 21, 26

d) 3, 5, 11, 21, 43, 85, 171, 341, 683, 1365

23. Find three different formulas or rules for the terms of a sequence $\{a_n\}$ if the first three terms of this sequence are 1, 2, 4.

24. Find three different formulas or rules for the terms of a sequence $\{a_n\}$ if the first three terms of this sequence are 2, 3, 6.

25. Show that the set of all integers greater than -100 is countable.

26. Show that the set of all rational numbers of the form $n/5$, where n is an integer, is countable.

27. Show that the set of all numbers of the form $a + b\sqrt{2}$, where a and b are integers, is countable.

* **28.** Show that the union of two countable sets is countable.

* **29.** Show that the union of a countable number of countable sets is countable.

30. Using a computational aid, if needed, find integers a and b such that $1 \leq a \leq 8$ and $|a\alpha - b| < 1/8$, where α has these values:

a) $\sqrt{2}$ b) $\sqrt[3]{2}$ c) π d) e

31. Using a computational aid, if needed, find integers a and b such that $1 \leq a \leq 10$ and $|a\alpha - b| < 1/10$, where α has these values:

a) $\sqrt{3}$ b) $\sqrt[3]{3}$ c) π^2 d) e^3

32. Prove the following stronger version of Dirichlet's approximation. If α is a real number and n is a positive integer, there are integers a and b such that $1 \leq a \leq n$ and $|a\alpha - b| \leq 1/(n+1)$. (*Hint:* Consider the $n+2$ numbers $0, \ldots, \{j\alpha\}, \ldots, 1$ and the $n+1$ intervals $(k-1)/(n+1) \leq x < k/(n+1)$ for $k = 1, \ldots, n+1$.)

33. Show that if α is a real number and n is a positive integer, then there is an integer k such that $|\alpha - n/k| \leq 1/2k$.

34. Use Dirichlet's approximation theorem to show that if α is an irrational number, then there are infinitely many positive integers q for which there is an integer p such that $|\alpha - p/q| \leq 1/q^2$.

35. Find four rational numbers p/q with $|\sqrt{2} - p/q| \leq 1/q^2$.

36. Find five rational numbers p/q with $|\sqrt[3]{5} - p/q| \leq 1/q^2$.

37. Show that if $\alpha = a/b$ is a rational number, then there are only finitely many rational numbers p/q such that $|p/q - a/b| < 1/q^2$.

The *spectrum sequence* of a real number α is the sequence that has $[n\alpha]$ as its nth term.

38. Find the first ten terms of the spectrum sequence of each of the following numbers.

a) 2 b) $\sqrt{2}$ c) $2 + \sqrt{2}$ d) e e) $(1 + \sqrt{5})/2$

39. Find the first ten terms of the spectrum sequence of each of the following numbers.

a) 3 b) $\sqrt{3}$ c) $(3 + \sqrt{3})/2$ d) π

40. Prove that if $\alpha \neq \beta$, then the spectrum sequence of α is different from the spectrum sequence of β.

** **41.** Show that every positive integer occurs exactly once in the spectrum sequence of α or in the spectrum sequence of β if and only if α and β are positive irrational numbers such that $1/\alpha + 1/\beta = 1$.

The *Ulam numbers* u_n, $n = 1, 2, 3, \ldots$ are defined as follows. We specify that $u_1 = 1$ and $u_2 = 2$. For each successive integer m, $m > 2$, this integer is an Ulam number if and only if it can be written uniquely as the sum of two distinct Ulam numbers. These numbers are named for *Stanislaw Ulam,* who first described them in 1964.

42. Find the first ten Ulam numbers.

* **43.** Show that there are infinitely many Ulam numbers.

* **44.** Prove that e is irrational. (*Hint:* Use the fact that $e = 1 + 1/1! + 1/2! + 1/3! + \cdots$.)

* **45.** Show that the set of real numbers is uncountable. (*Hint:* Suppose it is possible to list the real numbers between 0 and 1. Show that the number whose ith decimal digit is 4 when the ith decimal digit of the ith real number in the list is 5 and is 5 otherwise is not on the list.)

Computations and Explorations

1. Find 10 rational numbers p/q such that $|\pi - p/q| \leq 1/q^2$.

2. Find 20 rational numbers p/q such that $|e - p/q| \leq 1/q^2$.

3. Find as many terms as you can of the spectrum sequence of $\sqrt{2}$. (See the preamble to Exercise 38 for the definition of spectrum.)

STANISLAW M. ULAM (1909–1984) was born in Lvov, Poland. He became interested in astronomy and physics at age 12, after receiving a telescope from his uncle. He decided to learn the mathematics required to understand relativity theory, and at the age of 14 he used textbooks to learn calculus and other mathematics.

Ulam received his Ph.D. from the Polytechnic Institute in Lvov in 1933, completing his degree under the mathematician Banach, in the area of real analysis. In 1935, he was invited to spend several months at the Institute for Advanced Study; in 1936, he joined Harvard University as a member of the Society of Fellows, remaining in this position until 1940. During these years he returned each summer to Poland where he spent time in cafes, such as the Scottish Cafe, intensely doing mathematics with his fellow Polish mathematians.

Luckily for Ulam, he left Poland in 1939, just one month before the outbreak of World War II. In 1940, he was appointed to a position as an assistant professor at the University of Wisconsin, and in 1943, he was enlisted to work in Los Alamos on the development of the first atomic bomb, as part of the Manhattan Project. Ulam made several key contributions that led to the creation of thermonuclear bombs. At Los Alamos, Ulam also developed the Monte Carlo method, which uses a sampling technique with random numbers to find solutions of mathematical problems.

Ulam remained at Los Alamos after the war until 1965. He served on the faculties of the University of Southern California, the University of Colorado, and the University of Florida. Ulam had a fabulous memory and was an extremely verbal person. His mind was a repository of stories, jokes, puzzles, quotations, formulas, problems, and many other types of information. He wrote several books, including *Sets, Numbers, and Universes* and *Adventures of a Mathematician.* He was interested in and contributed to many areas of mathematics, including number theory, real analysis, probability theory, and mathematical biology.

4. Find as many terms as you can of the spectrum sequence of π. (See the preamble to Exercise 38 for the definition of spectrum.)
5. Find the first 1000 Ulam numbers.
6. How many pairs of consecutive integers can you find where both are Ulam numbers?
7. Can the sum of any two consecutive Ulam numbers, other than 1 and 2, be another Ulam number? If so, how many examples can you find?
8. How large are the gaps between consecutive Ulam numbers? Do you think that these gaps can be arbitrarily long?
9. What conjectures can you make about the number of Ulam numbers less than an integer n? Do your computations support these conjectures?

Programming Projects

1. Given a number α, find rational numbers p/q such that $|\alpha - p/q| \leq 1/q^2$.
2. Given a number α, find its spectrum sequence.
3. Find the first n Ulam numbers, where n is a positive integer.

1.2 Sums and Products

Because summations and products arise so often in the study of number theory, we now introduce notation for summations and products. The following notation represents the sum of the numbers $a_1, a_2, \ldots, a_n$:

$$\sum_{k=1}^{n} a_k = a_1 + a_2 + \cdots + a_n.$$

The letter k, the *index of summation,* is a "dummy variable" and can be replaced by any letter. For instance,

$$\sum_{k=1}^{n} a_k = \sum_{j=1}^{n} a_j = \sum_{i=1}^{n} a_i, \text{ and so forth.}$$

Example 1.13. We see that $\sum_{j=1}^{5} j = 1+2+3+4+5 = 15$, $\sum_{j=1}^{5} 2 = 2+2+2+2+2 = 10$, and $\sum_{j=1}^{5} 2^j = 2+2^2+2^3+2^4+2^5 = 62$.

We also note that, in summation notation, the index of summation may range between any two integers, as long as the lower limit does not exceed the upper limit. If m and n are integers such that $m \leq n$, then $\sum_{k=m}^{n} a_k = a_m + a_{m+1} + \cdots + a_n$. For instance, we have $\sum_{k=3}^{5} k^2 = 3^2 + 4^2 + 5^2 = 50$, $\sum_{k=0}^{2} 3^k = 3^0 + 3^1 + 3^2 = 13$, and $\sum_{k=-2}^{1} k^3 = (-2)^3 + (-1)^3 + 0^3 + 1^3 = -8$. ◄

We will often need to consider sums in which the index of summation ranges over all those integers that possess a particular property. We can use summation notation to specify the particular property or properties the index must have for a term with that index to be included in the sum. This use of notation is illustrated in the following example.

Example 1.14. We see that

$$\sum_{\substack{j \le 10 \\ j \in \{n^2 | n \in \mathbf{Z}\}}} 1/(j+1) = 1/1 + 1/2 + 1/5 + 1/10 = 9/5,$$

because the terms in the sum are all those for which j is an integer not exceeding 10 that is a perfect square. ◀

The following three properties for summations are often useful. We leave their proofs to the reader.

$$\sum_{j=m}^{n} ca_j = c \sum_{j=m}^{n} a_j \tag{1.1}$$

$$\sum_{j=m}^{n} (a_j + b_j) = \sum_{j=m}^{n} a_j + \sum_{j=m}^{n} b_j \tag{1.2}$$

$$\sum_{i=m}^{n} \sum_{j=p}^{q} a_i b_j = \left(\sum_{i=m}^{n} a_i \right) \left(\sum_{j=p}^{q} b_j \right) = \sum_{j=p}^{q} \sum_{i=m}^{n} a_i b_j \tag{1.3}$$

Next, we develop several useful summation formulas. We often need to evaluate sums of consecutive terms of a geometric series. The following example shows how a formula for such sums can be derived.

Example 1.15. To evaluate

$$S = \sum_{j=0}^{n} ar^j,$$

the sum of the first $n + 1$ terms of the geometric series $a, ar, \ldots, ar^k, \ldots$, we multiply both sides by r and manipulate the resulting sum to find:

$$\begin{aligned} rS &= r\sum_{j=0}^{n} ar^j \\ &= \sum_{j=0}^{n} ar^{j+1} \\ &= \sum_{k=1}^{n+1} ar^k && \text{(shifting the index of summation, taking } k = j+1) \\ &= \sum_{k=0}^{n} ar^k + (ar^{n+1} - a) && \text{(removing the term with } k = n+1 \\ &&& \quad \text{from the set and adding the term with } k = 0) \\ &= S + (ar^{n+1} - a). \end{aligned}$$

It follows that

$$rS - S = (ar^{n+1} - a).$$

Solving for S shows that when $r \neq 1$,

$$S = \frac{ar^{n+1} - a}{r - 1}.$$

Note that when $r = 1$, we have $\sum_{j=0}^{n} ar^j = \sum_{j=0}^{n} a = (n+1)a$. ◀

Example 1.16. Taking $a = 3$, $r = -5$, and $n = 6$ in the formula found in Example 1.15, we see that $\sum_{j=0}^{6} 3(-5)^j = \frac{3(-5)^7 - 3}{-5-1} = 39{,}063$. ◀

The following example shows that the sum of the first n consecutive powers of 2 is 1 less than the next power of 2.

Example 1.17. Let n be a positive integer. To find the sum

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

we use Example 1.15, with $a = 1$ and $r = 2$, to obtain

$$1 + 2 + 2^2 + \cdots + 2^n = \frac{2^{n+1} - 1}{2 - 1} = 2^{n+1} - 1.$$ ◀

A summation of the form $\sum_{j=1}^{n}(a_j - a_{j-1})$, where $a_0, a_1, a_2, \ldots, a_n$ is a sequence of numbers, is said to be *telescoping*. Telescoping sums are easily evaluated because

$$\begin{aligned} \sum_{j=1}^{n} a_j - a_{j-1} &= (a_1 - a_0) + (a_2 - a_1) + \cdots + (a_n - a_{n-1}) \\ &= a_n - a_0. \end{aligned}$$

The ancient Greeks were interested in sequences of numbers that can be represented by regular arrangements of equally spaced points. The following example illustrates one such sequence of numbers.

Example 1.18. The *triangular numbers* $t_1, t_2, t_3, \ldots, t_k, \ldots$ is the sequence where t_k is the number of dots in the triangular array of k rows with j dots in the jth row. ◄

Figure 1.2 illustrates that t_k counts the dots in successively larger regular triangles for $k = 1, 2, 3, 4$, and 5.

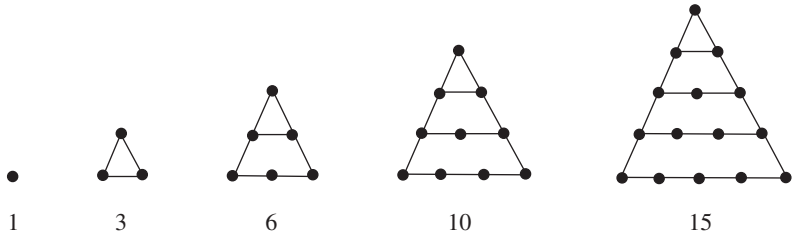


Figure 1.2 *The Triangular Numbers.*

Next, we will determine an explicit formula for the nth triangular number t_n.

Example 1.19. How can we find a formula for the nth triangular number? One approach is to use the identity $(k+1)^2 - k^2 = 2k + 1$. When we isolate the factor k, we find that $k = ((k+1)^2 - k^2)/2 - 1/2$. When we sum this expression for k over the values $k = 1, 2, \ldots, n$, we obtain

$$\begin{aligned}
t_n &= \sum_{k=1}^{n} k \\
&= \left(\sum_{k=1}^{n}((k+1)^2 - k^2)/2\right) - \sum_{k=1}^{n} 1/2 \quad (\textit{replacing } k \textit{ with } (((k+1)^2 - k^2)/2) - 1/2) \\
&= ((n+1)^2/2 - 1/2) - n/2 \qquad (\textit{simplifying a telescoping sum}) \\
&= (n^2 + 2n)/2 - n/2 \\
&= (n^2 + n)/2 \\
&= n(n+1)/2.
\end{aligned}$$

The second equality here follows by the formula for the sum of a telescoping series with $a_k = (k+1)^2 - k^2$. We conclude that the nth triangular number $t_n = n(n+1)/2$. (See Exercise 7 for another way to find t_n.) ◄

We also define a notation for products, analogous to that for summations. The product of the numbers $a_1, a_2, \ldots, a_n$ is denoted by

$$\prod_{j=1}^{n} a_j = a_1 a_2 \cdots a_n.$$

The letter j above is a "dummy variable," and can be replaced arbitrarily.

Example 1.20. To illustrate the notation for products, we have

$$\prod_{j=1}^{5} j = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 = 120,$$
$$\prod_{j=1}^{5} 2 = 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 = 2^5 = 32, \text{ and}$$
$$\prod_{j=1}^{5} 2^j = 2 \cdot 2^2 \cdot 2^3 \cdot 2^4 \cdot 2^5 = 2^{15}.$$

◀

The *factorial function* arises throughout number theory.

Definition. Let n be a positive integer. Then $n!$ (read as "n factorial") is the product of the integers $1, 2, \ldots, n$. We also specify that $0! = 1$. In terms of product notation, we have $n! = \prod_{j=1}^{n} j$.

Example 1.21. We have $1! = 1$, $4! = 1 \cdot 2 \cdot 3 \cdot 4 = 24$, and $12! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdot 8 \cdot 9 \cdot 10 \cdot 11 \cdot 12 = 479{,}001{,}600$. ◀

1.2 Exercises

1. Find each of the following sums.

a) $\sum_{j=1}^{5} j^2$ b) $\sum_{j=1}^{5} (-3)$ c) $\sum_{j=1}^{5} 1/(j+1)$

2. Find each of the following sums.

a) $\sum_{j=0}^{4} 3$ b) $\sum_{j=0}^{4} (j-3)$ c) $\sum_{j=0}^{4} (j+1)/(j+2)$

3. Find each of the following sums.

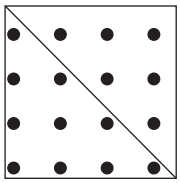
a) $\sum_{j=1}^{8} 2^j$ b) $\sum_{j=1}^{8} 5(-3)^j$ c) $\sum_{j=1}^{8} 3(-1/2)^j$

4. Find each of the following sums.

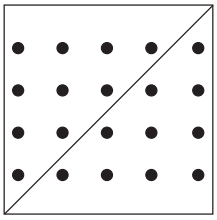
a) $\sum_{j=0}^{10} 8 \cdot 3^j$ b) $\sum_{j=0}^{10} (-2)^{j+1}$ c) $\sum_{j=0}^{10} (1/3)^j$

* **5.** Find and prove a formula for $\sum_{k=1}^{n} [\sqrt{k}]$ in terms of n and $[\sqrt{n}]$. (*Hint:* Use the formula $\sum_{k=1}^{t} k^2 = t(t+1)(2t+1)/6$.)

6. By putting together two triangular arrays, one with n rows and one with $n-1$ rows, to form a square (as illustrated for $n = 4$), show that $t_{n-1} + t_n = n^2$, where t_n is the nth triangular number.



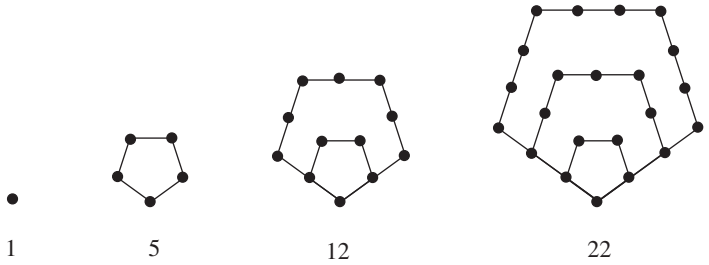
7. By putting together two triangular arrays, each with n rows, to form a rectangular array of dots of size n by $n+1$ (as illustrated for $n=4$), show that $2t_n = n(n+1)$. From this, conclude that $t_n = n(n+1)/2$.



8. Show that $3t_n + t_{n-1} = t_{2n}$, where t_n is the nth triangular number.

9. Show that $t_{n+1}^2 - t_n^2 = (n+1)^3$, where t_n is the nth triangular number.

The *pentagonal numbers* $p_1, p_2, p_3, \ldots, p_k, \ldots$, are the integers that count the number of dots in k nested pentagons, as shown in the following figure.



➢ **10.** Show that $p_1 = 1$ and $p_k = p_{k-1} + (3k-2)$ for $k \geq 2$. Conclude that $p_n = \sum_{k=1}^{n} (3k-2)$ and evaluate this sum to find a simple formula for p_n.

➢ **11.** Prove that the sum of the $(n-1)$st triangular number and the nth square number is the nth pentagonal number.

12. a) Define the hexagonal numbers h_n for $n = 1, 2, \ldots$ in a manner analogous to the definitions of triangular, square, and pentagonal numbers. (Recall that a hexagon is a six-sided polygon.)

b) Find a closed formula for hexagonal numbers.

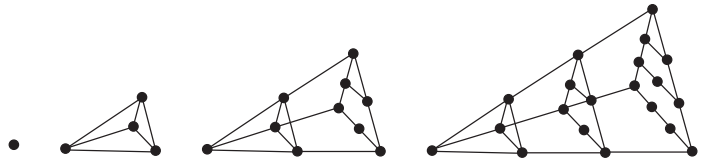
13. a) Define the heptagonal numbers in a manner analogous to the definitions of triangular, square, and pentagonal numbers. (Recall that a heptagon is a seven-sided polygon.)

b) Find a closed formula for heptagonal numbers.

14. Show that $h_n = t_{2n-1}$ for all positive integers n where h_n is the nth hexagonal number, defined in Exercise 12, and t_{2n-1} is the $(2n-1)$st triangular number.

15. Show that $p_n = t_{3n-1}/3$ where p_n is the nth pentagonal number and t_{3n-1} is the $(3n-1)$st triangular number.

The *tetrahedral numbers* $T_1, T_2, T_3, \ldots, T_k, \ldots$, are the integers that count the number of dots on the faces of k nested tetrahedra, as shown in the following figure.



16. Show that the nth tetrahedral number is the sum of the first n triangular numbers.

17. Find and prove a closed formula for the nth tetrahedral number.

18. Find $n!$ for n equal to each of the first ten positive integers.

19. List the integers $100!$, 100^{100}, 2^{100}, and $(50!)^2$ in order of increasing size. Justify your answer.

20. Express each of the following products in terms of $\prod_{i=1}^{n} a_i$, where k is a constant.

a) $\prod_{i=1}^{n} ka_i$ b) $\prod_{i=1}^{n} ia_i$ c) $\prod_{i=1}^{n} a_i^k$

21. Use the identity $\frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}$ to evaluate $\sum_{k=1}^{n} \frac{1}{k(k+1)}$.

22. Use the identity $\frac{1}{k^2-1} = \frac{1}{2}\left(\frac{1}{k-1} - \frac{1}{k+1}\right)$ to evaluate $\sum_{k=2}^{n} \frac{1}{k^2-1}$.

23. Find a formula for $\sum_{k=1}^{n} k^2$ using a technique analogous to that in Example 1.21 and the formula found there.

24. Find a formula for $\sum_{k=1}^{n} k^3$ using a technique analogous to that in Example 1.19, and the results of that example and Exercise 21.

25. Without multiplying all the terms, verify these equalities.

a) $10! = 6!\,7!$ b) $10! = 7!\,5!\,3!$ c) $16! = 14!\,5!\,2!$ d) $9! = 7!\,3!\,3!\,2!$

26. Let $a_1, a_2, \ldots, a_n$ be positive integers. Let $b = (a_1!\,a_2! \ldots a_n!) - 1$, and $c = a_1!\,a_2! \ldots a_n!$. Show that $c! = a_1!\,a_2! \cdots a_n!b!$.

27. Find all positive integers x, y, and z such that $x! + y! = z!$.

28. Find the values of the following products.

a) $\prod_{j=2}^{n}(1 - 1/j)$ b) $\prod_{j=2}^{n}(1 - 1/j^2)$

Computations and Explorations

1. What are the largest values of n for which $n!$ has fewer than 100 decimal digits, fewer than 1000 decimal digits, and fewer than 10,000 decimal digits?

2. Find as many triangular numbers that are perfect squares as you can. (We will study this question in the Exercises in Section 13.4.)

3. Find as many tetrahedral numbers that are perfect squares as you can.

Programming Projects

1. Given the terms of a sequence $a_1, a_2, \ldots, a_n$, compute $\sum_{j=1}^{n} a_j$ and $\prod_{j=1}^{n} a_j$.

2. Given the terms of a geometric progression, find the sum of its terms.

3. Given a positive integer n, find the nth triangular number, the nth perfect square, the nth pentagonal number, and the nth tetrahedral number.

1.3 Mathematical Induction

By examining the sums of the first n odd positive integers for small values of n, we can conjecture a formula for this sum. We have

$$\begin{aligned} 1 &= 1, \\ 1+3 &= 4, \\ 1+3+5 &= 9, \\ 1+3+5+7 &= 16, \\ 1+3+5+7+9 &= 25, \\ 1+3+5+7+9+11 &= 36. \end{aligned}$$

From these values, we conjecture that $\sum_{j=1}^{n}(2j-1) = 1+3+5+7+\cdots+2n-1 = n^2$ for every positive integer n.

How can we prove that this formula holds for all positive integers n?

The *principle of mathematical induction* is a valuable tool for proving results about the integers—such as the formula just conjectured for the sum of the first n odd positive integers. First, we will state this principle, and then we will show how it is used. Subsequently, we will use the well-ordering principle to show that mathematical induction is a valid proof technique. We will use the principle of mathematical induction, and the well-ordering property, many times in our study of number theory.

We must accomplish two things to prove by mathematical induction that a particular statement holds for every positive integer. Letting S be the set of positive integers for which we claim the statement to be true, we must show that 1 belongs to S; that is, that the statement is true for the integer 1. This is called the *basis step*.

Second, we must show, for each positive integer n, that $n+1$ belongs to S if n does; that is, that the statement is true for $n+1$ if it is true for n. This is called the *inductive step*. Once these two steps are completed, we can conclude by the principle of mathematical induction that the statement is true for all positive integers.

Theorem 1.5. ***The Principle of Mathematical Induction.*** A set of positive integers that contains the integer 1, and that has the property that, if it contains the integer k, then it also contains $k+1$, must be the set of all positive integers.

We illustrate the use of mathematical induction by several examples; first, we prove the conjecture made at the start of this section.

Example 1.22. We will use mathematical induction to show that

$$\sum_{j=1}^{n}(2j-1) = 1+3+\cdots+(2n-1) = n^2$$

for every positive integer n. (By the way, if our conjecture for the value of this sum was incorrect, mathematical induction would fail to produce a proof!)

We begin with the basis step, which follows because

$$\sum_{j=1}^{1}(2j-1) = 2\cdot 1 - 1 = 1 = 1^2.$$

For the inductive step, we assume the inductive hypothesis that the formula holds for n; that is, we assume that $\sum_{j=1}^{n}(2j-1) = n^2$. Using the inductive hypothesis, we have

$$\begin{aligned}
\sum_{j=1}^{n+1}(2j-1) &= \sum_{j=1}^{n}(2j-1) + (2(n+1)-1) && (\textit{splitting off the term with } j = n+1)\\
&= n^2 + 2(n+1) - 1 && (\textit{using the inductive hypothesis})\\
&= n^2 + 2n + 1\\
&= (n+1)^2.
\end{aligned}$$

Because both the basis and the inductive steps have been completed, we know that the result holds. ◀

Next, we prove an inequality via mathematical induction.

Example 1.23. We can show by mathematical induction that $n! \le n^n$ for every positive integer n. The basis step, namely, the case where $n = 1$, holds because $1! = 1 \le 1^1 = 1$. Now, assume that $n! \le n^n$; this is the inductive hypothesis. To complete the proof, we must show, under the assumption that the inductive hypothesis is true, that $(n+1)! \le (n+1)^{n+1}$. Using the inductive hypothesis, we have

The Origin of Mathematical Induction

The first known use of mathematical induction appears in the work of the sixteenth-century mathematician Francesco Maurolico (1494–1575). In his book *Arithmeticorum Libri Duo,* Maurolico presented various properties of the integers, together with proofs. He devised the method of mathematical induction so that he could complete some of the proofs. The first use of mathematical induction in his book was in the proof that the sum of the first n odd positive integers equals n^2.

$$\begin{aligned}(n+1)! &= (n+1)\cdot n! \\ &\le (n+1)n^n \\ &< (n+1)(n+1)^n \\ &\le (n+1)^{n+1}.\end{aligned}$$

This completes both the inductive step and the proof. ◄

We now show that the principle of mathematical induction follows from the well-ordering principle.

Proof. Let S be a set of positive integers containing the integer 1, and the integer $n+1$ whenever it contains n. Assume (for the sake of contradiction) that S is not the set of all positive integers. Therefore, there are some positive integers not contained in S. By the well-ordering property, because the set of positive integers not contained in S is nonempty, there is a least positive integer n that is not in S. Note that $n \neq 1$, because 1 is in S.

Now, because $n > 1$ (as there is no positive integer n with $n < 1$), the integer $n-1$ is a positive integer smaller than n, and hence must be in S. But because S contains $n-1$, it must also contain $(n-1)+1 = n$, which is a contradiction, as n is supposedly the smallest positive integer not in S. This shows that S must be the set of all positive integers. ■

A slight variant of the principle of mathematical induction is also sometimes useful in proofs.

Theorem 1.6. ***The Second Principle of Mathematical Induction.*** A set of positive integers that contains the integer 1, and that has the property that, for every positive integer n, if it contains all the positive integers $1, 2, \ldots, n$, then it also contains the integer $n+1$, must be the set of all positive integers.

The second principle of mathematical induction is sometimes called *strong induction* to distinguish it from the principle of mathematical induction, which is also called *weak induction*.

Before proving that the second principle of mathematical induction is valid, we will give an example to illustrate its use.

Example 1.24. We will show that any amount of postage more than one cent can be formed using just two-cent and three-cent stamps. For the basis step, note that postage of two cents can be formed using one two-cent stamp and postage of three cents can be formed using one three-cent stamp.

For the inductive step, assume that every amount of postage not exceeding n cents, $n \ge 3$, can be formed using two-cent and three-cent stamps. Then a postage amount of $n+1$ cents can be formed by taking stamps of $n-1$ cents together with a two-cent stamp. This completes the proof. ◄

We will now show that the second principle of mathematical induction is a valid technique.

Proof. Let T be a set of integers containing 1 and such that for every positive integer n, if it contains $1, 2, \ldots, n$, it also contains $n + 1$. Let S be the set of all positive integers n such that all the positive integers less than or equal to n are in T. Then 1 is in S, and by the hypotheses, we see that if n is in S, then $n + 1$ is in S. Hence, by the principle of mathematical induction, S must be the set of all positive integers, so clearly T is also the set of all positive integers, because S is a subset of T. ■

Recursive Definitions

The principle of mathematical induction provides a method for defining the values of functions at positive integers. Instead of explicitly specifying the value of the function at n, we give the value of the function at 1 and give a rule for finding, for each positive integer n, the value of the function at $n + 1$ from the value of the function at n.

Definition. We say that the function f is *defined recursively* if the value of f at 1 is specified and if for each positive integer n a rule is provided for determining $f(n + 1)$ from $f(n)$.

The principle of mathematical induction can be used to show that a function that is defined recursively is defined uniquely at each positive integer (see Exercise 25 at the end of this section). We illustrate how to define a function recursively with the following definition.

Example 1.25. We will recursively define the *factorial function* $f(n) = n!$. First, we specify that

$$f(1) = 1.$$

Then we give a rule for finding $f(n + 1)$ from $f(n)$ for each positive integer, namely,

$$f(n + 1) = (n + 1) \cdot f(n).$$

These two statements uniquely define $n!$ for the set of positive integers.

To find the value of $f(6) = 6!$ from the recursive definition, use the second property successively, as follows:

$$f(6) = 6 \cdot f(5) = 6 \cdot 5 \cdot f(4) = 6 \cdot 5 \cdot 4 \cdot f(3) = 6 \cdot 5 \cdot 4 \cdot 3 \cdot f(2) = 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot f(1).$$

Then use the first statement of the definition to replace $f(1)$ by its stated value 1, to conclude that

$$6! = 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 720.$$ ◀

The second principle of mathematical induction also serves as a basis for recursive definitions. We can define a function whose domain is the set of positive integers by specifying its value at 1 and giving a rule, for each positive integer n, for finding $f(n)$

from the values $f(j)$ for each integer j with $1 \leq j \leq n-1$. This will be the basis for the definition of the sequence of Fibonacci numbers discussed in Section 1.4.

1.3 Exercises

1. Use mathematical induction to prove that $n < 2^n$ whenever n is a positive integer.

2. Conjecture a formula for the sum of the first n even positive integers. Prove your result using mathematical induction.

3. Use mathematical induction to prove that $\sum_{k=1}^{n} \frac{1}{k^2} = \frac{1}{1^2} + \frac{1}{2^2} + \cdots + \frac{1}{n^2} \leq 2 - \frac{1}{n}$ whenever n is a positive integer.

4. Conjecture a formula for $\sum_{k=1}^{n} \frac{1}{k(k+1)} = \frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \cdots + \frac{1}{n(n+1)}$ from the value of this sum for small integers n. Prove that your conjecture is correct using mathematical induction. (Compare this to Exercise 17 in Section 1.2.)

5. Conjecture a formula for $\mathbf{A}^n$ where $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Prove your conjecture using mathematical induction.

6. Use mathematical induction to prove that $\sum_{j=1}^{n} j = 1 + 2 + 3 + \cdots + n = n(n+1)/2$ for every positive integer n. (Compare this to Example 1.19 in Section 1.2.)

7. Use mathematical induction to prove that $\sum_{j=1}^{n} j^2 = 1^2 + 2^2 + 3^2 + \cdots + n^2 = n(n+1)(2n+1)/6$ for every positive integer n.

8. Use mathematical induction to prove that $\sum_{j=1}^{n} j^3 = 1^3 + 2^3 + 3^3 + \cdots + n^3 = [n(n+1)/2]^2$ for every positive integer n.

9. Use mathematical induction to prove that $\sum_{j=1}^{n} j(j+1) = 1 \cdot 2 + 2 \cdot 3 + \cdots + n \cdot (n+1) = n(n+1)(n+2)/3$ for every positive integer n.

10. Use mathematical induction to prove that $\sum_{j=1}^{n} (-1)^{j-1} j^2 = 1^2 - 2^2 + 3^2 - \cdots + (-1)^{n-1} n^2 = (-1)^{n-1} n(n+1)/2$ for every positive integer n.

11. Find a formula for $\prod_{j=1}^{n} 2^j$.

12. Show that $\sum_{j=1}^{n} j \cdot j! = 1 \cdot 1! + 2 \cdot 2! + \cdots + n \cdot n! = (n+1)! - 1$ for every positive integer n.

13. Show that any amount of postage that is an integer number of cents greater than 11 cents can be formed using just 4-cent and 5-cent stamps.

14. Show that any amount of postage that is an integer number of cents greater than 53 cents can be formed using just 7-cent and 10-cent stamps.

Let H_n be the nth partial sum of the harmonic series, that is, $H_n = \sum_{j=1}^{n} 1/j$.

* **15.** Use mathematical induction to show that $H_{2^n} \geq 1 + n/2$.

* **16.** Use mathematical induction to show that $H_{2^n} \leq 1 + n$.

17. Show by mathematical induction that if n is a positive integer, then $(2n)! < 2^{2n}(n!)^2$.

18. Use mathematical induction to prove that $x - y$ is a factor of $x^n - y^n$, where x and y are variables.

➢ **19.** Use the principle of mathematical induction to show that a set of integers that contains the integer k, such that this set contains $n + 1$ whenever it contains n, contains the set of integers that are greater than or equal to k.

20. Use mathematical induction to prove that $2^n < n!$ for $n \geq 4$.

21. Use mathematical induction to prove that $n^2 < n!$ for $n \geq 4$.

22. Show by mathematical induction that if $h \geq -1$, then $1 + nh \leq (1 + h)^n$ for all nonnegative integers n.

23. A jigsaw puzzle is solved by putting its pieces together in the correct way. Show that exactly $n - 1$ moves are required to solve a jigsaw puzzle with n pieces, where a move consists of putting together two blocks of pieces, with a block consisting of one or more assembled pieces. (*Hint:* Use the second principle of mathematical induction.)

24. Explain what is wrong with the following proof by mathematical induction that all horses are the same color: Clearly all horses in any set of 1 horse are all the same color. This completes the basis step. Now assume that all horses in any set of n horses are the same color. Consider a set of $n + 1$ horses, labeled with the integers $1, 2, \ldots, n + 1$. By the induction hypothesis, horses $1, 2, \ldots, n$ are all the same color, as are horses $2, 3, \ldots, n, n + 1$. Because these two sets of horses have common members, namely, horses $2, 3, 4, \ldots, n$, all $n + 1$ horses must be the same color. This completes the induction argument.

25. Use the principle of mathematical induction to show that the value at each positive integer of a function defined recursively is uniquely determined.

26. What function $f(n)$ is defined recursively by $f(1) = 2$ and $f(n + 1) = 2f(n)$ for $n \geq 1$? Prove your answer using mathematical induction.

27. If g is defined recursively by $g(1) = 2$ and $g(n) = 2^{g(n-1)}$ for $n \geq 2$, what is $g(4)$?

28. Use the second principle of mathematical induction to show that if $f(1)$ is specified and a rule for finding $f(n + 1)$ from the values of f at the first n positive integers is given, then $f(n)$ is uniquely determined for every positive integer n.

29. We define a function recursively for all positive integers n by $f(1) = 1$, $f(2) = 5$, and for $n \geq 2$, $f(n + 1) = f(n) + 2f(n - 1)$. Show that $f(n) = 2^n + (-1)^n$, using the second principle of mathematical induction.

30. Show that $2^n > n^2$ whenever n is an integer greater than 4.

31. Suppose that $a_0 = 1, a_1 = 3, a_2 = 9$, and $a_n = a_{n-1} + a_{n-2} + a_{n-3}$ for $n \geq 3$. Show that $a_n \leq 3^n$ for every nonnegative integer n.

32. The tower of Hanoi was a popular puzzle of the late nineteenth century. The puzzle includes three pegs and eight rings of different sizes placed in order of size, with the largest on the bottom, on one of the pegs. The goal of the puzzle is to move all of the rings, one at a time, without ever placing a larger ring on top of a smaller ring, from the first peg to the second, using the third as an auxiliary peg.

a) Use mathematical induction to show that the minimum number of moves to transfer n rings from one peg to another, with the rules we have described, is $2^n - 1$.

b) An ancient legend tells of the monks in a tower with 64 gold rings and 3 diamond pegs. They started moving the rings, one move per second, when the world was created. When they finish transferring the rings to the second peg, the world will end. How long will the world last?

* **33.** The *arithmetic mean* and the *geometric mean* of the positive real numbers $a_1, a_2, \ldots, a_n$ are $A = (a_1 + a_2 + \cdots + a_n)/n$ and $G = (a_1 a_2 \cdots a_n)^{1/n}$, respectively. Use mathematical induction to prove that $A \geq G$ for every finite sequence of positive real numbers. When does equality hold?

34. Use mathematical induction to show that a $2^n \times 2^n$ chessboard with one square missing can be covered with L-shaped pieces, where each L-shaped piece covers three squares.

* **35.** A *unit fraction* is a fraction of the form $1/n$, where n is a positive integer. Because the ancient Egyptians represented fractions as sums of distinct unit fractions, such sums are called *Egyptian fractions*. Show that every rational number p/q, where p and q are integers with $0 < p < q$, can be written as a sum of distinct unit fractions, that is, as an Egyptian fraction. (*Hint:* Use strong induction on the numerator p to show that the greedy algorithm that adds the largest possible unit fraction at each stage always terminates. For example, running this algorithm shows that $5/7 = 1/2 + 1/5 + 1/70$.)

36. Using the algorithm in Exercise 35, write each of these numbers as Egyptian fractions.

a) $2/3$ b) $5/8$ c) $11/17$ d) $44/101$

Computations and Explorations

1. Complete the basis and inductive steps, using both numerical and symbolic computation, to prove that $\sum_{j=1}^{n} j = n(n+1)/2$ for all positive integers n.

2. Complete the basis and inductive steps, using both numerical and symbolic computation, to prove that $\sum_{j=1}^{n} j^2 = n(n+1)(2n+1)/6$ for all positive integers n.

3. Complete the basis and inductive steps, using both numerical and symbolic computation, to prove that $\sum_{j=1}^{n} j^3 = (n(n+1)/2)^2$ for all positive integers n.

4. Use the values $\sum_{j=1}^{n} j^4$ for $n = 1, 2, 3, 4, 5, 6$ to conjecture a formula for this sum that is a polynomial of degree 5 in n. Attempt to prove your conjecture via mathematical induction using numerical and symbolic computation.

5. Paul Erdős and E. Strauss have conjectured that the fraction $4/n$ can be written as the sum of three unit fractions, that is, $4/n = 1/x + 1/y + 1/z$, where x, y, and z are distinct positive integers for all integers n with $n > 1$. Find such representation for as many positive integers n as you can.

6. It is conjectured that the rational number p/q, where p and q are integers with $0 < p < q$ and q is odd, can be expressed as an Egyptian fraction that is the sum of unit fractions with odd denominators. Explore this conjecture using the greedy algorithm that successively adds the unit fraction with the least positive odd denominator q at each stage. (For example, $2/7 = 1/5 + 1/13 + 1/115 + 1/10{,}465$.)

Programming Projects

* **1.** List the moves in the tower of Hanoi puzzle (see Exercise 32). If you can, animate these moves.

** **2.** Cover a $2^n \times 2^n$ chessboard that is missing one square using L-shaped pieces (see Exercise 34).

3. Given a rational number p/q, express p/q as an Egyptian fraction using the algorithm described in Exercise 35.

1.4 The Fibonacci Numbers

In his book *Liber Abaci,* written in 1202, the mathematician *Fibonacci* posed a problem concerning the growth of the number of rabbits in a certain area. This problem can be phrased as follows: A young pair of rabbits, one of each sex, is placed on an island. Assuming that rabbits do not breed until they are two months old and after they are two months old, each pair of rabbits produces another pair each month, how many pairs are there after n months?

Let f_n be the number of pairs of rabbits after n months. We have $f_1 = 1$ because only the original pair is on the island after one month. As this pair does not breed during the second month, $f_2 = 1$. To find the number of pairs after n months, add the number on the island the previous month, f_{n-1}, to the number of newborn pairs, which equals f_{n-2}, because each newborn pair comes from a pair at least two months old. This leads to the following definition.

Definition. The *Fibonacci sequence* is defined recursively by $f_1 = 1$, $f_2 = 1$, and $f_n = f_{n-1} + f_{n-2}$ for $n \geq 3$. The terms of this sequence are called the *Fibonacci numbers*.

The mathematician Edouard Lucas named this sequence after Fibonacci in the nineteenth century when he established many of its properties. The answer to Fibonacci's question is that there are f_n rabbits on the island after n months.

Examining the initial terms of the Fibonacci sequence will be useful as we study their properties.

Example 1.26. We compute the first ten Fibonacci numbers as follows:

FIBONACCI (c. 1180–1228) (short for *filus Bonacci*, son of Bonacci), also known as Leonardo of Pisa, was born in the Italian commercial center of Pisa. Fibonacci was a merchant who traveled extensively throughout the Mideast, where he came into contact with mathematical works from the Arabic world. In his *Liber Abaci* Fibonacci introduced Arabic notation for numerals and their algorithms for arithmetic into the European world. It was in this book that his famous rabbit problem appeared. Fibonacci also wrote *Practica geometriae,* a treatise on geometry and trigonometry, and *Liber quadratorum,* a book on diophantine equations.

$$
\begin{aligned}
f_3 &= f_2 + f_1 = 1 + 1 = 2,\\
f_4 &= f_3 + f_2 = 2 + 1 = 3,\\
f_5 &= f_4 + f_3 = 3 + 2 = 5,\\
f_6 &= f_5 + f_4 = 5 + 3 = 8,\\
f_7 &= f_6 + f_5 = 8 + 5 = 13,\\
f_8 &= f_7 + f_6 = 13 + 8 = 21,\\
f_9 &= f_8 + f_7 = 21 + 13 = 34,\\
f_{10} &= f_9 + f_8 = 34 + 21 = 55.
\end{aligned}
$$

◀

We can define the value of $f_0 = 0$, so that $f_2 = f_1 + f_0$. We can also define f_n where n is a negative number so that the equality in the recursive definition is satisfied (see Exercise 37).

The Fibonacci numbers occur in an amazing variety of applications. For example, in botany the number of spirals in plants with a pattern known as phyllotaxis is always a Fibonacci number. They occur in the solution of a tremendous variety of counting problems, such as counting the number of bit strings with no two consecutive 1s (see [Ro07]).

The Fibonacci numbers also satisfy an extremely large number of identities. For example, we can easily find an identity for the sum of the first n consecutive Fibonacci numbers.

Example 1.27. The sum of the first n Fibonacci numbers for $3 \leq n \leq 8$ equals 1, 2, 4, 7, 12, 20, 33, and 54. Looking at these numbers, we see that they are all just 1 less than the Fibonacci number f_{n+2}. This leads us to the conjecture that

$$\sum_{k=1}^{n} f_k = f_{n+2} - 1.$$

Can we prove this identity for all positive integers n?

We will show, in two different ways, that this identity does hold for all integers n. We provide two different demonstrations, to show that there is often more than one way to prove that an identity is true.

First, we use the fact that $f_n = f_{n-1} + f_{n-2}$ for $n = 2, 3, \ldots$ to see that $f_k = f_{k+2} - f_{k+1}$ for $k = 1, 2, 3, \ldots$. This means that

$$\sum_{k=1}^{n} f_k = \sum_{k=1}^{n} (f_{k+2} - f_{k+1}).$$

We can easily evaluate this sum because it is telescoping. Using the formula for a telescoping sum found in Section 1.2, we have

$$\sum_{k=1}^{n} f_k = f_{n+2} - f_2 = f_{n+2} - 1.$$

This proves the result.

We can also prove this identity using mathematical induction. The basis step holds because $\sum_{k=1}^{1} f_k = 1$ and this equals $f_{1+2} - 1 = f_3 - 1 = 2 - 1 = 1$. The inductive hypothesis is

$$\sum_{k=1}^{n} f_k = f_{n+2} - 1.$$

We must show that, under this assumption,

$$\sum_{k=1}^{n+1} f_k = f_{n+3} - 1.$$

To prove this, note that by the inductive hypothesis we have

$$\begin{aligned}\sum_{k=1}^{n+1} f_k &= \left(\sum_{k=1}^{n} f_k\right) + f_{n+1} \\ &= (f_{n+2} - 1) + f_{n+1} \\ &= (f_{n+1} + f_{n+2}) - 1 \\ &= f_{n+3} - 1.\end{aligned}$$

◀

The exercise set at the end of this section asks you to prove many other identities of the Fibonacci numbers.

How Fast Do the Fibonacci Numbers Grow?

The following inequality, which shows that the Fibonacci numbers grow faster than a geometric series with common ratio $\alpha = (1 + \sqrt{5})/2$, will be used in Chapter 3.

Example 1.28. We can use the second principle of mathematical induction to prove that $f_n > \alpha^{n-2}$ for $n \geq 3$ where $\alpha = (1 + \sqrt{5})/2$. The basis step consists of verifying this inequality for $n = 3$ and $n = 4$. We have $\alpha < 2 = f_3$, so the theorem is true for $n = 3$. Because $\alpha^2 = (3 + \sqrt{5})/2 < 3 = f_4$, the theorem is true for $n = 4$.

The inductive hypothesis consists of assuming that $\alpha^{k-2} < f_k$ for all integers k with $k \leq n$. Because $\alpha = (1 + \sqrt{5})/2$ is a solution of $x^2 - x - 1 = 0$, we have $\alpha^2 = \alpha + 1$. Hence,

$$\alpha^{n-1} = \alpha^2 \cdot \alpha^{n-3} = (\alpha + 1) \cdot \alpha^{n-3} = \alpha^{n-2} + \alpha^{n-3}.$$

By the inductive hypothesis, we have the inequalities

$$\alpha^{n-2} < f_n, \quad \alpha^{n-3} < f_{n-1}.$$

By adding these two inequalities, we conclude that

$$\alpha^{n-1} < f_n + f_{n-1} = f_{n+1}.$$

This finishes the proof. ◀

We conclude this section with an explicit formula for the nth Fibonacci number. We will not provide a proof in the text, but Exercises 41 and 42 at the end of this section outline how this formula can be found using linear homogeneous recurrence relations and generating functions, respectively. Furthermore, Exercise 40 asks that you prove this identity by showing that the terms satisfy the same recursive definition as the Fibonacci numbers do, and Exercise 45 asks for a proof via mathematical induction. The advantage of the first two approaches is that they can be used to find the formula, while the second two approaches cannot.

Theorem 1.7. Let n be a positive integer and let $\alpha = \frac{1+\sqrt{5}}{2}$ and $\beta = \frac{1-\sqrt{5}}{2}$. Then the nth Fibonacci number f_n is given by

$$f_n = \frac{1}{\sqrt{5}}(\alpha^n - \beta^n).$$

We have presented a few important results involving the Fibonacci numbers. There is a vast literature concerning these numbers and their many applications to botany, computer science, geography, physics, and other areas (see [Va89]). There is even a scholarly journal, *The Fibonacci Quarterly,* devoted to their study.

1.4 Exercises

1. Find the following Fibonacci numbers.

a) f_{10} c) f_{15} e) f_{20}
b) f_{13} d) f_{18} f) f_{25}

2. Find each of the following Fibonacci numbers.

a) f_{12} c) f_{24} e) f_{32}
b) f_{16} d) f_{30} f) f_{36}

3. Prove that $f_{n+3} + f_n = 2f_{n+2}$ whenever n is a positive integer.

4. Prove that $f_{n+3} - f_n = 2f_{n+1}$ whenever n is a positive integer.

5. Prove that $f_{2n} = f_n^2 + 2f_{n-1}f_n$ whenever n is a positive integer. (Recall that $f_0 = 0$.)

6. Prove that $f_{n-2} + f_{n+2} = 3f_n$ whenever n is an integer with $n \geq 2$. (Recall that $f_0 = 0$.)

7. Find and prove a simple formula for the sum of the first n Fibonacci numbers with odd indices when n is a positive integer. That is, find a simple formula for $f_1 + f_3 + \cdots + f_{2n-1}$.

8. Find and prove a simple formula for the sum of the first n Fibonacci numbers with even indices when n is a positive integer. That is, find a simple formula for $f_2 + f_4 + \cdots + f_{2n}$.

9. Find and prove a simple formula for the expression $f_n - f_{n-1} + f_{n-2} - \cdots + (-1)^{n+1}f_1$ when n is a positive integer.

10. Prove that $f_{2n+1} = f_{n+1}^2 + f_n^2$ whenever n is a positive integer.

11. Prove that $f_{2n} = f_{n+1}^2 - f_{n-1}^2$ whenever n is a positive integer. (Recall that $f_0 = 0$.)

12. Prove that $f_n + f_{n-1} + f_{n-2} + 2f_{n-3} + 4f_{n-4} + 8f_{n-5} + \cdots + 2^{n-3} = 2^{n-1}$ whenever n is an integer with $n \geq 3$.

13. Prove that $\sum_{j=1}^{n} f_j^2 = f_1^2 + f_2^2 + \cdots + f_n^2 = f_n f_{n+1}$ for every positive integer n.

14. Prove that $f_{n+1}f_{n-1} - f_n^2 = (-1)^n$ for every positive integer n.

15. Prove that $f_{n+1}f_n - f_{n-1}f_{n-2} = f_{2n-1}$ for every positive integer n, $n > 2$.

16. Prove that $f_1f_2 + f_2f_3 + \cdots + f_{2n-1}f_{2n} = f_{2n}^2$ if n is a positive integer.

17. Prove that $f_{m+n} = f_mf_{n+1} + f_nf_{m-1}$ whenever m and n are positive integers.

The *Lucas numbers,* named after *François-Eduoard-Anatole Lucas* (see Chapter 7 for a biography), are defined recursively by

$$L_n = L_{n-1} + L_{n-2}, \quad n \geq 3,$$

with $L_1 = 1$ and $L_2 = 3$. They satisfy the same recurrence relation as the Fibonacci numbers, but the two initial values are different.

18. Find the first 12 Lucas numbers.

19. Find and prove a formula for the sum of the first n Lucas numbers when n is a positive integer.

20. Find and prove a formula for the sum of the first n Lucas numbers with odd indices when n is a positive integer.

21. Find and prove a formula for the sum of the first n Lucas numbers with even indices when n is a positive integer.

22. Prove that $L_n^2 - L_{n+1}L_{n-1} = 5(-1)^n$ when n is an integer with $n \geq 2$.

23. Prove that $L_1^2 + L_2^2 + \cdots + L_n^2 = L_nL_{n+1} - 2$ when n is an integer with $n \geq 1$.

24. Show that the nth Lucas number L_n is the sum of the $(n+1)$st and $(n-1)$st Fibonacci numbers, f_{n+1} and f_{n-1}, respectively.

25. Show that $f_{2n} = f_nL_n$ for all integers n with $n \geq 1$, where f_n is the nth Fibonacci number and L_n is the nth Lucas number.

26. Prove that $5f_{n+1} = L_n + L_{n+2}$ whenever n is a positive integer, f_n is the nth Fibonacci number, and L_n is the nth Lucas number.

* **27.** Prove that $L_{m+n} = f_{m+1}L_n + f_mL_{n-1}$ whenever m and n are positive integers with $n > 1$, f_n is the nth Fibonacci number, and L_n is the nth Lucas number.

28. Show that L_n, the nth Lucas number, is given by

$$L_n = \alpha^n + \beta^n,$$

where $\alpha = (1 + \sqrt{5})/2$ and $\beta = (1 - \sqrt{5})/2$.

The *Zeckendorf representation* of a positive integer is the unique expression of this integer as the sum of distinct Fibonacci numbers, where no two of these Fibonacci numbers are consecutive terms in the Fibonacci sequence and where the term $f_1 = 1$ is not used (but the term $f_2 = 1$ may be used).

29. Find the Zeckendorf representation of each of the integers 50, 85, 110, and 200.

* **30.** Show that every positive integer has a unique Zeckendorf representation.

31. Show that $f_n \leq \alpha^{n-1}$ for every integer n with $n \geq 2$, where $\alpha = (1 + \sqrt{5})/2$.

32. Show that

$$\binom{n}{0} + \binom{n-1}{1} + \binom{n-2}{2} + \cdots = f_{n+1},$$

where n is a nonnegative integer and f_{n+1} is the $(n+1)$st Fibonacci number. (See Appendix B for a review of binomial coefficients. Here, the sum ends with the term $\binom{1}{n-1}$.)

33. Prove that whenever n is a nonegative integer, $\sum_{j=1}^{n} \binom{n}{j} f_j = f_{2n}$, where f_j is the jth Fibonacci number.

34. Let $\mathbf{F} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$. Show that $\mathbf{F}^n = \begin{pmatrix} f_{n+1} & f_n \\ f_n & f_{n-1} \end{pmatrix}$ when $n \in \mathbf{Z}^+$.

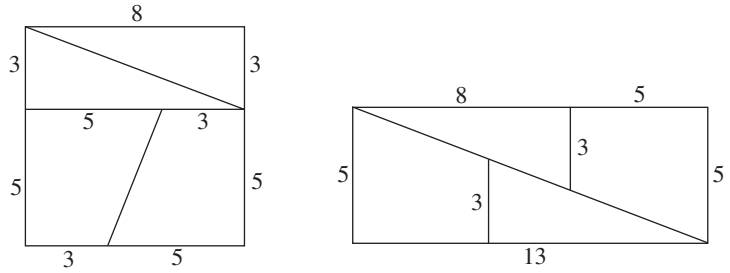
35. By taking determinants of both sides of the result of Exercise 34, prove the identity in Exercise 14.

36. Define the *generalized Fibonacci numbers* recursively by $g_1 = a$, $g_2 = b$, and $g_n = g_{n-1} + g_{n-2}$ for $n \geq 3$. Show that $g_n = af_{n-2} + bf_{n-1}$ for $n \geq 3$.

37. Give a recursive definition of the Fibonacci number f_n when n is a negative integer. Use your definition to find f_n for $n = -1, -2, -3, \ldots, -10$.

38. Use the results of Exercise 37 to formulate a conjecture that relates the values of f_{-n} and f_n when n is a positive integer. Prove this conjecture using mathematical induction.

39. What is wrong with the claim that an 8×8 square can be broken into pieces that can be reassembled to form a 5×13 rectangle as shown?



(*Hint:* Look at the identity in Exercise 14. Where is the extra square unit?)

40. Show that if $a_n = \frac{1}{\sqrt{5}}(\alpha^n - \beta^n)$, where $\alpha = (1+\sqrt{5})/2$ and $\beta = (1-\sqrt{5})/2$, then $a_n = a_{n-1} + a_{n-2}$ and $a_1 = a_2 = 1$. Conclude that $f_n = a_n$, where f_n is the nth Fibonacci number.

A *linear homogeneous recurrence relation of degree 2 with constant coefficients* is an equation of the form

$$a_n = c_1 a_{n-1} + c_2 a_{n-2},$$

where c_1 and c_2 are real numbers with $c_2 \neq 0$. It is not difficult to show (see [Ro07]) that if the equation $r^2 - c_1 r - c_2 = 0$ has two distinct roots r_1 and r_2, then the sequence $\{a_n\}$ is a solution of the linear homogeneous recurrence relation $a_n = c_1 a_{n-1} + c_2 a_{n-2}$ if and only if $a_n = C_1 r_1^n + C_2 r_2^n$ for $n = 0, 1, 2, \ldots$, where C_1 and C_2 are constants. The values of these constants can be found using the two initial terms of the sequence.

41. Find an explicit formula for f_n, proving Theorem 1.7, by solving the recurrence relation $f_n = f_{n-1} + f_{n-2}$ for $n = 2, 3, \ldots$ with initial conditions $f_0 = 0$ and $f_1 = 1$.

The *generating function* for the sequence $a_0, a_1, \ldots, a_k, \ldots$ is the infinite series

$$G(x) = \sum_{k=0}^{\infty} a_k x^k.$$

42. Use the generating function $G(x) = \sum_{k=0}^{\infty} f_k x^k$ where f_k is the kth Fibonacci number to find an explicit formula for f_k, proving Theorem 1.7. (*Hint:* Use the fact that $f_k = f_{k-1} + f_{k-2}$ for $k = 2, 3, \ldots$ to show that $G(x) - xG(x) - x^2G(x) = x$. Solve this to show that $G(x) = x/(1 - x - x^2)$ and then write $G(x)$ in terms of partial fractions, as is done in calculus.) (See [Ro07] for information on using generating functions.)

43. Find an explicit formula for the Lucas numbers using the technique of Exercise 41.

44. Find an explicit formula for the Lucas numbers using the technique of Exercise 42.

45. Use mathematical induction to prove Theorem 1.7.

Computations and Explorations

1. Find the Fibonacci numbers f_{100}, f_{200}, and f_{500}.

2. Find the Lucas numbers L_{100}, L_{200}, and L_{500}.

3. Examine as many Fibonacci numbers as possible to determine which are perfect squares. Formulate a conjecture based on your evidence.

4. Examine as many Fibonacci numbers as possible to determine which are triangular numbers. Formulate a conjecture based on your evidence.

5. Examine as many Fibonacci numbers as possible to determine which are perfect cubes. Formulate a conjecture based on your evidence.

6. Find the largest Fibonacci number less than 10,000, less than 100,000, and less than 1,000,000.

7. A surprising theorem states that the Fibonacci numbers are the positive values of the polynomial $2xy^4 + x^2y^3 - 2x^3y^2 - y^5 - x^4y + 2y$ as x and y range over all nonnegative integers. Verify this conjecture for the values of x and y where x and y are nonnegative integers with $x + y \leq 100$.

Programming Projects

1. Given a positive integer n, find the first n terms of the Fibonacci sequence.

2. Given a positive integer n, find the first n terms of the Lucas sequence.

3. Give a positive integer n, find its Zeckendorf representation (defined in the preamble to Exercise 29).

1.5 Divisibility

The concept of the divisibility of one integer by another is central in number theory.

Definition. If a and b are integers with $a \neq 0$, we say that *a divides b* if there is an integer c such that $b = ac$. If a divides b, we also say that a is a *divisor* or *factor* of b and that b is a *multiple* of a.

If a divides b we write $a \mid b$, and if a does not divide b we write $a \nmid b$. (Be careful not to confuse the notations $a \mid b$, which denotes that a divides b, and a/b, which is the quotient obtained when a is divided by b.)

Example 1.29. The following statements illustrate the concept of the divisibility of integers: $13 \mid 182$, $-5 \mid 30$, $17 \mid 289$, $6 \nmid 44$, $7 \nmid 50$, $-3 \mid 33$, and $17 \mid 0$. ◀

Example 1.30. The divisors of 6 are ± 1, ± 2, ± 3, ± 6. The divisors of 17 are ± 1, ± 17. The divisors of 100 are ± 1, ± 2, ± 4, ± 5, ± 10, ± 20, ± 25, ± 50, ± 100. ◀

In subsequent chapters, we will need some simple properties of divisibility, which we now state and prove.

Theorem 1.8. If a, b, and c are integers with $a \mid b$ and $b \mid c$, then $a \mid c$.

Proof. Because $a \mid b$ and $b \mid c$, there are integers e and f such that $ae = b$ and $bf = c$. Hence, $c = bf = (ae)f = a(ef)$, and we conclude that $a \mid c$. ■

Example 1.31. Because $11 \mid 66$ and $66 \mid 198$, Theorem 1.8 tells us that $11 \mid 198$. ◀

Theorem 1.9. If a, b, m, and n are integers, and if $c \mid a$ and $c \mid b$, then $c \mid (ma + nb)$.

Proof. Because $c \mid a$ and $c \mid b$, there are integers e and f such that $a = ce$ and $b = cf$. Hence, $ma + nb = mce + ncf = c(me + nf)$. Consequently, we see that $c \mid (ma + nb)$. ■

Example 1.32. As $3 \mid 21$ and $3 \mid 33$, Theorem 1.9 tells us that 3 divides

$$5 \cdot 21 - 3 \cdot 33 = 105 - 99 = 6.$$ ◀

The following theorem states an important fact about division.

Theorem 1.10. ***The Division Algorithm.*** If a and b are integers such that $b > 0$, then there are unique integers q and r such that $a = bq + r$ with $0 \leq r < b$. ■

In the equation given in the division algorithm, we call q the *quotient* and r the *remainder*. We also call a the *dividend* and b the *divisor*. (*Note:* We use the traditional name for this theorem even though the division algorithm is not actually an algorithm. We discuss algorithms in Section 2.2.)

We note that a is divisible by b if and only if the remainder in the division algorithm is 0. Before we prove the division algorithm, consider the following examples.

Example 1.33. If $a = 133$ and $b = 21$, then $q = 6$ and $r = 7$, because $133 = 21 \cdot 6 + 7$ and $0 \leq 7 < 21$. Likewise, if $a = -50$ and $b = 8$, then $q = -7$ and $r = 6$, because $-50 = 8(-7) + 6$ and $0 \leq 6 < 8$. ◀

We now prove the division algorithm using the well-ordering property.

Proof. Consider the set S of all integers of the form $a - bk$ where k is an integer, that is, $S = \{a - bk \mid k \in \mathbf{Z}\}$. Let T be the set of all nonnegative integers in S. T is nonempty, because $a - bk$ is positive whenever k is an integer with $k < a/b$.

By the well-ordering property, T has a least element $r = a - bq$. (These are the values for q and r specified in the theorem.) We know that $r \geq 0$ by construction, and it is easy to see that $r < b$. If $r \geq b$, then $r > r - b = a - bq - b = a - b(q + 1) \geq 0$, which contradicts the choice of $r = a - bq$ as the least nonnegative integer of the form $a - bk$. Hence, $0 \leq r < b$.

To show that these values for q and r are unique, assume that we have two equations $a = bq_1 + r_1$ and $a = bq_2 + r_2$, with $0 \leq r_1 < b$ and $0 \leq r_2 < b$. By subtracting the second of these equations from the first, we find that

$$0 = b(q_1 - q_2) + (r_1 - r_2).$$

Hence, we see that

$$r_2 - r_1 = b(q_1 - q_2).$$

This tells us that b divides $r_2 - r_1$. Because $0 \leq r_1 < b$ and $0 \leq r_2 < b$, we have $-b < r_2 - r_1 < b$. Hence, b can divide $r_2 - r_1$ only if $r_2 - r_1 = 0$ or, in other words, if $r_1 = r_2$. Because $bq_1 + r_1 = bq_2 + r_2$ and $r_1 = r_2$, we also see that $q_1 = q_2$. This shows that the quotient q and the remainder r are unique. ∎

We now use the greatest integer function (defined in Section 1.1) to give explicit formulas for the quotient and remainder in the division algorithm. Because the quotient q is the largest integer such that $bq \leq a$, and $r = a - bq$, it follows that

$$q = [a/b], \quad r = a - b[a/b]. \tag{1.4}$$

The following examples display the quotient and remainder of a division.

Example 1.34. Let $a = 1028$ and $b = 34$. Then $a = bq + r$ with $0 \leq r < b$, where $q = [1028/34] = 30$ and $r = 1028 - [1028/34] \cdot 34 = 1028 - 30 \cdot 34 = 8$. ◀

Example 1.35. Let $a = -380$ and $b = 75$. Then $a = bq + r$ with $0 \leq r < b$, where $q = [-380/75] = -6$ and $r = -380 - [-380/75] \cdot 75 = -380 - (-6)75 = 70$. ◀

We can use Equation (1.4) to prove a useful property of the greatest integer function.

Example 1.36. Show that if n is a positive integer, then $[x/n] = [[x]/n]$ whenever x is a real number. To prove this identity, suppose that $[x] = m$. By the division algorithm, we have integers q and r such that $m = nq + r$, where $0 \leq r \leq n - 1$. By Equation (1.4), we have $q = [[x]/n]$. Because $[x] \leq x < [x] + 1$, it follows that $x = [x] + \epsilon$, where $0 \leq \epsilon < 1$. We see that $[x/n] = [([x] + \epsilon)/n] = [(m + \epsilon)/n] = [((nq + r) + \epsilon)/n] = [q + (r + \epsilon)/n]$. Because $0 \leq \epsilon < 1$, we have $0 \leq r + \epsilon < (n - 1) + 1 = n$. It follows that $[x/n] = [q]$. ◀

Given a positive integer d, we can classify integers according to their remainders when divided by d. For example, with $d = 2$, we see from the division algorithm that

every integer when divided by 2 leaves a remainder of either 0 or 1. This leads to the following definition of some common terminology.

Definition. If the remainder when n is divided by 2 is 0, then $n = 2k$ for some integer k, and we say that n is *even,* whereas if the remainder when n is divided by 2 is 1, then $n = 2k + 1$ for some integer k, and we say that n is *odd.*

Similarly, when $d = 4$, we see from the division algorithm that when an integer n is divided by 4, the remainder is either 0, 1, 2, or 3. Hence, every integer is of the form $4k$, $4k + 1$, $4k + 2$, or $4k + 3$, where k is a positive integer.

We will pursue these matters further in Chapter 4.

Greatest Common Divisors

If a and b are integers, not both 0, then the set of common divisors of a and b is a finite set of integers, always containing the integers $+1$ and -1. We are interested in the largest integer among the common divisors of the two integers.

Definition. The *greatest common divisor* of two integers a and b, which are not both 0, is the largest integer that divides both a and b.

The greatest common divisor of a and b is written as (a, b). (Note that the notation $\gcd(a, b)$ is also used, especially outside of number theory. We will use the traditional notation (a, b) here, even though it is the same notation used for ordered pairs.) Note that $(0, n) = (n, 0) = n$ whenever n is a positive integer. Even though every positive integer divides 0, we define $(0, 0) = 0$. This is done to ensure that the results we prove about greatest common divisors hold in all cases.

Example 1.37. The common divisors of 24 and 84 are ± 1, ± 2, ± 3, ± 4, ± 6, and ± 12. Hence, $(24, 84) = 12$. Similarly, looking at sets of common divisors, we find that $(15, 81) = 3$, $(100, 5) = 5$, $(17, 25) = 1$, $(0, 44) = 44$, $(-6, -15) = 3$, and $(-17, 289) = 17$. ◀

We are particularly interested in pairs of integers sharing no common divisors greater than 1. Such pairs of integers are called *relatively prime.*

Definition. The integers a and b, with $a \neq 0$ and $b \neq 0$, are *relatively prime* if a and b have greatest common divisor $(a, b) = 1$.

Example 1.38. Because $(25, 42) = 1$, 25 and 42 are relatively prime. ◀

We will study greatest common divisors at length in Chapter 4. In that chapter, we will give an algorithm for computing greatest common divisors. We will also prove many important results about them that lead to key theorems in number theory.

1.5 Exercises

1. Show that $3 \mid 99$, $5 \mid 145$, $7 \mid 343$, and $888 \mid 0$.

2. Show that 1001 is divisible by 7, by 11, and by 13.

3. Decide which of the following integers are divisible by 7.

a) 0 c) 1717 e) $-285{,}714$
b) 707 d) 123,321 f) $-430{,}597$

4. Decide which of the following integers are divisible by 22.

a) 0 c) 1716 e) $-32{,}516$
b) 444 d) 192,544 f) $-195{,}518$

5. Find the quotient and remainder in the division algorithm, with divisor 17 and dividend

a) 100. b) 289. c) -44. d) -100.

6. Find all positive integers that divide each of these integers.

a) 12 b) 22 c) 37 d) 41

7. Find all positive integers that divide each of these integers.

a) 13 b) 21 c) 36 d) 44

8. Find these greatest common divisors by finding all positive integers that divide each integer in the pair and selecting the largest that divides both.

a) (8, 12) b) (7, 9) c) (15, 25) d) (16, 27)

9. Find these greatest common divisors by finding all positive integers that divide each integer in the pair and selecting the largest that divides both.

a) (11, 22) b) (36, 42) c) (21, 22) d) (16, 64)

10. Find all positive integers less than 10 that are relatively prime to it.

11. Find all positive integers less than 11 that are relatively prime to it.

12. Find all pairs of positive integers not exceeding 10 that are relatively prime.

13. Find all pairs of positive integers between 10 and 20, inclusive, that are relatively prime.

14. What can you conclude if a and b are nonzero integers such that $a \mid b$ and $b \mid a$?

15. Show that if a, b, c, and d are integers with a and c nonzero, such that $a \mid b$ and $c \mid d$, then $ac \mid bd$.

16. Are there integers a , b, and c such that $a \mid bc$, but $a \nmid b$ and $a \nmid c$?

17. Show that if a, b, and $c \neq 0$ are integers, then $a \mid b$ if and only if $ac \mid bc$.

18. Show that if a and b are positive integers and $a \mid b$, then $a \leq b$.

19. Show that if a and b are integers such that $a \mid b$, then $a^k \mid b^k$ for every positive integer k.

20. Show that the sum of two even or of two odd integers is even, whereas the sum of an odd and an even integer is odd.

21. Show that the product of two odd integers is odd, whereas the product of two integers is even if either of the integers is even.

22. Show that if a and b are odd positive integers and $b \nmid a$, then there are integers s and t such that $a = bs + t$, where t is odd and $|\, t \,| < b$.

23. When the integer a is divided by the integer b, where $b > 0$, the division algorithm gives a quotient of q and a remainder of r. Show that if $b \nmid a$, when $-a$ is divided by b, the division algorithm gives a quotient of $-(q+1)$ and a remainder of $b - r$, whereas if $b \mid a$, the quotient is $-q$ and the remainder is 0.

24. Show that if a, b, and c are integers with $b > 0$ and $c > 0$, such that when a is divided by b the quotient is q and the remainder is r, and when q is divided by c the quotient is t and the remainder is s, then when a is divided by bc, the quotient is t and the remainder is $bs + r$.

25. a) Extend the division algorithm by allowing negative divisors. In particular, show that whenever a and $b \neq 0$ are integers, there are unique integers q and r such that $a = bq + r$, where $0 \leq r < |\, b \,|$.

b) Find the remainder when 17 is divided by -7.

➢ **26.** Show that if a and b are positive integers, then there are unique integers q and r such that $a = bq + r$, where $-b/2 < r \leq b/2$. This result is called the *modified division algorithm.*

27. Show that if m and $n > 0$ are integers, then

$$\left\lfloor \frac{m+1}{n} \right\rfloor = \begin{cases} \left\lfloor \frac{m}{n} \right\rfloor & \text{if } m \neq kn - 1 \text{ for some integer } k; \\ \left\lfloor \frac{m}{n} \right\rfloor + 1 & \text{if } m = kn - 1 \text{ for some integer } k. \end{cases}$$

28. Show that the integer n is even if and only if $n - 2\lfloor n/2 \rfloor = 0$.

29. Show that the number of positive integers less than or equal to x, where x is a positive real number, that are divisible by the positive integer d equals $\lfloor x/d \rfloor$.

30. Find the number of positive integers not exceeding 1000 that are divisible by 5, by 25, by 125, and by 625.

31. How many integers between 100 and 1000 are divisible by 7? by 49?

32. Find the number of positive integers not exceeding 1000 that are not divisible by 3 or 5.

33. Find the number of positive integers not exceeding 1000 that are not divisible by 3, 5, or 7.

34. Find the number of positive integers not exceeding 1000 that are divisible by 3 but not by 4.

35. In early 2010, to mail a first-class letter in the United States of America it cost 44 cents for the first ounce and 17 cents for each additional ounce or fraction thereof. Find a formula involving the greatest integer function for the cost of mailing a letter in early 2010. Could it possibly have cost \$1.81 or \$2.65 to mail a first-class letter in the United States of America in early 2010?

36. Show that if a is an integer, then 3 divides $a^3 - a$.

37. Show that the product of two integers of the form $4k + 1$ is again of this form, whereas the product of two integers of the form $4k + 3$ is of the form $4k + 1$.

38. Show that the square of every odd integer is of the form $8k + 1$.

39. Show that the fourth power of every odd integer is of the form $16k + 1$.

40. Show that the product of two integers of the form $6k + 5$ is of the form $6k + 1$.

41. Show that the product of any three consecutive integers is divisible by 6.

42. Use mathematical induction to show that $n^5 - n$ is divisible by 5 for every positive integer n.

43. Use mathematical induction to show that the sum of the cubes of three consecutive integers is divisible by 9.

In Exercises 44–48, let f_n denote the nth Fibonacci number.

44. Show that f_n is even if and only if n is divisible by 3.

45. Show that f_n is divisible by 3 if and only if n is divisible by 4.

46. Show that f_n is divisible by 4 if and only if n is divisible by 6.

47. Show that $f_n = 5f_{n-4} + 3f_{n-5}$ whenever n is a positive integer with $n > 5$. Use this result to show that f_n is divisible by 5 whenever n is divisible by 5.

* **48.** Show that $f_{n+m} = f_m f_{n+1} + f_{m-1} f_n$ whenever m and n are positive integers with $m > 1$. Use this result to show that $f_n \mid f_m$ when m and n are positive integers with $n \mid m$.

Let n be a positive integer. We define

$$T(n) = \begin{cases} n/2 & \text{if } n \text{ is even;} \\ (3n+1)/2 & \text{if } n \text{ is odd.} \end{cases}$$

We then form the sequence obtained by iterating T: n, $T(n)$, $T(T(n))$, $T(T(T(n)))$, ... For instance, starting with $n = 7$, we have 7, 11, 17, 26, 13, 20, 10, 5, 8, 4, 2, 1, 2, 1, 2, 1, ... A well-known conjecture, sometimes called the *Collatz conjecture,* asserts that the sequence obtained by iterating T always reaches the integer 1 no matter which positive integer n begins the sequence.

49. Find the sequence obtained by iterating T starting with $n = 39$.

50. Show that the sequence obtained by iterating T starting with $n = (2^{2k} - 1)/3$, where k is a positive integer greater than 1, always reaches the integer 1.

51. Show that the Collatz conjecture is true if it can be shown that for every positive integer n with $n \geq 2$ there is a term in the sequence obtained by iterating T that is less than n.

52. Verify that there is a term in the sequence obtained by iterating T, starting with the positive integer n, that is less than n for all positive integers n with $2 \leq n \leq 100$. (*Hint:* Begin by considering sets of positive integers for which it is easy to show that this is true.)

* **53.** Show that $[(2 + \sqrt{3})^n]$ is odd whenever n is a nonnegative integer.

* **54.** Determine the number of positive integers n such that $[a/2] + [a/3] + [a/5] = a$, where, as usual, $[x]$ is the greatest integer function.

55. Prove the divison algorithm using the second principle of mathematical induction.

Computations and Explorations

1. Find the quotient and remainder when 111,111,111,111 is divided by 987,654,321.

2. Verify the Collatz conjecture described in the preamble to Exercise 49 for all integers n not exceeding 10,000.

3. Using numerical evidence, what sort of conjectures can you make concerning the number of iterations needed before the sequence of iterations $T(n)$ reaches 1, where n is a given positive integer?

4. Using numerical evidence, make conjectures about the divisibility of Fibonacci numbers by 7, by 8, by 9, by 11, and by 13.

Programming Projects

1. Decide whether an integer is divisible by a given integer.
2. Find the quotient and remainder in the division algorithm.
3. Find the quotient, remainder, and sign in the modified division algorithm given in Exercise 26.
4. Compute the terms of the sequence n, $T(n)$, $T(T(n))$, $T(T(T(n)))$, . . . for a given positive integer n, as defined in the preamble to Exercise 49.

2 Integer Representations and Operations

The way in which integers are represented has a major impact on how easily people and computers can do arithmetic with these integers. The purpose of this chapter is to explain how integers are represented using base b expansions, and how basic arithmetic operations can be carried out using these expansions. In particular, we will show that when b is a positive integer, every positive integer has a unique base b expansion. For example, when b is 10, we have the decimal expansion of an integer; when b is 2, we have the binary expansion of this integer; and when b is 16, we have the hexadecimal expansion. We will describe a procedure for finding the base b expansion of an integer, and describe the basic algorithms used to carry out integer arithmetic with base b expansions. Finally, after introducing big-O notation, we will analyze the computational complexity of these basic operations in terms of big-O estimates of the number of bit operations that they use.

2.1 Representations of Integers

In daily life, we use decimal notation to represent integers. We write out numbers using digits to represent powers of ten. For instance, when we write out the integer 37,465, we mean

$$3 \cdot 10^4 + 7 \cdot 10^3 + 4 \cdot 10^2 + 6 \cdot 10 + 5.$$

Decimal notation is an example of a *positional number system,* in which the position a digit occupies in a representation determines the quantity it represents. Throughout ancient and modern history, many other notations for integers have been used. For example, Babylonian mathematicians who lived more than 3000 years ago expressed integers using sixty as a base. The Romans employed Roman numerals, which are used even today to represent years. The ancient Mayans used a positional notation with twenty as a base. Many other systems of integer notation have been invented and used over time.

There is no special reason for using ten as the base in a fixed positional number system, other than that we have ten fingers. As we will see, any positive integer greater than 1 can be used as a base. With the invention and proliferation of computers, bases other than ten have become increasingly important. In particular, base 2, base 8, and base 16 representations of integers are used extensively by computers for various purposes.

In this section, we will demonstrate that no matter which positive integer b is chosen as a base, every positive integer can be expressed uniquely in base b notation. In Section

2.2, we will show how these expansions can be used to do arithmetic with integers. (See the exercise set at the end of this section to learn about one's and two's complement notations, which are used by computers to represent both positive and negative integers.)

For more information about the fascinating history of positional number systems, the reader is referred to [Or88] or [Kn97], where extensive surveys and numerous references may be found.

We now show that every positive integer greater than 1 may be used as a base.

Theorem 2.1. Let b be a positive integer with $b > 1$. Then every positive integer n can be written uniquely in the form

$$n = a_k b^k + a_{k-1} b^{k-1} + \cdots + a_1 b + a_0,$$

where k is a nonnegative integer, a_j is an integer with $0 \le a_j \le b - 1$ for $j = 0, 1, \ldots, k$, and the initial coefficient $a_k \neq 0$.

Proof. We obtain an expression of the desired type by successively applying the division algorithm in the following way. We first divide n by b to obtain

$$n = bq_0 + a_0, \quad 0 \le a_0 \le b - 1.$$

If $q_0 \neq 0$, we continue by dividing q_0 by b to find that

$$q_0 = bq_1 + a_1, \quad 0 \le a_1 \le b - 1.$$

We continue this process to obtain

$$\begin{aligned}
q_1 &= bq_2 + a_2, \quad 0 \le a_2 \le b - 1, \\
q_2 &= bq_3 + a_3, \quad 0 \le a_3 \le b - 1, \\
&\vdots \\
q_{k-2} &= bq_{k-1} + a_{k-1}, \quad 0 \le a_{k-1} \le b - 1, \\
q_{k-1} &= b \cdot 0 + a_k, \quad 0 \le a_k \le b - 1.
\end{aligned}$$

The last step of the process occurs when a quotient of 0 is obtained. To see that we must reach such a step, first note that the sequence of quotients satisfies

$$n > q_0 > q_1 > q_2 > \cdots \ge 0.$$

Because the sequence $q_0, q_1, q_2, \ldots$ is a decreasing sequence of nonnegative integers that continues as long as its terms are positive, there are at most q_0 terms in this sequence, and the last term equals 0.

From the first equation above, we find that

$$n = bq_0 + a_0.$$

We next replace q_0 using the second equation, to obtain

$$n = b(bq_1 + a_1) + a_0 = b^2 q_1 + a_1 b + a_0.$$

Successively substituting for $q_1, q_2, \ldots, q_{k-1}$, we have

$$\begin{aligned} n &= b^3 q_2 + a_2 b^2 + a_1 b + a_0, \\ &\vdots \\ &= b^{k-1} q_{k-2} + a_{k-2} b^{k-2} + \cdots + a_1 b + a_0, \\ &= b^k q_{k-1} + a_{k-1} b^{k-1} + \cdots + a_1 b + a_0 \\ &= a_k b^k + a_{k-1} b^{k-1} + \cdots + a_1 b + a_0, \end{aligned}$$

where $0 \le a_j \le b-1$ for $j = 0, 1, \ldots, k$ and $a_k \neq 0$, given that $a_k = q_{k-1}$ is the last nonzero quotient. Consequently, we have found an expansion of the desired type.

To see that the expansion is unique, assume that we have two such expansions equal to n, that is,

$$\begin{aligned} n &= a_k b^k + a_{k-1} b^{k-1} + \cdots + a_1 b + a_0 \\ &= c_k b^k + c_{k-1} b^{k-1} + \cdots + c_1 b + c_0, \end{aligned}$$

where $0 \le a_k < b$ and $0 \le c_k < b$ (and where, if necessary, we have added initial terms with zero coefficients to one of the expansions to have the number of terms agree). Subtracting one expansion from the other, we have

$$(a_k - c_k) b^k + (a_{k-1} - c_{k-1}) b^{k-1} + \cdots + (a_1 - c_1) b + (a_0 - c_0) = 0.$$

If the two expansions are different, there is a smallest integer j, $0 \le j \le k$, such that $a_j \neq c_j$. Hence,

$$b^j \big((a_k - c_k) b^{k-j} + \cdots + (a_{j+1} - c_{j+1}) b + (a_j - c_j) \big) = 0,$$

so that

$$(a_k - c_k) b^{k-j} + \cdots + (a_{j+1} - c_{j+1}) b + (a_j - c_j) = 0.$$

Solving for $a_j - c_j$, we obtain

$$\begin{aligned} a_j - c_j &= (c_k - a_k) b^{k-j} + \cdots + (c_{j+1} - a_{j+1}) b \\ &= b\big((c_k - a_k) b^{k-j-1} + \cdots + (c_{j+1} - a_{j+1}) \big). \end{aligned}$$

Hence, we see that

$$b \mid (a_j - c_j).$$

But because $0 \le a_j < b$ and $0 \le c_j < b$, we know that $-b < a_j - c_j < b$. Consequently, $b \mid (a_j - c_j)$ implies that $a_j = c_j$. This contradicts the assumption that the two expansions are different. We conclude that our base b expansion of n is unique. ∎

For $b = 2$, we see by Theorem 2.1 that the following corollary holds.

Corollary 2.1.1. Every positive integer may be represented as the sum of distinct powers of 2. ∎

Proof. Let n be a positive integer. From Theorem 2.1 with $b = 2$, we know that $n = a_k 2^k + a_{k-1}2^{k-1} + \cdots + a_1 2 + a_0$, where each a_j is either 0 or 1. Hence, every positive integer is the sum of distinct powers of 2. ∎

In the expansions described in Theorem 2.1, b is called the *base* or *radix* of the expansion. We call base 10 notation, our conventional way of writing integers, *decimal* notation. Base 2 expansions are called *binary* expansions, base 8 expansions are called *octal* expansions, and base 16 expansions are called *hexadecimal*, or *hex* for short. The coefficients a_j are called the *digits* of the expansion. Binary digits are called *bits* (*b*inary dig*its*) in computer terminology.

To distinguish representations of integers with different bases, we use a special notation. We write $(a_k a_{k-1} \ldots a_1 a_0)_b$ to represent the number $a_k b^k + a_{k-1}b^{k-1} + \cdots + a_1 b + a_0$.

Example 2.1. To illustrate base b notation, note that $(236)_7 = 2 \cdot 7^2 + 3 \cdot 7 + 6 = 125$ and $(10010011)_2 = 1 \cdot 2^7 + 1 \cdot 2^4 + 1 \cdot 2^1 + 1 = 147$. ◀

The proof of Theorem 2.1 provides a method of finding the base b expansion $(a_k a_{k-1} \ldots a_1 a_0)_b$ of any positive integer n. Specifically, to find the base b expansion of n, we first divide n by b. The remainder is the digit a_0. Then, we divide the quotient $[n/b] = q_0$ by b. The remainder is the digit a_1. We continue this process, successively dividing the quotient obtained by b, to obtain the digits in the base b expansion of n. The process stops once a quotient of 0 is obtained. In other words, to find the base b expansion of n, we perform the division algorithm repeatedly, replacing the dividend each time with the quotient, and stop when we come to a quotient that is 0. We then read up the list of remainders to find the base b expansion. We illustrate this procedure in Example 2.2.

Example 2.2. To find the base 2 expansion of 1864, we use the division algorithm successively:

$$\begin{aligned}
1864 &= 2 \cdot 932 + 0,\\
932 &= 2 \cdot 466 + 0,\\
466 &= 2 \cdot 233 + 0,\\
233 &= 2 \cdot 116 + 1,\\
116 &= 2 \cdot 58 + 0,\\
58 &= 2 \cdot 29 + 0,\\
29 &= 2 \cdot 14 + 1,\\
14 &= 2 \cdot 7 + 0,\\
7 &= 2 \cdot 3 + 1,\\
3 &= 2 \cdot 1 + 1,\\
1 &= 2 \cdot 0 + 1.
\end{aligned}$$

To obtain the base 2 expansion of 1864, we simply take the remainders of these divisions. This shows that $(1864)_{10} = (11101001000)_2$. ◀

Computers represent numbers internally by using a series of "switches" that may be either "on" or "off." (This may be done electrically or mechanically, or by other means.) Hence, we have two possible states for each switch. We can use "on" to represent the digit 1 and "off" to represent the digit 0; this is why computers use binary expansions to represent integers internally.

Computers use base 8 or base 16 for display purposes. In base 16 (hexadecimal) notation there are 16 digits, usually denoted by 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F. The letters A, B, C, D, E, and F are used to represent the digits that correspond to 10, 11, 12, 13, 14, and 15 (written in decimal notation). The following example demonstrates the conversion from hexadecimal to decimal notation.

Example 2.3. To convert $(\text{A35B0F})_{16}$ from hexadecimal to decimal notation, we write

$$\begin{aligned}(\text{A35B0F})_{16} &= 10 \cdot 16^5 + 3 \cdot 16^4 + 5 \cdot 16^3 + 11 \cdot 16^2 + 0 \cdot 16 + 15 \\ &= (10705679)_{10}.\end{aligned}$$

◀

A simple conversion is possible between binary and hexadecimal notation. We can write each hex digit as a block of four binary digits according to the correspondences given in Table 2.1.

Example 2.4. An example of conversion from hex to binary is $(\text{2FB3})_{16} = (10111110110011)_2$. Each hex digit is converted to a block of four binary digits (the initial zeros in the initial block $(0010)_2$ corresponding to the digit $(2)_{16}$ are omitted).

To convert from binary to hex, consider $(11110111101001)_2$. We break this into blocks of four, starting from the right. The blocks are, from right to left, 1001, 1110, 1101, and 0011 (with two initial zeros added). Translating each block to hex, we obtain $(\text{3DE9})_{16}$. ◀

Hex Digit	Binary Digits	Hex Digit	Binary Digits
0	0000	8	1000
1	0001	9	1001
2	0010	A	1010
3	0011	B	1011
4	0100	C	1100
5	0101	D	1101
6	0110	E	1110
7	0111	F	1111

Table 2.1 *Conversion from hex digits to blocks of binary digits.*

We note that a conversion between two different bases is as easy as binary–hex conversion whenever one of the bases is a power of the other.

2.1 Exercises

1. Convert $(1999)_{10}$ from decimal to base 7 notation. Convert $(6105)_7$ from base 7 to decimal notation.
2. Convert $(89156)_{10}$ from decimal to base 8 notation. Convert $(706113)_8$ from base 8 to decimal notation.
3. Convert $(10101111)_2$ from binary to decimal notation and $(999)_{10}$ from decimal to binary notation.
4. Convert $(101001000)_2$ from binary to decimal notation and $(1984)_{10}$ from decimal to binary notation.
5. Convert $(100011110101)_2$ and $(11101001110)_2$ from binary to hexadecimal.
6. Convert $(\text{ABCDEF})_{16}$, $(\text{DEFACED})_{16}$, and $(\text{9A0B})_{16}$ from hexadecimal to binary.
7. Explain why we really are using base 1000 notation when we break large decimal integers into blocks of three digits, separated by commas.
8. Show that if b is a negative integer less than -1, then every nonzero integer n can be uniquely written in the form
$$n = a_k b^k + a_{k-1} b^{k-1} + \cdots + a_1 b + a_0,$$
where $a_k \neq 0$ and $0 \leq a_j < |\, b \,|$ for $j = 0, 1, 2, \ldots, k$. We write $n = (a_k a_{k-1} \ldots a_1 a_0)_b$, just as we do for positive bases.
9. Find the decimal representation of $(101001)_{-2}$ and $(12012)_{-3}$.
10. Find the base -2 representations of the decimal numbers -7, -17, and 61.
11. Show that any weight not exceeding $2^k - 1$ may be measured using weights of $1, 2, 2^2, \ldots, 2^{k-1}$, when all the weights are placed in one pan.
12. Show that every nonzero integer can be uniquely represented in the form
$$e_k 3^k + e_{k-1} 3^{k-1} + \cdots + e_1 3 + e_0,$$
where $e_j = -1$, 0, or 1 for $j = 0, 1, 2, \ldots, k$ and $e_k \neq 0$. This expansion is called a *balanced ternary expansion.*
13. Use Exercise 12 to show that any weight not exceeding $(3^k - 1)/2$ may be measured using weights of $1, 3, 3^2, \ldots, 3^{k-1}$, when the weights may be placed in either pan.
14. Explain how to convert from base 3 to base 9 notation, and from base 9 to base 3 notation.
15. Explain how to convert from base r to base r^n notation, and from base r^n to base r notation, when $r > 1$ and n are positive integers.
16. Show that if $n = (a_k a_{k-1} \ldots a_1 a_0)_b$, then the quotient and remainder when n is divided by b^j are $q = (a_k a_{k-1} \ldots a_j)_b$ and $r = (a_{j-1} \ldots a_1 a_0)_b$, respectively.
17. If the base b expansion of n is $n = (a_k a_{k-1} \ldots a_1 a_0)_b$, what is the base b expansion of $b^m n$?

One's complement representations of integers are used to simplify computer arithmetic. To represent positive and negative integers with absolute value less than 2^n, a total of $n + 1$ bits is used.

The leftmost bit is used to represent the sign. A 0 in this position is used for positive integers, and a 1 in this position is used for negative integers.

For positive integers, the remaining bits are identical to the binary expansion of the integer. For negative integers, the remaining bits are obtained by first finding the binary expansion of the absolute value of the integer, and then taking the complement of each of these bits, where the complement of a 1 is a 0 and the complement of a 0 is a 1.

18. Find the one's complement representations, using bit strings of length six, of the following integers.

a) 22 b) 31 c) -7 d) -19

19. What integer does each of the following one's complement representations of length five represent?

a) 11001 b) 01101 c) 10001 d) 11111

20. How is the one's complement representation of $-m$ obtained from the one's complement of m, when bit strings of length n are used?

21. Show that if m is an integer with one's complement representation $a_{n-1}a_{n-2}\ldots a_1a_0$, then $m = -a_{n-1}(2^{n-1} - 1) + \sum_{i=0}^{n-2} a_i 2^i$.

Two's complement representations of integers also are used to simplify computer arithmetic (in fact, they are used much more commonly than one's complement representations). To represent an integer x with $-2^{n-1} \leq x \leq 2^{n-1} - 1$, n bits are used.

The leftmost bit represents the sign, with a 0 used for positive integers and a 1 for negative integers.

For a positive integer, the remaining $n - 1$ bits are identical to the binary expansion of the integer. For a negative integer, the remaining bits are the bits of the binary expansion of $2^{n-1} - |x|$.

22. Find the two's complement representations, using bit strings of length six, of the integers in Exercise 18.

23. What integers do the representations in Exercise 19 represent if each is the two's complement representation of an integer?

24. Show that if m is an integer with two's complement representation $a_{n-1}a_{n-2}\ldots a_1a_0$, then $m = -a_{n-1} \cdot 2^{n-1} + \sum_{i=0}^{n-2} a_i 2^i$.

25. How is the two's complement representation of $-m$ obtained from the two's complement representation of m, when bit strings of length n are used?

26. How can the two's complement representation of an integer be found from its one's complement representation?

27. Sometimes integers are encoded by using four-digit binary expansions to represent each decimal digit. This produces the *binary coded decimal* form of the integer. For instance, 791 is encoded in this way by 011110010001. How many bits are required to represent a number with n decimal digits using this type of encoding?

A *Cantor expansion* of a positive integer n is a sum

$$n = a_m m! + a_{m-1}(m-1)! + \cdots + a_2 2! + a_1 1!,$$

where each a_j is an integer with $0 \le a_j \le j$ and $a_m \neq 0$.

28. Find Cantor expansions of 14, 56, and 384.

∗ **29.** Show that every positive integer has a unique Cantor expansion. (*Hint:* For each positive integer n there is a positive integer m such that $m! \le n < (m+1)!$. For a_m, take the quotient from the division algorithm when n is divided by $m!$, then iterate.)

The Chinese game of *nim* is played as follows. There are several piles of matches, each containing an arbitrary number of matches at the start of the game. To make a move, a player removes one or more matches from one of the piles. The players take turns, and the player who removes the last match wins the game.

A *winning position* is an arrangement of matches in piles such that if a player can move to this position, then (no matter what the second player does) the first player can continue to play in a way that will win the game. An example is the position where there are two piles, each containing one match; this is a winning position, because the second player must remove a match, leaving the first player the opportunity to win by removing the last match.

30. Show that the position in nim where there are two piles, each with two matches, is a winning position.

31. For each arrangement of matches into piles, write the number of matches in each pile in binary notation, and then line up the digits of these numbers into columns (adding initial zeros where necessary). Show that a position is a winning one if and only if the number of 1s in each column is even. (For example: Three piles of 3, 4, and 7 give

$$\begin{array}{ccc} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{array}$$

where each column has exactly two 1s.) (*Hint:* Show that any move from a winning position produces a nonwinning one. Show that there is a move from any nonwinning position to a winning one.)

Let a be an integer with a four-digit decimal expansion, where not all digits are the same. Let a' be the integer with a decimal expansion obtained by writing the digits of a in descending order, and let a'' be the integer with a decimal expansion obtained by writing the digits of a in ascending order. Define $T(a) = a' - a''$. For instance, $T(7318) = 8731 - 1378 = 7353$.

∗ **32.** Show that the only integer with a four-digit decimal expansion (where not all digits are the same) such that $T(a) = a$ is $a = 6174$. The integer 6174 is called *Kaprekar's constant,* after the Indian mathematician *D. R. Kaprekar*, because it is the only integer with this property.

∗∗ **33.** a) Show that if a is a positive integer with a four-digit decimal expansion where not all digits are the same, then the sequence a, $T(a)$, $T(T(a))$, $T(T(T(a)))$, ..., obtained by iterating T, eventually reaches the integer 6174.

b) Determine the maximum number of steps required for the sequence defined in part (a) to reach 6174.

Let b be a positive integer and let a be an integer with a four-digit base b expansion, with not all digits the same. Define $T_b(a) = a' - a''$, where a' is the integer with base b expansion obtained

by writing the base b digits of a in descending order, and a'' is the integer with base b expansion obtained by writing the base b digits of a in ascending order.

** **34.** Let $b = 5$. Find the unique integer a_0 with a four-digit base 5 expansion such that $T_5(a_0) = a_0$. Show that this integer a_0 is a Kaprekar constant for base 5; in other words, that $a, T(a), T(T(a)), T(T(T(a))), \ldots$ eventually reaches a_0, whenever a is an integer with a four-digit base 5 expansion where not all digits are the same.

* **35.** Show that no Kaprekar constant exists for four-digit numbers to the base 6.

* **36.** Determine whether there is a Kaprekar constant for three-digit integers to the base 10. Prove that your answer is correct.

37. A sequence a_j, $j = 1, 2, \ldots$ is called a *Sidon sequence*, after the Hungarian mathematician Simon Sidon, if all the pairwise sums $a_i + a_j$ where $i \leq j$ are different. Use Theorem 2.1 to show that the sequence a_j, $j = 1, 2, \ldots$ is a Sidon sequence when $a_j = 2^j$.

Computations and Explorations

1. Find the binary, octal, and hexadecimal expansions of each of the following integers.
a) 9876543210 b) 1111111111 c) 10000000001

2. Find the decimal expansion of each of the following integers.
a) $(1010101010101)_2$ b) $(765432101234567)_8$ c) $(\text{ABBAFADACABA})_{16}$

3. Evaluate each of the following sums, expressing your answer in the same base used to represent the summands.
a) $(11011011011011011)_2 + (10010010010010010010)_2$
b) $(12345670123456)_8 + (765432107654321)_8$
c) $(123456789\text{ABCD})_{16} + (\text{BABACACADADA})_{16}$

4. Find the Cantor expansions of the integers 100,000, 10,000,000, and 1,000,000,000. (See the preamble to Exercise 28 for the definition of Cantor expansions.)

5. Verify the result described in Exercise 33 for several different four-digit integers, in which not all digits are the same.

6. Use numerical evidence to make conjectures about the behavior of the sequence $a, T(a), T(T(a)), \ldots$ where a is a five-digit integer in base 10 notation in which not all digits are the same, and $T(a)$ is defined as in the preamble to Exercise 32.

D. R. KAPREKAR (1905–1986) was born in Dahanu, India, and was interested in numbers even as a small child. He received his secondary school education in Thana and studied at Ferguson College in Poona. Kaprekar attended the University of Bombay, receiving his bachelor's degree in 1929. From 1930 until his retirement in 1962, he worked as a schoolteacher in Devlali, India. Kaprekar discovered many interesting properties in recreational number theory. He published extensively, writing about such topics as recurring decimals, magic squares, and integers with special properties.

7. Explore the behavior for different bases b of the sequence a, $T(a)$, $T(T(a))$, ... where a is a three-digit integer in base b notation. What conjectures can you make? Repeat your exploration using four-digit and then five-digit integers in base b notation.

Programming Projects

1. Find the binary expansion of an integer from the decimal expansion of this integer, and vice versa.

2. Convert from base b_1 notation to base b_2 notation, where b_1 and b_2 are arbitrary positive integers greater than 1.

3. Convert from binary notation to hexadecimal notation, and vice versa.

4. Find the base (-2) notation of an integer from its decimal notation (see Exercise 8).

5. Find the balanced ternary expansion of an integer from its decimal expansion (see Exercise 12).

6. Find the Cantor expansion of an integer from its decimal expansion (see the preamble to Exercise 28).

7. Play a winning strategy in the game of nim (see the preamble to Exercise 30).

* **8.** Investigate the sequence a, $T(a)$, $T(T(a))$, $T(T(T(a)))$, ... (defined in the preamble to Exercise 32), where a is a positive integer, to discover the minimum number of iterations required to reach 6174.

2.2 Computer Operations with Integers

Before computers were invented, mathematicians did computations either by hand or by using mechanical devices. Either way, they were only able to work with integers of rather limited size. Many number theoretic problems, such as factoring and primality testing, require computations with integers of as many as 100 or even 200 digits. In this section, we will study some of the basic algorithms for doing computer arithmetic. In the following section, we will study the number of basic computer operations required to carry out these algorithms.

We have mentioned that computers internally represent numbers using bits, or binary digits. Computers have a built-in limit on the size of integers that can be used in machine arithmetic. This upper limit is called the *word size,* which we denote by w. The word size is usually a power of 2, such as 2^{32} for Pentium machines or 2^{35}, although sometimes the word size is a power of 10.

To do arithmetic with integers larger than the word size, it is necessary to devote more than one word to each integer. To store an integer $n > w$, we express n in base w notation, and for each digit of this expansion we use one computer word. For instance, if the word size is 2^{35}, using ten computer words we can store integers as large as $2^{350} - 1$, because integers less than 2^{350} have no more than ten digits in their base 2^{35} expansions. Also note that to find the base 2^{35} expansion of an integer, we need only group together blocks of 35 bits.

The first step in discussing computer arithmetic with large integers is to describe how the basic arithmetic operations are methodically performed.

We will describe the classical methods for performing the basic arithmetic operations with integers in base r notation, where $r > 1$ is an integer. These methods are examples of *algorithms*.

Definition. An *algorithm* is a finite set of precise instructions for performing a computation or for solving a problem.

We will describe algorithms for performing addition, subtraction, and multiplication of two n-digit integers $a = (a_{n-1}a_{n-2} \ldots a_1a_0)_r$ and $b = (b_{n-1}b_{n-2} \ldots b_1b_0)_r$, where initial digits of zero are added if necessary to make both expansions the same length. The algorithms described are used for both binary arithmetic with integers less than the word size of a computer, and *multiple precision* arithmetic with integers larger than the word size w, using w as the base.

Addition When we add a and b, we obtain the sum

$$a + b = \sum_{j=0}^{n-1} a_j r^j + \sum_{j=0}^{n-1} b_j r^j = \sum_{j=0}^{n-1} (a_j + b_j) r^j.$$

To find the base r expansion of $a + b$, first note that by the division algorithm, there are integers C_0 and s_0 such that

$$a_0 + b_0 = C_0 r + s_0, \quad 0 \le s_0 < r.$$

Because a_0 and b_0 are positive integers not exceeding r, we know that $0 \le a_0 + b_0 \le 2r - 2$, so that $C_0 = 0$ or 1; here, C_0 is the *carry* to the next place. Next, we find that there are integers C_1 and s_1 such that

$$a_1 + b_1 + C_0 = C_1 r + s_1, \quad 0 \le s_1 < r.$$

Because $0 \le a_1 + b_1 + C_0 \le 2r - 1$, we know that $C_1 = 0$ or 1. Proceeding inductively, we find integers C_i and s_i for $1 \le i \le n - 1$ by

$$a_i + b_i + C_{i-1} = C_i r + s_i, \quad 0 \le s_i < r,$$

with $C_i = 0$ or 1. Finally, we let $s_n = C_{n-1}$, because the sum of two integers with n digits has $n + 1$ digits when there is a carry in the nth place. We conclude that the base r expansion for the sum is $a + b = (s_n s_{n-1} \ldots s_1 s_0)_r$.

When performing base r addition by hand, we can use the same familiar technique as is used in decimal addition.

Example 2.5. To add $(1101)_2$ and $(1001)_2$, we write

$$\begin{array}{cccccc} {\scriptstyle 1} & & & {\scriptstyle 1} & & \\ & 1 & 1 & 0 & 1 & \\ + & 1 & 0 & 0 & 1 & \\ \hline 1 & 0 & 1 & 1 & 0 & \end{array}$$

where we have indicated carries by 1s in italics written above the appropriate column. We found the binary digits of the sum by noting that $1 + 1 = 1 \cdot 2 + 0$, $0 + 0 + 1 = 0 \cdot 2 + 1$, $1 + 0 + 0 = 0 \cdot 2 + 1$, and $1 + 1 + 0 = 1 \cdot 2 + 0$. ◀

Subtraction Assume that $a > b$. Consider

$$a - b = \sum_{j=0}^{n-1} a_j r^j - \sum_{j=0}^{n-1} b_j r^j = \sum_{j=0}^{n-1} (a_j - b_j) r^j.$$

Note that by the division algorithm, there are integers B_0 and d_0 such that

$$a_0 - b_0 = B_0 r + d_0, \quad 0 \leq d_0 < r,$$

and because a_0 and b_0 are positive integers less than r, we have

$$-(r - 1) \leq a_0 - b_0 \leq r - 1.$$

When $a_0 - b_0 \geq 0$, we have $B_0 = 0$. Otherwise, when $a_0 - b_0 < 0$, we have $B_0 = -1$; B_0 is the *borrow* from the next place of the base r expansion of a. We use the division algorithm again to find integers B_1 and d_1 such that

$$a_1 - b_1 + B_0 = B_1 r + d_1, \quad 0 \leq d_1 < r.$$

From this equation, we see that the borrow $B_1 = 0$ as long as $a_1 - b_1 + B_0 \geq 0$, and that $B_1 = -1$ otherwise, because $-r \leq a_1 - b_1 + B_0 \leq r - 1$. We proceed inductively to find integers B_i and d_i, such that

$$a_i - b_i + B_{i-1} = B_i r + d_i, \quad 0 \leq d_i < r$$

with $B_i = 0$ or -1, for $1 \leq i \leq n - 1$. We see that $B_{n-1} = 0$, because $a > b$. We can conclude that

$$a - b = (d_{n-1} d_{n-2} \ldots d_1 d_0)_r.$$

Where the Word "Algorithm" Comes From

"Algorithm" is a corruption of the original term "algorism," which originally comes from the name of the author of the ninth-century book *Kitab al-jabr w'al-muqabala* (*Rules of Restoration and Reduction*), *Abu Ja'far Mohammed ibn Mûsâ al-Khwârizmî* (see his biography included on the next page). The word "algorism" originally referred only to the rules of performing arithmetic using Hindu-Arabic numerals, but evolved into "algorithm" by the eighteenth century. With growing interest in computing machines, the concept of an algorithm became more general, to include all definite procedures for solving problems, not just the procedures for performing arithmetic with integers expressed in Arabic notation.

When performing base r subtraction by hand, we use the familiar technique used in decimal subtraction.

Example 2.6. To subtract $(10110)_2$ from $(11011)_2$, we have

$$\begin{array}{cccccc} & & \textit{-1} & & & \\ 1 & 1 & 0 & 1 & 1 & \\ -1 & 0 & 1 & 1 & 0 & \\ \hline & & 1 & 0 & 1 & \end{array}$$

where the -1 in italics above a column indicates a borrow. We found the binary digits of the difference by noting that $1-0=0\cdot 2+1$, $1-1+0=0\cdot 2+0$, $0-1+0=-1\cdot 2+1$, $1-0-1=0\cdot 2+0$, and $1-1+0=0\cdot 2+0$. ◀

Multiplication Before discussing multiplication, we describe *shifting*. To multiply $(a_{n-1}\ldots a_1a_0)_r$ by r^m, we need only shift the expansion left m places, appending the expansion with m zero digits.

Example 2.7. To multiply $(101101)_2$ by 2^5, we shift the digits to the left five places and append the expansion with five zeros, obtaining $(10110100000)_2$. ◀

We first discuss the multiplication of an n-place integer by a one-digit integer. To multiply $(a_{n-1}\ldots a_1a_0)_r$ by $(b)_r$, we first note that

$$a_0b=q_0r+p_0,\quad 0\le p_0<r,$$

and $0\le q_0\le r-2$, because $0\le a_0b\le (r-1)^2$. Next, we have

$$a_1b+q_0=q_1r+p_1,\quad 0\le p_1<r,$$

and $0\le q_1\le r-1$. In general, we have

$$a_ib+q_{i-1}=q_ir+p_i,\quad 0\le p_i<r,$$

ABU JA'FAR MOHAMMED IBN MÛSÂ AL-KHWÂRIZMÎ (c. 780–c. 850), an astronomer and mathematician, was a member of the House of Wisdom, an academy of scientists in Baghdad. The name al-Khwârizmî means "from the town of Kowarzizm," now known as Khiva in modern Uzbekistan. Al-Khwârizmî was the author of books on mathematics, astronomy, and geography. People in the West first learned about algebra from his works; the word "algebra" comes from *al-jabr,* part of the title of his book *Kitab al-jabr w'al muqabala,* which was translated into Latin and widely used as a text. Another book describes procedures for arithmetic operations using Hindu-Arabic numerals.

and $0 \le q_i \le r-1$. Furthermore, we have $p_n = q_{n-1}$. This yields $(a_{n-1} \dots a_1 a_0)_r (b)_r = (p_n p_{n-1} \cdots p_1 p_0)_r$.

To perform a multiplication of two n-place integers, we write

$$ab = a\left(\sum_{j=0}^{n-1} b_j r^j\right) = \sum_{j=0}^{n-1} (ab_j) r^j.$$

For each j, we first multiply a by the digit b_j, then shift j places to the left, and finally add all of the n integers we have obtained to find the product.

When multiplying two integers with base r expansions, we use the familiar method of multiplying decimal integers by hand.

Example 2.8. To multiply $(1101)_2$ and $(1110)_2$, we write

$$\begin{array}{cccccccc}
 & & & & 1 & 1 & 0 & 1 \\
 & & & \times & 1 & 1 & 1 & 0 \\
\hline
 & & & & 0 & 0 & 0 & 0 \\
 & & & 1 & 1 & 0 & 1 & \\
 & & 1 & 1 & 0 & 1 & & \\
 & 1 & 1 & 0 & 1 & & & \\
\hline
1 & 0 & 1 & 1 & 0 & 1 & 1 & 0
\end{array}$$

Note that we first multiplied $(1101)_2$ by each digit of $(1110)_2$, shifting each time by the appropriate number of places, and then we added the appropriate integers to find our product. ◀

Division We wish to find the quotient q in the division algorithm

$$a = bq + R, \quad 0 \le R < b.$$

If the base r expansion of q is $q = (q_{n-1} q_{n-2} \dots q_1 q_0)_r$, then we have

$$a = b\left(\sum_{j=0}^{n-1} q_j r^j\right) + R, \quad 0 \le R < b.$$

To determine the first digit q_{n-1} of q, notice that

$$a - bq_{n-1} r^{n-1} = b\left(\sum_{j=0}^{n-2} q_j r^j\right) + R.$$

The right-hand side of this equation is not only positive, but also less than br^{n-1}, because $\sum_{j=0}^{n-2} q_j r^j \le \sum_{j=0}^{n-2} (r-1) r^j = \sum_{j=1}^{n-1} r^j - \sum_{j=0}^{n-2} r^j = r^{n-1} - 1$. Therefore, we know that

$$0 \le a - bq_{n-1} r^{n-1} < br^{n-1}.$$

This tells us that

$$q_{n-1} = \left\lfloor \frac{a}{br^{n-1}} \right\rfloor .$$

We can obtain q_{n-1} by successively subtracting br^{n-1} from a until we obtain a negative result; q_{n-1} is then one less than the number of subtractions.

To find the other digits of q, we define the sequence of *partial remainders* R_i by

$$R_0 = a$$

and

$$R_i = R_{i-1} - bq_{n-i}r^{n-i}$$

for $i = 1, 2, \ldots, n$. By mathematical induction, we show that

$$R_i = \left(\sum_{j=0}^{n-i-1} q_j r^j \right) b + R. \tag{2.1}$$

For $i = 0$, this is clearly correct, because $R_0 = a = qb + R$. Now assume that

$$R_k = \left(\sum_{j=0}^{n-k-1} q_j r^j \right) b + R.$$

Then

$$\begin{aligned} R_{k+1} &= R_k - bq_{n-k-1}r^{n-k-1} \\ &= \left(\sum_{j=0}^{n-k-1} q_j r^j \right) b + R - bq_{n-k-1}r^{n-k-1} \\ &= \left(\sum_{j=0}^{n-(k+1)-1} q_j r^j \right) b + R, \end{aligned}$$

establishing (2.1).

By (2.1), we see that $0 \le R_i < r^{n-i}b$, for $i = 1, 2, \ldots, n$, because $\sum_{j=0}^{n-i-1} q_j r^j \le r_{n-i} - 1$. Consequently, because $R_i = R_{i-1} - bq_{n-i}r^{n-i}$ and $0 \le R_i < r^{n-1}b$, we see that the digit q_{n-i} is given by $[R_{i-1}/(br^{n-i})]$ and can be obtained by successively subtracting br^{n-i} from R_{i-1} until a negative result is obtained, and then q_{n-i} is one less than the number of subtractions. This is how we find the digits of q.

Example 2.9. To divide $(11101)_2$ by $(111)_2$, we let $q = (q_2q_1q_0)_2$. We subtract $2^2(111)_2 = (11100)_2$ once from $(11101)_2$ to obtain $(1)_2$, and once more to obtain a negative result, so that $q_2 = 1$. Now, $R_1 = (11101)_2 - (11100)_2 = (1)_2$. We find that $q_1 = 0$, because $R_1 - 2(111)_2$ is less than zero, and likewise $q_0 = 0$. Hence, the quotient of the division is $(100)_2$ and the remainder is $(1)_2$. ◀

2.2 EXERCISES

1. Add $(101111011)_2$ and $(1100111011)_2$.
2. Add $(10001000111101)_2$ and $(11111101011111)_2$.
3. Subtract $(11010111)_2$ from $(1111000011)_2$.
4. Subtract $(101110101)_2$ from $(1101101100)_2$.
5. Multiply $(11101)_2$ and $(110001)_2$.
6. Multiply $(1110111)_2$ and $(10011011)_2$.
7. Find the quotient and remainder when $(110011111)_2$ is divided by $(1101)_2$.
8. Find the quotient and remainder when $(110100111)_2$ is divided by $(11101)_2$.
9. Add $(1234321)_5$ and $(2030104)_5$.
10. Subtract $(434421)_5$ from $(4434201)_5$.
11. Multiply $(1234)_5$ and $(3002)_5$.
12. Find the quotient and remainder when $(14321)_5$ is divided by $(334)_5$.
13. Add $(\text{ABAB})_{16}$ and $(\text{BABA})_{16}$.
14. Subtract $(\text{CAFE})_{16}$ from $(\text{FEED})_{16}$.
15. Multiply $(\text{FACE})_{16}$ and $(\text{BAD})_{16}$.
16. Find the quotient and remainder when $(\text{BEADED})_{16}$ is divided by $(\text{ABBA})_{16}$.
17. Explain how to add, subtract, and multiply the integers 18235187 and 22135674 on a computer with word size 1000.
18. Write algorithms for the basic operations with integers in base (-2) notation (see Exercise 8 of Section 2.1).
19. How is the one's complement representation of the sum of two integers obtained from the one's complement representations of those integers?
20. How is the one's complement representation of the difference of two integers obtained from the one's complement representations of those integers?
21. Give an algorithm for adding and an algorithm for subtracting Cantor expansions (see the preamble to Exercise 28 of Section 2.1).
22. A *dozen* equals 12, and a *gross* equals 12^2. Using base 12, or *duodecimal* arithmetic, answer the following questions.
    a) If 3 gross, 7 dozen, and 4 eggs are removed from a total of 11 gross and 3 dozen eggs, how many eggs are left?
    b) If 5 truckloads of 2 gross, 3 dozen, and 7 eggs each are delivered to the supermarket, how many eggs are delivered?
    c) If 11 gross, 10 dozen, and 6 eggs are divided in 3 groups of equal size, how many eggs are in each group?
23. A well-known rule used to find the square of an integer with decimal expansion $(a_n a_{n-1} \ldots a_1 a_0)_{10}$ and final digit $a_0 = 5$ is to find the decimal expansion of the product $(a_n a_{n-1} \ldots a_1)_{10}$ $[(a_n a_{n-1} \ldots a_1)_{10} + 1]$, and append this with the digits $(25)_{10}$. For instance, we see that the decimal expansion of $(165)^2$ begins with $16 \cdot 17 = 272$, so that $(165)^2 = 27{,}225$. Show that this rule is valid.

24. In this exercise, we generalize the rule given in Exercise 23 to find the squares of integers with final base *2B* digit B, where B is a positive integer. Show that the base *2B* expansion of the integer $(a_n a_{n-1} \ldots a_1 a_0)_{2B}$ starts with the digits of the base *2B* expansion of the integer $(a_n a_{n-1} \ldots a_1)_{2B}$ $[(a_n a_{n-1} \ldots a_1)_{2B} + 1]$ and ends with the digits $B/2$ and 0 when B is even, and the digits $(B-1)/2$ and B when B is odd.

Computations and Explorations

1. Verify the rules given in Exercises 23 and 24 for examples of your choice.

Programming Projects

1. Perform addition with arbitrarily large integers.

2. Perform subtraction with arbitrarily large integers.

3. Multiply two arbitrarily large integers using the conventional algorithm.

4. Divide arbitrarily large integers, finding the quotient and remainder.

2.3 Complexity of Integer Operations

Once an algorithm has been specified for an operation, we can consider the amount of time required to perform this algorithm on a computer. We will measure the amount of time in terms of *bit operations*. By a bit operation we mean the addition, subtraction, or multiplication of two binary digits, the division of a two-bit by a one-bit integer (obtaining a quotient and a remainder), or the shifting of a binary integer one place. (The actual amount of time required to carry out a bit operation on a computer varies depending on the computer architecture and capacity.) When we describe the number of bit operations needed to perform an algorithm, we are describing the *computational complexity* of this algorithm.

In describing the number of bit operations needed to perform calculations, we will use *big-O* notation. Big-O notation provides an upper bound on the size of a function in terms of a particular well-known reference function whose size at large values is easily understood.

To motivate the definition of this notation, consider the following situation. Suppose that to perform a specified operation on an integer n requires at most $n^3 + 8n^2 \log n$ bit operations. Because $8n^2 \log n < 8n^3$ for every positive integer, less than $9n^3$ bit operations are required for this operation for every integer n. Because the number of bit operations required is always less than a constant times n^3, namely, $9n^3$, we say that $O(n^3)$ bit operations are needed. In general, we have the following definition.

Definition. If f and g are functions taking positive values, defined for all $x \in S$, where S is a specified set of real numbers, then f is $O(g)$ on S if there is a positive constant K such that $f(x) < Kg(x)$ for all sufficiently large $x \in S$. (Normally, we take S to be the set of positive integers, and we drop all reference to S.)

Big-O notation is used extensively throughout number theory and in the analysis of algorithms. *Paul Bachmann* introduced big-O notation in 1892 ([Ba94]). The big-O notation is sometimes called a Landau symbol, after *Edmund Landau*, who used this notation throughout his work in the estimation of various functions in number theory. The use of big-O notation in the analysis of algorithms was popularized by renowned computer scientist *Donald Knuth*.

We illustrate this concept of big-O notation with several examples.

Example 2.10. We can show on the set of positive integers that $n^4 + 2n^3 + 5$ is $O(n^4)$. To do this, note that $n^4 + 2n^3 + 5 \leq n^4 + 2n^4 + 5n^4 = 8n^4$ for all positive integers. (We take $K = 8$ in the definition.) The reader should also note that n^4 is $O(n^4 + 2n^3 + 5)$. ◀

Example 2.11. We can easily give a big-O estimate for $\sum_{j=1}^{n} j$. Noting that each summand is less than n tells us that $\sum_{j=1}^{n} j \leq \sum_{j=1}^{n} n = n \cdot n = n^2$. Note that we could also derive this estimate easily from the formula $\sum_{j=1}^{n} j = n(n+1)/2$. ◀

We now will give some useful results for working with big-O estimates for combinations of functions.

Theorem 2.2. If f is $O(g)$ and c is a positive constant, then cf is $O(g)$.

PAUL GUSTAV HEINRICH BACHMANN (1837–1920), the son of a pastor, shared his father's pious lifestyle, as well as his love of music. His talent for mathematics was discovered by one of his early teachers. After recovering from tuberculosis, he studied at the University of Berlin and later in Göttingen, where he attended lectures presented by Dirichlet. In 1862, he received his doctorate under the supervision of the number theorist Kummer. Bachmann became a professor at Breslau and later at Münster. After retiring, he continued mathematical research, played the piano, and served as a music critic for newspapers. His writings include a five-volume survey of number theory, a two-volume work on elementary number theory, a book on irrational numbers, and a book on Fermat's last theorem (this theorem is discussed in Chapter 13). Bachmann introduced big-O notation in 1892.

EDMUND LANDAU (1877–1938) was the son of a Berlin gynecologist, and attended high school in Berlin. He received his doctorate in 1899 under the direction of Frobenius. Landau first taught at the University of Berlin and then moved to Göttingen, where he was full professor until the Nazis forced him to stop teaching. His main contributions to mathematics were in the field of analytic number theory; he established several important results concerning the distribution of primes. He authored a three-volume work on number theory and many other books on mathematical analysis and analytic number theory.

Proof. If f is $O(g)$, then there is a constant K with $f(x) < Kg(x)$ for all x under consideration. Hence $cf(x) < (cK)g(x)$, so cf is $O(g)$. ∎

Theorem 2.3. If f_1 is $O(g_1)$ and f_2 is $O(g_2)$, then $f_1 + f_2$ is $O(g_1 + g_2)$, and $f_1 f_2$ is $O(g_1 g_2)$.

Proof. If f is $O(g_1)$ and f_2 is $O(g_2)$, then there are constants K_1 and K_2 such that $f_1(x) < K_1 g_1(x)$ and $f_2(x) < K_2 g_2(x)$ for all x under consideration. Hence,

$$\begin{aligned} f_1(x) + f_2(x) &< K_1 g_1(x) + K_2 g_2(x) \\ &\leq K(g_1(x) + g_2(x)), \end{aligned}$$

where K is the maximum of K_1 and K_2. Hence, $f_1 + f_2$ is $O(g_1 + g_2)$.

Also,

$$\begin{aligned} f_1(x) f_2(x) &< K_1 g_1(x) K_2 g_2(x) \\ &= (K_1 K_2)(g_1(x) g_2(x)), \end{aligned}$$

so $f_1 f_2$ is $O(g_1 g_2)$. ∎

Corollary 2.3.1. If f_1 and f_2 are $O(g)$, then $f_1 + f_2$ is $O(g)$.

Proof. Theorem 2.3 tells us that $f_1 + f_2$ is $O(2g)$. But if $f_1 + f_2 < K(2g)$, then $f_1 + f_2 < (2K)g$, so $f_1 + f_2$ is $O(g)$. ∎

DONALD KNUTH (b. 1938) grew up in Milwaukee, where his father owned a small printing business and taught bookkeeping. He was an excellent student who also applied his intelligence in unconventional ways, such as finding more than 4500 words that could be spelled from the letters in "Ziegler's Giant Bar," winning a television set for his school and candy bars for everyone in his class.

Knuth graduated from Case Institute of Technology in 1960 with B.S. and M.S. degrees in mathematics, by special award of the faculty who considered his work outstanding. At Case, he managed the basketball team and applied his mathematical talents by evaluating each player using a formula he developed (receiving coverage on CBS television and in *Newsweek*). Knuth received his doctorate in 1963 from the California Institute of Technology.

Knuth taught at the California Institute of Technology and Stanford University, retiring in 1992 to concentrate on writing. He is especially interested in updating and adding to his famous series, *The Art of Computer Programming*. This series has had a profound influence on the development of computer science. Knuth is the founder of the modern study of computational complexity and has made fundamental contributions to the theory of compilers. Knuth has also invented the widely used TeX and Metafont systems used for mathematical (and general) typography. TeX played an important role in the development of HTML and the Internet. He popularized the big-O notation in his work on the analysis of algorithms.

Knuth has written for a wide range of professional journals in computer science and mathematics. However, his first publication, in 1957, when he was a college freshman, was the "The Potrzebie System of Weights and Measures," a parody of the metric system, which appeared in *MAD Magazine*.

The goal in using big-O estimates is to give the best big-O estimate possible while using the simplest reference function possible. Well-known reference functions used in big-O estimates include 1, $\log n$, n, $n \log n$, $n \log n \log \log n$, n^2, and 2^n, as well as some other important functions. Calculus can be used to show that each function in this list is smaller than the next function in the list, in the sense that the ratio of the function and the next function tends to 0 as n grows without bound. Note that more complicated functions than these occur in big-O estimates, as you will see in later chapters.

We illustrate how to use theorems for working with big-O estimates with the following example.

Example 2.12. To give a big-O estimate for $(n + 8 \log n)\ (10n \log n + 17n^2)$, first note that $n + 8 \log n$ is $O(n)$ and $10n \log n + 17n^2$ is $O(n^2)$ (because $\log n$ is $O(n)$ and $n \log n$ is $O(n^2)$) by Theorems 2.2 and 2.3 and Corollary 2.3.1. By Theorem 2.3, we see that $(n + 8 \log n)(10n \log n + 17n^2)$ is $O(n^3)$. ◀

Using big-O notation, we can see that to add or subtract two n-bit integers takes $O(n)$ bit operations, whereas to multiply two n-bit integers in the conventional way takes $O(n^2)$ bit operations (see Exercises 12 and 13 at the end of this section). Surprisingly, there are faster algorithms for multiplying large integers. To develop one such algorithm, we first consider the multiplication of two $2n$-bit integers, say, $a = (a_{2n-1}a_{2n-2} \ldots a_1a_0)_2$ and $b = (b_{2n-1}b_{2n-2} \ldots b_1b_0)_2$. We write

$$a = 2^n A_1 + A_0 \quad b = 2^n B_1 + B_0,$$

where

$$A_1 = (a_{2n-1}a_{2n-2} \ldots a_{n+1}a_n)_2 \quad A_0 = (a_{n-1}a_{n-2} \ldots a_1a_0)_2$$
$$B_1 = (b_{2n-1}b_{2n-2} \ldots b_{n+1}b_n)_2 \quad B_0 = (b_{n-1}b_{n-2} \ldots b_1b_0)_2.$$

We will use the identity

$$ab = (2^{2n} + 2^n)A_1B_1 + 2^n(A_1 - A_0)(B_0 - B_1) + (2^n + 1)A_0B_0. \tag{2.2}$$

To find the product of a and b using (2.2) requires that we perform three multiplications of n-bit integers (namely, A_1B_1, $(A_1 - A_0)(B_0 - B_1)$, and A_0B_0), as well as a number of additions and shifts. This is illustrated by the following example.

Example 2.13. We can use (2.2) to multiply $(1101)_2$ and $(1011)_2$. We have $(1101)_2 = 2^2(11)_2 + (01)_2$ and $(1011)_2 = 2^2(10)_2 + (11)_2$. Using (2.2), we find that

$$\begin{aligned}(1101)_2(1011)_2 &= (2^4 + 2^2)(11)_2(10)_2 + 2^2((11)_2 - (01)_2) \cdot ((11)_2 - (10)_2) + \\ &\qquad (2^2 + 1)(01)_2(11)_2 \\ &= (2^4 + 2^2)(110)_2 + 2^2(10)_2(01)_2 + (2^2 + 1)(11)_2 \\ &= (1100000)_2 + (11000)_2 + (1000)_2 + (1100)_2 + (11)_2 \\ &= (10001111)_2.\end{aligned}$$

◀

We will now estimate the number of bit operations required to multiply two n-bit integers by using (2.2) repeatedly. If we let $M(n)$ denote the number of bit operations needed to

multiply two n-bit integers, we find from (2.2) that

$$M(2n) \le 3M(n) + Cn, \tag{2.3}$$

where C is a constant, because each of the three multiplications of n-bit integers takes $M(n)$ bit operations, whereas the number of additions and shifts needed to compute ab via (2.2) does not depend on n, and each of these operations takes $O(n)$ bit operations.

From (2.3), using mathematical induction, we can show that

$$M(2^k) \le c(3^k - 2^k), \tag{2.4}$$

where c is the maximum of the quantities $M(2)$ and C (the constant in (2.3)). To carry out the induction argument, we first note that with $k = 1$, we have $M(2) \le c(3^1 - 2^1) = c$, because c is the maximum of $M(2)$ and C.

As the induction hypothesis, we assume that

$$M(2^k) \le c(3^k - 2^k).$$

Then, using (2.3), we have

$$\begin{aligned} M(2^{k+1}) &\le 3M(2^k) + C2^k \\ &\le 3c(3^k - 2^k) + C2^k \\ &\le c3^{k+1} - c \cdot 3 \cdot 2^k + c2^k \\ &\le c(3^{k+1} - 2^{k+1}). \end{aligned}$$

This establishes that (2.4) is valid for all positive integers k.

Using inequality (2.4), we can prove the following theorem.

Theorem 2.4. Multiplication of two n-bit integers can be performed using $O(n^{\log_2 3})$ bit operations. (*Note:* $\log_2 3$ is approximately 1.585, which is considerably less than the exponent 2 that occurs in the estimate of the number of bit operations needed for the conventional multiplication algorithm.)

Proof. From (2.4), we have

$$\begin{aligned} M(n) = M(2^{\log_2 n}) &\le M(2^{[\log_2 n]+1}) \\ &\le c(3^{[\log_2 n]+1} - 2^{[\log_2 n]+1}) \\ &\le 3c \cdot 3^{[\log_2 n]} \le 3c \cdot 3^{\log_2 n} = 3cn^{\log_2 3} \quad (\textit{because } 3^{\log_2 n} = n^{\log_2 3}). \end{aligned}$$

Hence, $M(n)$ is $O(n^{\log_2 3})$. ∎

We now state, without proof, two pertinent theorems. Proofs may be found in [Kn97] or [Kr79].

Theorem 2.5. Given a positive number $\epsilon > 0$, there is an algorithm for multiplication of two n-bit integers using $O(n^{1+\epsilon})$ bit operations.

Note that Theorem 2.4 is a special case of Theorem 2.5 with $\epsilon = \log_2 3 - 1$, which is approximately 0.585.

Theorem 2.6. There is an algorithm to multiply two n-bit integers using $O(n \log_2 n \log_2 \log_2 n)$ bit operations.

Because $\log_2 n$ and $\log_2 \log_2 n$ are much smaller than n^ϵ for large numbers n, Theorem 2.6 is an improvement over Theorem 2.5. Although we know that $M(n)$ is $O(n \log_2 n \log_2 \log_2 n)$, for simplicity we will use the obvious fact that $M(n)$ is $O(n^2)$ in our subsequent discussions.

The conventional algorithm described in Section 2.2 performs a division of a $2n$-bit integer by an n-bit integer with $O(n^2)$ bit operations. However, the number of bit operations needed for integer division can be related to the number of bit operations needed for integer multiplication. We state the following theorem, which is based on an algorithm discussed in [Kn97].

Theorem 2.7. There is an algorithm to find the quotient $q = [a/b]$, when the $2n$-bit integer a is divided by the integer b (having no more than n bits), using $O(M(n))$ bit operations, where $M(n)$ is the number of bit operations needed to multiply two n-bit integers.

2.3 Exercises

1. Determine whether each of the following functions is $O(n)$ on the set of positive integers.

a) $2n + 7$ c) 10 e) $\sqrt{n^2 + 1}$
b) $n^2/3$ d) $\log(n^2 + 1)$ f) $(n^2 + 1)/(n + 1)$

2. Show that $2n^4 + 3n^3 + 17$ is $O(n^4)$ on the set of positive integers.

3. Show that $(n^3 + 4n^2 \log n + 101n^2)(14n \log n + 8n)$ is $O(n^4 \log n)$.

4. Show that $n!$ is $O(n^n)$ on the set of positive integers.

5. Show that $(n! + 1)(n + \log n) + (n^3 + n^n)((\log n)^3 + n + 7)$ is $O(n^{n+1})$.

6. Suppose that m is a positive real number. Show that $\sum_{j=1}^{n} j^m$ is $O(n^{m+1})$.

* **7.** Show that $n \log n$ is $O(\log n!)$ on the set of positive integers.

8. Show that if f_1 and f_2 are $O(g_1)$ and $O(g_2)$, respectively, and c_1 and c_2 are constants, then $c_1 f_1 + c_2 f_2$ is $O(g_1 + g_2)$.

9. Show that if f is $O(g)$, then f^k is $O(g^k)$ for all positive integers k.

10. Let r be a positive real number greater than 1. Show that a function f is $O(\log_2 n)$ if and only if f is $O(\log_r n)$. (*Hint:* Recall that $\log_a n / \log_b n = \log_a b$.)

11. Show that the base b expansion of a positive integer n has $[\log_b n] + 1$ digits.

12. Analyzing the conventional algorithms for subtraction and addition, show that these operations require $O(n)$ bit operations with n-bit integers.