

ELEMENTARY NUMBER THEORY



Mike Barrus & W. Edwin Clark
University of Rhode Island

University of Rhode Island
Elementary Number Theory

Mike Barrus & W. Edwin Clark

This text is disseminated via the Open Education Resource (OER) LibreTexts Project (<https://LibreTexts.org>) and like the thousands of other texts available within this powerful platform, it is freely available for reading, printing, and "consuming."

The LibreTexts mission is to bring together students, faculty, and scholars in a collaborative effort to provide an accessible, and comprehensive platform that empowers our community to develop, curate, adapt, and adopt openly licensed resources and technologies; through these efforts we can reduce the financial burden born from traditional educational resource costs, ensuring education is more accessible for students and communities worldwide.

Most, but not all, pages in the library have licenses that may allow individuals to make changes, save, and print this book. Carefully consult the applicable license(s) before pursuing such effects. Instructors can adopt existing LibreTexts texts or Remix them to quickly build course-specific resources to meet the needs of their students. Unlike traditional textbooks, LibreTexts' web based origins allow powerful integration of advanced features and new technologies to support learning.



LibreTexts is the adaptable, user-friendly non-profit open education resource platform that educators trust for creating, customizing, and sharing accessible, interactive textbooks, adaptive homework, and ancillary materials. We collaborate with individuals and organizations to champion open education initiatives, support institutional publishing programs, drive curriculum development projects, and more.

The LibreTexts libraries are Powered by [NICE CXone Expert](#) and was supported by the Department of Education Open Textbook Pilot Project, the California Education Learning Lab, the UC Davis Office of the Provost, the UC Davis Library, the California State University Affordable Learning Solutions Program, and Merlot. This material is based upon work supported by the National Science Foundation under Grant No. 1246120, 1525057, and 1413739.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation nor the US Department of Education.

Have questions or comments? For information about adoptions or adaptations contact info@LibreTexts.org or visit our main website at <https://LibreTexts.org>.

This text was compiled on 06/09/2025

TABLE OF CONTENTS

Licensing

Preface

1: Chapters

- 1.1: What Is Number Theory?
- 1.2: Basic Axioms for $\mathbb{Z}$
- 1.3: Proof by Induction
- 1.4: Elementary Divisibility Properties
- 1.5: The Division Algorithm
- 1.6: The Base b Representation of n
- 1.7: Greatest Common Divisor and Least Common Multiple
- 1.8: The Euclidean Algorithm
- 1.9: Bezout's Lemma
- 1.10: Computing Coefficients for Bezout's Lemma
- 1.11: Prime Numbers
- 1.12: Unique Factorization
- 1.13: The Gaussian Integers
- 1.14: Fermat Primes and Mersenne Primes
- 1.15: Number Theoretic Functions
- 1.16: Perfect Numbers and Mersenne Primes
- 1.17: Congruences
- 1.18: Divisibility Tests for 2, 3, 5, 9, 11
- 1.19: Divisibility Tests for 7 and 13
- 1.20: More Properties of Congruences
- 1.21: Residue Classes and the Integers Modulo m
- 1.22: The Groups U_m
- 1.23: Chinese Remainder Theorem
- 1.24: Theorems of Wilson, Euler, and Fermat
- 1.25: Primality Tests
- 1.26: Computation of $a^n \bmod m$
- 1.27: The RSA Scheme
- 1.28: Sum of Squares
- 1.29: Epilogue

2: Appendices

- 2.1: Prime Numbers less than 200
- 2.2: Prime Numbers Up to 400
- 2.3: A Rings and Groups
- 2.4: Bibliography

Index

Glossary

Licensing

A detailed breakdown of this resource's licensing can be found in [Back Matter/Detailed Licensing](#).

Preface

This version CC BY-NC-SA 2021 by Michael D. Barrus

Under the Creative Commons license BY-NC-SA, users of this text may freely use, transform, and distribute this work for non-commercial purposes as long as they give credit to the authors and share their creations under this same license. For details, see <https://creativecommons.org/licenses/by-nc-sa/4.0/> .

This text is an adaptation of W. Edwin Clark's text from 2002, which appears with a similar license that Clark referred to as "copyleft," at

www.math.usf.edu/%7Eeclark/elem_num_th_book.pdf .

Preface

The following text is an extensive update of an original manuscript by Professor W. Edwin Clark (now Emeritus) of the University of South Florida, written in 2002 and made freely available on his website. Professor Clark's gracious use of what he termed "copyleft" status for his book, wherein I and other instructors have been able to distribute his text freely to our students and edit it to suit our needs, has been a great boon to many.

As I have taught out of Professor Clark's text for multiple semesters, I have made several changes in my own presentation of the material, and because of Professor Clark's generous permission, I have undertaken to adapt the book to reflect these changes.

The text is designed for a one-semester course and for self study. As Clark did, I have assumed that students have some familiarity with basic set theory and high school-level algebra (including imaginary numbers, if the chapter on Gaussian integers is to be taught), as well as the notion of a limit in one chapter. As Clark wrote in his original preface,

The text requires only a certain amount of mathematical maturity. And, hopefully, the student's level of mathematical maturity will increase as the course progresses.

I have tried to keep the same focus.

Clark's book was largely self-contained, and I have tried to maintain that quality as I have made changes and added material to the text. Following is a brief description of most of the changes made to the text.

- Textual changes to improve flow were added to various chapters.
- I have moved parts of the original book's preface to various locations in text.
- An introductory chapter (which appears as Chapter 1) was added to the text to engage student interest and convey the excitement of experimentation in number theory.
- Some exercises were reordered, and the wording in and around various exercises was changed to correct grammatical mistakes and/or make the intent clearer. Exercises were collected at the ends of the chapters, with additional exercises included, in several chapters of the text.
- Major changes were made to what is now Chapter in an attempt to motivate and improve the presentation on mathematical induction.
- Material on the floor and ceiling functions has been moved to the initial chapter discussing the integers, where its relation to the Well-Ordering Principle is mentioned. I have substituted an alternate proof of the Division Algorithm in Chapter and reduced the use of the floor function in that chapter to a mention and an exercise.
- The chapter on base b representations of n has been moved earlier in the text (it is now Chapter) to appear just after the chapter on the Division Algorithm. In this way students see two distinct uses of the Division Algorithm in the chapters that follow (the Euclidean Algorithm being the other).
- The concept of the least common multiple was introduced along with the greatest common denominator and developed through a few chapters' exercises.
- The chapter on Bezout's Lemma (currently Chapter) was supplemented with additional motivation and exercises.
- The chapter on Blankinship's Method (currently Chapter) was retitled and now includes a brief description of the Extended Euclidean Algorithm; exercises were modified to adopt this change.
- As a means of commenting on prime factorizations and providing some context for the sums-of-squares discussion at the end of the text, a new chapter is devoted to the Gaussian integers.

- The chapter on Fermat and Mersenne primes has a broader theme, with a new introduction about functions that generate prime numbers (including a mention of the Green–Tao Theorem on arithmetic progressions in the primes).
- Information on the prime counting function and Mersenne primes found has been updated to include the latest records as of July 2021.
- The material on Euler’s totient function has been expanded upon and moved to the chapter with other number theoretic functions.
- The introduction of perfect numbers was moved from the chapter on number theoretic functions to Chapter .
- For the sake of pacing, I have compressed three chapters from Clark’s original text into one (currently Chapter), omitting the notion of a complete system of residues, as it was not needed anywhere else in the text.
- Clark’s original text had the virtue of distinguishing between $\mathbb{Z}_m$ and the sets $\{0, 1, \dots, m - 1\}$ through the use of the notation J_m , $\oplus$, and $\odot$. However, in light of student difficulties due to competing notation used in other mathematics classes (notably abstract algebra classes), and because the formalism is not built upon very much in the chapters that follow, I have blurred the lines between these two rings by renaming Clark’s J_m as Z_m (without the blackboard bold font) and not circling the symbols of the associated binary operations.
- A new chapter discusses the Chinese Remainder Theorem.
- A discussion of Wilson’s Theorem was added to Chapters and , and content in Chapter was reordered.
- The chapter on RSA (Chapter) has been reorganized, with the addition of supporting examples and text and several exercises.
- A new chapter discusses representation of integers by sums of squares, connecting the end of the course with Chapter 1 and pointing interested students towards results on quadratic residues.
- References to Clark’s supplementary worksheets in Maple (eg., in Chapter) have been removed. I recommend that instructors supplement the text with computational experiments using technology that best suits their students’ needs.

My goals for students are identified as “three wishes” in Chapter 1. I am happy to hear of feedback on the text (including the correction of errors); I may be contacted via email at barrus@uri.edu . In accordance with the terms of the copyleft status Clark originally imposed, this text, including my changes, may be used under the Creative Commons BY-NC-SA 4.0 license. As such it may be freely modified and distributed with attribution (please mention both W. Edwin Clark and myself), and later adapters are expected to license their editions the same permissions they have found with this one.

Michael D. Barrus

University of Rhode Island

August 2021

CHAPTER OVERVIEW

1: Chapters

- 1.1: What Is Number Theory?
- 1.2: Basic Axioms for $\mathbb{Z}$
- 1.3: Proof by Induction
- 1.4: Elementary Divisibility Properties
- 1.5: The Division Algorithm
- 1.6: The Base b Representation of n
- 1.7: Greatest Common Divisor and Least Common Multiple
- 1.8: The Euclidean Algorithm
- 1.9: Bezout's Lemma
- 1.10: Computing Coefficients for Bezout's Lemma
- 1.11: Prime Numbers
- 1.12: Unique Factorization
- 1.13: The Gaussian Integers
- 1.14: Fermat Primes and Mersenne Primes
- 1.15: Number Theoretic Functions
- 1.16: Perfect Numbers and Mersenne Primes
- 1.17: Congruences
- 1.18: Divisibility Tests for 2, 3, 5, 9, 11
- 1.19: Divisibility Tests for 7 and 13
- 1.20: More Properties of Congruences
- 1.21: Residue Classes and the Integers Modulo m
- 1.22: The Groups U_m
- 1.23: Chinese Remainder Theorem
- 1.24: Theorems of Wilson, Euler, and Fermat
- 1.25: Primality Tests
- 1.26: Computation of $a^n \bmod m$
- 1.27: The RSA Scheme
- 1.28: Sum of Squares
- 1.29: Epilogue

This page titled [1: Chapters](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Mike Barrus & W. Edwin Clark](#).

1.1: What Is Number Theory?

Welcome to number theory!

In this chapter we will see a bit of what number theory is about and why you might enjoy studying it. Carl Friedrich Gauss (1777–1855), one of the greatest mathematicians of all time, had this to say about number theory (which he called *arithmetic*):

Mathematics is the queen of sciences and arithmetic the queen of mathematics.

- Quoted by Sartorius von Waltershausen in Gauss zum Gedachtniss (1856)

So what is "arithmetic," or number theory?

Simply stated, number theory is concerned with questions about and properties of the integers

$$\dots, -4, -3, -2, -1, 0, 1, 2, 3, 4, \dots$$

and closely-related numbers. Since you've been dealing with whole numbers of one kind or another for almost your whole life, some of what we'll see in the text will seem familiar, and much may seem simple and easy at first glance. Still, number theory is a surprisingly deep subject, and though this text only delves into what is known as *elementary number theory*, you will see new and different sides to a few things you may have thought you already knew.

Whetting your appetite

To give you a taste of what number theory is like, look at the following three questions:

Writing numbers as sums of squares

A *perfect square* is a number obtained by squaring an integer. For example, the four smallest perfect squares are

$$0 = 0^2, \quad 1 = 1^2 = (-1)^2, \quad 4 = 2^2 = (-2)^2, \quad 9 = 3^2 = (-3)^2.$$

A list of 21 perfect squares is found in [Appendix B](#).

As you can see, not every integer is a perfect square, and in fact the perfect squares get farther apart the larger they get. However, more numbers can be made by adding perfect squares together, which is what this question is about.

Which numbers can be written as the sum of two perfect squares? If we try putting 0, 1, 4, 9 together in pairs (possibly taking two of the same number), we can create 0, 1, 2, 4, 5, 8, 9, 10, 13, 18. As we use later squares to do the building, we see that the complete list of numbers that can be written as perfect squares begins with

$$0, 1, 2, 4, 5, 8, 9, 10, 13, 16, 17, 18, 20, 25, 26, 29, 32, 34, 36, 37, 40, \dots$$

Is there any pattern to which whole numbers do or do not appear in the list?

Similarly, which numbers can be written as the sum of three perfect squares? The list begins with

$$0, 1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 16, 17, 18, 19, 20, 21, 22, 24, \dots$$

It looks like most nonnegative whole numbers show up here, but 7, 15, and 23 do not. Is it a coincidence that these non-appearing numbers are 8 apart—would 31 be the next non-appearing number?

Which numbers can be written as the sum of four perfect squares? Allowing an extra square allows us to produce some numbers we couldn't before; for example, $23 = 9 + 9 + 4 + 1$. Making a list of those numbers that can be written with four squares, we begin with

$$0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, \dots;$$

in fact, checking all numbers between 0 and 100 shows that *every one* can be written as a sum of 4 perfect squares. Is it true that every positive integer is the sum of four squares? If so, why is this?

Finding a formula for prime numbers

You probably remember that a prime number is a number greater than 1 for which the only positive integers that evenly divide it are itself and 1. The first several prime numbers, beginning with 2, 3, 5, 7, 11, 13, ..., are listed in [Appendix A](#).

As we will see in the text, understanding prime numbers is a very important part of understanding many of the properties of integers. Is there a nice way to generate prime numbers? Perhaps a formula?

Consider the function $f(n) = n^3 + n^2 + 17$. Plugging in some integer values of n , we find

$$f(0) = 17, \quad f(1) = 19, \quad f(2) = 29, \quad \text{and} \quad f(3) = 53.$$

Each of these is prime, and in fact, plugging in each of the numbers from 0 to 10 always produces a prime number. Could $f(n)$ be a prime number for all n ? And are there other functions or techniques for producing prime numbers?

Integer solutions to equations

Can you find a pair of integers x, y such that $3x + 17y = 10$? A little trial and error might lead you to $x = -8$ and $y = 2$, but is this the only pair of integers that works? And is there a way besides trial and error that can systematically produce one or all of the solutions?

Other equations, like $154x - 33y = 10$, seem to have no solution where x and y are both integers. Why is this, and can we recognize equations like this from patterns in the numbers involved?

What about just slightly more complicated equations, like $x^2 + 3y = 1$? Why does this have several solutions in the integers (like $x = 1, y = 0$, and $x = 2, y = -1$, and $x = 4, y = -5$), while $x^2 + 3y = 2$ seems to have none?

Here's a famous equation along the same lines: Can you find positive integers x, y, z, n with n greater than 2 such that $x^n + y^n = z^n$?

We'll touch at least briefly on each of these questions throughout the text. Specifically:

- Clues about the answers to the sums-of-squares questions will be answered in a few places, culminating in a complete answer in [Section 1.28](#).
- Prime numbers will be studied in depth in [Sections 1.11, 1.14, 1.24, and 1.25](#), and one example of a valid prime-generating formula will be presented in the exercises of [Section 1.24](#). Incidentally, the function $f(n)$ above cannot always generate prime numbers—what happens if $n = 17$?
- The keys to finding integer solutions to linear equations with integer constants will be developed in the text and exercises of [Sections 1.8, 1.9, and 1.10](#). We will not be able to present solution techniques to some of the more complicated equations mentioned above, but with a little library work or searching online, you will be able to find out and, perhaps with a little patience and effort, digest and appreciate what's been solved. (You might start by doing a little reading on Fermat's Last Theorem.)

A practical application

In addition to answering these theoretical curiosities, we'll develop the tools used in the RSA cryptosystem, a widely used modern scheme for encrypting sensitive digital information.

As a simplified example, an online seller might publicly instruct a web browser using RSA to encrypt part of a credit card number, say, the number '1234' (perhaps part of a credit card number) by raising it to the exponent 43 and finding the remainder when the answer is divided by 1517. (We abbreviate the "find the remainder after dividing by" instruction by "mod".) The browser computes

$$1234^{43} \bmod 1517 = 1253$$

and, instead of transmitting the sensitive number '1234', transmits the encrypted number '1253'. The online seller then uses a secret decryption exponent (67) and the same "remainder process" to change the encrypted message back into the original:

$$1253^{67} \bmod 1517 = 1234.$$

So how does this work? How are the numbers 1517, 43, and 67 chosen? And how do we find division's remainders when the exponents involved produce such large numbers?¹ These questions are answered in [Sections 1.26 and 1.27](#), using results developed all throughout the text. As we will see, the strengths of the RSA system (and the codebreaking attacks that can break it if the numbers are not chosen wisely) have everything to do with the number theoretic properties of the numbers used to do the encryption and decryption.

The themes

The questions and application above illustrate some themes we'll run into throughout the text. We will look at various forms in which integers can sometimes be written. We will look multiple times at producing or identifying prime numbers. We will look for integer solutions to equations, and we will use properties of the integers to design algorithms to accomplish many different tasks. As you skim the table of contents now, and perhaps return to this chapter occasionally as you go through the text, you'll see these themes (and others) played out again and again.

There's another happy feature of number theory: many of the results we'll discuss won't be necessarily hard to recognize when you see them in action—in fact, several of the results will pop up quite easily as we look for patterns among multiple examples. (Of course, as mathematicians we're never satisfied until we can rigorously justify our observations through proof, but I hope you'll find the proofs in this text pleasant to digest, as well.)

Because number theory is about patterns in the integers, you will be well served to work out several numerical examples of ideas you encounter in the text and exercises. If you are familiar with writing simple programs in a computer algebra system (eg., CoCalc, Maple, Mathematica, MATLAB) or in a programming language, please try often to turn what you see in your studies into programs. You will be able to see many more examples this way, and the thought processes involved in writing your programs will enhance your understanding of number theory. As contemporary number theorist William Stein has said,

A computer is to a number theorist, like a telescope is to an astronomer. It would be a shame to teach an astronomy class without touching a telescope; likewise, it would be a shame to teach this class without telling you how to look at the integers through the lens of a computer.

Because there is a wide variety in the programming/computing environments with which students may be familiar, this book will not focus on any one computational system; however, you are heartily encouraged to pick one and dig deeply into number theory with it.

Three wishes

This book will have succeeded if it helps you do the following (not necessarily in order of importance):

- appreciate the beauty of patterns found in the integers;
- appreciate some of the practical applications of number theory;
- continue your growth in mathematical maturity and skill.

Here's to our success...Let's get started!

Footnotes

[1] The number 1253^{67} has 208 digits, and the numbers used in this example are much smaller than those used in practice!

This page titled [1.1: What Is Number Theory?](#) is shared under a [not declared](#) license and was authored, remixed, and/or curated by [Mike Barrus & W. Edwin Clark](#).

1.2: Basic Axioms for $\mathbb{Z}$

Basic Properties of $\mathbb{Z}$

Since number theory is concerned with properties of the integers, we begin by setting up some notation and reviewing some basic properties of the integers that will be needed later. We begin with our fundamental sets and their notation. Recall that $x \in S$ means that x is an element of the set S , and $S \subset T$ means that set S is a subset of set T .

$\mathbb{N} = \{1, 2, 3, \dots\}$ (the set of **natural numbers** or positive integers)

$\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$ (the set of **integers**)

$\mathbb{Q} = \left\{ \frac{n}{m} \mid n, m \in \mathbb{Z} \text{ and } m \neq 0 \right\}$ (the set of **rational numbers**)

$\mathbb{R}$ = the set of **real numbers**

$\mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}\}$ (the set of **complex numbers**)

In the last line, recall from previous mathematics classes that i is a (non-real, *imaginary*) number satisfying $i^2 = -1$. Note that $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$.

We assume a knowledge of the basic rules of high school algebra which apply to $\mathbb{R}$ and therefore to $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{Q}$. This means things like $ab = ba$ and $ab + ac = a(b + c)$. (Most of these properties also apply to numbers in $\mathbb{C}$.) We will not list all of these properties here. However, below we list some particularly important properties of $\mathbb{Z}$ that will be needed. We call these *axioms* since we will not prove them in this course.¹

Some Basic Axioms for $\mathbb{Z}$

1. If $a, b \in \mathbb{Z}$, then $a + b$, $a - b$ and $ab \in \mathbb{Z}$. ($\mathbb{Z}$ is closed under addition, subtraction and multiplication.)

2. If $a \in \mathbb{Z}$ then there is no $x \in \mathbb{Z}$ such that $a < x < a + 1$.

3. If $a, b \in \mathbb{Z}$ and $ab = 1$, then either $a = b = 1$ or $a = b = -1$.

4. **Laws of Exponents:** For n, m in $\mathbb{N}$ and a, b in $\mathbb{R}$ we have

a. $(a^n)^m = a^{nm}$

b. $(ab)^n = a^n b^n$

c. $a^n a^m = a^{n+m}$.

These rules hold for all $n, m \in \mathbb{Z}$ if a and b are not zero.

5. **Properties of Inequalities:** For a, b, c in $\mathbb{R}$ the following hold:

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

b. If $a < b$ then $a + c < b + c$.

c. If $a < b$ and $0 < c$ then $ac < bc$.

d. If $a < b$ and $c < 0$ then $bc < ac$.

e. (*Trichotomy*) Given a and b , one and only one of the following holds:

$$a = b, \quad a < b, \quad b < a.$$

6. **The Archimedean Property:** For every real number r there exists a natural number n such that $n > r$. In other words, the set $\mathbb{N}$ is a subset of $\mathbb{R}$ that has no upper bound.

7. **The Well-Ordering Principle:** Every non-empty subset of $\mathbb{N}$ contains a least element.

8. **The Principle of Mathematical Induction:** Let $P(n)$ be a statement concerning the integer variable n . Let n_0 be any fixed integer. $P(n)$ is true for all integers $n \geq n_0$ if one can establish both of the following statements:

a. $P(n)$ is true if $n = n_0$.

b. Whenever $P(n)$ is true for $n_0 \leq n \leq k$ then $P(n)$ is true for $n = k + 1$.

The use of the Principle of Mathematical Induction in proofs will be discussed in the next chapter. We illustrate a use of the Archimedean Property momentarily.

We use the usual conventions:

1. $a \leq b$ means $a < b$ or $a = b$,

2. $a > b$ means $b < a$, and

3. $a \geq b$ means $b \leq a$.

Floors and ceilings of real numbers

Since this chapter discusses basic properties of integers and their relationships with real numbers, we take this opportunity to define the floor, a.k.a., the greatest integer, and the ceiling, a.k.a., the least integer, functions. According to Donald Knuth [6], who popularized the notation presented below, Kenneth Iverson introduced the notation, as well as the terms floor and ceiling, in the early 1960s. Since then the notation has become standard in most areas of mathematics.

Definition 1.2.1

If x is any real number we define

$\lfloor x \rfloor$ = the greatest integer less than or equal to x , and

$\lceil x \rceil$ = the least integer greater than or equal to x .

Here $\lfloor x \rfloor$ is called the *floor* of x and $\lceil x \rceil$ is called the *ceiling* of x . (Note: the *floor* $\lfloor x \rfloor$ is in some texts denoted $[x]$ and called the greatest integer function.) Here are a few simple examples:

1. $\lfloor 3.1 \rfloor = 3$ and $\lceil 3.1 \rceil = 4$
2. $\lfloor 3 \rfloor = 3$ and $\lceil 3 \rceil = 3$
3. $\lfloor -3.1 \rfloor = -4$ and $\lceil -3.1 \rceil = -3$

For a more detailed treatment of both the floor and ceiling see the book *Concrete Mathematics* [5].

By the definitions, we have

$$\lfloor x \rfloor = \max\{n \in \mathbb{Z} | n \leq x\} \text{ and } \lceil x \rceil = \min\{n \in \mathbb{Z} | n \geq x\}.$$

The fact that $\lfloor x \rfloor$ exists for every real number x follows from the Archimedean Property and the Well-Ordering Principle. By a similar argument (see Exercise 1.2.6), $\lceil x \rceil$ also exists for every real number x .

By definition, $\lfloor x \rfloor \leq x$ for all x . Going further, note that for an integer n ,

$$\lfloor x \rfloor = n \iff n \leq x < n + 1. \quad (1.2.1)$$

By [Basic Axiom 2](#) above we also have that

$$\lfloor x \rfloor = x \iff x \in \mathbb{Z}.$$

The following lemma is helpful in proving facts involving floors.

Lemma 1.2.1

For all $x \in \mathbb{R}$

$$x - 1 < \lfloor x \rfloor \leq x.$$

Proof

Let $n = \lfloor x \rfloor$. Then by (1.2.1) we have $n \leq x < n + 1$. This gives immediately that $\lfloor x \rfloor \leq x$, as already noted above. It also gives $x < n + 1$ which implies that $x - 1 < n$, that is, $x - 1 < \lfloor x \rfloor$.

Important Convention

Since in this course we will be almost exclusively concerned with integers we shall assume from now on, unless otherwise stated, that all lower case roman letters $a, b, \dots, z$ denote integers.

Exercises

? Exercise 1.2.1

Using only the properties of inequalities listed in this chapter, and stating which ones you use (in other words, do *not* assume *anything else* about how inequalities behave when you operate on them), carefully prove that if x and y are positive real numbers and $x < y$, then

- a. $x^2 < y^2$;
- b. $x < y + 5$;
- c. $x^3 + 2x + 3 < y^3 + 2y + 4$.

? Exercise 1.2.2

Find $\lfloor \pi \rfloor$, $\lceil \pi \rceil$, $\lfloor \sqrt{2} \rfloor$, $\lceil \sqrt{2} \rceil$, $\lfloor -\pi \rfloor$, $\lceil -\pi \rceil$, $\lfloor -\sqrt{2} \rfloor$, and $\lceil -\sqrt{2} \rceil$.

? Exercise 1.2.3

Sketch the graph of the function $f(x) = \lfloor x \rfloor$ for $-3.5 \leq x \leq 3.5$.

? Exercise 1.2.4

Sketch the graph of $y = \lceil x \rceil - \lfloor x \rfloor$ for $-3.5 \leq x \leq 3.5$, and describe in words how the function $f(x) = \lceil x \rceil - \lfloor x \rfloor$ behaves for all $x \in \mathbb{R}$.

? Exercise 1.2.5

If x is a real number, are $\lfloor 2x \rfloor$ and $2\lfloor x \rfloor$ always the same? If not, then which one is bigger than the other? Does the answer depend on whether x is positive, negative, or zero? Once you think you know the answer, state your answer as an inequality, and carefully prove it by using [Lemma 1.2.1](#) and properties of inequalities or equations.

? Exercise 1.2.6

Prove that $\lfloor x \rfloor$, as it has been defined in this chapter, exists for every real number x .

(Hint: either think about what would happen if $\lfloor x \rfloor$ didn't exist, and explain why this violates one or more of the Basic Axioms in this chapter, or show that $\lfloor x \rfloor = -\lceil -x \rceil$ for all real numbers x and rely on the fact that ceilings always exist, as explained in this chapter.)

Footnotes

[1] These are not the simplest axioms we could use. Indeed, the Archimedean Property follows from a stronger property of $\mathbb{R}$, and strictly speaking, the Well-Ordering Principle and Principle of Mathematical Induction shouldn't both be axioms, since each can be derived from the other.

This page titled [1.2: Basic Axioms for \$\mathbb{Z}\$](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Mike Barrus & W. Edwin Clark](#).

1.3: Proof by Induction

A brief exploration

Suppose you were playing around with the smallest positive perfect squares (a list is found in [Appendix B](#)), and you decided to look for patterns in their sums. You might compute

$$\begin{aligned} 1 &= 1; \\ 5 &= 1 + 4; \\ 14 &= 1 + 4 + 9; \\ 30 &= 1 + 4 + 9 + 16; \\ 55 &= 1 + 4 + 9 + 16 + 25; \\ 91 &= 1 + 4 + 9 + 16 + 25 + 36; \\ 140 &= 1 + 4 + 9 + 16 + 25 + 36 + 49; \\ 204 &= 1 + 4 + 9 + 16 + 25 + 36 + 49 + 64. \end{aligned}$$

and wonder what patterns exist in this "sums of squares" sequence

$$1, 5, 14, 30, 55, 91, 140, 204, \dots$$

Plotting these numbers as points in the coordinate plane, i.e., plotting $(1, 1)$, $(2, 5)$, $(3, 14)$, $(4, 30)$ and so on yields the following picture:

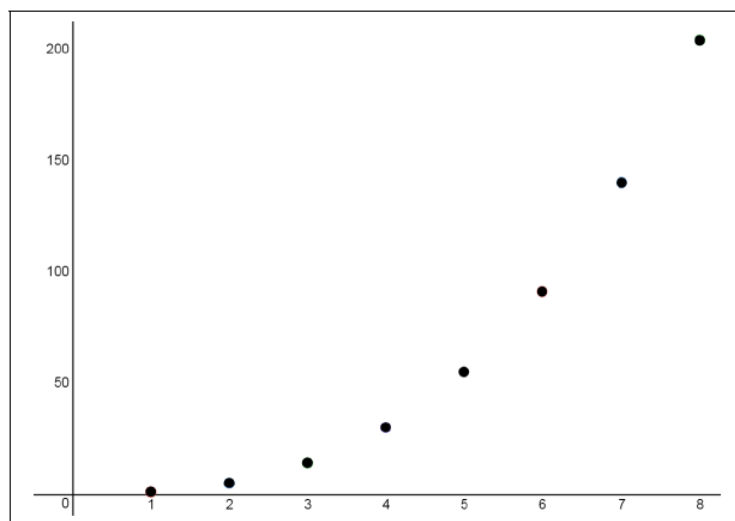


Figure 1.3.1

From a first glance at the picture, it looks like the the points certainly fit a smooth curve—perhaps a parabola or other polynomial curve, or perhaps the graph of an exponential function. One thing we don't know is the equation of this curve.

Looking at the sums-of-squares numbers 1, 5, 14, 30, 55, 91, 140, 204, . . . again, and heading in perhaps a different direction, we might look at their prime factorizations:¹

$$\begin{aligned} 1 & \\ 5 &= 5; \\ 14 &= 2 \cdot 7; \\ 30 &= 2 \cdot 3 \cdot 5; \\ 55 &= 5 \cdot 11; \\ 91 &= 7 \cdot 13; \\ 140 &= 2^2 \cdot 5 \cdot 7; \\ 204 &= 2^2 \cdot 3 \cdot 17. \end{aligned}$$

One thing that stands out about the factorizations are the numbers 5, 7, 11, 13, and 17. Other than in the factorizations of 30 and 140 (where we don't see 9 or 15), it looks as though the last number in each successive factorization is just the next odd number. In fact, if we forgot about prime factorizations and rewrote the factorizations of 1, 30, and 140 above, we would see the pattern even more clearly:

$$\begin{aligned}1 &= \frac{1}{3} \cdot 3; \\5 &= 5; \\14 &= 2 \cdot 7; \\30 &= \frac{1}{3} \cdot 2 \cdot 5 \cdot 9; \\55 &= 5 \cdot 11; \\91 &= 7 \cdot 13; \\140 &= \frac{1}{3} \cdot 2^2 \cdot 7 \cdot 15; \\204 &= 2^2 \cdot 3 \cdot 17.\end{aligned}$$

Note that in order to make 3, 9, and 15 appear in the expected places, we had to introduce the fraction $\frac{1}{3}$ in front of a few of the lines.

But now can we explain the patterns in the other factors appearing before the final odd factors? Can we explain why some lines need the fraction in the front while the others don't seem to? Through a little clever manipulation and perhaps some luck, we might happen on the following ways of writing our sums-of-squares numbers:

$$\begin{aligned}1 &= \frac{1}{6} \cdot 1 \cdot 2 \cdot 3; \\5 &= \frac{1}{6} \cdot 2 \cdot 3 \cdot 5; \\14 &= \frac{1}{6} \cdot 3 \cdot 4 \cdot 7; \\30 &= \frac{1}{6} \cdot 4 \cdot 5 \cdot 9; \\55 &= \frac{1}{6} \cdot 5 \cdot 6 \cdot 11; \\91 &= \frac{1}{6} \cdot 6 \cdot 7 \cdot 13; \\140 &= \frac{1}{6} \cdot 7 \cdot 8 \cdot 15; \\204 &= \frac{1}{6} \cdot 8 \cdot 9 \cdot 17.\end{aligned}$$

There's *definitely* a pattern here. In fact, looking at

$$\begin{aligned}1^2 &= \frac{1}{6} \cdot 1 \cdot 2 \cdot 3, \\1^2 + 2^2 &= \frac{1}{6} \cdot 2 \cdot 3 \cdot 5, \\1^2 + 2^2 + 3^2 &= \frac{1}{6} \cdot 3 \cdot 4 \cdot 7,\end{aligned}$$

and so on, we might guess that the following is true:

If n is any positive integer, then

$$1^2 + 2^2 + \cdots + n^2 = \frac{1}{6}n(n+1)(2n+1).$$

Plotting $y = \frac{1}{6}x(x+1)(2x+1)$ on the coordinate plane, we get a curve that passes perfectly through the points on our graph from before, and as we expected, our conjectured formula for the numbers is a polynomial—it's a cubic function.

At this point we believe we've discovered a formula for the sum of the first n positive square numbers. However, other than by just appealing to our observations (which is rarely good enough for a mathematician), *how would we prove that our formula is correct?*

Proofs by mathematical induction

We now discuss a powerful tool for answering questions like the one above and for proving statements about integers. This tool will reappear at various places throughout this text. It is the Principle of Mathematical Induction introduced in the previous chapter, which we will refer to by *PMI* or simply *induction*. Here it is again, slightly restated:²

The Principle of Mathematical Induction

If a statement about integers satisfies both

- a. the statement is true for the number n_0 , and
- b. whenever the statement is true for each of the numbers

$$n_0, n_0 + 1, \dots, k,$$

the statement is true for $k + 1$ as well,

then the statement is true for integer greater than or equal to n_0 .

There is a lot to digest in this principle. We illustrate it with a few examples, pointing out some key features of induction along the way.

We begin with our conjectured statement about the sums of squares. We call the statement we want to prove a *proposition*. It might also be called a *theorem*, *lemma* or *corollary* depending on the situation. Here again is our statement.

Proposition 1.3.1

If n is any positive integer, then

$$1^2 + 2^2 + \dots + n^2 = \frac{1}{6}n(n+1)(2n+1).$$

Proof

As we think about a proof, this statement seems like a good candidate for a proof using PMI. Why is this?

Key Idea

PMI is a proof technique designed specifically for statements of the form

$$P(n) \text{ for all integers } n \geq n_0.$$

where $P(n)$ is a statement depending on an integer n , and n_0 is a specific integer.

Now the proposition above is written as an "if/then" statement, but it can be rephrased as

$$1^2 + 2^2 + \dots + n^2 = \frac{1}{6}n(n+1)(2n+1) \text{ for all integers } n \geq 1.$$

This makes it a good candidate for PMI, and, using the notation of the key idea above, it also tells us two things:

- $P(n)$ is the statement " $1^2 + 2^2 + \dots + n^2 = \frac{1}{6}n(n+1)(2n+1)$ ";
- $n_0 = 1$.

Before proceeding, a few things are worth noting. First, the notation $P(n)$ is used to suggest that values of n can be substituted into the $P(n)$. If $n = 1$ then our $P(n)$ becomes $P(1)$, i.e., the statement $1^2 = \frac{1}{6} \cdot 1(1+1)(2 \cdot 1 + 1)$, and if $n = 2$ then $P(n)$ becomes $P(2)$, which is the statement $1^2 + 2^2 = \frac{1}{6} \cdot 2(2+1)(2 \cdot 2 + 1)$.

Next, it is important to note that $P(n)$ is the full equation, not simply $1^2 + 2^2 + \cdots + n^2$ or n^2 or $\frac{1}{6}n(n+1)(2n+1)$; instead of an expression, it is a *statement*. This will virtually always be the case: if you've correctly identified what $P(n)$ should be, then $P(n)$ will be a complete mathematical sentence, such as an equation or inequality, though it may sometimes be a more complicated English sentence,³ perhaps spanning multiple lines.

Carrying on towards a proof of our proposition, now that we have recognized that the statement has a suitable form for using PMI, and we have identified $P(n)$ and n_0 , we are ready to start assembling a proof. A proof by induction shows that $P(n)$ satisfies both the conditions (a) and (b) in the Principle:

Simplified outline of a proof using PMI:

PMI(a) Explain why $P(n_0)$ is true.⁴

PMI(b) Carefully justify the following: If $P(n)$ is known to be true for all values of n such that $n_0 \leq n \leq k$ and k is some integer, then these statements imply that $P(k+1)$ is true. Said another way, what we need to prove is that

if $P(n_0); P(n_0 + 1); P(n_0 + 2); \dots P(k)$ are all true, then $P(k+1)$ must be true, too.

What does this outline look like for our proposition?

Step PMI(a) requires a demonstration that $P(n_0)$ is true. For us this means that we must explain why $P(1)$ is true, i.e., why

$$1^2 = \frac{1}{6} \cdot 1(1+1)(2 \cdot 1 + 1).$$

Since both sides of the equation equal 1, this can immediately be seen to be true; we just need to state it in the proof, and PMI(a) will be complete.

Now, since PMI(b) is an if/then statement, to prove it we begin by *assuming* or *supposing*, just for the sake of argument, that

$$P(n) \text{ is true for } 1 \leq n \leq k.$$

That is, we assume

$$1^2 + 2^2 + \cdots + n^2 = \frac{1}{6}n(n+1)(2n+1) \text{ for } 1 \leq n \leq k. \quad (1.3.1)$$

Another way to say this is that we suppose that

$$\begin{aligned} 1^2 &= \frac{1}{6} \cdot 1(1+1)(2 \cdot 1 + 1), \\ 1^2 + 2^2 &= \frac{1}{6} \cdot 2(2+1)(2 \cdot 2 + 1), \\ &\vdots \\ 1^2 + 2^2 + \cdots + k^2 &= \frac{1}{6} \cdot k(k+1)(2k+1) \end{aligned}$$

are all true statements, and we would like to see what these would imply.

The assumption (1.3.1) is called the *induction hypothesis*; it is the "if" part of the "if/then" we are proving in PMI(b). Remember, to prove this implication, we suppose that the statement $P(n)$ is known to be true for every choice of n between n_0 and k , and we then show how to use these facts to prove that $P(n)$ holds when $n = k+1$. Proving that $P(k+1)$ is true is sometimes called the *induction step*.

Here is one way to carry out the inductive step in our example. We take the equation

$$1^2 + 2^2 + \cdots + k^2 = \frac{1}{6}k(k+1)(2k+1)$$

(which is $P(k)$, one of the equations we agreed to suppose was true in the induction hypothesis) and add $(k+1)^2$ to both sides to get

$$1^2 + 2^2 + \cdots + k^2 + (k+1)^2 = \frac{1}{6}k(k+1)(2k+1) + (k+1)^2. \quad (1.3.2)$$

Note that we are trying to prove $P(k+1)$, so we would *like* to show that

$$1^2 + 2^2 + \cdots + (k+1)^2 = \frac{1}{6}(k+1)(k+1+1)(2(k+1)+1);$$

we are not quite there yet. The last two equations above do agree on the left-hand side, which is good, but the right-hand sides are not yet the same; how will we justify $P(k+1)$?

Here, all that is needed is simple algebraic manipulation. Starting from equation (1.3.2), we can write

$$\begin{aligned} 1^2 + 2^2 + \cdots + k^2 + (k+1)^2 &= \frac{1}{6}k(k+1)(2k+1) + (k+1)^2 \\ &= \frac{1}{6}k(k+1)(2k+1) + \frac{1}{6} \cdot 6(k+1)^2 \\ &= \frac{1}{6}(k+1)[k(2k+1) + 6(k+1)] \\ &= \frac{1}{6}(k+1)(2k^2 + 7k + 6) \\ &= \frac{1}{6}(k+1)(k+2)(2k+3) \\ &= \frac{1}{6}(k+1)(k+1+1)(2(k+1)+1), \end{aligned}$$

which shows that $P(k+1)$ can be proved if we know that $P(1), \dots, P(k)$ are true. Thus we have established PMI(b), and PMI tells us that since PMI(a) and PMI(b) hold, the statement $P(n)$ is true for $n \geq 1$. We have finished!

Summarizing the discussion above, we now give a less pedagogical, more streamlined proof, which is more in line with the kinds of proofs you should strive for.

Proposition 1.3.2

If n is any positive integer, then

$$1^2 + 2^2 + \cdots + n^2 = \frac{1}{6}n(n+1)(2n+1).$$

Proof

We prove the proposition by induction on the variable n .

When $n = 1$ we find

$$1^2 = 1 = \frac{1}{6} \cdot 1(1+1)(2 \cdot 1 + 1),$$

so the claimed equation is true when $n = 1$.

Assume that

$$1^2 + 2^2 + \cdots + n^2 = \frac{1}{6}n(n+1)(2n+1) \quad \text{for } 1 \leq n \leq k \quad (\text{the induction hypothesis}).$$

Taking $n = k$ we have

$$1^2 + 2^2 + \cdots + k^2 = \frac{1}{6}k(k+1)(2k+1).$$

Then

$$\begin{aligned}
 1^2 + 2^2 + \cdots + (k+1)^2 &= [1^2 + 2^2 + \cdots + k^2] + (k+1)^2 \\
 &= \frac{1}{6}k(k+1)(2k+1) + (k+1)^2 \quad (\text{by the induction hypothesis}) \\
 &= \frac{1}{6}k(k+1)(2k+1) + \frac{1}{6} \cdot 6(k+1)^2 \\
 &= \frac{1}{6}(k+1)[k(2k+1) + 6(k+1)] \\
 &= \frac{1}{6}(k+1)(2k^2 + 7k + 6) \\
 &= \frac{1}{6}(k+1)(k+2)(2k+3) \\
 &= \frac{1}{6}(k+1)(k+1+1)(2(k+1)+1).
 \end{aligned}$$

Hence by PMI we conclude that

$$1^2 + 2^2 + \cdots + n^2 = \frac{1}{6}n(n+1)(2n+1).$$

for $n \geq 1$.

One of the hallmarks of a correctly written proof by induction is that if we check the claim by letting n equal every integer from n_0 on, in turn, in $P(n)$, the proof should give us convincing justification through a "domino" effect. For example, in the proposition above, we identified n_0 as 1; does the proof justify $P(1)$? Yes—this is what second sentence explained.

Now does the proof justify $P(2)$? Consider the remainder of the proof. We start by taking as a fact that $P(1), \dots, P(k)$ are all true. Well, at the moment we only know that $P(1)$ is true, so we can only suppose that $P(1), \dots, P(k)$ are all true if $k = 1$ and the list $P(1), \dots, P(k)$ really just contains $P(1)$. So let's read the rest of the proof, substituting 1 everywhere we see k . The remaining sentences in the proof give us a procedure for proving $P(k+1)$, which for us means $P(2)$. So yes, the proof does justify $P(2)$.

Now does the proof justify $P(3)$? Let's return to midway through the proof and reread the proof beginning at the induction hypothesis. There we are supposed to start from the claim that $P(1), \dots, P(k)$ are all true. Since at this point we have been convinced that $P(1)$ and $P(2)$ are both true, we can now let $k = 2$ and substitute 2 everywhere we see k . The remaining sentences in the proof then prove $P(k+1)$, which for us at this point means $P(3)$, so the proof does justify $P(3)$.

Continuing this process, rereading the proof over and over, allows us to justifiably and logically let k be larger and larger in the induction hypothesis, and we see that if we reapplied the portion of the proof beginning the induction hypothesis enough times, we would have a proof of $P(n)$ no matter which finite integer (past n_0) n happened to be. This is why induction proves that $P(n)$ is true for any $n \geq n_0$.

Following is another example of a proof using PMI. This time the statement $P(n)$ is an inequality, and the value n_0 is not 1. Try coming up with a proof of your own before reading the proof presented here.

Proposition 1.3.3

If $n \geq 5$ then $2^n > 5n$.

Proof

We prove the proposition by induction on the variable n . If $n = 5$ we have $2^5 > 5 \cdot 5$ or $32 > 25$ which is true.

Assume

$$2^n > 5n \quad \text{for } 5 \leq n \leq k \quad (\text{the induction hypothesis}).$$

Taking $n = k$ we have

$$2^k > 5k.$$

Multiplying both sides by 2 gives

$$2^{k+1} > 10k.$$

Now $10k = 5k + 5k$ and $k \geq 5$ so $k \geq 1$ and therefore $5k \geq 5$. Hence

$$10k = 5k + 5k \geq 5k + 5 = 5(k + 1).$$

It follows that

$$2^{k+1} > 10k \geq 5(k + 1)$$

and therefore

$$2^{k+1} > 5(k + 1).$$

Hence by PMI we conclude that $2^n > 5n$ for $n \geq 5$.

Note that in [Proposition 1.3.3](#) the condition $n \geq 5$ (which leads to us having $n_0 = 5$) is necessary, since $2^5 > 5 \cdot 5$, but $2^4 < 5 \cdot 4$.

Considering both the previous propositions, a few general comments may help. First, note that the finished version of a proof by induction is often much more concise and straightforward than the calculations and notes that went into finding it. For that reason induction proofs in texts like this one may seem a bit more mysterious than necessary, but to most mathematicians, the conciseness is preferable, and you would do well to try to include only the necessities in your proofs when you are polishing them up.

Next, and more importantly, how did we figure out what to do to prove $P(k + 1)$ once we had supposed that $P(n)$ was true for $n_0 \leq n \leq k$? This is the big challenge of mathematical induction, and the one place where proofs by induction require problem solving and sometimes some creativity or ingenuity. Different steps were required at this stage of the proofs of the two propositions above, and figuring out how to show that $P(k + 1)$ automatically happens if $P(n_0), \dots, P(k)$ do is not an exact science.

One of the best things to try is to pick out a part of the statement $P(k + 1)$ that looks like something from one of the earlier statements $P(n_0), \dots, P(k)$, and try to conclude something about that part of $P(k + 1)$.

In the proof of [Proposition 1.3.2](#), for example, the left-hand side of $P(k + 1)$ included all the terms of the left-hand side of $P(k)$ and one more term besides, and it was this connection to the equation $P(k)$ that made proving $P(k + 1)$ possible. In the proof of [Proposition 1.3.3](#), the inequality $P(k + 1)$ had a left-hand side that included 2^{k+1} . This didn't match the left-hand side of any other $P(5), \dots, P(k)$, but it was pretty close to the 2^k appearing on the left-hand side of $P(k)$. Using the inequality $2^k > 5k$ from $P(k)$ to try to say something about the 2^{k+1} in $P(k + 1)$ is what motivated us to multiply both sides of $P(k)$ by 2, getting $2^{k+1} > 10k$ as we did.

This may still not completely clarify how to use earlier $P(n)$ statements to prove $P(k + 1)$, but it's a starting point, and proving $P(k + 1)$ in PMI(b) will become much clearer and easier with practice. Again, *pick out a part of the statement $P(k + 1)$ that looks like something from one of the earlier statements $P(n)$, and try to conclude something about the rest of $P(k + 1)$.*

Finally, you'll notice that there were a few things common to the proofs of [Propositions 1.3.2](#) and [1.3.3](#) that were not spelled out in the simplified outline for induction proofs from before. We now give a more detailed outline for proofs by induction, followed by exercises where you can practice applying it.

Eight major parts of a proof by induction:

1. First state what proposition you are going to prove. Precede the statement by *Proposition*, *Theorem*, *Lemma*, *Corollary*, *Fact*, or *To Prove*.
2. Write the *Proof* or *Pf.* at the very beginning of your proof.
3. Say that you are going to use induction (not every mathematical proof uses induction!) and if it is not obvious from the statement of the proposition, clearly identify $P(n)$, i.e., the statement to be proved and the variable it depends upon, and the starting value n_0 . Even though this is usually clear, sometimes these things may not be obvious. And, of course, the variable need not be n .
4. Prove that $P(n)$ holds when $n = n_0$.
5. Assume that $P(n)$ holds for $n_0 \leq n \leq k$. (State this assumption in writing, closely following the form in the examples here!) This assumption will be referred to as the *induction hypothesis*.
6. Use the induction hypothesis and anything else that is known to be true to prove that $P(n)$ holds when $n = k + 1$.
7. Conclude, perhaps in writing, that since the conditions of the PMI have been met, $P(n)$ holds for $n \geq n_0$.
8. Write QED, $\square$, $\blacksquare$, or $//$ or something to indicate that you have completed your proof.

Exercises

? Exercise 1.3.1

Consider the following statement:

Every nonnegative integer can be written as the sum of four perfect squares.

In this exercise we'll imagine what a proof of this statement by induction would look like, if one were possible.

- For this statement, answer these questions (coming from Part 3 of the eight major parts at the end of the section): What would n_0 be? What would $P(n)$ be? (Remember that $P(n)$ should be a statement—a complete mathematical sentence.)
- In order to get some practice with working with $P(n)$, use your answer in part (a) to write out the statements $P(2)$ and $P(5)$. (Do a thorough job—be sure to write out the complete statements.)
- Now for Part 4 of the list: Write out the statement $P(n_0)$ using your answer from part (a), and then prove that $P(n_0)$ is true.
- Carry out Part 5: Carefully write out the induction hypothesis we would use in an induction proof of the statement above, starting with the word "Assume."
- What is the "goal" of Part 6? In other words, what is the specific statement about numbers and perfect squares we would like to show holds in this part?
- Yes or no—is it easy to you to see how to use the facts

$$\begin{aligned} 0 &= 0^2 + 0^2 + 0^2 + 0^2; & 3 &= 1^2 + 1^2 + 1^2 + 0^2; \\ 1 &= 1^2 + 0^2 + 0^2 + 0^2; & 4 &= 1^2 + 1^2 + 1^2 + 1^2; \\ 2 &= 1^2 + 1^2 + 0^2 + 0^2; \end{aligned}$$

to write 5 as a sum of four perfect squares? Do you see how to complete Part 6 of the eight major parts, for any k ? How likely do you think it is that an induction proof is possible for the statement we started with?⁵

? Exercise 1.3.2

Prove that $2^n > 6n$ for $n \geq 5$.

? Exercise 1.3.3

Prove that $1 + 2 + \cdots + n = \frac{n(n+1)}{2}$ for $n \geq 1$.

? Exercise 1.3.4

Using only PMI and properties of inequalities from the previous chapter, prove that if $0 < a < b$ then $0 < a^n < b^n$ for all $n \in \mathbb{N}$.

? Exercise 1.3.5

Prove that $n! < n^n$ for $n \geq 2$.

(For positive integers n , we define n factorial, written $n!$, as the product of all integers between and including n and 1. For example, $1! = 1$ and $2! = 2 \cdot 1 = 2$ and $5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$.)

? Exercise 1.3.6

Prove that $1 + 2 + 2^2 + \cdots + 2^n = 2^{n+1} - 1$ for $n \geq 1$.

? Exercise 1.3.7

Use PMI to prove that $\underbrace{111 \cdots 1}_{n \text{ 1's}} = \frac{10^n - 1}{9}$ for $n \geq 1$.

? Exercise 1.3.8

Prove that if a and r are real numbers and $r \neq 1$, then for $n \geq 1$,

$$a + ar + ar^2 + \cdots + ar^n = \frac{a(r^{n+1} - 1)}{r - 1}.$$

? Exercise 1.3.9

Prove that $1^3 + 2^3 + 3^3 + \cdots + n^3 = \frac{n^2(n+1)^2}{4}$ if $n \geq 1$. (Compare this result to [Exercise 1.3.3](#).)

? Exercise 1.3.10

Prove that $\frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} + \cdots + \frac{1}{n(n+1)} = \frac{n}{n+1}$ for all $n \geq 1$.

? Exercise 1.3.11

Perhaps imitating the steps from the "brief exploration" in this chapter, conjecture a formula for each of the sums below. Then using PMI, prove that your conjectured formula is correct.

- $1 \cdot 2 + 2 \cdot 3 + 3 \cdot 4 + \cdots + n(n+1)$;
- $1 + 3 + 5 + \cdots + (2n-1)$.

? Exercise 1.3.12

Using induction, prove that

$$a^n - b^n = (a - b)(a^{n-1} + a^{n-2}b + \cdots + ab^{n-2} + b^{n-1})$$

for all integers $n \geq 2$. (Hint: during the induction step, write $a^{k+1} - b^{k+1} = a^{k+1} - ab^k + ab^k - b^{k+1}$.)

? Exercise 1.3.13

Prove that if $n \geq 12$ then n can be written as a sum of 4's and 5's. For example, $23 = 5 + 5 + 5 + 4 + 4 = 3 \cdot 5 + 2 \cdot 4$.
(Hint: In this case it will help to do the cases $n = 12, 13, 14$, and 15 separately. Then use induction to handle $n \geq 16$.)

? Exercise 1.3.14

- For $n \geq 1$, the *triangular number* t_n is the number of dots in a triangular array that has n rows with i dots in the i -th row. Find a formula for t_n , $n \geq 1$.
- Suppose that for $n \geq 1$, s_n denotes the number of dots in a square array that has n rows with n dots in each row. Find a formula for s_n . The numbers s_n are usually called *squares*.

? Exercise 1.3.15

- Find the first 10 triangular numbers and the first 10 squares.
- Which of the triangular numbers in part (a) are also squares?
- Can you find the next triangular number which is a square?

- d. After completing parts (a)–(c), read the Wikipedia.org article "Square triangular number" and write one or more detailed sentences on what you thought of it.

? Exercise 1.3.16

Some propositions that can be proved by induction can also be proved without induction. Prove the propositions in [Exercises 1.3.3](#) and [1.3.8](#) and [1.3.11\(b\)](#) without induction.

(Hints: For [Exercise 1.3.3](#) write $s = 1 + 2 + \cdots + (n-1) + n$. Directly under this equation write $s = n + (n-1) + \cdots + 2 + 1$. Add these equations to obtain $2s = n(n+1)$. Solve for s . For [Exercise 1.3.8](#) write $p = a + ar + ar^2 + \cdots + ar^n$. Then multiply both sides of this equation by r to get a new equation with rp as the left hand side. Subtract these two equations to obtain $rp - p = ar^{n+1} - a$. Now solve for p . For [Exercise 1.3.11\(b\)](#), draw rows of dots as in [Exercise 1.3.14\(b\)](#) and break them up into sets of size 1, 3, 5, ...

Footnotes

- [1] Here we rely a bit on your previous knowledge. We will discuss primes and prime factorizations in more depth later. Note that 1 is not a prime.
- [2] In other math courses you may have seen this principle called strong induction, with "ordinary" induction referring to a principle with a different condition (b) than the one appearing here. This should not matter too much; the PMI used here is chosen for simplicity.
- [3] An example may be found in [Exercise 1.3.1](#) at the end of the chapter.
- [4] Depending on the details of PMI(b), sometimes in PMI(a) not only $P(n_0)$ but also a few more specific statements $P(i)$ (such as $P(n_0 + 1)$ or $P(n_0 + 2)$) are necessary to check in order to make the logic work; see, for example, [Exercise 1.3.13](#) at the end of the chapter. Most of the induction proofs in this text will require just checking $P(n_0)$; when more early cases are required, this will be pointed out.
- [5] We leave it as an unanswered question whether the statement this exercise is based on is true, though we will come back to this question later in the text. The point of the exercise is to show that you can carry out most of the eight major parts even before you figure out how to carry out Part 6 (or even before you know whether the statement you are attempting to prove is even true!).

This page titled [1.3: Proof by Induction](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Mike Barrus & W. Edwin Clark](#).

1.4: Elementary Divisibility Properties

Definition 1.4.1

We write $d \mid n$ if there is an integer k such that $n = dk$. The notation $d \nmid n$ means that $d \mid n$ is false.

Note that $a \mid b$ does not mean the same thing as a/b . Recall that a/b represents the fraction $\frac{a}{b}$, a mathematical “noun.” The notation $a \mid b$ represents a complete mathematical statement or sentence.

The expression $d \mid n$ may be read in any of the following ways:

1. d divides n .
2. d is a divisor of n .
3. d is a factor of n .
4. n is a multiple of d .

Thus, the following five statements are equivalent, that is, they are all different ways of saying the same thing.

1. $2 \mid 6$.
2. 2 divides 6.
3. 2 is a divisor of 6.
4. 2 is a factor of 6.
5. 6 is a multiple of 2.

Definitions will play an important role in this course. Students should learn all definitions and be able to state them precisely. An alternative way to state the definition of $d \mid n$ is as follows.

Definition 1.4.2

$d \mid n \iff n = dk$ for some k .

or maybe

Definition 1.4.3

$d \mid n$ if and only if $n = dk$ for some k .

Keep in mind that we are assuming that all letters $a, b, \dots, z$ represent integers. Otherwise we would have to add this fact to our definitions. One might also see the following definition sometimes.

Definition 1.4.4

$d \mid n$ if $n = dk$ for some k .

Note that $\iff$ and *if and only if* mean the same thing. In definitions (such as [Definition 1.4.4](#)), *if* is interpreted to mean *if and only if*. It should be emphasized that all the above definitions are acceptable. Take your pick. But be careful about making up your own definitions.

Theorem 1.4.1: Divisibility Properties

If n , m , and d are integers then the following statements hold:

1. $n \mid n$ (everything divides itself)
2. $d \mid n$ and $n \mid m \implies d \mid m$ (transitivity)
3. $d \mid n$ and $d \mid m \implies d \mid (an + bm)$ for all a and b (linearity property)
4. $d \mid n \implies ad \mid an$ (multiplication property)
5. $ad \mid an$ and $a \neq 0 \implies d \mid n$ (cancellation property)
6. $1 \mid n$ (one divides everything)

7. $n \mid 1 \implies n = \pm 1$ (1 and -1 are the only divisors of 1.)
8. $d \mid 0$ (everything divides zero)
9. $0 \mid n \implies n = 0$ (zero divides only zero)
10. If d and n are positive and $d \mid n$ then $d \leq n$ (comparison property)

We end this chapter with one more definition, that will be important in the chapters that follow.

 Definition 1.4.5

If $c = as + bt$ for some integers s and t we say that c is a *linear combination* of a and b .

Thus, statement 3 in [Theorem 1.4.1](#) says that if d divides a and b , then d divides all linear combinations of a and b . In particular, d divides $a + b$ and $a - b$. This will turn out to be a useful fact, so we emphasize it in the exercises.

Exercises

? Exercise 1.4.1

Prove each of the properties 1 through 10 in [Theorem 1.4.1](#).

? Exercise 1.4.2

Prove that if $b \neq 0$ then $b \mid a$ if and only if $a/b \in \mathbb{Z}$.

(As said earlier, one difficulty some students have with the “divides” notation is confusing its vertical bar with a fraction bar—the two bars are *not* the same. One (namely, $\mid$) represents a relationship; the other (namely, $/$ or $-$) represents a binary operation. In your answer, pay particular care that you are using the correct bar in each place.)

? Exercise 1.4.3

Prove that if $d \mid a$ and $d \mid b$ then $d \mid (a - b)$.

? Exercise 1.4.4

Prove that if $a \in \mathbb{Z}$ then the only positive divisor of both a and $a + 1$ is 1. (Hint: use [Exercise 1.4.3](#).)

? Exercise 1.4.5

Prove that $x + y$ is divisor of $x^n + y^n$ for all *odd* positive integers n .

(Hint: it might be helpful to consider the result proved in [Exercise 1.3.12](#); you may use it here, if desired, without supplying the proof.)

? Exercise 1.4.6

- a. Use induction to prove that $(6n + 1)^2 - 1$ is divisible by 4 for every integer $n \geq 0$.
- b. Prove the same statement from part (a) without using induction by expanding $(6n + 1)^2 - 1$.

? Exercise 1.4.7

- a. Prove by induction, not relying on any of the exercises from the previous chapter, that $7^n - 1$ is divisible by 6 for every positive integer n .
- b. Now substitute one or more values for n in an exercise from the previous chapter to find a simpler proof of the statement in part (a). (Something to think about: how does your proof in part (a) compare to the proof in the exercise you used from the previous chapter?)

This page titled [1.4: Elementary Divisibility Properties](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Mike Barrus & W. Edwin Clark](#).

1.5: The Division Algorithm

The goal of this chapter is to introduce and prove the following important result.

Theorem 1.5.1: The Division Algorithm

If a and b are integers and $b > 0$ then there exist unique integers q and r satisfying the two conditions:

$$a = bq + r \quad \text{and} \quad 0 \leq r < b. \quad (1.5.1)$$

In this situation q is called the *quotient* and r is called the *remainder when a is divided by b* . We sometimes refer to a as the *dividend* and b as the *divisor*.

(Some care must be taken with this last term, in light of how the divisor of an integer was defined in [Section 1.4](#). It is correct both to say that 2 is not a divisor of 7, since $2 \nmid 7$, though 2 is the divisor when 7 is divided by 2, yielding $7 = 2 \cdot 3 + 1$. Context should make it clear which meaning is intended.)

Note that there are two parts to the Division Algorithm's statement. One part is the *existence* of integers q and r satisfying [\(1.5.1\)](#), and the second part is the *uniqueness* of the integers q and r satisfying [\(1.5.1\)](#).

Proof

We first prove the existence of q and r .

If b divides a then $a = kb$ for some integer k , and we may write $q = k$ and $r = 0$.

Suppose instead that b does not divide a , and consider the collection

$$\dots, a - 3b, a - 2b, a - b, a, a + b, a + 2b, a + 3b, \dots$$

of integers. Let S be the set of all these numbers that are non-negative. In symbols,

$$S = \{a - kb \mid k \in \mathbb{Z} \text{ and } a - kb \geq 0\}.$$

The set S is non-empty, since by the Archimedean Property there exists a natural number n such that $n > -a$ and so $nb \geq n > -a$ and hence $a - (-n)b > 0$. Now by the Well-Ordering Principle S contains a smallest element; call it r . By our definition of S , this means that for some particular integer q we have $r = a - qb$. Thus $a = bq + r$, and it only remains to show that $0 \leq r < b$.

All the elements of S are positive, so it is clear that $0 \leq r$. If it is *not* true that $r < b$, then $a - bq \geq b$ and so $r - b = a - (q + 1)b \geq 0$; thus either $r - b$ is an element of S or $r - b = 0$. Since r was defined as the smallest element of S , it cannot be true that $r - b$ is in S . However, if $r - b = 0$, then $a = bq + b = (q + 1)b$, which contradicts our assumption that b does not divide a . We are forced to conclude that $0 \leq r < b$.

We still have to prove that q and r are uniquely determined. To do this we suppose that

$$a = bq_1 + r_1 \quad \text{and} \quad 0 \leq r_1 < b,$$

and

$$a = bq_2 + r_2 \quad \text{and} \quad 0 \leq r_2 < b$$

for integers q_1, q_2, r_1, r_2 . We must show that $r_1 = r_2$ and $q_1 = q_2$. If $r_1 \neq r_2$ without loss of generality we can assume that $r_2 > r_1$. Subtracting these two equations we obtain

$$0 = a - a = (bq_1 + r_1) - (bq_2 + r_2) = b(q_1 - q_2) + (r_1 - r_2).$$

This implies that

$$r_2 - r_1 = b(q_1 - q_2). \quad (1.5.2)$$

This implies that $b \mid r_2 - r_1$. By [Theorem 1.4.1\(10\)](#) this implies that $b \leq r_2 - r_1$. But since

$$0 \leq r_1 < r_2 < b$$

we have $r_2 - r_1 < b$. This contradicts $b \leq r_2 - r_1$. So we must conclude that $r_1 = r_2$. Now from (1.5.2) we have $0 = b(q_1 - q_2)$. Since $b > 0$ this tells us that $q_1 - q_2 = 0$, that is, $q_1 = q_2$. This completes the proof of the uniqueness of r and q in (1.5.1).

In number theory we usually prefer to work with integers, but if we are willing to consider division of real numbers, it can be shown that when a is divided by b , the quotient and remainder satisfy

$$q = \left\lfloor \frac{a}{b} \right\rfloor \quad \text{and} \quad r = a - bq = a - b \left\lfloor \frac{a}{b} \right\rfloor$$

(see Exercise 1.5.2).

We now turn to some familiar concepts.

Definition 1.5.1

An integer n is *even* if $n = 2k$ for some k , and is *odd* if $n = 2k + 1$ for some k .

Definition 1.5.2

By the *parity* of an integer we mean whether it is even or odd.

Sometimes a problem in number theory can be solved by dividing the integers into various classes depending on their remainders when divided by some number b . For example, this is helpful in solving Exercises 1.5.5 and 1.5.6 at the end of this chapter.

Definition 1.5.3

For $b > 0$ define $a \bmod b$ as the value of r in the pair q, r satisfying the expressions $a = bq + r$ and $0 \leq r < b$.

In other words, $a \bmod b$ is the remainder given by the Division Algorithm when a is divided by b .

For example $23 \bmod 7 = 2$ since $23 = 7 \cdot 3 + 2$ and $-4 \bmod 5 = 1$ since $-4 = 5 \cdot (-1) + 1$.

Note that some calculators and most programming languages have a function often denoted by $MOD(a, b)$ or $mod(a, b)$ whose value is what we have just defined as $a \bmod b$. When this is the case the values r and q in the Division Algorithm for given a and $b > 0$ are given by

$$\begin{aligned} r &= a \bmod b \\ q &= \frac{a - (a \bmod b)}{b} \end{aligned}$$

If also the floor function is available we have

$$\begin{aligned} r &= a \bmod b \\ q &= \lfloor a/b \rfloor \end{aligned}$$

Exercises

Exercise 1.5.1

Find the values of q and r in the Division Algorithm for the following values of a and b :

- Let $b = 3$ and $a = 0, 1, -1, 10, -10$.
- Let $b = 345$ and $a = 0, -1, 1, 344, 7863, -7863$

Exercise 1.5.2

Given integers $b > 0$ and a , let q and r be the quotient and remainder when b is divided by a , so $a = bq + r$. Carefully prove that

$$q = \left\lfloor \frac{a}{b} \right\rfloor.$$

? Exercise 1.5.3

Using the Division Algorithm, prove that every integer is either even or odd, but never both.

? Exercise 1.5.4

Prove n and n^2 always have the same parity. That is, n is even if and only if n^2 is even.

? Exercise 1.5.5

Show that for all integers n the number $n^3 - n$ always has 3 as a factor.

(Hint: by the Division Algorithm, every integer n falls into one of three cases: $n = 3k$, $n = 3k + 1$, $n = 3k + 2$. Prove that $n^3 - n$ is divisible by 3 for each of these cases separately.)

? Exercise 1.5.6

Show that the product of any three consecutive integers has 6 as a factor. (How many cases should you use here?)

? Exercise 1.5.7

Prove the following:

1. If $b > 0$ and $b \mid a$, then $a \bmod b = 0$.
2. If $b > 0$ and $a \bmod b = 0$, then $b \mid a$.

? Exercise 1.5.8

Calculate the following:

- a. $0 \bmod 10$
- b. $123 \bmod 10$
- c. $10 \bmod 123$
- d. $457 \bmod 33$
- e. $(-7) \bmod 3$
- f. $(-3) \bmod 7$
- g. $(-5) \bmod 5$

? Exercise 1.5.9

- a. Use the Division Algorithm to prove the following similar theorem: If $b > 0$ then for any a there exists unique integers p and s such that

$$a = bp + s \quad \text{and} \quad \lfloor -b/2 \rfloor < s \leq \lfloor b/2 \rfloor.$$

- b. Find integers p and s such that $a = bp + s$ and $\lfloor -b/2 \rfloor < s \leq \lfloor b/2 \rfloor$ for
 - i. $a = 36, b = 13$
 - ii. $a = 75, b = 50$
 - iii. $a = 75, b = 49$.

1.6: The Base b Representation of n

In this chapter we show how the Division Algorithm is related to a concept touched on since grade school mathematics.

Definition 1.6.1: Decimal Representation

Recall that the *decimal representation* of a positive integer a is given by $a = a_{n-1}a_{n-2} \cdots a_1a_0$ where

$$a = a_{n-1}10^{n-1} + a_{n-2}10^{n-2} + \cdots + a_110 + a_0 \quad (1.6.1)$$

and the *digits* $a_{n-1}, a_{n-2}, \dots, a_1, a_0$ are in the set $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ with $a_{n-1} \neq 0$. In this case we say that *the integer a is an n -digit number* or that *a is n digits long*.

For example, the numerals in “682” are understood to mean six hundred eighty-two, because the locations of 6, 8, and 2 (in the “hundreds’ place”, “tens’ place”, and “ones’ place”) indicate the number $6 \cdot 100 + 8 \cdot 10 + 2$.

In this chapter we will justify the claim that every positive integer has a decimal representation. However, we will do so in a more general setting.

Definition 1.6.2: Base b Representation of n

Let $b \geq 2$ and $n > 0$. We write

$$n = [a_k, a_{k-1}, \dots, a_1, a_0]_b \quad (1.6.2)$$

if and only if for some $k \geq 0$

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

where $a_i \in \{0, 1, \dots, b-1\}$ for $i = 0, 1, \dots, k$. $[a_k, a_{k-1}, \dots, a_1, a_0]$ is called a *base b representation of n* .

Remark 1.6.1

Base b is referred to as

$$\begin{array}{ll} \text{binary} & \text{if } b = 2, \\ \text{ternary} & \text{if } b = 3, \\ \text{octal} & \text{if } b = 8, \\ \text{decimal} & \text{if } b = 10, \\ \text{hexadecimal} & \text{if } b = 16. \end{array}$$

If b is understood, especially if $b = 10$, we write $a_k a_{k-1} \cdots a_1 a_0$ in place of $[a_k, a_{k-1}, \dots, a_1, a_0]_b$. In the case of $b = 16$, which is used frequently in computer science, the “digits” 10, 11, 12, 13, 14 and 15 are replaced by A, B, C, D, E and F , respectively.

For a fixed base $b \geq 2$, the numbers $a_i \in \{0, 1, 2, \dots, b-1\}$ in equation (1.6.2) are called the *digits* of the base b representation of n . In the binary case $a_i \in \{0, 1\}$ and the a_i ’s are called *bits* (binary digits).

Here are a few examples:

- $267 = [5, 3, 1]_7$ since $267 = 5 \cdot 7^2 + 3 \cdot 7 + 1$.
- $147 = [1, 0, 0, 1, 0, 0, 1, 1]_2$ since $147 = 1 \cdot 2^7 + 0 \cdot 2^6 + 0 \cdot 2^5 + 1 \cdot 2^4 + 0 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2 + 1$.
- $4879 = [4, 8, 7, 9]_{10}$ since $4879 = 4 \cdot 10^3 + 8 \cdot 10^2 + 7 \cdot 10 + 9$.
- $10705679 = [A, 3, 5, B, 0, F]_{16}$ since $10705679 = 10 \cdot 16^5 + 3 \cdot 16^4 + 5 \cdot 16^3 + 11 \cdot 16^2 + 0 \cdot 16 + 15$.
- $107056791 = [107, 56, 791]_{1000}$ since $107056791 = 107 \cdot 1000^2 + 56 \cdot 1000 + 791$.

Theorem 1.6.1: Base b Representation

If $b \geq 2$, then every $n > 0$ has a unique base b representation of the form $n = [a_k, \dots, a_1, a_0]_b$ with $a_k > 0$.

Proof

Apply repeatedly the Division Algorithm as follows:

$$\begin{aligned} n &= bq_0 + r_0, & 0 \leq r_0 < b \\ q_0 &= bq_1 + r_1, & 0 \leq r_1 < b \\ q_1 &= bq_2 + r_2, & 0 \leq r_2 < b \\ &\vdots \\ q_{k-1} &= bq_k + r_k, & 0 \leq r_k < b \\ q_k &= bq_{k+1} + r_{k+1}, & 0 \leq r_{k+1} < b. \end{aligned}$$

Note that in every step we always divide by b . In the first line, we divide n by b , and then in each line after the first, what we divide by b is the quotient of the previous line; we then use the resulting quotient for the next line.

From the repeated division it is not difficult to see that as long as $q_k > 0$, it is true that

$$n > q_0 > q_1 > \dots > q_k.$$

Since this cannot go on forever we eventually obtain $q_\ell = 0$ for some ℓ . Then we have

$$q_{\ell-1} = b \cdot 0 + r_\ell.$$

We claim that $n = [r_\ell, r_{\ell-1}, \dots, r_0]_b$ if ℓ is the smallest integer such that $q_\ell = 0$. To see this, note that

$$n = bq_0 + r_0$$

and

$$q_0 = bq_1 + r_1.$$

Hence

$$\begin{aligned} n &= b(bq_1 + r_1) + r_0 \\ n &= b^2q_1 + br_1 + r_0. \end{aligned}$$

Continuing in this way we find that

$$n = b^{\ell+1}q_\ell + b^\ell r_\ell + \dots + br_1 + r_0.$$

And, since $q_\ell = 0$ we have

$$n = b^\ell r_\ell + \dots + br_1 + r_0, \tag{1.6.3}$$

which shows that

$$n = [r_\ell, \dots, r_1, r_0]_b.$$

To see that this representation is unique, note that from equation (1.6.3) we have

$$n = b(b^{\ell-1}r_\ell + \dots + r_1) + r_0, \quad 0 \leq r_0 < b.$$

By the Division Algorithm it follows that r_0 is uniquely determined by n , as is the quotient $q = b^{\ell-1}r_\ell + \dots + r_1$. A similar argument shows that r_1 is uniquely determined. Continuing in this way we see that all the digits $r_\ell, r_{\ell-1}, \dots, r_0$ are uniquely determined.

Example 1.6.1: Base Conversions

We can use the idea from the proof of Theorem 1.6.1, repeated applications of the Division Algorithm, to express numbers in an arbitrary base. Following are some examples.

1. We find the base 7 representation of 1,749.

$$\begin{aligned} 1749 &= 249 \cdot 7 + 6 \\ 249 &= 35 \cdot 7 + 4 \\ 35 &= 5 \cdot 7 + 0 \\ 5 &= 0 \cdot 7 + 5 \end{aligned}$$

Hence $1749 = [5, 0, 4, 6]_7$.

2. We find the base 12 representation of 19,151.

$$\begin{aligned} 19,151 &= 1595 \cdot 12 + 11 \\ 1,595 &= 132 \cdot 12 + 11 \\ 132 &= 11 \cdot 12 + 0 \\ 11 &= 0 \cdot 12 + 11 \end{aligned}$$

$\therefore 19,151 = [11, 0, 11, 11]_{12}$

3. Find the base 10 representation of 1,203.

$$\begin{aligned} 1203 &= 120 \cdot 10 + 3 \\ 120 &= 12 \cdot 10 + 0 \\ 12 &= 1 \cdot 10 + 2 \\ 1 &= 0 \cdot 10 + 1 \end{aligned}$$

$\therefore 1203 = [1, 2, 0, 3]_{10}$

4. Find the base 2 (*binary*) representation of 137.

$$\begin{aligned} 137 &= 2 \cdot 68 + 1 \\ 68 &= 2 \cdot 34 + 0 \\ 34 &= 2 \cdot 17 + 0 \\ 17 &= 2 \cdot 8 + 1 \\ 8 &= 2 \cdot 4 + 0 \\ 4 &= 2 \cdot 2 + 0 \\ 2 &= 2 \cdot 1 + 0 \\ 1 &= 2 \cdot 0 + 1 \end{aligned}$$

$\therefore 137 = [1, 0, 0, 0, 1, 0, 0, 1]_2$

Remark 1.6.2

To find the binary representation of a small number, the following method is often easier than the above method:

Given $n > 0$ let 2^{n_1} be the largest power of 2 satisfying $2^{n_1} \leq n$. Let 2^{n_2} be the largest power of 2 satisfying

$$2^{n_2} \leq n - 2^{n_1}.$$

Let 2^{n_3} be the largest power of 2 satisfying

$$2^{n_3} \leq n - 2^{n_1} - 2^{n_2}.$$

Note that at this point we have

$$0 \leq n - (2^{n_1} + 2^{n_2} + 2^{n_3}) < n - (2^{n_1} + 2^{n_2}) < n - 2^{n_1} < n.$$

Continuing in this way, eventually we get

$$0 = n - (2^{n_1} + 2^{n_2} + \cdots + 2^{n_k}).$$

Then $n = 2^{n_1} + 2^{n_2} + \cdots + 2^{n_k}$, and this gives the binary representation of n .

✓ Example 1.6.2

Take $n = 137$. Note that $2^1 = 2$, $2^2 = 4$, $2^3 = 8$, $2^4 = 16$, $2^5 = 32$, $2^6 = 64$, $2^7 = 128$, and $2^8 = 256$. Using the above method we compute:

$$\begin{aligned} 137 - 2^7 &= 137 - 128 = 9, \\ 9 - 2^3 &= 1, \\ 1 - 2^0 &= 0. \end{aligned}$$

So we have

$$\begin{aligned} 137 &= 2^7 + 9 = 2^7 + 2^3 + 1, \\ \therefore 137 &= 2^7 + 0 \cdot 2^6 + 0 \cdot 2^5 + 0 \cdot 2^4 + 2^3 + 0 \cdot 2^2 + 0 \cdot 2 + 1. \end{aligned}$$

So $137 = [1, 0, 0, 0, 1, 0, 0, 1]_2$

Exercises

? Exercise 1.6.1

Generalize the following observations

$$\begin{aligned} 1 &= [1]_2 \\ 3 &= [1, 1]_2 \\ 7 &= [1, 1, 1]_2 \\ 15 &= [1, 1, 1, 1]_2 \\ 31 &= [1, 1, 1, 1, 1]_2 \\ 63 &= [1, 1, 1, 1, 1, 1]_2 \end{aligned}$$

by noticing a pattern in the numbers on the left-hand sides and conjecturing a formula for the number that has n 1's in its binary representation. Prove your generalization.

(Hint: See [Exercise 1.3.8](#).)

? Exercise 1.6.2

Generalize the following observation:

$$\begin{aligned} 2 &= [2]_3 \\ 8 &= [2, 2]_3 \\ 26 &= [2, 2, 2]_3 \\ 80 &= [2, 2, 2, 2]_3 \\ 242 &= [2, 2, 2, 2, 2]_3 \end{aligned}$$

by noticing a pattern in the numbers on the left-hand sides and conjecturing a formula for the number that has n 2's in its ternary representation. Prove your generalization.

(Hint: See [Exercise 1.3.8](#).)

? Exercise 1.6.3

Generalize [Exercises 1.6.1](#) and [1.6.2](#) to an arbitrary base $b \geq 2$.

? Exercise 1.6.4

Show how to use *both methods* to find the binary representation of 455.

? Exercise 1.6.5

Make a vertical list of the binary representation of the integers 1 to 16.

? Exercise 1.6.6

Find the decimal representation of each of the following numbers.

- $[5, 3, 6, 2]_7$
- hexadecimal numbers FEEDAFACE and BADBEEF.

? Exercise 1.6.7

Following the technique of [Example 1.6.1](#), find the representations of 12345 in (i) base 3 and (ii) base 8. For both parts (i) and (ii), show your work with the same level of detail as [Example 1.6.1](#).

? Exercise 1.6.8

Find the base-7 representation of the number $[1, 2, 3, 4, 5, 6, 7, 8]_9$

? Exercise 1.6.9

Prove that if $a \in \mathbb{N}$ is an n -digit number, then $\lfloor \log(a) \rfloor + 1$. Here $\log$ means logarithm to base 10. Thus $\lfloor \log(a) \rfloor + 1$ is a formula for the number of digits in a .

(Hint: to prove the equation, show that if equation (1.6.1) holds with $a_{n-1} \neq 0$ then $10^{n-1} \leq a < 10^n$. Then apply the log to all terms of this inequality. This is allowed since $\log$ is an increasing function on the set of positive real numbers.)

? Exercise 1.6.10

Use the previous exercise to determine the number of digits in the decimal representation of the number $2^{3321928}$.

(Hint: recall that $\log(x^y) = y \log(x)$ when x and y are positive.)

? Exercise 1.6.11

Recall that $\log_b(a)$ means the logarithm of a with base b . Can you adapt the ideas from [Exercise 1.6.9](#) to find a formula for the number of digits in the base b representation of a ?

This page titled [1.6: The Base \$b\$ Representation of \$n\$](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Mike Barrus & W. Edwin Clark](#).

1.7: Greatest Common Divisor and Least Common Multiple

In the last few chapters we have discussed divisibility and the Division Algorithm when a single number is divided by another. In this chapter we begin to look at divisors and multiples that two numbers have in common.

Definition 1.7.1: Greatest Common Divisor

Let $a, b \in \mathbb{Z}$. If $a \neq 0$ or $b \neq 0$, we define the *greatest common divisor of a and b* , denoted $\gcd(a, b)$, to be the largest integer d such that $d \mid a$ and $d \mid b$. We define $\gcd(0, 0) = 0$.

Discussion.

If $e \mid a$ and $e \mid b$ we call e a *common divisor of a and b* . Let

$$C(a, b) = \{e : e \mid a \text{ and } e \mid b\},$$

that is, $C(a, b)$ is the set of *all* common divisors of a and b . Note that since everything divides 0

$$C(0, 0) = \mathbb{Z}$$

so there is no largest common divisor of 0 with 0. This is why we must *define* $\gcd(0, 0) = 0$.

✓ Example 1.7.1

$$C(18, 30) = \{-1, 1, -2, 2, -3, 3, -6, 6\}.$$

So $\gcd(18, 30) = 6$.

Lemma 1.7.1

If $e \mid a$ then $-e \mid a$.

Proof

If $e \mid a$ then $a = ek$ for some k . Then $a = (-e)(-k)$. Since $-e$ and $-k$ are also integers $-e \mid a$.

Lemma 1.7.2

If $a \neq 0$, the largest positive integer that divides a is $|a|$.

Proof

Recall that

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

First note that $|a|$ actually divides a : If $a > 0$, since we know $a \mid a$ we have $|a| \mid a$. If $a < 0$, $|a| = -a$. In this case $a = (-a)(-1) = |a|(-1)$ so $|a|$ is a factor of a . So, in either case $|a|$ divides a , and in either case $|a| > 0$, since $a \neq 0$.

Now suppose $d \mid a$ and d is positive. Then $a = dk$ some k so $-a = d(-k)$ for some k . So $d \mid |a|$. So by [Theorem 1.4.1](#) (10) we have $d \leq |a|$.

The following lemma shows that in computing $\gcd$'s we may restrict ourselves to the case where both integers are positive.

Lemma 1.7.3

For any integers a and b ,

$$\gcd(a, b) = \gcd(|a|, |b|).$$

Proof

If $a = 0$ and $b = 0$, we have $|a| = a$ and $|b| = b$. So $\gcd(a, b) = \gcd(|a|, |b|)$. Suppose one of a or b is not 0. Note that $d \mid a \Leftrightarrow d \mid |a|$. See [Exercise 1.7.1](#). It follows that

$$C(a, b) = C(|a|, |b|).$$

So the largest common divisor of a and b is also the largest common divisor of $|a|$ and $|b|$.

The last lemma above showed that to find the greatest common divisor of any two numbers, it is enough to be able to do so for nonnegative numbers. [Exercise 1.7.1](#) at the end of the chapter proves a more general statement about divisibility.

Lemma 1.7.4

For any integers a and b ,

$$\gcd(a, b) = \gcd(b, a).$$

Proof

Clearly $C(a, b) = C(b, a)$. It follows that the largest integer in $C(a, b)$ is the largest integer in $C(b, a)$, that is, $\gcd(a, b) = \gcd(b, a)$.

Lemma 1.7.5

If $a \neq 0$ and $b \neq 0$, then $\gcd(a, b)$ exists and satisfies

$$0 < \gcd(a, b) \leq \min\{|a|, |b|\}.$$

Proof

Note that $\gcd(a, b)$ is the largest integer in the set $C(a, b)$ of common division of a and b . Since $1 \mid a$ and $1 \mid b$ we know that $1 \in C(a, b)$. So the largest common divisor must be at least 1 and is therefore positive. On the other hand $d \in C(a, b) \Rightarrow d \mid |a|$ and $d \mid |b|$ so d is no larger than $|a|$ and no larger than $|b|$. So d is at most the smaller of $|a|$ and $|b|$. Hence $\gcd(a, b) \leq \min\{|a|, |b|\}$.

Example 1.7.2

From the above lemmas we have

$$\begin{aligned} \gcd(48, 732) &= \gcd(-48, 732) \\ &= \gcd(-48, -732) \\ &= \gcd(48, -732). \end{aligned}$$

We also know that

$$0 < \gcd(48, 732) \leq 48.$$

Since if $d = \gcd(48, 732)$, then $d \mid 48$, to find d we may check only which positive divisors of 48 also divide 732.

In the next chapter we will discuss a more computationally practical way to compute $\gcd(a, b)$. Right now we discuss a complementary concept.

Definition 1.7.2

Let $a, b \in \mathbb{N}$. We define the *least common multiple* of a and b , denoted $\text{lcm}(a, b)$, to be the smallest positive integer m such that $a \mid m$ and $b \mid m$.

✓ Example 1.7.3

If we wish to find $\text{lcm}(6, 10)$, we may start by listing positive multiples of 10 and looking for the first one that is divisible by 6. Since neither 10 nor 20 is divisible by 6 but $30 = 10 \cdot 3 = 6 \cdot 5$, we see both that 30 is a multiple of both 6 and 10, and no smaller positive multiple of 10 is divisible by 6. Hence $\text{lcm}(6, 10) = 30$.

Exercises 1.7.5 and 1.7.6 in this chapter invite you to discover a few basic properties of $\text{lcm}(a, b)$. Like the greatest common divisor, the least common multiple will be further developed over the next few chapters.

Exercises

? Exercise 1.7.1

Prove that

$$d \mid a \Leftrightarrow d \mid |a|.$$

(Hint: recall that

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

so it may help to consider two cases. Also remember the definition of “divides,” and use it here.)

? Exercise 1.7.2

Find $\text{gcd}(48, 732)$ using Example 1.7.2.

? Exercise 1.7.3

Find $\text{gcd}(a, b)$ for each of the following values of a and b :

1. $a = -14, b = 14$
2. $a = -1, b = 78654$
3. $a = 0, b = -78$
4. $a = 2, b = -786541$

? Exercise 1.7.4

Find both $\text{gcd}(a, b)$ and $\text{lcm}(a, b)$ for each of the following values of a and b . (Explain your reasoning, using only the definitions and properties discussed in this chapter; do not assume other properties that have not been proved yet.)

1. $a = 3, b = 14$
2. $a = 9, b = 15$
3. $a = 58, b = 406$
4. $a = 74, b = 111$

? Exercise 1.7.5

The definition given in this chapter requires that $\text{lcm}(a, b)$ be a positive integer and that a and b be natural numbers. To examine why this might be convenient, answer the following:

- a. What is the smallest common multiple (not necessarily positive) of 0 and b , where b is any integer?
- b. If neither a nor b is 0, is there a smallest common multiple of a and b among the negative integers? Illustrate your answer with an example, choosing a specific pair a, b of nonzero integers and discussing their common multiples.

? Exercise 1.7.6

Each of the lemmas in this chapter discusses divisors and/or $\gcd(a, b)$. For each lemma, can you find a similar, true statement about multiples and/or $\text{lcm}(a, b)$? If so, write these statements down (proving them is encouraged but optional); if not, explain why not.

This page titled [1.7: Greatest Common Divisor and Least Common Multiple](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Mike Barrus & W. Edwin Clark](#).

1.8: The Euclidean Algorithm

The Euclidean Algorithm is named after Euclid of Alexandria, who lived about 300 BCE. The algorithm¹ described in this chapter was recorded and proved to be successful in Euclid's *Elements*, so this algorithm is over two thousand years old. It provides a simple method to compute $\gcd(a, b)$, even if we do not know much about the divisors of a and b .

Since, as already noted, $\gcd(0, 0) = 0$ and $\gcd(a, b) = \gcd(|a|, |b|)$ and $\gcd(a, b) = \gcd(b, a)$, in order to compute $\gcd(a, b)$ it is enough to find a method for computing $\gcd(a, b)$ when $a \geq b \geq 0$, so we will assume that a and b satisfy these inequalities.

Before discussing the Euclidean Algorithm, we establish a few useful facts. First of all, as [Exercise 1.8.1](#) establishes, if $a > 0$, then $\gcd(a, a) = a$. The next result is in a similar vein.

Lemma 1.8.1

If $a > 0$, then $\gcd(a, 0) = a$.

Proof

Since every integer divides 0, $C(a, 0)$ is just the set of divisors of a . By [Lemma 1.7.2](#) the largest divisor of a is $|a|$. Since $a > 0$, $|a| = a$. This shows that $\gcd(a, 0) = a$.

With these facts we are now reduced to the problem of finding $\gcd(a, b)$ when $a > b > 0$.

Lemma 1.8.2

Let $a > b > 0$. If $a = bq + r$, then

$$\gcd(a, b) = \gcd(b, r).$$

Proof

It suffices to show that $C(a, b) = C(b, r)$, that is, the common divisors of a and b are the same as the common divisors of b and r . To show this first let $d \mid a$ and $d \mid b$. Note that $r = a - bq$, which is a linear combination of a and b . So by [Theorem 1.4.1\(3\)](#) $d \mid r$. Thus $d \mid b$ and $d \mid r$. Next assume $d \mid b$ and $d \mid r$. Using [Theorem 1.4.1\(3\)](#) again and the fact that $a = bq + r$ is a linear combination of b and r , we have $d \mid a$. So $d \mid a$ and $d \mid b$. We have thus shown that $C(a, b) = C(b, r)$. So $\gcd(a, b) = \gcd(b, r)$.

Remark 1.8.1

The *Euclidean Algorithm* is the process of using [Lemmas 1.8.2](#) and [1.8.1](#) to compute $\gcd(a, b)$ when $a > b > 0$.

Rather than give a precise statement of the algorithm we will start with an example to show how it goes.

Example 1.8.1

Let's compute $\gcd(803, 154)$

$\gcd(803, 154)$	=	$\gcd(154, 33)$	since $803 = 154 \cdot 5 + 33$;
$\gcd(154, 33)$	=	$\gcd(33, 22)$	since $154 = 33 \cdot 4 + 22$;
$\gcd(33, 22)$	=	$\gcd(22, 11)$	since $33 = 22 \cdot 1 + 11$;
$\gcd(22, 11)$	=	$\gcd(11, 0)$	since $22 = 11 \cdot 2 + 0$;
$\gcd(11, 0)$	=	11.	

Hence $\gcd(803, 154) = 11$.

By way of commentary, observe that in our repeated application of the Division Algorithm, both the dividend and divisor change at each step, with some numbers (like 33 and 22 above) appearing first as a remainder, then moving to be the divisor in the next division statement, and then finally serving as the dividend in the statement after that. Note that the process ends once a remainder of 0 is reached (this must happen, since at each step the remainders decrease). We always arrive at $\gcd(a, b)$ as the last nonzero remainder.

Remark 1.8.2

Note that in our example we have formed the gcd of 803 and 154 *without* factoring 803 and 154. This method is generally much faster than factoring and can find gcd's when factoring is not feasible.

Exercises

Exercise 1.8.1

Prove that if $a > 0$ then $\gcd(a, a) = a$.

Exercise 1.8.2

Let $a > b > 0$. Show that $\gcd(a, b) = \gcd(b, a \bmod b)$.

(Hint: compare this statement to [Lemma 1.8.2](#).)

Remark 1.8.3

So if your calculator can compute $a \bmod b$ you may use it when executing the Euclidean Algorithm.

Exercise 1.8.3

Find $\gcd(a, b)$ using the Euclidean Algorithm for each of the values below:

1. $a = 37, b = 60$
2. $a = 793, b = 3172$
3. $a = 25174, b = 42722$
4. $a = 377, b = 233$

Exercise 1.8.4

Use the Euclidean Algorithm to show that if n is an odd integer, then each fraction of the form

$$\frac{2n+2}{3n+2},$$

like $\frac{4}{5}$ and $\frac{8}{11}$, is already in lowest terms. (Hint: What is the greatest common divisor of the numerator and denominator?)²

Exercise 1.8.5

Write a paragraph describing the similarities and differences of the Euclidean Algorithm and the algorithm for converting a number into its base b representation.

Following are two small projects you may enjoy:

Exercise 1.8.6

- a. Carry out the Euclidean Algorithm to find $\gcd(a, b)$ when $a = 13$ (always) and b is each number in the set $\{1, 2, \dots, 12\}$ in turn. For which b 's does the Euclidean Algorithm take the most steps to produce $\gcd(13, b)$?

- b. The numbers that led to longer procedures in the last part have something in common. If you're not yet familiar with them, do an internet search to learn about "Fibonacci numbers." How are Fibonacci numbers connected to what you observed in Part 1 of this exercise?
- c. Do you have a prediction for which b 's will lead to the longest Euclidean Algorithm procedures if we compute $\gcd(21, b)$? Test it out!
- d. Now do an internet search for webpages containing both the phrases "Euclidean algorithm" and "Fibonacci numbers". What do you find out?

? Exercise 1.8.7

Here is a simple game: Starting with distinct positive integers a and b on a sheet of paper, two players take turns trying to write a new number on the sheet, subject to the conditions that (1) the number does not already appear on the paper, and (2) the number is a positive number that is the difference of two numbers already on the paper.

For example, if the starting numbers were 21 and 6, the first player could add "15" to the paper, because $21 - 6 = 15$. The second player could then write "9" (which equals $15 - 6$), and the first player could next write "12" (which is $21 - 9$) or "3" (which is $9 - 6$), and the game would continue.

The game ends when a player cannot find another number to add to the sheet of paper; that player loses the game. Answer the following questions.

- a. If the game described above continued, who would win the game? Does it matter what choices are made by the players on their turns?
- b. When the game above ends, which numbers are listed on the sheet of paper? What patterns do you see in this collection of numbers?
- c. If the same game is played starting with any two distinct positive integers a and b , the smallest integer on the paper at the end of the game has a special relationship to a and b . What is this relationship, and why does it happen? (Hint: think of the results of this chapter!)
- d. Can you devise a strategy of deciding which player to be (the first or the second player) if you wanted to be certain of winning the game, if you knew what a and b were? What would it be? Why would it work?

Footnotes

[1] Unlike the Division Algorithm, the Euclidean Algorithm really is an algorithm!

[2] This problem is adapted from a 2017 monthly contest of the Berkeley Math Circle. Accessed at <https://mathcircle.berkeley.edu/site...files/handouts/mc/problems/2017/mc4.pdf> on August 17, 2021.

This page titled 1.8: The Euclidean Algorithm is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Mike Barrus & W. Edwin Clark.

1.9: Bezout's Lemma

In [Definition 1.4.5](#) we defined a linear combination of two integers a and b to an integer of the form $as + bt$. If we keep track of the linear combinations of two integers like 9 and 24, what do we see?

Evaluating $9s + 24t$ for various choices of integers s and t , and keeping track of the results, shows that the numbers between 0 and 20 that are linear combinations of 9 and 24 are the following:

$$0, 3, 6, 9, 12, 15, 18.$$

One way to begin to convince yourself of this might be to notice that $3 = 9(-5) + 24(2)$, and then multiply both sides of this equation by integers to get the following:

$$\begin{aligned}0 &= 9(0) + 24(0); \\3 &= 9(-5) + 24(2); \\6 &= 9(-10) + 24(4); \\9 &= 9(-15) + 24(6); \\12 &= 9(-20) + 24(8); \\15 &= 9(-25) + 24(10); \\18 &= 9(-30) + 24(12).\end{aligned}$$

Continuing in this way will show you that every multiple of 3 is a linear combination of 9 and 24. (Can you see why things that are *not* multiples of 3 are *not* linear combinations of 9 and 24?)

Doing similar work for other pairs of 9 and another number, we might see that

- every multiples of 9 are the linear combinations of 0 and 9;
- every integer is a linear combination of 1 and 9;
- every integer is a linear combination of 2 and 9;
- the multiples of 3 are the linear combinations of 3 and 9;
- every integer is a linear combination of 4 and 9;
- every integer is a linear combination of 5 and 9;
- the multiples of 3 are the linear combinations of 6 and 9;
- every integer is a linear combination of 7 and 9;
- every integer is a linear combination of 8 and 9;
- the multiples of 9 are the linear combinations of 9 and 9.

There are some definite patterns here. First, the linear combinations of a and b always seem to be exactly the set of multiples of some positive integer (notice that the set of *every* integer is exactly the set of multiples of 1).

Next, the number whose multiples form the linear combinations of a and b appears to have a special relationships with a and b —looking at the examples above, something may stand out to you: the linear combinations of a and 9 turn out to be exactly the

- multiples of 9 if a is a multiple of 9,
- multiples of 3 if a is divisible by 3 but not by 9, and
- multiples of 1 if a is not divisible by 3.

The following important theorem explains what we have noticed. It will also be crucial to the proofs of some very important theorems in the coming chapters.

Lemma 1.9.1: Bezout's Lemma

For all integers a and b there exist integers s and t such that

$$\gcd(a, b) = sa + tb.$$

Furthermore, if a or b is nonzero, then $\gcd(a, b)$ is the smallest positive integer that can be written as $sa + tb$, and the linear combinations of a and b are precisely the multiples of $\gcd(a, b)$.

Proof

If $a = b = 0$ then s and t may be anything since

$$\gcd(0, 0) = 0 = s \cdot 0 + t \cdot 0.$$

So we may assume that $a \neq 0$ or $b \neq 0$. Let

$$J = \{na + mb : n, m \in \mathbb{Z}\}.$$

Note that J contains a , $-a$, b and $-b$ since

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

Since $a \neq 0$ or $b \neq 0$ one of the elements a , $-a$, b , $-b$ is positive. So we can say that J contains some positive integers. Let S denote the set of positive integers in J . That is,

$$S = \{na + mb : na + mb > 0, n, m \in \mathbb{Z}\}.$$

By the Well-Ordering Principle for $\mathbb{N}$, we know that S contains a smallest positive integer; call this integer d . Let's show that $d = \gcd(a, b)$.

First, we show that d is a common divisor of a and b . Note that since $d \in S$ we have $d = sa + tb$ for some integers, s and t . Note also that $d > 0$. To show $d \mid a$ using the Division Algorithm we write $a = dq + r$ where $0 \leq r < d$. Now

$$\begin{aligned} r &= a - dq \\ &= a - (sa + tb)q \\ &= (1 - sq)a + (-tq)b. \end{aligned}$$

Hence r is a linear combination of a and b , and $r \in J$. If $r > 0$ then $r \in S$. But this cannot be since $r < d$ and d is the smallest integer in S . So we must have $r = 0$. That is, $a = dq$. Hence $d \mid a$. By a similar argument we can show that $d \mid b$. Thus, d is indeed a common divisor of a and b .

We now show that d is the *greatest* common divisor. Let $e = \gcd(a, b)$. Then $e \mid a$ and $e \mid b$, so by [Theorem 1.4.1\(3\)](#), $e \mid (sa + tb)$, that is, $e \mid d$. Since e and d are positive, by [Theorem 1.4.1\(10\)](#), we have $e \leq d$. On the other hand, since d is a common divisor of a and b , we also know that $e \geq d$. Hence $d = \gcd(a, b)$.

Now suppose that c is any linear combination of a and b . By [Theorem 1.4.1\(3\)](#), we know that $d \mid c$, so c is a multiple of d . On the other hand, if h is any multiple of d , then $h = dk$ for some integer k , and multiplying both sides of the equation $d = sa + tb$ by k yields

$$h = dk = (sk)a + (tk)b,$$

so h is indeed a linear combination of a and b . We have thus shown that the linear combinations of a and b are precisely the multiples of $\gcd(a, b)$.

✓ Example 1.9.1

$1 = \gcd(2, 3)$ and we have $1 = (-1)2 + 1 \cdot 3$. Also we have $1 = 2 \cdot 2 + (-1)3$. So the numbers s and t in Bezout's Lemma are *not* uniquely determined. In fact, as we will see later there are *infinitely* many choices for s and t for each pair a, b .

⚙ Remark 1.9.1

The above proof is an *existence* theorem. It asserts the existence of s and t , but does not provide a way to actually find s and t . Also *the proof* does not give any clue about how to go about calculating s and t . We will give two algorithms in the next chapter for finding s and t .

We end this chapter with the first two of several consequences of Bezout's Lemma, one about the greatest common divisor and the other about the least common multiple. As in [Section 1.7](#), let $C(a, b)$ be the set of all common divisors of a and b .

Corollary 1.9.1

Every element of $C(a, b)$ divides $\gcd(a, b)$.

Proof

If c is a common divisor of a and b , then by [Theorem 1.4.1\(3\)](#), c divides every linear combination of a and b . By Bezout's Lemma, d is a linear combination of a and b , so $c \mid d$.

Proposition 1.9.1

If a and b are positive integers, then their least common multiple satisfies

$$\text{lcm}(a, b) = \frac{ab}{\gcd(a, b)}.$$

Proof

Let $q = \frac{ab}{\gcd(a, b)}$; our goal is show that $\text{lcm}(a, b) = q$.

Since $\gcd(a, b)$ is a common divisor of a and of b , there exist integers v and w such that $a = v \gcd(a, b)$ and $b = w \gcd(a, b)$. Then

$$q = \frac{ab}{\gcd(a, b)} = bv = aw,$$

which shows that q is a common multiple of a and b . Since $\text{lcm}(a, b)$ is the smallest of the positive common multiples, we have that $\text{lcm}(a, b) \leq q$.

Since $\text{lcm}(a, b)$ is a common multiple of a and b , there exist integers j and k such that

$$\text{lcm}(a, b) = ja = kb.$$

By Bezout's Lemma, there exist integers s and t such

$$\gcd(a, b) = as + bt.$$

Multiplying both sides by $\text{lcm}(a, b)$ and recalling how j , k , and q are defined, we find

$$\begin{aligned} \text{lcm}(a, b) \gcd(a, b) &= as \text{lcm}(a, b) + bt \text{lcm}(a, b) \\ &= as(kb) + bt(ja) \\ &= (sk + tj)ab \\ &= (sk + tj)q \gcd(a, b). \end{aligned}$$

Dividing both sides by $\gcd(a, b)$ yields $\text{lcm}(a, b) = (sk + tj)q$, so q divides $\text{lcm}(a, b)$. This implies that $q \leq \text{lcm}(a, b)$.

Since $\text{lcm}(a, b) \leq q$ and $\text{lcm}(a, b) \geq q$, we conclude that $\text{lcm}(a, b) = q$, and the proof is complete.

Exercises

Exercise 1.9.1

By trial and error, find coefficients s and t to make the following equations true:

1. $60s + 35t = \gcd(60, 35)$
2. $42s + 28t = \gcd(42, 28)$

? Exercise 1.9.2

Find three different pairs of integers s, t that all make the equation $3s + 11t = 1$.

? Exercise 1.9.3

1. True or false, and why? If s and t are integers and $3s + 11t = 1$, then $\gcd(s, t) = 1$.
2. If $s = 9$ and $t = -24$ and $19s + 7t = 3$, does this mean that $\gcd(19, 7) = 3$? Whether the answer is yes or no, explain what Bezout's Lemma says about the integers 19 and 7.

? Exercise 1.9.4

Using [Proposition 1.9.1](#), compute the following.

- a. $\text{lcm}(803, 154)$
- b. $\text{lcm}(11, 123)$
- c. $\text{lcm}(2185, 3059)$

? Exercise 1.9.5

According to the Summer 2017 issue of *QuadAngles*, a publication of the University of Rhode Island Alumni Association, knowing your least common multiples and greatest common divisors could help you get into college in the University's early days:

Here are the requirements for admission to the Rhode Island College of Agriculture and Mechanic Arts for September 1893. The College had opened the year before; this is the earliest admission test we could find.

ARITHMETIC

1. Find the L. C. M. and G. C. D. of 724 and 896. ...

Find the answers to this question.

? Exercise 1.9.6

Let $a > b > 0$. Show that if $a = bq + r$, then

$$\text{lcm}(a, b) = \frac{a}{r} \text{lcm}(b, r).$$

(Hint: Recall [Lemma 1.8.2](#).)

This page titled [1.9: Bezout's Lemma](#) is shared under a [not declared](#) license and was authored, remixed, and/or curated by [Mike Barrus & W. Edwin Clark](#).

1.10: Computing Coefficients for Bezout's Lemma

Though the proof of Bezout's Lemma in the last chapter simply showed that, given integers a and b , coefficients s and t exist such that $\gcd(a, b) = sa + tb$, suitable modifications of the Euclidean Algorithm give us ways of computing these coefficients. In this chapter we discuss two such ways, known as the Extended Euclidean Algorithm and Blankinship's Method.

The Extended Euclidean Algorithm

With the Extended Euclidean Algorithm, the idea is to rearrange and substitute in the equations arising from the Division Algorithm during the Euclidean Algorithm, so as to arrive at an expression of $\gcd(a, b)$ as a linear combination of a and b .

Suppose that the Euclidean Algorithm produces the following equations:

$$\begin{aligned} a &= bq_0 + r_0; \\ b &= r_0q_1 + r_1; \\ r_0 &= r_1q_2 + r_2; \\ &\vdots \\ r_{k-2} &= r_{k-1}q_k + r_k; \\ r_{k-1} &= r_kq_{k+1}. \end{aligned}$$

Assuming that r_k is the last nonzero remainder, we know that $\gcd(a, b) = r_k$. Suppose that we omit the last equation above and solve the rest of these equations for the remainder term. If we then order them from last to first, we get a new sequence of equations (we show a few more of the last equations in the list):

$$\begin{aligned} r_k &= r_{k-1}q_k - r_{k-2}; \\ r_{k-1} &= r_{k-2}q_{k-1} - r_{k-3}; \\ r_{k-2} &= r_{k-3}q_{k-2} - r_{k-4}; \\ &\vdots \\ r_2 &= r_1q_2 - r_0; \\ r_1 &= r_0q_1 - b; \\ r_0 &= bq_0 - a; \end{aligned}$$

Now look at the first of these new equations. Since the Euclidean Algorithm produces $\gcd(a, b)$, we can actually write

$$\gcd(a, b) = r_{k-1}q_k - r_{k-2}. \quad (1.10.1)$$

Here we see that $\gcd(a, b)$ is a linear combination of r_{k-1} and r_{k-2} . Since we'd rather have $\gcd(a, b)$ as a linear combination of something else (namely, a and b), it would be nice to replace r_{k-1} and r_{k-2} by expressions involving a and b . It will take some steps, but this is exactly what we will do.

Notice that the second equation, $r_{k-1} = r_{k-2}q_{k-1} - r_{k-3}$, tells us what r_{k-1} equals. Let's use this in equation (1.10.1), substituting for r_{k-1} ; this yields

$$\gcd(a, b) = (r_{k-2}q_{k-1} - r_{k-3})q_k - r_{k-2}, \quad (1.10.2)$$

which we can rearrange as

$$\gcd(a, b) = (q_{k-1}q_k - 1)r_{k-2} - q_k r_{k-3}.$$

At this point $\gcd(a, b)$ is a linear combination of r_{k-2} and r_{k-3} (rather than a combination of r_{k-1} and r_{k-2} , as it was previously). If we make another substitution using the equation $r_{k-2} = r_{k-3}q_{k-2} - r_{k-4}$, we will get

$$\begin{aligned} \gcd(a, b) &= (q_{k-1}q_k - 1)(r_{k-3}q_{k-2} - r_{k-4}) - q_k r_{k-3} \\ &= (q_{k-2}q_{k-1}q_k - q_{k-2} - q_k)r_{k-3} - (q_{k-1}q_k - 1)r_{k-4}, \end{aligned}$$

in which $\gcd(a, b)$ is now in terms of r_{k-3} and r_{k-4} .

At this point the equations are getting messy, so we won't continue the substitutions (instead, we will work an example below). The idea, however, is simple; by working our way backwards through the equations produced by the Euclidean Algorithm, we can express $\gcd(a, b)$ as linear combinations that get progressively closer to the linear combination of a and b that we need.

✓ Example 1.10.1

For our example, we will continue with the numbers used in [Example 1.8.1](#) from [Section 1.8](#). There we found that $\gcd(803, 154) = 11$. Bezout's Lemma guarantees that $11 = 803s + 154t$ for some integers s, t . Let us determine a possible choice of s and t .

In [Example 1.8.1](#) we found the following:

$$\begin{aligned} 803 &= 154 \cdot 5 + 33 \\ 154 &= 33 \cdot 4 + 22 \\ 33 &= 22 \cdot 1 + 11 \\ 22 &= 11 \cdot 2 + 0. \end{aligned}$$

Rearranging and reordering all but the last of these equations, we get

$$\begin{aligned} 11 &= 33 - 22 \cdot 1 \\ 22 &= 154 - 33 \cdot 4 \\ 33 &= 803 - 154 \cdot 5 \end{aligned}$$

We now set about to find an expression of 11 as a linear combination of 803 and 154. Starting with the equation $11 = 33 - 22 \cdot 1$, we substitute $22 = 154 - 33 \cdot 4$ (the next equation) to obtain

$$\begin{aligned} 11 &= 33 - (154 - 33 \cdot 4) \cdot 1 \\ &= 5 \cdot 33 - 1 \cdot 154. \end{aligned}$$

Now we will substitute for 33, using the equation $33 = 803 - 154 \cdot 5$ above:

$$11 = 5 \cdot (803 - 154 \cdot 5) - 1 \cdot 154$$

and finally obtain

$$\boxed{11 = 5 \cdot 803 - 26 \cdot 154.}$$

We have arrived at an expression for $\gcd(803, 154)$ as a linear combination of 803 and 154, as described by Bezout's Lemma, with $s = 5$ and $t = -26$ in the linear combination.

Blankinship's Method

One feature of the Extended Euclidean Algorithm just discussed is that in order to find the coefficients s and t in Bezout's Lemma, we must keep track of quotients and remainders used in carrying out the (original) Euclidean Algorithm, so that we can find s and t by "working backwards" through this information once the basic Euclidean Algorithm has ended.

In this section we see a simple approach to finding s and t that appeared in an article by W.A. Blankinship in the August-September 1963 issue of the *American Mathematical Monthly*.¹ As opposed to the Extended Euclidean Algorithm, Blankinship's method produces the integers s and t in Bezout's Lemma at the same time it produces $\gcd(a, b)$.

Given $a > b > 0$ we start with the array

$$\begin{bmatrix} a & 1 & 0 \\ b & 0 & 1 \end{bmatrix}$$

Then we continue to add multiples of one row to another row, alternating choice of rows until we reach an array of the form

$$\begin{bmatrix} 0 & x_1 & x_2 \\ d & y_1 & y_2 \end{bmatrix}$$

or

$$\begin{bmatrix} d & y_1 & y_2 \\ 0 & x_1 & x_2 \end{bmatrix}$$

In other words, the goal is to get a 0 in the first column. Once we reach one of these forms, it will be true that $d = \gcd(a, b) = y_1 a + y_2 b$.

✓ Example 1.10.2

First take $a = 35, b = 15$.

$$\begin{bmatrix} 35 & 1 & 0 \\ 15 & 0 & 1 \end{bmatrix}$$

Note $35 = 15 \cdot 2 + 5$, hence

$$35 + 15(-2) = 5.$$

So we multiply row 2 by -2 and add it to row 1, getting

$$\begin{bmatrix} 5 & 1 & -2 \\ 15 & 0 & 1 \end{bmatrix}$$

Now $3 \cdot 5 = 15$ or $15 + (-3)5 = 0$, so we multiply row 1 by -3 and add it to row 2, getting

$$\begin{bmatrix} 5 & 1 & -2 \\ 0 & -3 & 7 \end{bmatrix}.$$

Now we can say that

$$\boxed{\gcd(35, 15) = 5}$$

and

$$\boxed{5 = 1 \cdot 35 + (-2) \cdot 15.}$$

Let's now consider a more complicated example: Take $a = 1876, b = 365$.

$$\begin{bmatrix} 1876 & 1 & 0 \\ 365 & 0 & 1 \end{bmatrix}$$

Now $1876 = 365 \cdot 5 + 51$ so we add -5 times the second row to the first row, getting:

$$\begin{bmatrix} 51 & 1 & -5 \\ 365 & 0 & 1 \end{bmatrix}$$

Now $365 = 51 \cdot 7 + 8$, so we add -7 times row 1 to row 2, getting:

$$\begin{bmatrix} 51 & 1 & -5 \\ 8 & -7 & 36 \end{bmatrix}$$

Now $51 = 8 \cdot 6 + 3$, so we add -6 times row 2 to row 1, getting:

$$\begin{bmatrix} 3 & 43 & -221 \\ 8 & -7 & 36 \end{bmatrix}$$

Now $8 = 3 \cdot 2 + 2$, so we add -2 times row 1 to row 2, getting:

$$\begin{bmatrix} 3 & 43 & -221 \\ 2 & -93 & 478 \end{bmatrix}$$

Then $3 = 2 \cdot 1 + 1$, so we add -1 times row 2 to row 1, getting:

$$\begin{bmatrix} 1 & 136 & -699 \\ 2 & -93 & 478 \end{bmatrix}$$

Finally, $2 = 1 \cdot 2$ so if we add -2 times row 1 to row 2 we get:

$$\begin{bmatrix} 1 & 136 & -699 \\ 0 & -365 & 1876 \end{bmatrix}. \quad (1.10.3)$$

This tells us that

$$\gcd(1876, 365) = 1$$

and

$$1 = 136 \cdot 1876 + (-699)365. \quad (1.10.4)$$

Note that it was not necessary to compute the last two entries -365 and 1876 in equation (1.10.3). It is a good idea however to check that equation (1.10.4) holds. In this case we have:

$$\begin{array}{rcl} 136 \cdot 1876 & = & 255136 \\ \underline{(-699) \cdot 365} & = & \underline{-255135} \\ & & 1 \end{array}$$

So it is correct.

Why Blankinship's Method works: Looking just at what happens in the first column, note that we are carrying out the Euclidean Algorithm, so when one element in column 1 is 0, the other is, in fact, the gcd. Note that at the start we have

$$\begin{bmatrix} a & 1 & 0 \\ b & 0 & 1 \end{bmatrix}$$

and

$$\begin{aligned} a &= 1 \cdot a + 0 \cdot b \\ b &= 0 \cdot a + 1 \cdot b. \end{aligned}$$

One can show that at every intermediate step

$$\begin{bmatrix} a_1 & x_1 & x_2 \\ b_1 & y_1 & y_2 \end{bmatrix}$$

we always have

$$\begin{aligned} a_1 &= x_1 a + x_2 b \\ b_1 &= y_1 a + y_2 b, \end{aligned}$$

and the result follows. We will omit the details.

Exercises

? Exercise 1.10.1

Use the Extended Euclidean Algorithm to compute the s and t in Bezout's Lemma for each of the following values of a and b .

1. $a = 267, b = 112$
2. $a = 216, b = 135$
3. $a = 11312, b = 11321$

? Exercise 1.10.2

Use Blankinship's Method to compute the s and t in Bezout's Lemma for each of the pairs a, b of integers in the previous exercise.

? Exercise 1.10.3

Show that if $1 = as + bt$ then $\gcd(a, b) = 1$.

? Exercise 1.10.4

Show that whenever s, t are integers making $as + bt = \gcd(a, b)$, then $\gcd(s, t) = 1$.

? Exercise 1.10.5

- a. Find integers a, b, d, s, t such that all of the following hold
 - i. $a > 0, b > 0$,
 - ii. $d = sa + tb$, and
 - iii. $d \neq \gcd(a, b)$.
- b. Explain why the d you chose in part (a) cannot be 1.
- c. What does your example (i.e., the integers a, b, d, s, t) in part (a) show you about linear combinations of a and b ?

Footnotes

[1] Thanks to Chris Miller for bringing this method to my attention|W.E. Clark.

This page titled [1.10: Computing Coefficients for Bezout's Lemma](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Mike Barrus & W. Edwin Clark](#).

1.11: Prime Numbers

Definition 1.11.1

An integer p is *prime* if $p \geq 2$ and the only positive divisors of p are 1 and p . An integer n is *composite* if $n \geq 2$ and n is not prime.

Remark 1.11.1

The number 1 is neither prime nor composite.

Lemma 1.11.1

An integer $n \geq 2$ is composite if and only if there are integers a and b such that $n = ab$, $1 < a < n$, and $1 < b < n$.

Proof

Let $n \geq 2$. If n is composite there is a positive integer a such that $a \neq 1$, $a \neq n$ and $a \mid n$. This means that $n = ab$ for some b . Since n and a are positive so is b . Hence $1 \leq a$ and $1 \leq b$. By [Theorem 1.4.1\(10\)](#) $a \leq n$ and $b \leq n$. Since $a \neq 1$ and $a \neq n$ we have $1 < a < n$. If $b = 1$ then $a = n$, which is not possible, so $b \neq 1$. If $b = n$ then $a = 1$, which is also not possible. So $1 < b < n$. The converse is true by the definition of composite numbers.

Lemma 1.11.2

If $n > 1$, there is a prime p such that $p \mid n$.

Proof

Assume there is some integer $n > 1$ which has no prime divisor. Let S denote the set of all such integers. By the Well-Ordering Principle there is a smallest such integer, call it m . Now $m > 1$ and has no prime divisor. So m cannot be prime. Hence m is composite. Therefore by [Lemma 1.11.1](#)

$$m = ab, \quad 1 < a < m, \quad 1 < b < m.$$

Since $1 < a < m$ then a is not in the set S . So a must have a prime divisor, call it p . Then $p \mid a$ and $a \mid m$ so by [Theorem 1.4.1](#), $p \mid m$. This contradicts the fact that m has no prime divisor. So the set S must be empty and this proves the lemma.

Theorem 1.11.1: Euclid's Theorem¹

There are infinitely many prime numbers.

Proof

Assume, by way of contradiction, that there are only a finite number of prime numbers, say:

$$p_1, p_2, \dots, p_n.$$

Define

$$N = p_1 p_2 \cdots p_n + 1.$$

Since $p_1 \geq 2$, clearly $N \geq 3$. So by Lemma 10.2 N has a prime divisor p . By assumption $p = p_i$ for some $i = 1, \dots, n$. Let $a = p_1 \cdots p_n$. Note that

$$a = p_i (p_1 p_2 \cdots p_{i-1} p_{i+1} \cdots p_n),$$

so $p_i \mid a$. Now $N = a + 1$ and by assumption $p_i \mid (a + 1)$. So by [Exercise 1.4.3](#) $p_i \mid ((a + 1) - a)$, that is $p_i \mid 1$. By Basic Axiom 3 in [Section 1.2](#) this implies that $p_i = 1$. This contradicts the fact that primes are > 1 . It follows that the assumption that there are only finitely many primes is not true.

Theorem 1.11.2

If $n > 1$ is composite then n has a prime divisor $p \leq \sqrt{n}$.

Proof

Let $n > 1$ be composite. Then $n = ab$ where $1 < a < n$ and $1 < b < n$. We claim that one of a or b is $\leq \sqrt{n}$. If not, then $a > \sqrt{n}$ and $b > \sqrt{n}$. Hence $n = ab > \sqrt{n} \sqrt{n} = n$. This implies that $n > n$, a contradiction. So $a \leq \sqrt{n}$ or $b \leq \sqrt{n}$. Suppose $a \leq \sqrt{n}$. Since $1 < a$, by [Lemma 1.11.2](#) there is a prime p such that $p \mid a$. Hence, by [Theorem 1.4.1](#) since $a \mid n$ we have $p \mid n$. Also by [Theorem 1.4.1](#) since $p \mid a$ we have $p \leq a \leq \sqrt{n}$.

The proof of Euclid's Theorem hints at a way of generating a new prime number, given a list of already-known prime numbers.² However, this method does not produce all primes in the order in which they occur among the integers. A more systematic way of producing primes, known as the *sieve of Eratosthenes*.³ The sieve of Eratosthenes works in this way:

To generate all the prime numbers less than or equal to n , we begin by writing out all positive integers from 2 through n . For instance, if we wish to find all prime numbers that are at most 100, we begin by writing

	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

(here the grid-like arrangement we have used is convenient but not necessary for the sieve to work).

We next repeatedly apply the following step:

Circle the first non-crossed out number; then cross out any other multiple of this number appearing in the list that hasn't already been crossed out.

We repeat this process until we cannot, i.e., until there are no more non-crossed out numbers to circle. The prime numbers are then the numbers that have been circled.

For example, in our example above we would begin by circling 2, the first number that has not been circled or crossed out. We would then cross out all multiples of 2 in the list other than 2. This produces the following figure:

	<u>2</u>	3	~~4~~	5	~~6~~	7	~~8~~	9	~~10~~
11	~~12~~	13	~~14~~	15	~~16~~	17	~~18~~	19	~~20~~
21	~~22~~	23	~~24~~	25	~~26~~	27	~~28~~	29	~~30~~
31	~~32~~	33	~~34~~	35	~~36~~	37	~~38~~	39	~~40~~
41	~~42~~	43	~~44~~	45	~~46~~	47	~~48~~	49	~~50~~
51	~~52~~	53	~~54~~	55	~~56~~	57	~~58~~	59	~~60~~
61	~~62~~	63	~~64~~	65	~~66~~	67	~~68~~	69	~~70~~

71	~~72~~	73	~~74~~	75	~~76~~	77	~~78~~	79	~~80~~
81	~~82~~	83	~~84~~	85	~~86~~	87	~~88~~	89	~~90~~
91	~~92~~	93	~~94~~	95	~~96~~	97	~~98~~	99	~~100~~

The next step is to circle 3 (the first number not already circled or crossed out) and cross out any multiples of 3 (like 9, 15, etc.) not already crossed out. When this step is finished, our list looks like this:

	<u>2</u>	<u>3</u>	~~4~~	5	~~6~~	7	~~8~~	~~9~~	~~10~~
11	~~12~~	13	~~14~~	~~15~~	~~16~~	17	~~18~~	19	~~20~~
~~21~~	~~22~~	23	~~24~~	25	~~26~~	~~27~~	~~28~~	29	~~30~~
31	~~32~~	~~33~~	~~34~~	35	~~36~~	37	~~38~~	~~39~~	~~40~~
41	~~42~~	43	~~44~~	~~45~~	~~46~~	47	~~48~~	49	~~50~~
~~51~~	~~52~~	53	~~54~~	55	~~56~~	~~57~~	~~58~~	59	~~60~~
61	~~62~~	~~63~~	~~64~~	65	~~66~~	67	~~68~~	~~69~~	~~70~~
71	~~72~~	73	~~74~~	~~75~~	~~76~~	77	~~78~~	79	~~80~~
~~81~~	~~82~~	83	~~84~~	85	~~86~~	~~87~~	~~88~~	89	~~90~~
91	~~92~~	~~93~~	~~94~~	95	~~96~~	97	~~98~~	~~99~~	~~100~~

If we keep going until all numbers have either been circled or crossed out, our list will look like this:

	<u>2</u>	<u>3</u>	~~4~~	<u>5</u>	~~6~~	<u>7</u>	~~8~~	~~9~~	~~10~~
<u>11</u>	~~12~~	<u>13</u>	~~14~~	~~15~~	~~16~~	<u>17</u>	~~18~~	<u>19</u>	~~20~~
~~21~~	~~22~~	<u>23</u>	~~24~~	~~25~~	~~26~~	~~27~~	~~28~~	<u>29</u>	~~30~~
<u>31</u>	~~32~~	~~33~~	~~34~~	~~35~~	~~36~~	<u>37</u>	~~38~~	~~39~~	~~40~~
<u>41</u>	~~42~~	<u>43</u>	~~44~~	~~45~~	~~46~~	<u>47</u>	~~48~~	~~49~~	~~50~~
~~51~~	~~52~~	<u>53</u>	~~54~~	~~55~~	~~56~~	~~57~~	~~58~~	<u>59</u>	~~60~~
<u>61</u>	~~62~~	~~63~~	~~64~~	~~65~~	~~66~~	<u>67</u>	~~68~~	~~69~~	~~70~~
<u>71</u>	~~72~~	<u>73</u>	~~74~~	~~75~~	~~76~~	~~77~~	~~78~~	<u>79</u>	~~80~~
~~81~~	~~82~~	<u>83</u>	~~84~~	~~85~~	~~86~~	~~87~~	~~88~~	<u>89</u>	~~90~~
~~91~~	~~92~~	~~93~~	~~94~~	~~95~~	~~96~~	<u>97</u>	~~98~~	~~99~~	~~100~~

Observe how the name of this technique is descriptive—by crossing out all the multiples of the circled numbers, we let these numbers slip away from us as if through a sieve, while we can imagine the circled primes getting caught in the sieve. In our example, the numbers preserved by the sieve of Eratosthenes, that is, the prime numbers less than 100, are the numbers

2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97.

Definition 1.11.2

The *primality* of an integer $n \geq 2$ refers to its membership (or not) in the set of prime numbers. To check the primality of n means to determine whether n is prime.

Remark 1.11.2

We can use [Theorem 1.11.2](#) to help in checking the primality of an integer: To check whether or not $n > 1$ is prime we need only try to divide it by all primes $p \leq \sqrt{n}$. If none of these primes divides n then n must be prime.

Example 1.11.1

Consider the number 97. Note that $\sqrt{97} < \sqrt{100} = 10$. The primes ≤ 10 are 2, 3, 5, and 7. One easily checks that $97 \bmod 2 = 1$, $97 \bmod 3 = 1$, $97 \bmod 5 = 2$, $97 \bmod 7 = 6$. So none of the primes 2, 3, 5, 7 divide 97 and 97 is prime by [Theorem 1.11.2](#).

This gives us a slight shortcut to finding primes with the Sieve of Eratosthenes: in our example above, once we have circled 7 and crossed out its multiples in the example above, every other number currently in the list that has not yet been circled or crossed out is *guaranteed to be prime* and can immediately be circled, since 7 is the largest prime number that is less than or equal to the square root of 100.

Definition 1.11.3

Let $x \in \mathbb{R}$, $x > 0$. The *prime-counting function* $\pi(x)$ is defined as the number of primes p such that $p \leq x$.

For example, since the only primes $p \leq 10$ are 2, 3, 5, and 7 we have $\pi(10) = 4$.

Here is a table of values of $\pi(10^i)$ for $i = 2, \dots, 10$. We also include known approximations to $\pi(x)$. Note that the formulas for the approximations do not give integer values, but for the table we have rounded each to the nearest integer. The values in the table were computed using the computer algebra system Maple.

x	$\pi(x)$	$\frac{x}{\ln(x)}$	$\int_2^x \frac{1}{\ln(t)} dt$
10^2	25	22	29
10^3	168	145	177
10^4	1229	1086	1245
10^5	9592	8686	9629
10^6	78498	72382	78627
10^7	664579	620421	664917
10^8	5761455	5428681	5762208
10^9	50847534	48254942	50849234
10^{10}	455052511	434294482	455055614

You may judge for yourself which approximations appear to be the best. This table has been continued⁴ up to 10^{28} , but undoubtedly people are still working on finding the value of $\pi(10^n)$ for larger n . Of course, the approximations are easy to compute with Maple or a similar computational tool, but the exact value of $\pi(10^{28})$ is difficult to find.

Why do the functions defined by $\frac{x}{\ln x}$ and $\int_2^x \frac{dt}{\ln t}$ do as well as they do in approximating $\pi(x)$? This is not an easy question to answer. The relationship between $\pi(x)$ and $\int_2^x \frac{dt}{\ln t}$ was conjectured by Gauss in 1793. More than 100 years later, Hadamard and de la Vallée Poussin, independently of each other, successfully completed proofs of this result. Here is an equivalent statement.

Theorem 1.11.3: The Prime Number Theorem

$$\pi(x) \sim \frac{x}{\ln(x)} \quad \text{for } x > 0.$$

Remark 1.11.3

The $\sim$ in the Prime Number Theorem means that

$$\lim_{x \rightarrow \infty} \frac{\pi(x)}{\frac{x}{\ln(x)}} = 1.$$

For a connection between $\frac{x}{\ln x}$ and $\int_2^x \frac{dt}{\ln t}$, see [Exercise 1.11.11](#).

The distribution of the prime numbers among the integers is in some respects still mysterious. Although there are infinitely many primes, there are long stretches of consecutive integers containing no primes.

Theorem 1.11.4

For any positive integer n there is an integer a such that the n consecutive integers

$$a, a+1, a+2, \dots, a+(n-1)$$

are all composite.

Proof

Given $n \geq 1$ let $a = (n+1)! + 2$. (Factorials were defined in [Exercise 1.3.5](#).) We claim that all the numbers

$$a, a+1, a+2, \dots, a+(n-1)$$

are composite. Since $(n+1) \geq 2$ clearly $2 \mid (n+1)!$ and $2 \mid 2$. Hence $2 \mid ((n+1)! + 2)$. Since $(n+1)! + 2 > 2$, $(n+1)! + 2$ is composite. Consider

$$a+i = (n+1)! + i + 2$$

where $0 \leq i \leq n-1$ so $2 \leq i+2 \leq n+1$. Thus $i+2 \mid (n+1)!$ and $i+2 \mid (i+2)$. Therefore $i+2 \mid (a+i)$. Now $a+i > i+2 > 1$, so $a+i$ is composite.

On the other hand, statements such as the *Twin Prime Conjecture* describe how *short* a list of consecutive composite numbers might be guaranteed to be. Twin primes are prime numbers that differ by exactly 2. Examples of “twin primes” include 5 and 7, 29 and 31, and 1427 and 1429. Between twin primes, there is only one composite number, but do twin primes eventually stop happening?

The Twin Prime Conjecture

There are infinitely many pairs of twin primes.

Though this conjecture remains unproven, in April 2013 mathematician Yitang Zhang announced a proof that there are infinitely many pairs of primes that differ by 70 million or less.⁵ Following up on these ideas, one year later The Polymath Project⁶ improved the result to show that there are infinitely many pairs of primes that differ by no more than 246.

Another conjecture that has ties to the locations of prime numbers is Goldbach’s Conjecture.

Goldbach’s Conjecture

Every even integer $n > 2$ is the sum of two primes.

This conjecture arose in correspondence from 1742 between Christian Goldbach and Leonhard Euler. It has been verified by T. Oliveira e Silva, S. Herzog, and S. Pardi⁷ for all even integers not larger than $4 \cdot 10^{18}$ but remains unproven.

There is even a connection between the distribution of the prime numbers (in particular, how rough of an approximation the functions in the Prime Number Theorem are for $\pi(x)$) and the question that is perhaps most famous among all currently unsolved problems of mathematics: is the Riemann Hypothesis true? We will not discuss the Riemann Hypothesis here, but you can read about it (and its connection to approximations of $\pi(x)$) with a bit of research.⁸

We have covered a lot of ground in this chapter. Truly, even thousands of years after Euclid and Eratosthenes studied them, the prime numbers remain a rich source of understandable but as-yet-unsolved mysteries.

Exercises

? Exercise 1.11.1

Prove that 2 is the only even prime number. (Joke: Hence it is said that 2 is the “oddest” prime.)

? Exercise 1.11.2

Show that every prime number other than 2 or 3 is either one less than, or one more than, a multiple of 6. ([Appendix A](#) organizes the primes less than 200 into cases based on this fact.)

? Exercise 1.11.3

Show that if p and q are primes and $p \mid q$, then $p = q$.

? Exercise 1.11.4

Use the idea of the proof of Euclid’s Theorem ([Theorem 1.11.1](#)) to show that if $q_1, q_2, \dots, q_n$ are *any* primes (not necessarily the smallest primes, or consecutive ones), then there is a prime $q \notin \{q_1, \dots, q_n\}$.

(Hint: Take $N = q_1 \cdots q_n + 1$; by [Lemma 1.11.2](#) there is a prime q such that $q \mid N$. Prove that $q \notin \{q_1, \dots, q_n\}$.)

? Exercise 1.11.5

Let $p_1 = 2, p_2 = 3, p_3 = 5, \dots$ and, in general, $p_i =$ the i -th prime. Prove or disprove that

$$p_1 p_2 \cdots p_n + 1$$

is prime for all $n \geq 1$.

(Hint: If $n = 1$ we have $2 + 1 = 3$ is prime. If $n = 2$ we have $2 \cdot 3 + 1 = 7$ is prime. If $n = 3$ we have $2 \cdot 3 \cdot 5 + 1 = 31$ is prime. Try the next few values of n . You may want to use [Theorem 1.11.2](#) to check primality.)

? Exercise 1.11.6

By using [Theorem 1.11.2](#) as in [Example 1.11.1](#), determine the primality of the following integers:

$$143, \quad 221, \quad 199, \quad 223, \quad 3521.$$

? Exercise 1.11.7

- Imitating the proof of [Theorem 1.11.2](#), show that if $n = abc$ where $a > 1$, $b > 1$, and $c > 1$, then one of a, b, c is less than or equal to $\sqrt[3]{n}$.
- Comparing [Theorem 1.11.2](#) and part (a), come up with a generalization for when n is the product of k integers, each greater than 1.

? Exercise 1.11.8

Use the sieve of Eratosthenes to find all prime numbers less than 300, showing your work.

? Exercise 1.11.9

Prove that while the sieve of Eratosthenes is performed on a list $2, \dots, n$ of integers, when a number k is circled, the first multiple of k that then needs to be crossed out (because it hasn't already been crossed out) is k^2 .

? Exercise 1.11.1

Use the Prime Number Theorem and a calculator to approximate the number of primes $\leq 10^8$. Note $\ln(10^8) = 8 \ln(10)$.

? Exercise 1.11.11

Use l'Hospital's rule and the Fundamental Theorem of Calculus (with integration by parts) to show that

$$\lim_{x \rightarrow \infty} \frac{\int_2^x \frac{dt}{\ln t}}{\frac{x}{\ln x}} = 1,$$

and use this and [Theorem 1.11.3](#) to show that

$$\pi(x) \sim \int_2^x \frac{dt}{\ln t}.$$

? Exercise 1.11.12

Find 10 consecutive composite numbers.

? Exercise 1.11.13

- Is $2^n - 1$ always prime if $n \geq 2$? Explain.
- Is $2^n - 1$ always prime if n is prime? Explain.

Footnotes

[1] This theorem appears as Proposition 20 in Book IX of Euclid's Elements, so a proof of this fact has been passed down for over 2000 years.

[2] If we repeatedly apply the proof of Euclid's Theorem, starting with $p_1 = 2$ and each time letting p_{n+1} be the smallest prime divisor when N is created from $p_1, \dots, p_n$, then the order in which the prime numbers are generated, beginning with 2, 3, 7, 43 etc., is not well understood. For some notes on this ordering of primes, visit the entry on it at the Online Encyclopedia of Integer Sequences: see <http://oeis.org/A000945>.

[3] This method was attributed by an ancient writer to the Greek mathematician Eratosthenes, who lived in the 3rd century BC (also in Alexandria, though perhaps a little after Euclid).

[4] The value $\pi(10^{28})$ was announced to be 157,589,269,275,973,410,412,739,598 by David Baugh and Kim Walisch in August 2020 (see <https://mersenneforum.org/showthread.php?p=555442>).

[5] For an engaging video on the breakthrough, see Numberphile's "Gaps between Primes" at <https://www.youtube.com/watch?v=vkMXdShDdtY>.

[6] <https://polymathprojects.org/>

[7] "Empirical verification of the even Goldbach Conjecture and computation of prime gaps up to $4 \cdot 10^{18}$," *Mathematics of Computation*, Vol. 83, No. 288 (July 2014), pp. 2033-2060.

[8] One source to consult is at <https://www.simonsfoundation.org/202...nn-hypothesis/>.

This page titled [1.11: Prime Numbers](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Mike Barrus & W. Edwin Clark](#).

1.12: Unique Factorization

Our goal in this chapter is to prove the following fundamental theorem.

Theorem 1.12.1: The Fundamental Theorem of Arithmetic

Every integer $n > 1$ can be written uniquely in the form

$$n = p_1 p_2 \cdots p_s,$$

where s is a positive integer and $p_1, p_2, \dots, p_s$ are primes satisfying

$$p_1 \leq p_2 \leq \cdots \leq p_s.$$

Remark 1.12.1

If $n = p_1 p_2 \cdots p_s$ where each p_i is prime, we call this the *prime factorization* of n . [Theorem 1.12.1](#) is sometimes stated as follows:

Every integer $n > 1$ can be expressed as a product $n = p_1 p_2 \cdots p_s$, for some positive integer s , where each p_i is prime and this factorization is unique except for the order of the primes p_i .

Note for example that

$$\begin{aligned} 600 &= 2 \cdot 2 \cdot 2 \cdot 3 \cdot 5 \cdot 5 \\ &= 2 \cdot 3 \cdot 2 \cdot 5 \cdot 2 \cdot 5 \\ &= 3 \cdot 5 \cdot 2 \cdot 2 \cdot 2 \cdot 5 \\ &\text{etc.} \end{aligned}$$

Perhaps the nicest way to write the prime factorization of 600 is

$$600 = 2^3 \cdot 3 \cdot 5^2.$$

In general it is clear that $n > 1$ can be written uniquely in the form

$$n = p_1^{a_1} p_2^{a_2} \cdots p_s^{a_s}, \text{ some } s \geq 1,$$

where $p_1 < p_2 < \cdots < p_s$ and $a_i \geq 1$ for all i . Sometimes the product is written as

$$n = \prod_{i=1}^s p_i^{a_i}.$$

Here $\prod$ stands for product, just as $\sum$ stands for sum.

To prove [Theorem 1.12.1](#) we need to first establish a few lemmas.

Lemma 1.12.1

If $a \mid bc$ and $\gcd(a, b) = 1$ then $a \mid c$.

Proof

Since $\gcd(a, b) = 1$ by Bezout's Lemma there are s, t such that

$$1 = as + bt.$$

If we multiply both sides by c we get

$$c = cas + cbt = a(cs) + (bc)t.$$

By assumption $a \mid bc$. Clearly $a \mid a(cs)$ so, by [Theorem 1.4.1](#), a divides the linear combination $a(cs) + (bc)t = c$.

Definition 1.12.1

We say that a and b are *relatively prime* if $\gcd(a, b) = 1$.

So we may restate [Lemma 1.12.1](#) as follows: If $a \mid bc$ and a is relatively prime to b then $a \mid c$.

✓ Example 1.12.1

It is not true generally that when $a \mid bc$ then $a \mid b$ or $a \mid c$. For example, $6 \mid 4 \cdot 9$, but $6 \nmid 4$ and $6 \nmid 9$. Note that [Lemma 1.12.1](#) doesn't apply here since $\gcd(6, 4) \neq 1$ and $\gcd(6, 9) \neq 1$.

Lemma 1.12.2: Euclid's Lemma¹

If p is a prime and $p \mid ab$, then $p \mid a$ or $p \mid b$.

Proof

Assume that $p \mid ab$. If $p \mid a$ we are done. Suppose $p \nmid a$. Let $d = \gcd(p, a)$. Note that $d > 0$ and $d \mid p$ and $d \mid a$. Since $d \mid p$ we have $d = 1$ or $d = p$. If $d \neq 1$ then $d = p$. But this says that $p \mid a$, which we assumed was not true. So we must have $d = 1$. Hence $\gcd(p, a) = 1$ and $p \mid ab$. So by [Lemma 1.12.1](#), $p \mid b$.

Lemma 1.12.3

Let p be prime. Let $a_1, a_2, \dots, a_n$, $n \geq 1$, be integers. If $p \mid a_1 a_2 \cdots a_n$, then $p \mid a_i$ for at least one $i \in \{1, 2, \dots, n\}$.

Proof

We use induction on n , the number of integers multiplied in the product. The result is clear if $n = 1$. Assume that the lemma holds for n such that $1 \leq n \leq k$. Let's show it holds for $n = k + 1$. So assume p is a prime and $p \mid a_1 a_2 \cdots a_k a_{k+1}$. Let $a = a_1 a_2 \cdots a_k$ and $b = a_{k+1}$. Then $p \mid a$ or $p \mid b$ by [Lemma 1.12.2](#). If $p \mid a = a_1 \cdots a_k$, by the induction hypothesis, $p \mid a_i$ for some $i \in \{1, \dots, k\}$. If $p \mid b = a_{k+1}$ then $p \mid a_{k+1}$. So we can say $p \mid a_i$ for some $i \in \{1, 2, \dots, k + 1\}$. So the lemma holds for $n = k + 1$. Hence by PMI it holds for all $n \geq 1$.

Lemma 1.12.4: Existence Part of [Theorem 1.12.1](#)

If $n > 1$ then there exist primes $p_1, \dots, p_s$ for some $s \geq 1$ such that

$$n = p_1 p_2 \cdots p_s$$

and $p_1 \leq p_2 \leq \cdots \leq p_s$.

Proof

Proof by induction on n , with starting value $n = 2$: If $n = 2$ then since 2 is prime we can take $p_1 = 2$, $s = 1$. Assume the lemma holds for n such that $2 \leq n \leq k$. Let's show it holds for $n = k + 1$. If $k + 1$ is prime we can take $s = 1$ and $p_1 = k + 1$ and we are done. If $k + 1$ is composite we can write $k + 1 = ab$ where $1 < a < k + 1$ and $1 < b < k + 1$. By the induction hypothesis there are primes $p_1, \dots, p_u$ and $q_1, \dots, q_v$ such that

$$a = p_1 \cdots p_u \text{ and } b = q_1 \cdots q_v.$$

This gives us

$$k + 1 = ab = p_1 p_2 \cdots p_u q_1 q_2 \cdots q_v,$$

that is, $k + 1$ is a product of primes. Let $s = u + v$. By reordering and relabeling where necessary, we have

$$k + 1 = p_1 p_2 \cdots p_s,$$

where $p_1 \leq p_2 \leq \cdots \leq p_s$. So the lemma holds for $n = k + 1$. Hence by PMI, it holds for all $n > 1$.

Lemma 1.12.5: Uniqueness Part of [Theorem 1.12.1](#)

Let

$$n = p_1 p_2 \cdots p_s \text{ for some } s \geq 1,$$

and

$$n = q_1 q_2 \cdots q_t \text{ for some } t \geq 1,$$

where $p_1, \dots, p_s, q_1, \dots, q_t$ are primes satisfying

$$p_1 \leq p_2 \leq \cdots \leq p_s$$

and

$$q_1 \leq q_2 \leq \cdots \leq q_t.$$

Then, $t = s$ and $p_i = q_i$ for $i = 1, 2, \dots, t$.

Proof

Our proof is by induction on s . Suppose $s = 1$. Then $n = p_1$ is prime and we have

$$p_1 = n = q_1 q_2 \cdots q_t.$$

If $t > 1$, this contradicts the fact that p_1 is prime. So $t = 1$ and we have $p_1 = q_1$, as desired. Now assume the result holds for all s such that $1 \leq s \leq k$. We want to show that it holds for $s = k + 1$. So assume

$$n = p_1 p_2 \cdots p_k p_{k+1}$$

and

$$n = q_1 q_2 \cdots q_t$$

where $p_1 \leq p_2 \leq \cdots \leq p_{k+1}$ and $q_1 \leq q_2 \leq \cdots \leq q_t$. Clearly $p_{k+1} \mid n$ so $p_{k+1} \mid q_1 \cdots q_t$. So by [Lemma 1.12.3](#) $p_{k+1} \mid q_i$ for some $i \in \{1, 2, \dots, t\}$. It follows from [Exercise 1.11.3](#) that $p_{k+1} = q_i$. Hence $p_{k+1} = q_i \leq q_t$.

By a similar argument $q_t \mid n$ so $q_t \mid p_1 \cdots p_{k+1}$ and $q_t = p_j$ for some j . Hence $q_t = p_j \leq p_{k+1}$. This shows that

$$p_{k+1} \leq q_t \leq p_{k+1}$$

so $p_{k+1} = q_t$. Note that

$$p_1 p_2 \cdots p_k p_{k+1} = q_1 q_2 \cdots q_{t-1} q_t$$

Since $p_{k+1} = q_t$ we can cancel this prime from both sides and we have

$$p_1 p_2 \cdots p_k = q_1 q_2 \cdots q_{t-1}.$$

Now by the induction hypothesis $k = t - 1$ and $p_i = q_i$ for $i = 1, \dots, t - 1$. Thus we have $k + 1 = t$ and $p_i = q_i$ for $i = 1, 2, \dots, t$. So the lemma holds for $s = k + 1$ and by the PMI, it holds for all $s \geq 1$.

Now the proof of [Theorem 1.12.1](#) follows immediately from [Lemmas 1.12.4](#) and [1.12.5](#).

Remark 1.12.2

If a and b are positive integers we can find primes $p_1, \dots, p_k$ and integers $a_1, \dots, a_k, b_1, \dots, b_k$ each ≥ 0 such that

$$\begin{cases} a = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k} \\ b = p_1^{b_1} p_2^{b_2} \cdots p_k^{b_k} \end{cases} \quad (1.12.1)$$

For example, if $a = 600$ and $b = 252$ we have

$$600 = 2^3 \cdot 3^1 \cdot 5^2 \cdot 7^0$$

$$252 = 2^2 \cdot 3^2 \cdot 5^0 \cdot 7^1.$$

It follows that

$$\gcd(600, 252) = 2^2 \cdot 3^1 \cdot 5^0 \cdot 7^0 \quad \text{and} \quad \text{lcm}(600, 252) = 2^3 \cdot 3^2 \cdot 5^2 \cdot 7^1.$$

In general, if a and b are given by equations (1.12.1) it can be proved that

$$\gcd(a, b) = p_1^{\min(a_1, b_1)} p_2^{\min(a_2, b_2)} \cdots p_k^{\min(a_k, b_k)}$$

and

$$\text{lcm}(a, b) = p_1^{\max(a_1, b_1)} p_2^{\max(a_2, b_2)} \cdots p_k^{\max(a_k, b_k)}$$

(try it!). This gives one way to calculate the gcd or lcm *provided* you can factor both numbers. But generally speaking factorization is *very difficult*! On the other hand, the Euclidean algorithm is relatively fast.

Exercises

? Exercise 1.12.1

Find the prime factorizations of 1147 and 1716 by trying all primes $p \leq \sqrt{1147}$ ($p \leq \sqrt{1716}$) in succession.

? Exercise 1.12.2

Determine the prime factorization of $15! = 1307674368000$

? Exercise 1.12.3

- Using the prime factorization of $100!$, determine how many 0's appear at the end of the base 10 representation.
- How many 0's appear at the end of the binary representation of $100!$?

? Exercise 1.12.4

The prime factorization of an integer n expresses n as the product of one or more prime numbers, and by the Fundamental Theorem of Arithmetic, the factorization is unique when the primes are ordered by size.

Let's examine the idea of factorizations in a different set. Let S be the set of all positive multiples of 3. For each $n \in S$, define n to be *S-prime* if n cannot be written as the product of two smaller elements of S .

- Write down the 10 smallest *S*-prime integers.
- Find a factorization of 18 into *S*-prime numbers.
- Explain how we know that every element of S can be written as the product of one or more *S*-prime numbers.
- Is the factorization of an element of S into *S*-prime units unique? If so, explain why. If not, give an example of an element of S that can be factored into *S*-prime numbers in at least two different ways (along with the two factorizations).

? Exercise 1.12.5

Let $\mathbb{N}^*$ be the set $\{4, 5, 6, 7, \dots\}$. This set $\mathbb{N}^*$ would be the set of positive integers if for some reason we changed our mind and declared that 1, 2, 3 were not integers.

- If we define a number in $\mathbb{N}^*$ to be "prime*" if it cannot be written as a product of two or more smaller elements in $\mathbb{N}^*$, then which numbers from the set $\{4, 5, \dots, 25\}$ are prime*?
- Express 32 as the product of two prime* numbers.
- Find a number in $\mathbb{N}^*$ that can be written as a product of prime* numbers in at least two different ways. (As an additional challenge, try to find the smallest such number.)

? Exercise 1.12.6

Most students have seen the Fundamental Theorem of Arithmetic so long ago that it seems somehow obvious or unremarkable. Based on the previous two exercises, do you think that [Theorem 1.12.1](#) is at all surprising? Explain your answer in at least a short paragraph.

Footnotes

[1] This result is proved in Proposition 30 of Book VII of Euclid's *Elements*.

This page titled [1.12: Unique Factorization](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Mike Barrus & W. Edwin Clark](#).

1.13: The Gaussian Integers

In this section we study a special subset of the complex numbers known as the Gaussian integers.

Definition 1.13.1: Gaussian integer

A *Gaussian integer* is a complex number of the form $a + bi$ where both a and b are integers. We often denote the set of Gaussian integers by $\mathbb{Z}[i]$.

In order to simplify notation (and not confuse Gaussian integers with ordinary integers), we will sometimes use Greek letters α, β , etc. to represent Gaussian integers.

The Gaussian integers have many special properties that are similar to those of the integers. In this chapter, once we have a few fundamental concepts, we will see how the Gaussian integers satisfy a division algorithm and a version of unique factorization. We will also see the Gaussian integers pop up a few times in later chapters.

We start with a graphical representation. Just as the integers form points along the standard real line,



Figure 1.13.1

we can visualize the Gaussian integers as regularly spaced points, though in a new context.

The complex numbers can be represented by a 2-dimensional drawing called an *Argand plane*. Just like a normal Cartesian plane, an Argand plane contains a horizontal axis and a vertical axis for measurements. However, instead of representing an ordered pair of real numbers, each point in the Argand plane represents a single complex number—the point located at what would traditionally be labeled (x, y) in Cartesian coordinates instead represents the complex number $x + yi$, where i is a square root of -1 , as usual. Shown here is a picture of the Argand plane:

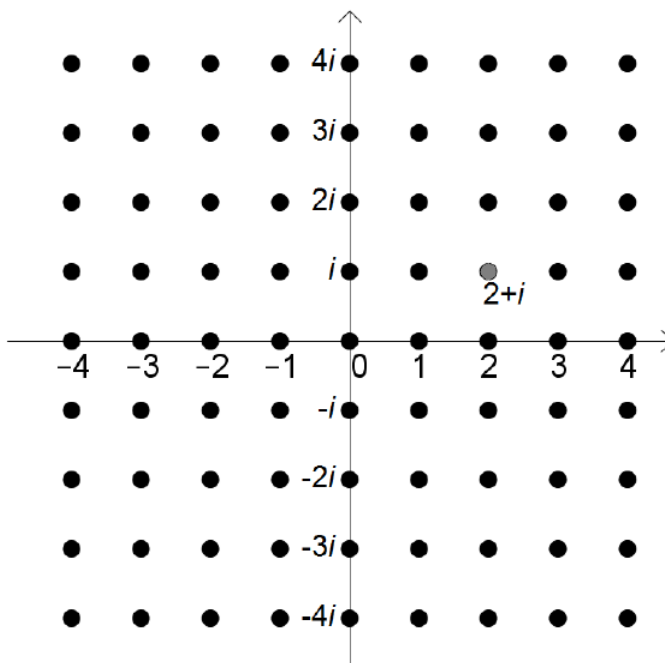


Figure 1.13.2

The solid points mark the locations of the Gaussian integers. Note that the horizontal axis measures the real part of each complex number, and the vertical axis measures the imaginary part. Hence the point $2 + i$ is located 2 units horizontally to the right of the center and one unit vertically up.

Since the (real) integers lie on a single line, it is easy to understand whether one integer is “bigger” than another. For the Gaussian integers, though, it is not immediately clear whether $2 + i$ is bigger or smaller than 3. Because being able to compare Gaussian integers will be important to us shortly, we will define a way to do this.

Definition 1.13.2: Norm

The *norm* of a Gaussian integer $\alpha = a + bi$, denoted $N(\alpha)$ or $N(a + bi)$, is a real number defined by

$$N(\alpha) = N(a + bi) = a^2 + b^2.$$

Since the norm of a Gaussian integer is a real number, we can use norms to compare “sizes” of Gaussian integers. For example, $N(2 + i) = 5$ and $N(3) = 9$ (since $3 = 3 + 0i$), so intuitively we can think of $2 + i$ as smaller than 3. Indeed, you might notice that $N(a + bi)$ is the square of the usual straight-line distance from 0 to $a + bi$ if we think of them as the Cartesian points $(0, 0)$ and (a, b) . For instance 1, -1 , i , and $-i$ are each one unit away from 0, and $N(1) = N(-1) = N(i) = N(-i) = 1 = 1^2$.

One idea you may remember from earlier exposure to complex numbers is that of a complex conjugate.

Definition 1.13.3: Complex Conjugate

Given a complex number $\alpha = a + bi$, the (*complex*) *conjugate* is the complex number $a - bi$, i.e., the complex number obtained by changing the sign on the imaginary part. We denote the conjugate of α with a horizontal bar above the number, so $\overline{\alpha} = \overline{a + bi} = a - bi$.

Note that $\overline{a - bi} = a + bi$, too, so the conjugate of a conjugate is again the original complex number. Note also that the conjugate of a real number is itself.

Some useful observations connect norms and conjugates. In the following, remember how complex numbers are multiplied:

$$\begin{aligned}(a + bi)(c + di) &= ac + adi + bci + bdi^2 \\ &= (ac - bd) + (ad + bc)i,\end{aligned}$$

since $i^2 = -1$.

Lemma 1.13.1: Conjugate Norm

For any $\alpha \in \mathbb{Z}[i]$ we have the following:

- i. $N(\alpha) = \alpha \cdot \overline{\alpha}$;
- ii. $N(\overline{\alpha}) = N(\alpha)$.

Proof

See [Exercise 1.13.1](#)

Perhaps the most important property about norms is the following.

Lemma 1.13.2: Multiplicative Norms

[lem: multiplicative norms] For any $\alpha, \beta \in \mathbb{Z}[i]$,

$$N(\alpha\beta) = N(\alpha)N(\beta).$$

Proof

See [Exercise 1.13.2](#)

Divisibility and division

We can make many of the same definitions for Gaussian integers as we have done in earlier chapters for integers.

Definition 1.13.4: Divisor

Given Gaussian integers α and β , we will say that α divides β if there exists another Gaussian integer γ such that $\alpha\gamma = \beta$. In this case we call α a *divisor* of β , and β a *multiple* of α .

If p and q are (real) integers, then it can be shown that p divides q under this new definition of “divides” if and only if $p|q$ as integers.

It may seem a little difficult at first to decide whether one Gaussian integer divides another. For example, how would you know whether $1 - i$ divides 4, or whether $2 + i$ divides 6? As a corollary to [Lemma 1.13.2](#), we can find a quick way to rule out certain possibilities.

Corollary 1.13.1: Norm Divisibility

If $N(\alpha)$ does not divide $N(\beta)$, then α is not a divisor of β .

✓ Example 1.13.1

Since $N(2 + i) = 5$ and $N(6) = 36$ and 5 is not a divisor of 36, we know that $2 + i$ does not divide 6.

However, is $1 - i$ a divisor of 4? Comparing their norms

$$\begin{aligned} N(1 - i) &= 1^2 + (-1)^2 = 2, \\ N(4) &= 4^2 + 0^2 = 16, \end{aligned}$$

we do see that 2 divides 16. Neither [Corollary 1.13.1](#) nor [Lemma 1.13.2](#) guarantees that $1 - i$ divides 4; it still might not. If it did, how could we find a number to multiply $1 - i$ by to produce 4?

At this point, though we have tried to avoid much use of fractions in this text, we will find it convenient to use them. The number we seek is

$$\frac{4}{1 - i},$$

but this complex number isn't written in our usual form of $a + bi$; how can we put it in this form?

You may remember the answer from past experience with complex numbers: to put a complex fraction with a non-real denominator in the form $a + bi$, we multiply both the numerator and denominator by the conjugate of the denominator.

In this case, we find

$$\frac{4}{1 - i} = \frac{4}{1 - i} \cdot \frac{1 + i}{1 + i} = \frac{4 + 4i}{1^2 - i^2} = \frac{4}{2} + \frac{4i}{2} = 2 + 2i.$$

Since $2 + 2i$ is a Gaussian integer and $(1 - i)(2 + 2i) = 4$, we conclude that yes, $1 - i$ is a divisor of 4. (If the quotient had turned out to be a complex number other than a Gaussian integer, then we would have concluded that $1 - i$ wasn't a divisor.)

Now that we have defined divisors, we need a few more definitions that weren't important when working with integers but will be useful now.

Definition 1.13.5: Unit

A *unit* in $\mathbb{Z}[i]$ is an element ω for which $N(\omega) = 1$. Two Gaussian integers α and β are *associates* if there exists a unit ω such that $\alpha\omega = \beta$.

Note that if $v + wi$ is a unit in $\mathbb{Z}[i]$, then v and w are integers, and $v^2 + w^2 = 1$. The only ways this can happen is if one of v or w is ± 1 and the other is 0. Thus the only units in $\mathbb{Z}[i]$ are 1, -1 , i , and $-i$. Consequently, each Gaussian unit has a total of four associates (counting itself).

✓ Example 1.13.2

The Gaussian integer $2 + 3i$ has four associates: itself, and

$$\begin{aligned}(2 + 3i)(-1) &= -2 - 3i, \\ (2 + 3i)i &= -3 + 2i, \text{ and} \\ (2 + 3i)(-i) &= 3 - 2i.\end{aligned}$$

Associates play an important role in statements about divisibility, because of the next lemma.

🔗 Lemma 1.13.3: Associates Divisibility

Let α , β , and γ be Gaussian integers.

- i. If α and β are associates, then γ divides one of α , β if and only if it divides the other.
- ii. If α and β are associates, then α divides γ if and only if β divides γ .

See [Exercise 1.13.3](#)

Now what can we say for Gaussian integers like $2 + i$ and 6 where neither divides the other? As we did for the integers, we introduce a Division Algorithm-type theorem for the Gaussian integers.

🔗 Theorem 1.13.1: Gaussian Division Algorithm

Given Gaussian integers α and β , where $\beta \neq 0$, there exist Gaussian integers κ and ρ such that

$$\alpha = \beta\kappa + \rho \quad \text{and} \quad 0 \leq N(\rho) < N(\beta).$$

For the sake of brevity, and because the details are in some respects analogous to those in the proof of the Division Algorithm, we will skip the proof of [Theorem 1.13.1](#). However, we will illustrate some of its ideas in the following example.

✓ Example 1.13.3: Gaussian Division

We will find κ and ρ given $\alpha = 6$ and $\beta = 2 + i$. We start with κ by computing the complex number

$$\frac{\alpha}{\beta} = \frac{6}{2+i} = \frac{6}{2+i} \cdot \frac{2-i}{2-i} = \frac{12-6i}{2^2-(-1)^2} = \frac{12}{5} - \frac{6}{5}i.$$

The answer is not a Gaussian integer, as κ should be, so we will round both the real and imaginary parts to the nearest integer and set κ equal to the result, obtaining here $\kappa = 2 - i$.

We now find $\rho = \alpha - \beta\kappa$:

$$\rho = 6 - (2+i)(2-i) = 6 - 5 = 1.$$

We check that the norms work as they should, and we see that indeed,

$$0 \leq N(\rho) = N(1) = 1 < 5 = N(2+i) = N(\beta).$$

In general, finding a quotient and remainder when one Gaussian integer is divided by another can be accomplished just as in the previous example—we treat the Gaussian integers as complex numbers and divide as normal, producing a complex number with possibly non-integer coefficients for real and imaginary parts. Rounding these numbers to the nearest integer will provide a quotient, which can then be used to compute the remainder. Note that as opposed to the Division Algorithm, here we don't require a nonnegative remainder, and we don't claim that the quotient and remainder are unique. The next example illustrates why.

✓ Example 1.13.4

If we attempt to find a quotient and remainder for the Gaussian integers $1 + i$ and $3i$ using the same process as in the last example, we find

$$3i = (1+i)(2+2i) - i,$$

$$0 \leq N(-i) < N(1+i);$$

or

$$3i = (1+i)(1+i) + i,$$

$$0 \leq N(i) < N(1+i);$$

or perhaps

$$3i = (1+i)(1+2i) + 1,$$

$$0 \leq N(1) < N(1+i);$$

or

$$3i = (1+i)(2+i) - 1,$$

$$0 \leq N(-1) < N(1+i);$$

since at the first step when we divide $3i$ by $1+i$ the fraction $3i/(1+i)$ is equal to $3/2 + (3/2)i$, and $3/2$ is equally close to 1 and to 2.

Greatest common divisors

Now that we have a Gaussian integer version of the Division Algorithm, can we adapt some of our earlier algorithms, like the Euclidean Algorithm? What would the outcome tell us? Let's try.

✓ Example 1.13.5: Gaussian Euclidean Algorithm

If we begin the Euclidean Algorithm with $2 - 3i$ and $3 + 5i$, we may obtain

$$3 + 5i = (2 - 3i)(-1 + i) + 2;$$

$$2 - 3i = 2(1 - 2i) + i;$$

$$2 = i(-2i) + 0.$$

The last nonzero remainder is i .

If the numbers involved in the Euclidean Algorithm were integers, we would interpret the last nonzero remainder to be the greatest common divisor. Can we do the same for Gaussian integers?

Yes, if we take the right meaning for the word "greatest". As in the proof of the Euclidean Algorithm, a Gaussian integer is a common divisor of $3 + 5i$ and $2 - 3i$ if and only if it is a common divisor of $2 - 3i$ and 2, which happens if and only if it is a common divisor of 2 and i , which happens if and only if it is a common divisor of i and 0. Thus the common divisors all divide i , so i is the "greatest" common divisor in the sense that every common divisor divides it.

By [Lemma 1.13.3](#), a Gaussian integer divides i if and only if it divides 1, -1 , or $-i$, so these four associates are *all* greatest common divisors of $3 + 5i$ and $2 - 3i$.

In general, carrying out the Euclidean Algorithm with Gaussian integers will always produce a greatest common divisor, and we may find others by considering associates.

Unique factorization

Studying the divisors of integers led us to think about prime numbers, those integers that could not be divided evenly by any smaller positive integers other than 1. We then saw that every positive integer greater than 1 could be written uniquely as a product of these primes, if we ordered the primes from smallest to largest. We can study similar questions for the Gaussian integers, though the answers will be different at times.

Definition 1.13.6: Gaussian Prime

An element in $\mathbb{Z}[i]$ is called a *Gaussian prime* if it cannot be written as the product of two or more other Gaussian numbers where none of the numbers in the product is a unit.

✓ Example 1.13.6: Showing Gaussian Prime

The numbers $11 - 10i$ and 27 are not Gaussian primes, since $11 - 10i = (2 - 3i)(4 + i)$ and $27 = 3 \cdot 9 = (3 + 0i)(9 + 0i)$.

In contrast, $2 + i$ is a Gaussian prime. To see why, suppose to the contrary that $2 + i = \alpha\beta$, where $\alpha, \beta \in \mathbb{Z}[i]$ and neither is a unit. By [Lemma 1.13.2](#)

$$N(\alpha)N(\beta) = N(2 + i) = 5.$$

Since $N(\alpha)$ and $N(\beta)$ are both nonnegative integers, one must equal 5 and the other must equal 1. However, if either α or β has a norm of 1 then it is a unit, contrary to our assumption. Thus $2 + i$ is a Gaussian prime.

By the same reasoning as in the example above, we can say the following.

Proposition 1.13.1: Prime Norm means Gaussian Prime

For every $\alpha \in \mathbb{Z}[i]$, if $N(\alpha)$ is a prime number, then α is a Gaussian prime.

Note that the converse of [Proposition 1.13.1](#) isn't true; for example, the norm of 3 is 9, is composite, but 3 is still a Gaussian prime (see [Exercise 1.13.6](#)).

✓ Example 1.13.7

The number 41 is a prime number; is it also a Gaussian prime?

If not, then there exist Gaussian integers α and β whose product is 41, such that neither is a unit. By [Lemma 1.13.2](#)

$$N(\alpha)N(\beta) = N(41) = 41^2.$$

This forces $N(\alpha) = 41$ and $N(\beta) = 41$. Can we find two Gaussian integers with norm 41 whose product is 41?

We can! Observe that $4 + 5i$ and $4 - 5i$ both have norm 41, and $(4 + 5i)(4 - 5i) = 41$. Hence 41, though a prime number, is *not* a Gaussian prime.

Many other prime numbers are not Gaussian primes (see [Exercise 1.13.6](#)), though some, like 3, are prime in both the integers and in the Gaussian integers. We will be able to say more about primes and Gaussian primes in the exercises of [Section 1.28](#).

Now that we have the notion of Gaussian primes, can we factor Gaussian integers uniquely into Gaussian primes? At first glance, the answer might seem to be no.

✓ Example 1.13.8

In the previous example we saw that

$$41 = (4 + 5i)(4 - 5i).$$

Since $N(4 + 5i)$ and $N(4 - 5i)$ are prime numbers, we know by [Proposition 1.13.1](#) that $4 + 5i$ and $4 - 5i$ are Gaussian primes.

However, we may also write

$$41 = (5 + 4i)(5 - 4i),$$

and $5 + 4i$ and $5 - 4i$ are also Gaussian primes. This means that 41 cannot be written uniquely as a product of non-unit Gaussian primes.

Still, the numbers $4 \pm 5i$ and $5 \pm 4i$ are related. Note that $(4 + 5i)(-i) = 5 - 4i$, and $(4 - 5i)i = 5 + 4i$, so $4 + 5i$ and $5 - 4i$ are associates, as are $4 - 5i$ and $5 + 4i$. We can also derive one factorization from the other:

$$\begin{aligned} 41 &= (4 + 5i)(4 - 5i) \cdot 1 \\ &= (4 + 5i)(4 - 5i) \cdot (-i)i \\ &= [(4 + 5i)(-i)][(4 - 5i)i] \\ &= (5 - 4i)(5 + 4i). \end{aligned}$$

Because $-i$ and i multiply to 1, or because $4 \pm 5i$ and $5 \mp 4i$ are associates, couldn't we say that the two factorizations are pretty much the same?

In fact, one can show that if α is a Gaussian prime, then $\omega\alpha$ is a Gaussian prime for any unit ω , so by using units whose product is 1 as in the example above, we can find *many* different-looking factorizations of a Gaussian integer into Gaussian primes.

Still, what if we relax our notion of "sameness" in factorizations so as to ignore differences in associates? Then the Gaussian integers *can* be uniquely factored into Gaussian primes.

Theorem 1.13.2

Any Gaussian integer α with norm greater than 1 can be factored into Gaussian primes, and the factorization is unique in the following sense: if

$$\alpha = \beta_1 \beta_2 \cdots \beta_k = \gamma_1 \gamma_2 \cdots \gamma_\ell,$$

where each β_i and each γ_i is a Gaussian prime, then $k = \ell$ (i.e., the numbers of primes involved are the same) and the primes may be reordered and renumbered so that β_i is an associate of γ_i for all i .

We will not prove this theorem here, though you are encouraged to attempt it. Proving the existence of a prime factorization can be done by induction in much the same way that [Lemma 1.12.4](#) uses; try to adapt [Lemma 1.12.5](#) for the rest.

We have seen how many of the key results we have encountered so far have analogues in the number system of the Gaussian integers. Though as the text progresses, we will mention the Gaussian integers only a few more times, you are invited to think about how each new result would look if the integers involved were replaced by Gaussian integers. Doing so will lead you to research questions; it may also give you a new perspective on things you've already long since known.

Exercises

? Exercise 1.13.1

Prove both statements in [Lemma 1.13.1](#).

? Exercise 1.13.2

Prove [Lemma 1.13.2](#).

? Exercise 1.13.3

Prove both statements in [Lemma 1.13.3](#).

(Hint: Recall that the only units in $\mathbb{Z}[i]$ are ± 1 and $\pm i$.)

? Exercise 1.13.4

Write i as a linear combination of $2 - 3i$ and $3 + 5i$.

(Hint: apply a method from [Section 1.10](#) and perhaps refer to [Example 1.13.5](#). Remember that the coefficients on $2 - 3i$ and $3 + 5i$ must be Gaussian integers; they cannot involve fractions.)

? Exercise 1.13.5

Use the Euclidean Algorithm to find a greatest common divisor of 5 and $1 - i$. Then write this greatest common divisor as a linear combination of 5 and $1 - i$.

(Hint: See hints from the previous exercise.)

? Exercise 1.13.6

- Show that none of 2, 5, 13, or $-3 + 4i$ is a Gaussian prime by writing each as a product of non-unit Gaussian integers. (Hint: $-3 + 4i$ is a perfect square.)
- Show that 3 is a Gaussian prime. (Hint: suppose that $a + bi$ and $c + di$ are non-unit Gaussian integers whose product is 3, and conclude that each must have norm 3; then show that no Gaussian integer has norm 3.)

? Exercise 1.13.7

- Prove that $1 + i$ is a Gaussian prime.
- Prove that $N(\alpha)$ is even if and only if α is a multiple of $1 + i$.

? Exercise 1.13.8

- Find a Gaussian prime factorization of $7 + 9i$. (Hint: get an idea from [Exercise 1.13.7](#)).
- Find at least two more factorizations of $7 + 9i$ into Gaussian primes, so that no Gaussian prime factor is repeated among the factorizations in part (a) and here.

? Exercise 1.13.9

The set $\mathbb{Z}[i]$ of Gaussian integers is a number system that, like $\mathbb{Z}$, has unique factorizations (up to associates). In this exercise we'll consider a slightly different number system.

Define

$$\mathbb{Z}[\sqrt{-5}] = \{a + b\sqrt{-5} \mid a, b \in \mathbb{Z}\}.$$

It turns out that if we define a new norm by $N(a + b\sqrt{-5}) = a^2 + 5b^2$, then $N(\alpha\beta) = N(\alpha)N(\beta)$ for every pair of elements $\alpha, \beta \in \mathbb{Z}[\sqrt{-5}]$.

- Show that 2, 23, and $1 + 3\sqrt{-5}$ are all “primes” in $\mathbb{Z}[\sqrt{-5}]$, i.e., that they cannot be written as products of factors all having smaller norms. (Hint: Use the multiplicative property of the norm and a proof by contradiction as in [Example 1.13.6](#). As part of your answer, show that $a^2 + 5b^2$ never equals 2 or 23 when a and b are integers.)
- Show that $1 + 3\sqrt{-5}$ divides 46 in $\mathbb{Z}[\sqrt{-5}]$.
- Show that $\mathbb{Z}[\sqrt{-5}]$ does not have a unique factorization property. (Hint: Show that 46 does not have a unique factorization into the primes in $\mathbb{Z}[\sqrt{-5}]$.)

1.14: Fermat Primes and Mersenne Primes

Finding large primes and proving that they are indeed prime is not easy. For a long time, people have looked for formulas for producing prime numbers, with varying degrees of success. In this chapter we'll learn about related questions and answers contributed by many people over the past several centuries—and even in the current one.

✓ Example 1.14.1

Let $f(n) = n^2 - n + 41$, where n is a positive integer. As we plug in $1, 2, \dots$, the first few numbers we get the sequence

$$41, 43, 47, 53, 61, \dots,$$

which seems to be producing a lot of prime numbers. In fact, the first *forty* terms are all prime, up to and including $40^2 - 40 + 41 = 1601$, so if we were to make a hasty generalization, we might suspect that $f(n)$ will *always* be prime, and we have found a simple formula for generating prime numbers.

This is false, however—observe what happens if we let $n = 41$:

$$f(41) = 41^2 - 41 + 41 = 41 \cdot 41.$$

Similar results occur whenever n is a multiple of 41, so unfortunately f does not always produce prime numbers.

The functions $f(n)$ and the similar polynomial $n^2 + n + 41$, both examples of polynomials that produce several prime numbers, were studied by the great mathematicians Leonhard Euler (1707–1783) and Adrien-Marie Legendre (1752–1833). It was shown by Christian Goldbach (1690–1764) in 1752 that no polynomial with integer coefficients can generate *only* prime numbers, so such polynomials as $f(n)$ cannot be used to consistently generate primes.

✓ Example 1.14.2

The function $g(n) = 6n + 5$ clearly does not always produce primes, by Goldbach's proof. It is also easy to see that $g(6) = 35 = 5 \cdot 7$, and 5 divides $g(n)$ whenever n is a multiple of 5.

However, [Exercise 1.11.2](#) indicates that every prime number other than 2 or 3 has the form $6n + 1$ or $6n + 5$. Since there are infinitely many primes, either $g(n) = 6n + 5$ or $h(n) = 6n + 1$ (or both) must produce a prime number for infinitely many n . [Exercise 1.14.2](#) leads you through showing that this is true for $g(n)$.

In fact, Johann Lejeune Dirichlet (1805–1859) showed in 1837 that whenever a and b are relatively prime, the function $f(n) = an + b$ produces a prime number for infinitely many values of n . In 2004, Ben Green and Terence Tao announced a proof that for any positive length k , there exist coefficients a and b such that $f(n) = an + b$ will produce at least k primes using *consecutive* values of n . (For example, $g(0) = 5, g(1) = 11, g(2) = 17, g(3) = 23, g(4) = 29$, so $6n + 5$ already produces primes for five consecutive values of n .)

Though functions like $n^2 - n + 41$ or $6n + 5$ cannot always produce primes, large numbers of primes (and sometimes infinitely many) *do* have this form, so one way to find primes, and perhaps very large primes, is to look at numbers that have some special form. For example, several of the smallest primes, like 3, 5, 7, 17, 31 are all either one unit less than or one unit more than a power of 2. With that in mind, the rest of this chapter will look at the following.

📌 Question

For which values of a and n might

$$a^n + 1 \quad \text{or} \quad a^n - 1$$

be a prime number?

It is easy to rule out some values of a and n . For example we have the following.

Theorem 1.14.1: Mersenne Prep

Let $a > 1$ and $n > 1$. The following are true.

1. If $a^n - 1$ is prime, then $a = 2$ and n is prime.
2. If $a^n + 1$ is prime, then a is even and $n = 2^k$ for some $k \geq 1$.

Proof of 1

We know from [Exercise 1.3.8](#) that

$$a^n - 1 = (a - 1)(a^{n-1} + \cdots + a + 1). \quad (1.14.1)$$

If $a > 2$ and $n > 1$, then $a - 1 > 1$ and $a^{n-1} + \cdots + a + 1 > a + 1 > 3$ so both factors in equation (1.14.1) are greater than 1 and thus $a^n - 1$ is not prime. Hence if $a^n - 1$ is prime we must have $a = 2$. Suppose this is the case, i.e., that $2^n - 1$ is prime. We claim that n is prime. If not, then $n = st$ for integers s and t such that $1 < s < n$ and $1 < t < n$. Then

$$2^n - 1 = 2^{st} - 1 = (2^s)^t - 1$$

is prime. But we just showed that if $a^n - 1$ is prime we must have $a = 2$. So we must have $2^s = 2$, forcing $s = 1$, a contradiction. Hence n must be prime, as claimed.

Proof of 2

Suppose first that n is odd. Replacing a by $-a$ in equation (1.14.1), we get

$$(-a)^n - 1 = (-a - 1)((-a)^{n-1} + (-a)^{n-2} + \cdots + (-a) + 1)$$

Since n is odd, $n - 1$ is even, $n - 2$ is odd, $\dots$, etc., we have $(-a)^n = -a^n$, $(-a)^{n-1} = a^{n-1}$, $(-a)^{n-2} = -a^{n-2}$, $\dots$, etc. Substituting these values yields

$$-(a^n + 1) = -(a + 1)(a^{n-1} - a^{n-2} + \cdots - a + 1).$$

Multiplying both sides by -1 we get

$$a^n + 1 = (a + 1)(a^{n-1} - a^{n-2} + \cdots - a + 1)$$

when n is odd. If $n \geq 2$, then $a^n + 1 > a + 1$, and the equation above shows that when n is odd and $a > 1$, the number $a^n + 1$ is not prime, since $a + 1$ is a divisor strictly between 1 and $a^n + 1$. Therefore, if $a^n + 1$ is supposed to be prime, it must be the case that n is even.

Suppose now that n is even, and express it as $n = 2^s t$, where t is odd (so $s \geq 1$). If $a^n + 1$ is prime, then $(a^{2^s})^t + 1$ is prime. But by what we just showed, this cannot be prime if t is odd and $t \geq 2$. Since t is odd, we must have $t = 1$ and so $n = 2^s$.

Finally, note that since n is positive, a and a^n have the same parity, so $a^n + 1$ is odd if and only if a is even. The number 2 (the only even prime) is too small to be written as $a^n + 1$ for any $a > 1$ and $n > 1$, so if $a^n + 1$ is prime, then a is even and n is a power of 2, as claimed.

Number theorists have studied the prime numbers satisfying the conclusions in [Theorem 1.14.1](#). We remember two in particular, namely Pierre de Fermat (1607–1665) and Marin Mersenne (1588–1648), through names given to these numbers.

Definition 1.14.1: Fermat Number

A number of the form $F_n = 2^{(2^n)} + 1$, $n \geq 0$, is called a *Fermat number*. If F_n is prime, it is called a *Fermat prime*. A number of the form $M_n = 2^n - 1$, $n \geq 2$, is said to be a *Mersenne number*. If M_n is prime, it is called a *Mersenne prime*.

As Fermat studied the numbers now named after him, he conjectured that all Fermat numbers are prime. Indeed, one may prove that $F_0 = 3$, $F_1 = 5$, $F_2 = 17$, $F_3 = 257$ and $F_4 = 65537$ are primes. As n increases the numbers $F_n = 2^{(2^n)} + 1$ increase in size very rapidly and are not easy to check for primality in a reasonable amount of time. However, it is now known that F_n is composite for many values of $n \geq 5$. This includes all n such that $5 \leq n \leq 30$ and a large number of other values of n including 18233954 (the largest one known as of a recent revision of this text)¹. It is now conjectured that F_n is composite for all $n \geq 5$. So Fermat's original thought that F_n is prime for $n \geq 0$ seems to be pretty far from reality.

Mersenne, a friend of mathematician and philosopher René Descartes, studied the Mersenne numbers and attempted to make a list of Mersenne primes with exponents up to 257, though the list contained some errors. To see some small Mersenne numbers, look at the numbers $M_2 = 2^2 - 1 = 3$ and $M_3 = 2^3 - 1 = 7$, which are both Mersenne primes, and at $M_4 = 2^4 - 1 = 15$, which is a Mersenne number but not a prime (this should not be surprising in light of [Theorem 1.14.1](#), since the exponent 4 is not prime). We pause here to reiterate this idea.

Lemma 1.14.1

If M_n is prime, then n is prime.

Proof

This is immediate from [Theorem 1.14.1\(1\)](#).

You might wonder whether the converse of this Lemma is true—whether $M_p = 2^p - 1$ is prime *whenever* p is prime. The answer is no; the Mersenne number $M_{11} = 2^{11} - 1 = 2047 = 23 \cdot 89$ is not prime.

Over the years people have continued to work on the problem of determining for which primes p , $M_p = 2^p - 1$ is prime. To date 51 Mersenne primes have been found. It is known that $2^p - 1$ is prime if p is one of the following 51 primes: 2, 3, 5, 7, 13, 17, 19, 31, 61, 89, 107, 127, 521, 607, 1279, 2203, 2281, 3217, 4253, 4423, 9689, 9941, 11213, 19937, 21701, 23209, 44497, 86243, 110503, 132049, 216091, 756839, 859433, 1257787, 1398269, 2976221, 3021377, 6972593, 13466917, 20996011, 24036583, 25964951, 30402457, 32582657, 37156667, 42643801, 43112609, 57885161, 74207281, 77232917, 82589933.

The largest Mersenne prime currently known, $M_{82589933} = 2^{82589933} - 1$, was found on December 7, 2018. The decimal representation of this number has 24,862,048 digits (about one and one-half *million* digits *more* than in the representation of $M_{77232917}$). It was found by a computer volunteered by Patrick Laroche as a part of the Great Internet Mersenne Prime Search (GIMPS); see www.mersenne.org for more about this. This prime could be the 51st Mersenne prime (in order of size), but we will only know this for sure when GIMPS completes testing all prime exponents between 43112609 and 82589933. Later we show the connection between Mersenne primes and a special class of numbers called *perfect numbers*.

The following primality test for Mersenne numbers makes it easier to check whether or not M_p is prime when p is a large prime.

Theorem 1.14.2: The Lucas-Lehmer Mersenne Prime Test

Let p be an *odd* prime. Define the sequence

$$r_1, r_2, r_3, \dots, r_{p-1}$$

by the rules

$$r_1 = 4$$

and for $k \geq 2$,

$$r_k = (r_{k-1}^2 - 2) \bmod M_p.$$

Then M_p is prime if and only if $r_{p-1} = 0$.

[The proof of this is not easy. One place to find a proof is the book “A Selection of Problems in the Theory of Numbers” by W. Sierpinski, Pergamon Press, 1964.]

Example 1.14.3

Let $p = 5$. Then $M_p = M_5 = 31$.

$$\begin{aligned} r_1 &= 4 \\ r_2 &= (4^2 - 2) \bmod 31 = 14 \bmod 31 = 14 \\ r_3 &= (14^2 - 2) \bmod 31 = 194 \bmod 31 = 8 \\ r_4 &= (8^2 - 2) \bmod 31 = 62 \bmod 31 = 0. \end{aligned}$$

Hence by the Lucas-Lehmer test, $M_5 = 31$ is prime.

Remark 1.14.1

Note that the Lucas-Lehmer test for $M_p = 2^p - 1$ takes only $p - 1$ steps. On the other hand, if one attempts to prove M_p prime by testing all primes $\leq \sqrt{M_p}$ one must consider about $2^{p/2}$ steps. This is MUCH larger than p in general.

This chapter has introduced Fermat and Mersenne numbers and primes and some of what we know about them. Hopefully you have noticed from the names and dates mentioned here that it has taken the work of many people over a long time to find out what we know now. Hopefully you have also gotten a sense that this work is not over—there is much more to learn! Some basic unsolved questions are here.

Open Questions

- Are there infinitely many primes of the form $2^{2^n} + 1$ (i.e., Fermat primes)?
- Are there infinitely many primes of the form $2^n - 1$ (i.e., Mersenne primes)?

When will we have answers? Who will provide them? Will *you* be part of that work?

Exercises

Exercise 1.14.1

Show that if

$$f(x) = a_p x^p + a_{p-1} x^{p-1} + \cdots + a_2 x^2 + a_1 x + a_0,$$

where $a_0, \dots, a_p$ are all integers, with $a_p \neq 0$ and $p \geq 1$ and $a_0 > 1$, then there exists a value of x such that $f(x)$ is guaranteed not to be prime. (Give an example of such a value.)

Exercise 1.14.2

In this exercise, we will show that $g(n) = 6n + 5$ will produce a prime number for infinitely many values of n .

- Show that other than 2 or 3, every prime number has the form $6n + 1$ or $6n + 5$.

(Compare this to [Exercise 1.11.2](#).)

- Show that multiplying any two numbers of the form $6n + 1$ creates another number of the form $6n + 1$.
- Show that multiplying two numbers of the form $6n + 5$ creates a number of the form $6n + 1$.
- Show that there are infinitely many primes of the form $6n + 5$. (Hint: imitate Euclid's proof in [Theorem 1.11.1](#), showing that if $q_1, q_2, \dots, q_k$ are all primes of the form $6n + 5$, then $q_1 q_2 \cdots q_k + 4$ or $5q_1 q_2 \cdots q_k + 4$ must have a prime divisor of the form $6n + 5$.)

Exercise 1.14.3

Use a computer to factor F_5 , describing any steps and tools you use.

Exercise 1.14.4

Prove that if m, n are integers with $m < n$, then $F_m \mid (F_n - 2)$.

(Hint: Recall how to factor a difference of squares.)

? Exercise 1.14.5

- Use the previous exercise to show that $\gcd(F_m, F_n) = 1$ whenever $m \neq n$.
- Explain why part (a) gives a new proof that there are infinitely many prime numbers.

? Exercise 1.14.6

- Use mathematical induction to show that

$$F_n = (F_{n-1} - 1)^2 + 1$$

for all $n \geq 2$.

- Use mathematical induction and part (a) to show that other than F_0 and F_1 , all Fermat numbers have 7 as their last digit.

? Exercise 1.14.7

Find all integers, if any, that are both a Mersenne prime and a Fermat prime. Justify your answer.

? Exercise 1.14.8

Determine which Mersenne numbers M_n are prime when $2 \leq n \leq 12$. You may use a computer or computer algebra system for this exercise, though in your answer you should describe all steps and tools you use.

? Exercise 1.14.9

Show using the Lucas-Lehmer test that $M_7 = 127$ is prime.

Footnotes

[1] This value was announced in October 2020 by user ryanp at <https://www.mersenneforum.org/showth...=15449&page=28>.

This page titled 1.14: Fermat Primes and Mersenne Primes is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Mike Barrus & W. Edwin Clark.

1.15: Number Theoretic Functions

We will return to the Mersenne primes in the next chapter. Before we do, we introduce a few concepts that will be helpful both then and later in the text.

The prime-counting function $\pi(x)$ appearing in the Prime Number Theorem ([Theorem 1.11.3](#)) and the prime-generating functions imagined and studied in [Section 1.14](#) are by no means the only functions studied in number theory. Mathematicians through history have profitably looked at several additional functions tied to our key questions about the integers. In this chapter we introduce three of these. Though their definitions may seem a bit strange at first, hopefully the results of this chapter and the exercises that follow will convince you that their properties are interesting enough to make studying these functions worthwhile.

Functions related to divisors

We start with two functions related to the positive divisors of an integer n .

Definition 1.15.1

For each integer $n > 0$, define

$\tau(n)$ = the number of positive divisors of n ,
 $\sigma(n)$ = the sum of the positive divisors of n .

✓ Example 1.15.1

The integer $12 = 3 \cdot 2^2$ has positive divisors

$1, 2, 3, 4, 6, 12$.

Hence

$$\tau(12) = 6$$

and

$$\sigma(12) = 1 + 2 + 3 + 4 + 6 + 12 = 28.$$

Definition 1.15.2: Proper Divisor

A positive divisor d of n is said to be a *proper divisor* of n if $d < n$. We denote the sum of all proper divisors of n by $\sigma^*(n)$. Note that if $n \geq 2$ then

$$\sigma^*(n) = \sigma(n) - n.$$

✓ Example 1.15.2

Carrying our last example further,

$$\sigma^*(12) = 16.$$

The next theorem shows a simple way to compute $\sigma(n)$ and $\tau(n)$ from the prime factorization of n .

Theorem 1.15.1

Let

$$n = p_1^{e_1} p_2^{e_2} \cdots p_r^{e_r}, \quad r \geq 1,$$

where $p_1 < p_2 < \cdots < p_r$ are primes and $e_i \geq 0$ for each $i \in \{1, 2, \dots, r\}$. Then

$$1. \tau(n) = (e_1 + 1)(e_2 + 1) \cdots (e_r + 1) ;$$

$$2. \sigma(n) = \left(\frac{p_1^{e_1+1} - 1}{p_1 - 1} \right) \left(\frac{p_2^{e_2+1} - 1}{p_2 - 1} \right) \cdots \left(\frac{p_r^{e_r+1} - 1}{p_r - 1} \right).$$

Proof

(1) From the Fundamental Theorem of Arithmetic, every positive factor d of n will have its prime factors coming from those of n . Hence $d \mid n$ if and only if $d = p_1^{f_1} p_2^{f_2} \cdots p_r^{f_r}$ for some integer exponents f_i where for each i ,

$$0 \leq f_i \leq e_i.$$

That is, for each f_i we can choose a value in the set of $e_i + 1$ numbers $\{0, 1, 2, \dots, e_i\}$. So, in all, there are $(e_1 + 1)(e_2 + 1) \cdots (e_r + 1)$ choices for the exponents $f_1, f_2, \dots, f_r$, and by the Fundamental Theorem of Arithmetic, each set of choices yields a different factor. So (1) holds.

To prove (2), we first establish two lemmas.

Before proving this let's look at an example. Take $n = 72 = 8 \cdot 9 = 2^3 \cdot 3^2$. The theorem says that

$$\tau(72) = (3 + 1)(2 + 1) = 12 \quad \text{and}$$

$$\sigma(72) = \left(\frac{2^4 - 1}{2 - 1} \right) \left(\frac{3^3 - 1}{3 - 1} \right) = 15 \cdot 13 = 195.$$

You can compute $\tau(72)$ and $\sigma(72)$ from their original definitions to verify that these results are correct (and much quicker to compute!).

Lemma 1.15.1

Let $n = ab$ where $a, b > 0$ and $\gcd(a, b) = 1$. Then $\sigma(n) = \sigma(a)\sigma(b)$.

Proof

Since a and b have only 1 as a positive common factor, using the Fundamental Theorem of Arithmetic it is easy to see that $d \mid ab \Leftrightarrow d = d_1 d_2$ where $d_1 \mid a$ and $d_2 \mid b$. That is, the divisors of ab are products of the divisors of a and the divisors of b . Let

$$1, a_1, \dots, a_s$$

denote the divisors of a and let

$$1, b_1, \dots, b_t$$

denote the divisors of b . Then

$$\sigma(a) = 1 + a_1 + a_2 + \cdots + a_s,$$

$$\sigma(b) = 1 + b_1 + b_2 + \cdots + b_t.$$

The divisors of $n = ab$ can be listed as follows

$$\begin{aligned} &1, b_1, b_2, \dots, b_t, \\ &a_1 \cdot 1, a_1 \cdot b_1, a_1 \cdot b_2, \dots, a_1 \cdot b_t, \\ &a_2 \cdot 1, a_2 \cdot b_1, a_2 \cdot b_2, \dots, a_2 \cdot b_t, \\ &\vdots \\ &a_s \cdot 1, a_s \cdot b_1, a_s \cdot b_2, \dots, a_s \cdot b_t. \end{aligned}$$

It is important to note that since $\gcd(a, b) = 1$, $a_i b_j = a_k b_\ell$ implies that $a_i = a_k$ and $b_j = b_\ell$. That is there are no repetitions in the above array.

If we sum each row we get

$$\begin{aligned} 1 + b_1 + \cdots + b_t &= \sigma(b) \\ a_1 1 + a_1 b_1 + \cdots + a_1 b_t &= a_1 \sigma(b) \\ &\vdots \\ a_s \cdot 1 + a_s b_1 + \cdots + a_s b_t &= a_s \sigma(b). \end{aligned}$$

By adding these partial sums together we get

$$\begin{aligned} \sigma(n) &= \sigma(b) + a_1 \sigma(b) + a_2 \sigma(b) + \cdots + a_s \sigma(b) \\ &= (1 + a_1 + a_2 + \cdots + a_s) \sigma(b) \\ &= \sigma(a) \sigma(b). \end{aligned}$$

This proves the lemma.

Lemma 1.15.2

If p is a prime and $k \geq 0$ we have

$$\sigma(p^k) = \frac{p^{k+1} - 1}{p - 1}.$$

Proof

Since p is prime, the divisors of p^k are $1, p, p^2, \dots, p^k$. A standard formula for geometric series yields

$$\sigma(p^k) = 1 + p + p^2 + \cdots + p^k = \frac{p^{k+1} - 1}{p - 1},$$

as desired.

Proof of Theorem 1.15.1

Proof

(2) Let $n = p_1^{e_1} p_2^{e_2} \cdots p_r^{e_r}$. Our proof is by induction on r . If $r = 1$, $n = p_1^{e_1}$ and the result follows from [Lemma 1.15.2](#). Suppose the result is true when $1 \leq r \leq k$. Consider now the case $r = k + 1$. That is, let

$$n = p_1^{e_1} \cdots p_k^{e_k} p_{k+1}^{e_{k+1}}$$

where the primes $p_1, \dots, p_k, p_{k+1}$ are distinct and $e_i \geq 0$. Let $a = p_1^{e_1} \cdots p_k^{e_k}$, $b = p_{k+1}^{e_{k+1}}$. Clearly $\gcd(a, b) = 1$. So by [Lemma 1.15.1](#) we have $\sigma(n) = \sigma(a) \sigma(b)$. By the induction hypothesis

$$\sigma(a) = \left(\frac{p_1^{e_1+1} - 1}{p_1 - 1} \right) \cdots \left(\frac{p_k^{e_k+1} - 1}{p_k - 1} \right)$$

and by [Lemma 1.15.2](#)

$$\sigma(b) = \frac{p_{k+1}^{e_{k+1}+1} - 1}{p_{k+1} - 1}$$

and it follows that

$$\sigma(n) = \left(\frac{p_1^{e_1+1} - 1}{p_1 - 1} \right) \cdots \left(\frac{p_{k+1}^{e_{k+1}+1} - 1}{p_{k+1} - 1} \right).$$

So the result holds for $r = k + 1$. By PMI it holds for $r \geq 1$.

The functions $\sigma(n)$ and $\sigma^*(n)$ will appear in the next chapter as we introduce perfect numbers. Additional properties of $\tau(n)$ and $\sigma(n)$ are discussed in the exercises.

Euler's totient function

The final function we will introduce in this chapter is known as Euler's phi-function, or the Euler totient function. As opposed to τ and σ , which dealt with divisors of their input, ϕ will deal with numbers that have *no* prime factor in common with n .

Definition 1.15.3

If X is a set, the number of elements in X is denoted by $|X|$.

For example, $|\{1\}| = 1$ and $|\{0, 1, 3, 9\}| = 4$

Definition 1.15.4

If $m \geq 1$, the *Euler totient function* of n is defined by

$$\phi(n) = |\{i \in \mathbb{Z} \mid 1 \leq i \leq n \text{ and } \gcd(i, n) = 1\}|.$$

At first glance, $\phi(n)$ may seem tedious to calculate. For instance, in order to compute $\phi(1000)$, would a person need to list all numbers relatively prime to 1000 and then count them?

Fortunately, the following theorems show that once the prime factorization of n is given, computing $\phi(n)$ is easy.

Theorem 1.15.2

If $a > 0$ and $b > 0$ and $\gcd(a, b) = 1$, then

$$\phi(ab) = \phi(a)\phi(b).$$

Theorem 1.15.3

If p is prime and $n > 0$ then

$$\phi(p^n) = p^n - p^{n-1}.$$

Theorem 1.15.4

Let $p_1, p_2, \dots, p_k$ be distinct primes and let $n_1, n_2, \dots, n_k$ be positive integers, then

$$\phi(p_1^{n_1} p_2^{n_2} \cdots p_k^{n_k}) = (p_1^{n_1} - p_1^{n_1-1}) \cdots (p_k^{n_k} - p_k^{n_k-1}).$$

Before discussing the proofs of these three theorems, let's illustrate their use:

$$\begin{aligned} \phi(12) &= \phi(2^2 \cdot 3) = (2^2 - 2^1)(3^1 - 3^0) = 2 \cdot 2 = 4 \\ \phi(9000) &= \phi(2^3 \cdot 5^3 \cdot 3^2) = (2^3 - 2^2)(5^3 - 5^2)(3^2 - 3^1) \\ &= 4 \cdot 100 \cdot 6 = 2400. \end{aligned}$$

Note that if p is any prime then

$$\phi(p) = p - 1.$$

We will postpone a proof of [Theorem 1.15.2](#) until [Section 1.23](#). Here we give the proof of [Theorem 1.15.3](#)

Proof of Theorem 1.15.3

Proof

We want to count the number of elements in the set $A = \{1, 2, \dots, p^n\}$ that are relatively prime to p^n . Let B be the set of elements of A that have a factor greater than 1 in common with A . Note that if $b \in B$ and $\gcd(b, p^n) = d > 1$, then d is a

factor of p^n and $d > 1$ so d has p as a factor. Hence $b = pk$, for some k , and $p \leq b \leq p^n$, so $p \leq kp \leq p^n$. It follows that $1 \leq k \leq p^{n-1}$. That is,

$$B = \{p, 2p, 3p, \dots, kp, \dots, p^{n-1}p\}.$$

We are interested in the number of elements of A not in B . Since $|A| = p^n$ and $|B| = p^{n-1}$, this number is $p^n - p^{n-1}$. That is, $\phi(p^n) = p^n - p^{n-1}$.

The proof of [Theorem 1.15.4](#) follows from [Theorems 1.15.2](#) and [1.15.3](#). The proof is by induction on n and is quite similar to the proof of [Theorem 1.15.1\(2\)](#), so we omit the details. Other properties of the ϕ -function are developed in the exercises.

Multiplicative functions

The functions τ , σ , and ϕ all have a common property, shown in [Theorem 1.15.1](#), [Lemma 1.15.1](#), and [Theorem 1.15.2](#).

Definition 1.15.5: Multiplicative

A function $f(n)$ defined for positive integers n is called *multiplicative* if $f(ab) = f(a)f(b)$ whenever $\gcd(a, b) = 1$.

From the results mentioned above, τ , σ , and ϕ are all multiplicative functions. If you are so inclined, do some additional reading on your own to learn about several pleasing properties that multiplicative functions have in common; it will be worth the effort!

Some final words

Later studies in number theory will lead you into greater depth with τ , σ , and ϕ , as well as with the prime-counting function π , and will introduce you to still other functions satisfying remarkable properties. Though we will say little more in this book about number theoretic functions,¹ we finish our discussion with an intriguing unsolved problem in number theory.

In 1907 Robert Carmichael announced that he had proved the following statement:

Carmichael's Conjecture

For every positive integer n there exists a different positive integer m for which $\phi(n) = \phi(m)$.

In other words, Carmichael's statement is that if we were to list $\phi(n)$ for all positive integers n , each value would show up at least twice; every integer shares its ϕ -value with at least one other integer.

Unfortunately, Carmichael's proof was faulty, and today the conjecture still has not been proved true (or disproved!). In 1998 Kevin Ford proved that if Carmichael's conjecture is *not* true, then the smallest counterexample (i.e., the smallest number n such that no other number has the same ϕ -value as n) must be larger than $10^{10^{10}}$. That's huge! On the other hand, mathematicians have shown that if any counterexample exists, then there are infinitely many counterexamples.

You can find more about Carmichael's conjecture with a quick internet or library search. Happy reading!

Exercises

Exercise 1.15.1

Find $\sigma(n)$ and $\tau(n)$ for the following values of n .

- $n = 900$
- $n = 496$
- $n = 32$
- $n = 128$
- $n = 1024$

? Exercise 1.15.2

Does [Lemma 1.15.1](#) hold if we replace σ by σ^* ?

(Hint: The answer is no, but find explicit numbers a and b such that the result fails yet $\gcd(a, b) = 1$.)

? Exercise 1.15.3

Prove that $\tau(n)$ is odd if and only if n is a square.

? Exercise 1.15.4

Prove that $\prod_{d|n} d = n^{\tau(n)/2}$, where the product is over all positive divisors d of n .

? Exercise 1.15.5

Observe that $n = 30$ can be written in multiple ways as the sum of one or more consecutive positive integers:

$$30, \quad 9 + 10 + 11, \quad 6 + 7 + 8 + 9, \quad 4 + 5 + 6 + 7 + 8.$$

Show that for every positive integer n , if $n = 2^p q$, where $p \geq 0$ and q is odd, then $\tau(q)$ is the number of ways that n can be written as the sum of a sequence of consecutive integers.

(Hint: For multiple values of n , try looking at all the ways that n can be written as a sum of consecutive positive integers. Try to match different sums up with different divisors of q in a way that always works, no matter what n is.)

? Exercise 1.15.6

Show that if

$$m = p_1^{n_1} p_2^{n_2} \cdots p_k^{n_k}$$

where $p_1, \dots, p_k$ are distinct primes and each $n_i \geq 1$, then

$$\phi(m) = m \left(1 - \frac{1}{p_1}\right) \left(1 - \frac{1}{p_2}\right) \cdots \left(1 - \frac{1}{p_k}\right).$$

? Exercise 1.15.7

Prove that $\phi(n)$ is even for all positive integers n other than 1 or 2.

? Exercise 1.15.8

Determine, with proof, all numbers n such that

- $\phi(n) = 2$
- $\phi(n) = 4$
- $\phi(n) = 6$

? Exercise 1.15.9

Prove that if $\phi(n) \mid (n-1)$, then no prime appears more than once in the prime factorization of n .

? Exercise 1.15.10

Let

$$S(n) = \sum_{d|n} \phi(d),$$

i.e., $S(n)$ is the sum of the values we get when we evaluate $\phi(d)$ for all positive divisors d of n . For example,

$$S(4) = \phi(1) + \phi(2) + \phi(4) = 1 + 1 + 2 = 4, \quad \text{and}$$

$$S(15) = \phi(1) + \phi(3) + \phi(5) + \phi(15) = 1 + 2 + 4 + 18 = 15.$$

Compute $S(n)$ for $1 \leq n \leq 10$ and conjecture² a formula for $S(n)$.

? Exercise 1.15.11

For each function below, decide whether the function is multiplicative or not. (Assume that functions are defined on the set of positive integers.)

- a. $f(n) = 1$
- b. $f(n) = n$
- c. $f(n) = 1 + n$
- d. $f(n) = \log(n)$
- e. $f(n) = 2^n$
- f. $f(n) = 1/n$

? Exercise 1.15.12

For positive integers n , define the function ρ by

$$\rho(n) = \text{the number of prime factors of } n;$$

for example, $\rho(4) = 1$ (since 4 is divisible by 2 and no other prime) and $\rho(100) = 2$ (since 100 is divisible by 2 and 5 and no other prime).

- a. Is the function $\rho(n)$ multiplicative? Explain why or why not.
- b. Is the function $f(n) = (-1)^{\rho(n)}$ multiplicative? Explain why or why not.

? Exercise 1.15.13

Show that for any positive integers a and b , if f is a multiplicative function, then $f(a) \cdot f(b) = f(\gcd(a, b)) \cdot f(\text{lcm}(a, b))$.

Footnotes

[1] You'll find more, though, in the exercises in this chapter.

[2] Though the proof of your conjecture, once you make it, is a bit beyond the scope of this text, you may find a proof in many other textbooks on number theory.

This page titled 1.15: Number Theoretic Functions is shared under a [not declared](#) license and was authored, remixed, and/or curated by [Mike Barrus & W. Edwin Clark](#).

1.16: Perfect Numbers and Mersenne Primes

In this chapter we discuss perfect numbers.¹

Definition 1.16.1: Perfect

An integer $n > 1$ is *perfect* if $\sigma^*(n) = n$; in other words, n is perfect if it equals the sum of its proper divisors.

An equivalent way to define perfect numbers is as those numbers n for which $\sigma(n) = 2n$.

Example 1.16.1

The proper divisors of 6 are 1, 2 and 3. So $\sigma^*(6) = 6$. Therefore 6 is perfect.

If you do a search for perfect numbers up to 10,000 you will find only the following perfect numbers:

$$\begin{aligned} 6 &= 2 \cdot 3, \\ 28 &= 2^2 \cdot 7, \\ 496 &= 2^4 \cdot 31, \\ 8128 &= 2^6 \cdot 127. \end{aligned}$$

Note that $2^2 = 4$, $2^3 = 8$, $2^5 = 32$, $2^7 = 128$ so we have:

$$\begin{aligned} 6 &= 2 \cdot (2^2 - 1), \\ 28 &= 2^2 \cdot (2^3 - 1), \\ 496 &= 2^4 \cdot (2^5 - 1), \\ 8128 &= 2^6 \cdot (2^7 - 1). \end{aligned}$$

Note also that $2^2 - 1$, $2^3 - 1$, $2^5 - 1$, $2^7 - 1$ are Mersenne primes. One might conjecture that all perfect numbers follow this pattern. We discuss to what extent this is known to be true. We start with the following result.

Theorem 1.16.1: Mersenne Yields Perfect

If $2^p - 1$ is a Mersenne prime, then $2^{p-1} \cdot (2^p - 1)$ is perfect.

Proof

Write $q = 2^p - 1$ and let $n = 2^{p-1}q$. If q is a Mersenne prime, then since it is odd and prime, by [Theorem 1.15.1\(2\)](#) we have $\sigma(n) = \sigma(2^{p-1}q) = \left(\frac{2^p-1}{2-1}\right) \left(\frac{q^2-1}{q-1}\right) = (2^p-1)(q+1) = (2^p-1)2^p = 2n$. That is, $\sigma(n) = 2n$ and n is perfect.

Now we show that all *even* perfect numbers have the conjectured form.

Theorem 1.16.2: Even Perfects

If n is even and perfect then there is a Mersenne prime $2^p - 1$ such that $n = 2^{p-1}(2^p - 1)$.

Proof

Let n be even and perfect. Since n is even, $n = 2m$ for some m . We take out as many powers of 2 as possible, obtaining

$$n = 2^k \cdot q, \quad k \geq 1, q \text{ odd.}$$

Since n is perfect, $\sigma^*(n) = n$, that is, $\sigma(n) = 2n$. Since q is odd, $\gcd(2^k, q) = 1$, so by [Lemmas 1.15.1](#) and [1.15.2](#),

$$\sigma(n) = \sigma(2^k)\sigma(q) = (2^{k+1} - 1)\sigma(q).$$

So we have

$$2^{k+1}q = 2n = \sigma(n) = (2^{k+1} - 1)\sigma(q),$$

hence

$$2^{k+1}q = (2^{k+1} - 1)\sigma(q). \quad (1.16.1)$$

Now $\sigma^*(q) = \sigma(q) - q$, so

$$\sigma(q) = \sigma^*(q) + q.$$

Putting this in equation (1.16.1) we get

$$2^{k+1}q = (2^{k+1} - 1)(\sigma^*(q) + q)$$

or

$$2^{k+1}q = (2^{k+1} - 1)\sigma^*(q) + 2^{k+1}q - q,$$

which implies

$$\sigma^*(q)(2^{k+1} - 1) = q. \quad (1.16.2)$$

In other words, $\sigma^*(q)$ is a divisor of q . Since $k \geq 1$ we have $2^{k+1} - 1 \geq 4 - 1 = 3$. So $\sigma^*(q)$ is a proper divisor of q . But $\sigma^*(q)$ is the sum of *all* proper divisors of q . This can only happen if q has only one proper divisor. This means that q must be prime and $\sigma^*(q) = 1$. Then equation (1.16.2) shows that $q = 2^{k+1} - 1$. So q must be a Mersenne prime and $k + 1 = p$ is prime. So $n = 2^{p-1}(2^p - 1)$, as desired.

Corollary 1.16.1

There is a one-to-one correspondence between *even* perfect numbers and Mersenne primes.

Even knowing of the strong connection between perfect numbers and Mersenne primes, a number of well-known questions remain.

Open questions

- Are there infinitely many perfect numbers?
- Are there any odd perfect numbers?

Of course showing that there are infinitely many Mersenne primes would answer the first question.

So far no one has found a single odd perfect number. It is known that if an odd perfect number exists, it must be $> 10^{50}$.

The idea of a perfect number is pretty old, as is the result of [Theorem 1.16.1](#). Euclid's *Elements*² defines perfect numbers at the beginning of Book VII, and a proof that Mersenne primes can be used to build the even perfect numbers appears as Proposition 36 in Book IX. Of course, Euclid's use of "Mersenne" primes dates to about 1800 years before Mersenne was born, so he didn't use exactly that terminology (or our modern algebraic notation), but the ideas are similar and worth a read.

Some even think that Euclid's knowledge that $2^{p-1}(2^p - 1)$ is perfect when $2^p - 1$ is prime may have been his motivation for defining prime numbers.

Exercises

Exercise 1.16.1

Use the definition of perfect numbers to determine which of the following numbers are perfect, showing your work.

- $n = 28$
- $n = 32$
- $n = 128$
- $n = 496$
- $n = 900$
- $n = 1024$
- $n = 8128$

? Exercise 1.16.2

Prove that if n is a perfect number, then $\sum_{d|n} \frac{1}{d} = 2$, where the sum is over positive divisors d of n .

(Hint: If you cleared the denominators in this equation, what would the equation look like?)

? Exercise 1.16.3

- Using results from this chapter, show that if M is a Mersenne prime, then $\frac{1}{2}M(M+1)$ is a perfect number.
- Show that if M is a Mersenne prime, then the sum

$$1 + 2 + 3 + \cdots + M$$

is a perfect number. (Hint: use part (a), and if necessary, look at the formula in [Exercise 1.3.3](#).)

The next two exercises show that the sets of *perfect squares* (numbers of the form n^2) and of *perfect numbers* do not overlap.

? Exercise 1.16.4

Show that if n is odd, then n^2 is not perfect.

(Hint: Compare the parities of n^2 and $\sigma^*(n^2)$.)

? Exercise 1.16.5

Show that if n is even, then n^2 is not perfect.

(Hint: use [Theorem 1.16.2](#).)

Footnotes

[1] Note that our usage of "perfect" here and the expression "perfect square" elsewhere in the text are unrelated; squares of integers are never perfect numbers (see [Exercises 1.16.4](#) and [1.16.5](#) at the end of the chapter), so there is no perfect perfect square!

[2] Remember that we've already come across *Elements* in [Sections 1.8](#) and [1.11](#) and [1.12](#). This ancient text definitely contains some gems!

This page titled [1.16: Perfect Numbers and Mersenne Primes](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Mike Barrus & W. Edwin Clark](#).

1.17: Congruences

Definition 1.17.1

Let $m \geq 0$. We write $a \equiv b \pmod{m}$ if $m \mid (a - b)$, and we say that a is *congruent to b modulo m* . Here m is said to be the *modulus of the congruence*. The notation $a \not\equiv b \pmod{m}$ means that it is false that $a \equiv b \pmod{m}$.

✓ Example 1.17.1

1. $25 \equiv 1 \pmod{4}$ since $4 \mid 24$
2. $25 \not\equiv 2 \pmod{4}$ since $4 \nmid 23$
3. $1 \equiv -3 \pmod{4}$ since $4 \mid 4$
4. $a \equiv b \pmod{1}$ for all a, b since “1 divides everything.”
5. $a \equiv b \pmod{0} \iff a = b$ for all a, b since “0 divides only 0.”

Remark 1.17.1

As you see, the cases $m = 1$ and $m = 0$ are not very interesting so mostly we will only be interested in the case $m \geq 2$.

Warning

Do not confuse the use of mod in [Definition 1.17.1](#) with that of [Definition 1.5.3](#). We shall see that the two uses of mod are related, but have different meanings: Recall

$a \pmod{b} = r$ where r is the remainder given by the Division Algorithm when a is divided by b ,

and by [Definition 1.17.1](#)

$a \equiv b \pmod{m}$ means $m \mid (a - b)$.

✓ Example 1.17.2

The statement

$$25 \equiv 5 \pmod{4} \text{ is true,}$$

since $4 \mid 20$ but

$$25 = 5 \bmod 4 \text{ is false,}$$

since the latter means $25 = 1$.

Remark 1.17.2

The mod in $a \equiv b \pmod{m}$ is, together with the $\equiv$, part of a **binary relation**, whereas the mod in $a \bmod b$ is a **binary operation**.

More terminology:

Expressions such as

$$\begin{aligned} x &= 2 \\ 4^2 &= 16 \\ x^2 + 2x &= \sin(x) + 3 \end{aligned}$$

are **equations**. By analogy, expressions such as

$$\begin{aligned}x &\equiv 2 \pmod{16} \\25 &\equiv 5 \pmod{5} \\x^3 + 2x &\equiv 6x^2 + 3 \pmod{27}\end{aligned}$$

are called **congruences**. Before discussing further the analogy between equations and congruences, we show the relationship between the two different definitions of mod.

Theorem 1.17.1

For $m > 0$ and for all a, b :

$$a \equiv b \pmod{m} \iff a \bmod m = b \bmod m.$$

Proof

“ $\Rightarrow$ ” Assume that $a \equiv b \pmod{m}$. Let $r_1 = a \bmod m$ and $r_2 = b \bmod m$. We want to show that $r_1 = r_2$. By definition we have

1. $m \mid (a - b)$,
2. $a = mq_1 + r_1$, $0 \leq r_1 < m$, and
3. $b = mq_2 + r_2$, $0 \leq r_2 < m$.

From (1) we obtain

$$a - b = mt$$

for some t . Hence

$$a = mt + b.$$

Using (2) and (3) we see that

$$a = mq_1 + r_1 = m(q_2 + t) + r_2.$$

Since $0 \leq r_1 < m$ and $0 \leq r_2 < m$ by the uniqueness part of the Division Algorithm we obtain $r_1 = r_2$, as desired.

“ $\Leftarrow$ ” Assume that $a \bmod m = b \bmod m$. We must show that $a \equiv b \pmod{m}$. If $r = a \bmod m = b \bmod m$, then by definition we have

$$a = mq_1 + r, \quad 0 \leq r < m,$$

and

$$b = mq_2 + r, \quad 0 \leq r < m.$$

Hence

$$a - b = m(q_1 - q_2).$$

This shows that $m \mid (a - b)$ and hence $a \equiv b \pmod{m}$, as desired.

The next two theorems show that congruences and equations share many similar properties.

Theorem 1.17.2

For all a, b, c and $m > 0$ we have

1. $a \equiv a \pmod{m}$ [**reflexivity**]
2. $a \equiv b \pmod{m} \Rightarrow b \equiv a \pmod{m}$ [**symmetry**]
3. $a \equiv b \pmod{m}$ and $b \equiv c \pmod{m} \Rightarrow a \equiv c \pmod{m}$ [**transitivity**]

Proof

of (1) See [Exercise 1.17.1](#).

of (2) If $a \equiv b \pmod{m}$, then $m \mid (a - b)$. Hence $a - b = mq$. Hence $b - a = m(-q)$, so $m \mid (b - a)$. Hence $b \equiv a \pmod{m}$.

of (3) If $a \equiv b \pmod{m}$ and $b \equiv c \pmod{m}$ then $m \mid (a - b)$ and $m \mid (b - c)$. By the linearity property $m \mid ((a - b) + (b - c))$. That is, $m \mid (a - c)$. Hence $a \equiv c \pmod{m}$.

Recall that a *polynomial* is an expression of the form

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0.$$

Here we will assume that the coefficients $a_n, \dots, a_0$ are integers and x also represents an integer variable. Here, of course, $n \geq 0$ and n is an integer.

Theorem 1.17.3

If $a \equiv b \pmod{m}$ and $c \equiv d \pmod{m}$, then

1. $a \pm c \equiv b \pm d \pmod{m}$
2. $ac \equiv bd \pmod{m}$
3. $a^n \equiv b^n \pmod{m}$ for all $n \geq 1$
4. $f(a) \equiv f(b) \pmod{m}$ for all polynomials $f(x)$ with integer coefficients.

Proof

of (1) To prove (1) since $a - c = a + (-c)$, it suffices to prove only the “+ case.” By assumption $m \mid (a - b)$ and $m \mid (c - d)$. By linearity, $m \mid ((a - b) + (c - d))$, that is $m \mid ((a + c) - (b + d))$. Hence

$$a + c \equiv b + d \pmod{m}.$$

of (2) Since $m \mid (a - b)$ and $m \mid (c - d)$ by linearity,

$$m \mid (c(a - b) + b(c - d)).$$

Now $c(a - b) + b(c - d) = ca - bd$, hence

$$m \mid (ca - bd),$$

and $ca \equiv bd \pmod{m}$, as desired.

of (3) We prove $a^n \equiv b^n \pmod{m}$ by induction on n . If $n = 1$, the result is true by our assumption that $a \equiv b \pmod{m}$. Assume it holds for $n = k$. Then we have $a^k \equiv b^k \pmod{m}$. This, together with $a \equiv b \pmod{m}$ using (2) above, gives $aa^k \equiv bb^k \pmod{m}$. Hence $a^{k+1} \equiv b^{k+1} \pmod{m}$. So it holds for all $n \geq 1$, by the PMI.

of (4) Let $f(x) = c_n x^n + \cdots + c_1 x + c_0$. We prove by induction on n that if $a \equiv b \pmod{m}$ then

$$c_n a^n + \cdots + c_0 \equiv c_n b^n + \cdots + c_0 \pmod{m}.$$

If $n = 0$ we have $c_0 \equiv c_0 \pmod{m}$ by Theorem 1.17.2(1). Assume the result holds for $n = k$. Then we have

$$c_k a^k + \cdots + c_1 a + c_0 \equiv c_k b^k + \cdots + c_1 b + c_0 \pmod{m}. \quad (1.17.1)$$

By part (3) above we have $a^{k+1} \equiv b^{k+1} \pmod{m}$. Since $c_{k+1} \equiv c_{k+1} \pmod{m}$ using (2) above we have

$$c_{k+1} a^{k+1} \equiv c_{k+1} b^{k+1} \pmod{m}. \quad (1.17.2)$$

Now we can apply Theorem 1.17.3(1) to equations (1.17.1) and (1.17.2) to obtain

$$c_{k+1} a^{k+1} + c_k a^k + \cdots + c_0 \equiv c_{k+1} b^{k+1} + c_k b^k + \cdots + c_0 \pmod{m}.$$

So by the PMI, the result holds for $n \geq 0$.

Before continuing to develop properties of congruences, we give the following example to show one way that congruences can be useful.

✓ Example 1.17.3

(This example was taken from [1] *Introduction to Analytic Number Theory*, by Tom Apostol.)

The first five Fermat numbers

$$F_0 = 3, \quad F_1 = 5, \quad F_2 = 17, \quad F_3 = 257, \quad F_4 = 65,537$$

are primes. We show using congruences *without explicitly calculating* F_5 that $F_5 = 2^{32} + 1$ is divisible by 641 and is therefore not prime: [apostle]

$$\begin{aligned} 2^2 &= 4 \\ 2^4 &= (2^2)^2 = 4^2 = 16 \\ 2^8 &= (2^4)^2 = 16^2 = 256 \\ 2^{16} &= (2^8)^2 = 256^2 = 65,536 \\ 65,536 &\equiv 154 \pmod{641}. \end{aligned}$$

So we have

$$2^{16} \equiv 154 \pmod{641}.$$

By [Theorem 1.17.3\(3\)](#):

$$(2^{16})^2 \equiv (154)^2 \pmod{641}.$$

That is,

$$2^{32} \equiv 23,716 \pmod{641}.$$

Since

$$23,716 \equiv 640 \pmod{641}$$

and

$$640 \equiv -1 \pmod{641}$$

we have

$$2^{32} \equiv -1 \pmod{641}$$

and hence

$$2^{32} + 1 \equiv 0 \pmod{641}.$$

So $641 \mid (2^{32} + 1)$, as claimed. Clearly $2^{32} + 1 \neq 641$, so $2^{32} + 1$ is composite. Of course, if you already did [Exercise 1.12.1](#) you will already know that

$$2^{32} + 1 = 4,294,967,297 = (641) \cdot (6,700,417)$$

and that 641 and 6,700,417 are indeed primes. Note that 641 is the 116th prime, so if you used trial division you would have had to divide by 115 primes before reaching one that divides $2^{32} + 1$, and this assumes that you have a list of the first 116 primes.

Theorem 1.17.4

If $m > 0$ and

$$a \equiv r \pmod{m} \text{ where } 0 \leq r < m,$$

then $a \bmod m = r$.

Exercises

? Exercise 1.17.1

Prove that for all $m > 0$ and for all a ,

$$a \equiv a \bmod m \pmod{m}.$$

? Exercise 1.17.2

Using [Definition 1.17.1](#), show that the following congruences are true:

$$\begin{aligned} 385 &\equiv 322 \pmod{3}; \\ -385 &\equiv -322 \pmod{3}; \\ 1 &\equiv -17 \pmod{3}; \\ 33 &\equiv 0 \pmod{3}. \end{aligned}$$

? Exercise 1.17.3

Use [Theorem 1.17.1](#) to show that the congruences in [Exercise 1.17.2](#) are valid.

? Exercise 1.17.4

- a. Show that a is even if and only if $a \equiv 0 \pmod{2}$, and a is odd if and only if $a \equiv 1 \pmod{2}$.
- b. Show that a is even if and only if $a \bmod 2 = 0$, and a is odd if and only if $a \bmod 2 = 1$.

(Note: your answers to parts (a) and (b) will be similar, but they should not be identical! Make every effort to use the vocabulary correctly.)

? Exercise 1.17.5

Show that if $m > 0$ and a is any integer, there is a *unique* integer $r \in \{0, 1, 2, \dots, m-1\}$ such that $a \equiv r \pmod{m}$.

? Exercise 1.17.6

Find integers a and b such that $0 < a < 15$, $0 < b < 15$ and $ab \equiv 0 \pmod{15}$.

? Exercise 1.17.7

Find **three** separate pairs a, b of integers such that $1 < a < 15$, $1 < b < 15$, and $ab \equiv 1 \pmod{15}$.

? Exercise 1.17.8

Show that if $d \mid m$ and $d > 0$, then

$$a \equiv b \pmod{m} \Rightarrow a \equiv b \pmod{d}.$$

? Exercise 1.17.9

Prove [Theorem 1.17.4](#)

(Hint: The Division Algorithm may be useful.)

? Exercise 1.17.10

Find the value of each of the following (*without using a computer!*).

1. $2^{32} \bmod 7$
2. $10^{35} \bmod 7$
3. $3^{35} \bmod 7$

(Hint: Use [Theorem 1.17.4](#) and the ideas used in [Example 1.17.3](#).)

? Exercise 1.17.11

Let $\gcd(m_1, m_2) = 1$. Prove that

$$a \equiv b \pmod{m_1} \quad \text{and} \quad a \equiv b \pmod{m_2}$$

if and only if

$$a \equiv b \pmod{m_1 m_2}.$$

(Hint: use [Lemma 1.12.1](#).)

This page titled [1.17: Congruences](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Mike Barrus & W. Edwin Clark](#).

1.18: Divisibility Tests for 2, 3, 5, 9, 11

Recall from [Definition 1.6.1](#) that the decimal representation of the positive integer a is given by

$$a = a_{n-1}a_{n-2} \cdots a_1a_0 \quad (1.18.1)$$

when

$$a = a_{n-1}10^{n-1} + a_{n-2}10^{n-2} + \cdots + a_110 + a_0$$

and $0 \leq a_i \leq 9$ for $i = 0, 1, \dots, n-1$.

Theorem 1.18.1

Let the decimal representation of a be given by $a = a_{n-1}a_{n-2} \cdots a_1a_0$. Then

- $a \bmod 2 = a_0 \bmod 2$,
- $a \bmod 5 = a_0 \bmod 5$,
- $a \bmod 3 = (a_{n-1} + \cdots + a_0) \bmod 3$,
- $a \bmod 9 = (a_{n-1} + \cdots + a_0) \bmod 9$,
- $a \bmod 11 = (a_0 - a_1 + a_2 - a_3 + \cdots) \bmod 11$.

Before proving this theorem, let's give some examples.

$$1457 \bmod 2 = 7 \bmod 2 = 1;$$

$$1457 \bmod 5 = 7 \bmod 5 = 2;$$

$$\begin{aligned} 1457 \bmod 3 &= (1 + 4 + 5 + 7) \bmod 3 = 17 \bmod 3 \\ &= (1 + 7) \bmod 3 = 8 \bmod 3 = 2; \end{aligned}$$

$$\begin{aligned} 1457 \bmod 9 &= (1 + 4 + 5 + 7) \bmod 9 \\ &= 17 \bmod 9 \\ &= (1 + 7) \bmod 9 \\ &= 8 \bmod 9 \\ &= 8; \end{aligned}$$

$$\begin{aligned} 1457 \bmod 11 &= (7 - 5 + 4 - 1) \bmod 11 \\ &= 5 \bmod 11 \\ &= 5. \end{aligned}$$

Proof of Theorem 1.18.1

Proof

Consider the polynomial

$$f(x) = a_{n-1}x^{n-1} + \cdots + a_1x + a_0.$$

Note that $10 \equiv 0 \pmod{2}$. So by [Theorem 1.17.3 \(4\)](#)

$$a_{n-1}10^{n-1} + \cdots + a_110 + a_0 \equiv a_{n-1}0^{n-1} + \cdots + a_10 + a_0 \pmod{2}.$$

That is,

$$a \equiv a_0 \pmod{2}.$$

This, together with [Theorem 1.17.1](#), proves part (a). Since $10 \equiv 0 \pmod{5}$, the proof of part (b) is similar.

Note that $10 \equiv 1 \pmod{3}$, so applying [Theorem 1.17.3 \(4\)](#) again, we have

$$a_{n-1}10^{n-1} + \cdots + a_110 + a_0 \equiv a_{n-1}1^{n-1} + \cdots + a_11 + a_0 \pmod{3}.$$

That is,

$$a \equiv a_{n-1} + \cdots + a_1 + a_0 \pmod{3}.$$

This, using [Theorem 1.17.1](#), proves part (c). Since $10 \equiv 1 \pmod{9}$, the proof of part (d) is similar.

Now $10 \equiv -1 \pmod{11}$ so

$$a_{n-1}10^{n-1} + \cdots + a_110 + a_0 \equiv a_{n-1}(-1)^{n-1} + \cdots + a_1(-1) + a_0 \pmod{11}.$$

That is,

$$a \equiv a_0 - a_1 + a_2 - \cdots \pmod{11}$$

and by [Theorem 1.17.1](#) we are done.

Remark 1.18.1

Note that

$$m \mid a \Leftrightarrow a \bmod m = 0,$$

so from [Theorem 1.18.1](#) we obtain immediately the following corollary.

Corollary 1.18.1

Let a be given by expression [\(1.18.1\)](#). Then

- a. $2 \mid a \Leftrightarrow a_0 = 0, 2, 4, 6 \text{ or } 8$
- b. $5 \mid a \Leftrightarrow a_0 = 0 \text{ or } 5$
- c. $3 \mid a \Leftrightarrow 3 \mid (a_0 + a_1 + \cdots + a_{n-1})$
- d. $9 \mid a \Leftrightarrow 9 \mid (a_0 + a_1 + \cdots + a_{n-1})$
- e. $11 \mid a \Leftrightarrow 11 \mid (a_0 - a_1 + a_2 - a_3 + \cdots)$.

Note that in applying (c), (d) and (e) we can use the fact that

$$(a + m) \bmod m = a$$

to “cast out” 3’s (for (c)) and 9’s (for (d)). Here’s an example of “casting out 9’s:”

$$\begin{aligned} 1487 \bmod 9 &= (1 + 4 + 8 + 7) \bmod 9 \\ &= (9 + 4 + 7) \bmod 9 \\ &= (4 + 7) \bmod 9 \\ &= (2 + 9) \bmod 9 \\ &= 2 \bmod 9 = 2. \end{aligned}$$

So $1487 \bmod 9 = 2$.

Note that if $0 \leq r < m$ then
 $r \bmod m = r$.

Exercises

Exercise 1.18.1

Let $a = 18726132117057$ Find $a \bmod m$ for $m = 2, 3, 5, 9$ and 11 .

Exercise 1.18.2

Use the divisibility test presented in this chapter to determine the next year after this one that will be divisible by 11.

? Exercise 1.18.3

Let $a = a_n \cdots a_1 a_0$ be the decimal representation of a . Then prove

- $a \bmod 10 = a_0$.
- $a \bmod 100 = a_1 a_0$.
- $a \bmod 1000 = a_2 a_1 a_0$.

? Exercise 1.18.4

Prove that if b is a positive square, i.e., $b = a^2$, $a > 0$, then the least significant digit of b is one of 0, 1, 4, 5, 6, 9.

(Hint: $b \bmod 10$ is the least significant digit of b . Write $a = a_{n-1} \cdots a_0$. Then $a \equiv a_0 \pmod{10}$ so $a^2 \equiv a_0^2 \pmod{10}$. For each digit $a_0 \in \{0, 1, 2, \dots, 9\}$ find $a_0^2 \bmod 10$. Use [Theorem 1.17.4](#), among other results.)

? Exercise 1.18.5

Write down, with proof, a statement like that in [Exercise 1.18.4](#) that tests whether b is a positive square by looking at the last two digits of its decimal representation.

? Exercise 1.18.6

Are any of the following numbers squares? Explain, without using a device to take a square root.

10, 11, 16, 19, 24, 25, 272, 2983, 11007, 1120378

This page titled [1.18: Divisibility Tests for 2, 3, 5, 9, 11](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Mike Barrus & W. Edwin Clark](#).

1.19: Divisibility Tests for 7 and 13

Theorem 1.19.1

Let $a = a_r a_{r-1} \cdots a_1 a_0$ be the decimal representation of a . Then

- a. $7 \mid a \Leftrightarrow 7 \mid (a_r \cdots a_1 - 2a_0)$.
- b. $13 \mid a \Leftrightarrow 13 \mid (a_r \cdots a_1 - 9a_0)$.

[Here $a_r \cdots a_1 = \frac{a - a_0}{10} = a_r 10^{r-1} + \cdots + a_2 10 + a_1$.]

Before proving this theorem we illustrate it with two examples.

$$\begin{aligned} 7 \mid 2481 &\Leftrightarrow 7 \mid (248 - 2) \\ &\Leftrightarrow 7 \mid 246 \\ &\Leftrightarrow 7 \mid (24 - 12) \\ &\Leftrightarrow 7 \mid 12 \end{aligned}$$

since $7 \nmid 12$ we have $7 \nmid 2481$.

$$\begin{aligned} 13 \mid 12987 &\Leftrightarrow 13 \mid (1298 - 63) \\ &\Leftrightarrow 13 \mid 1235 \\ &\Leftrightarrow 13 \mid (123 - 45) \\ &\Leftrightarrow 13 \mid 78 \end{aligned}$$

since $6 \cdot 13 = 78$, we have $13 \mid 78$. So, by [Theorem 1.19.1\(b\)](#), $13 \mid 12987$.

Proof of Theorem 1.19.1

Proof

of (a)

Let $c = a_r \cdots a_1$. So we have $a = 10c + a_0$. Hence $-2a = -20c - 2a_0$. Now $1 \equiv -20 \pmod{7}$ so we have

$$-2a \equiv c - 2a_0 \pmod{7}.$$

It follows from [Theorem 1.17.1](#) that

$$-2a \bmod 7 = c - 2a_0 \bmod 7.$$

Hence, $7 \mid -2a \Leftrightarrow 7 \mid (c - 2a_0)$. Since $\gcd(7, -2) = 1$ we have $7 \mid -2a \Leftrightarrow 7 \mid a$. Hence $7 \mid a \Leftrightarrow 7 \mid (c - 2a_0)$, which is what we wanted to prove.

of (b)

(This has a similar proof to that for [1.19.1\(a\)](#) and is left for the interested reader.)

Exercises

Exercise 1.19.1

Using [Theorem 1.19.1](#), determine which of the following, if any, is divisible by 7; also decide whether either is divisible by 13. Show all your computations.

- a. 6994
- b. 6993

? Exercise 1.19.2

Give multiple reasons, including an application of [Theorem 1.19.1](#), why no integer of the form 10^n can be divisible by 7 or 13.

? Exercise 1.19.3

Using [Theorem 1.19.1](#), determine the next year (after, possibly, the current year) that is

- a. divisible by 7.
- b. divisible by 13.

? Exercise 1.19.4

Show by example that, in the notation of [Theorem 1.19.1](#), $a \bmod 7$ need not be equal to $(a_r \cdots a_1 - 2a_0) \bmod 7$.

? Exercise 1.19.5

Show that the number $a = a_r a_{r-1} \cdots a_1 a_0$ is divisible by 13 if and only if $a_r a_{r-1} \cdots a_1 + 4a_0$ is divisible by 13. (This gives a slightly different test for divisibility by 13.) Once you have proved this, apply it to test whether 111,111 is divisible by 13.

This page titled [1.19: Divisibility Tests for 7 and 13](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Mike Barrus & W. Edwin Clark](#).

1.20: More Properties of Congruences

In this chapter we introduce an important idea in working with congruences. It will have useful consequences and will illustrate another reason Bezout's Lemma is important. We introduce it with a theorem.

Theorem 1.20.1

Let $m \geq 2$. If a and m are relatively prime, then there exists a unique integer a^* such that $aa^* \equiv 1 \pmod{m}$ and $0 < a^* < m$.

✓ Example 1.20.1

When $m = 9$, the relatively prime values for a are 1, 2, 4, 5, 7, 8. Indeed, observe that

$1 \cdot 1 \equiv 1 \pmod{9},$		$2 \cdot 5 \equiv 1 \pmod{9},$
$4 \cdot 7 \equiv 1 \pmod{9},$		$8 \cdot 8 \equiv 1 \pmod{9}.$

Hence

$$1^* = 1, \quad 2^* = 5, \quad 4^* = 7, \quad 5^* = 2, \quad 7^* = 4, \quad 8^* = 8.$$

Observe, too, that requiring that a be relatively prime to m is important in this case: note that a could not equal 6, since $\gcd(6, 9) \neq 1$, and indeed, any multiple of 6 is congruent to 0, 3, or 6 modulo 9—there is no integer 6^* that produces 1 (modulo 9) when 6 is multiplied by it.

We call a^* the *inverse of a modulo m* . Note that we choose here not to denote a^* by a^{-1} (though we could) since this might cause some confusion. Of course, if $c \equiv a^* \pmod{m}$ then $ac \equiv 1 \pmod{m}$ so a^* is not unique unless we specify that $0 < a^* < m$.

Proof of Theorem 1.20.1

Proof

If $\gcd(a, m) = 1$, then by Bezout's Lemma there exist s and t such that

$$as + mt = 1.$$

Hence

$$as - 1 = m(-t),$$

that is, $m \mid (as - 1)$ and so $as \equiv 1 \pmod{m}$. Let $a^* = s \pmod{m}$. Then $a^* \equiv s \pmod{m}$ so $aa^* \equiv 1 \pmod{m}$ and clearly $0 < a^* < m$.

To show uniqueness assume that $ac \equiv 1 \pmod{m}$ and $0 < c < m$. Then $ac \equiv aa^* \pmod{m}$. So if we multiply both sides of this congruence on the left by c and use the fact that $ca \equiv 1 \pmod{m}$ we obtain $c \equiv a^* \pmod{m}$. It follows from [Exercise 1.17.5](#) that $c = a^*$, and this proves the uniqueness.

Remark 1.20.1

From the above proof we see that Blankinship's Method may be used to compute the inverse of a when it exists, but for small m we may often find a^* by "guess and check." We can even search for it systematically. For example, if $m = 15$ take $a = 2$. Then we can check each element $0, 1, 2, \dots, 14$

$$\begin{aligned}
 2 \cdot 0 &\not\equiv 1 \pmod{15} \\
 2 \cdot 1 &\not\equiv 1 \pmod{15} \\
 2 \cdot 2 &\not\equiv 1 \pmod{15} \\
 2 \cdot 3 &\not\equiv 1 \pmod{15} \\
 2 \cdot 4 &\not\equiv 1 \pmod{15} \\
 2 \cdot 5 &\not\equiv 1 \pmod{15} \\
 2 \cdot 6 &\not\equiv 1 \pmod{15} \\
 2 \cdot 7 &\not\equiv 1 \pmod{15} \\
 2 \cdot 8 &\equiv 1 \pmod{15} \text{ since } 15 \mid (16 - 1).
 \end{aligned}$$

So we can take $2^* = 8$.

Theorem 1.20.1 tells us one way of knowing whether a has an inverse modulo m —we can check to see if a and m are relatively prime. But is it *necessary* that a and m be relative prime for a to have an inverse? The next results give us an answer.

Theorem 1.20.2

Let $m > 0$. If $ab \equiv 1 \pmod{m}$ then both a and b are relatively prime to m .

Proof

If $ab \equiv 1 \pmod{m}$, then $m \mid (ab - 1)$. So $ab - 1 = mt$ for some t . Hence,

$$ab + m(-t) = 1.$$

By [Exercise 1.10.3](#), this implies that $\gcd(a, m) = 1$ and $\gcd(b, m) = 1$, as claimed.

Corollary 1.20.1

The integer a has an inverse modulo m if and only if a and m are relatively prime.

Now that we know exactly which integers have inverses modulo m , let's see what we can do with inverses. First, let's use them in the next theorem to “simplify” a congruence.

Theorem 1.20.3: Cancellation

Let $m > 0$ and assume that $\gcd(c, m) = 1$. Then

$$ca \equiv cb \pmod{m} \Rightarrow a \equiv b \pmod{m}. \quad (1.20.1)$$

If $\gcd(c, m) = 1$, there is an integer c^* such that $c^*c \equiv 1 \pmod{m}$. Now since $c^*c \equiv c^* \pmod{m}$ and $ca \equiv cb \pmod{m}$ by [Theorem 1.17.3](#),

$$c^*ca \equiv c^*cb \pmod{m}.$$

But $c^*c \equiv 1 \pmod{m}$ so

$$c^*ca \equiv a \pmod{m}$$

and

$$c^*cb \equiv b \pmod{m}.$$

By reflexivity and transitivity this yields

$$a \equiv b \pmod{m}.$$

Note that in the proof above we “canceled” the c in the equation by multiplying by the inverse. This is very similar to what happens in earlier math courses where you solve $2x = 6$ by multiplying both sides of the equation by $1/2$ (the “multiplicative inverse” of 2). The difference above is that we are dealing with congruences, not equations, and when an integer has an inverse modulo m , the

inverse is another *integer*, not a fraction. (Note also that we could solve $2x = 6$ by dividing by 2; we don't divide when dealing with congruences—instead, we multiply by inverses.)

✓ Example 1.20.2: Solve Congruences

Sometimes we may solve for an unknown integer x in a congruence. For example, consider the congruence

$$7x \equiv 4 \pmod{25}.$$

Can we determine which integer(s) x can be?

Since $\gcd(7, 25) = 1$, [Corollary 1.20.1](#) implies that 7 has an inverse modulo 25. Indeed, a little trial and error shows that $7^* = 18$, since $7 \cdot 18 = 126 \equiv 1 \pmod{25}$.

Now, as in the proof of [Theorem 1.20.3](#), we use 7^* to “cancel” the 7 in the congruence. We multiply both sides by 7^* , i.e., by 18, and obtain the following:

$$\begin{aligned} 18(7x) &\equiv 18(4) \pmod{25}; \\ (18 \cdot 7)x &\equiv 72 \pmod{25}; \\ 1x &\equiv 22 \pmod{25}. \end{aligned}$$

We see that x must be congruent to 22 modulo 25—so x must belong to the set

$$\{25q + 22 : q \in \mathbb{Z}\} = \{\dots, -28, -3, 22, 47, \dots\}.$$

Indeed, substituting $25q + 22$ in for x in the left-hand side of the congruence produces

$$7(25q + 22) = 175q + 154 = 25(7q + 6) + 4,$$

and $25(7q + 6) + 4$ is congruent to 4 modulo 25 no matter what integer value q takes on. Hence *every* number in $\{25q + 22 : q \in \mathbb{Z}\}$ is a solution to the congruence $7x \equiv 4 \pmod{25}$.

Although equation [\(1.20.1\)](#) above is not generally true when $\gcd(c, m) > 1$, we do have the following other kinds of “cancellation:”

Theorem 1.20.4

If $c > 0$, $m > 0$ then

$$a \equiv b \pmod{m} \Leftrightarrow ca \equiv cb \pmod{cm}.$$

See [Exercise 1.20.3](#).

Theorem 1.20.5

Let $m > 0$ and let $d = \gcd(c, m)$. Then

$$ca \equiv cb \pmod{m} \Rightarrow a \equiv b \pmod{(m/d)}.$$

Proof

Since $d = \gcd(c, m)$ we can write $c = d(\frac{c}{d})$ and $m = d(\frac{m}{d})$. Then $\gcd(\frac{c}{d}, \frac{m}{d}) = 1$. Now rewriting $ca \equiv cb \pmod{m}$ we have

$$d \frac{c}{d} a \equiv d \frac{c}{d} b \pmod{d(m/d)}.$$

Since $m > 0$, $d > 0$, so by [Theorem 1.20.4](#) we have

$$\frac{c}{d} a \equiv \frac{c}{d} b \pmod{(m/d)}.$$

Now since $\gcd(\frac{c}{d}, \frac{m}{d}) = 1$, by [Theorem 1.20.3](#)

$$a \equiv b \pmod{(m/d)}.$$

We end this chapter by mentioning how the concept of inverses modulo m interacts with congruence modulo m . To do so, we first discuss the greatest common divisors of m and numbers that are congruent modulo m .

Theorem 1.20.6

If $m > 0$ and $a \equiv b \pmod{m}$ we have

$$\gcd(a, m) = \gcd(b, m).$$

Proof

Since $a \equiv b \pmod{m}$ we have $a - b = mt$ for some t . So we can write

$$a = mt + b \quad \text{and} \quad b = m(-t) + a.$$

Let $d = \gcd(m, a)$ and $e = \gcd(m, b)$. Since $e \mid m$ and $e \mid b$, the first equation above implies that $e \mid a$, so e is a common divisor of m and a . Hence $e \leq d$. From the second equation we see similarly that $d \leq e$. So $d = e$.

Corollary 1.20.2

Let $m > 0$. Let $a \equiv b \pmod{m}$. Then a has an inverse modulo m if and only if b does.

Proof

Immediate from [Theorems 1.20.1](#), [1.20.2](#) and [1.20.6](#)

Exercises

Exercise 1.20.1

Show that the inverse of 2 modulo 7 is not the inverse of 2 modulo 15.

Exercise 1.20.2

Find specific positive integers a, b, c and m such that $c \not\equiv 0 \pmod{m}$, $\gcd(c, m) > 0$, and $ca \equiv cb \pmod{m}$, but $a \not\equiv b \pmod{m}$. What does this show about the theorems from this chapter?

Exercise 1.20.3

Prove [Theorem 1.20.4](#)

Exercise 1.20.4

Determine whether or not each of the following is true. **Give reasons in each case.**

1. $x \equiv 3 \pmod{7} \Rightarrow \gcd(x, 7) = 1$
2. $\gcd(68019, 3) = 3$
3. $12x \equiv 15 \pmod{35} \Rightarrow 4x \equiv 5 \pmod{7}$
4. $x \equiv 6 \pmod{12} \Rightarrow \gcd(x, 12) = 6$
5. $3x \equiv 3y \pmod{17} \Rightarrow x \equiv y \pmod{17}$
6. $5x \equiv y \pmod{6} \Rightarrow 15x \equiv 3y \pmod{18}$
7. $12x \equiv 12y \pmod{15} \Rightarrow x \equiv y \pmod{5}$
8. $x \equiv 73 \pmod{75} \Rightarrow x \bmod 75 = 73$
9. $x \equiv 73 \pmod{75}$ and $0 \leq x < 75 \Rightarrow x = 73$
10. There is no integer x such that

$$12x \equiv 7 \pmod{33}.$$

? Exercise 1.20.5

- Determine the inverse of 8 modulo 13.
- Using your answer to (a), and perhaps imitating [Example 1.20.2](#) determine all integers x such that $8x \equiv 4 \pmod{13}$. (Note: there will be infinitely many such integers; list enough of them to make the pattern clear, and justify your answer.)

? Exercise 1.20.6

Try to use the results from this chapter to completely describe the integer solutions to the following congruences. (Here the coefficients on x are *not* relatively prime to the modulus.) If a congruence has no solution in the integers, say so and explain why.

- $4x \equiv 6 \pmod{10}$
- $30x \equiv 90 \pmod{200}$
- $6x \equiv 4 \pmod{12}$.

? Exercise 1.20.7

Is it possible to find integers a, b, m where $m \geq 2$ and $0 < a, b < m$ and $a \neq b$, and there are at least *three* solutions to the congruence

$$ax \equiv b \pmod{m}$$

in the set $\{0, \dots, m-1\}$? If so, demonstrate this for a specific choice of a, b , and m . If it is not possible, explain why not.

This page titled [1.20: More Properties of Congruences](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Mike Barrus & W. Edwin Clark](#).

1.21: Residue Classes and the Integers Modulo m

Definition 1.21.1

Let $m > 0$ be given. For each integer a we define

$$[a] = \{x : x \equiv a \pmod{m}\}. \quad (1.21.1)$$

In other words, $[a]$ is the set of all integers that are congruent to a modulo m . We call $[a]$ the *residue class of a modulo m* . Some people call $[a]$ the *congruence class* or *equivalence class of a modulo m* .

✓ Example 1.21.1

If $m = 3$ and $a = 7$, we see

$$\begin{aligned} [7] &= \{x : x \equiv 7 \pmod{3}\} \\ &= \{\dots, -5, -2, 1, 4, 7, 10, 13, 16, 19, \dots\}. \end{aligned}$$

Keeping $m = 3$, we also see that

$$\begin{aligned} [0] &= \{x : x \equiv 0 \pmod{3}\} \\ &= \{\dots, -9, -6, -3, 0, 3, 6, 9, \dots\}; \\ [1] &= \{x : x \equiv 1 \pmod{3}\} \\ &= \{\dots, -8, -5, -2, 1, 4, 7, 10, \dots\}; \\ [2] &= \{x : x \equiv 2 \pmod{3}\} \\ &= \{\dots, -7, -4, -1, 2, 5, 8, 11, \dots\}. \end{aligned}$$

So every integer belongs to one of the residue classes modulo 3. Note also that $[7] = [1]$, since $[7]$ and $[1]$ are both sets containing exactly the same elements; think of $[7]$ and $[1]$ as different *names* for the same set. (With this perspective, each residue class has infinitely many names.)

Theorem 1.21.1

For $m > 0$ and any integer a we have

$$[a] = \{mq + a \mid q \in \mathbb{Z}\}. \quad (1.21.2)$$

Proof

$x \in [a] \Leftrightarrow x \equiv a \pmod{m} \Leftrightarrow m \mid (x - a) \Leftrightarrow x - a = mq$ for some $q \in \mathbb{Z} \Leftrightarrow x = mq + a$ for some $q \in \mathbb{Z}$. So equation (1.21.2) follows from the definition in equation (1.21.1).

Note that $[a]$ really depends on m and it would be more accurate to write $[a]_m$ instead of $[a]$, but this would be too cumbersome. Nevertheless it should be kept clearly in mind that $[a]$ depends on some understood value of m .

Remark 1.21.1

Two alternative ways to write equation (1.21.2) are

$$[a] = \{mq + a \mid q = 0, \pm 1, \pm 2, \dots\}$$

and

$$[a] = \{\dots, -2m + a, -m + a, a, m + a, 2m + a, \dots\}.$$

Theorem 1.21.2

For a given modulus $m > 0$ we have:

$$[a] = [b] \Leftrightarrow a \equiv b \pmod{m}.$$

Proof

“ $\Rightarrow$ ” Assume $[a] = [b]$. Note that since $a \equiv a \pmod{m}$ we have $a \in [a]$. Since $[a] = [b]$ we have $a \in [b]$. By definition of $[b]$ this gives $a \equiv b \pmod{m}$, as desired.

“ $\Leftarrow$ ” Assume $a \equiv b \pmod{m}$. We must prove that the sets $[a]$ and $[b]$ are equal. To do this we prove that every element of $[a]$ is in $[b]$ and vice-versa. Let $x \in [a]$. Then $x \equiv a \pmod{m}$. Since $a \equiv b \pmod{m}$, by transitivity $x \equiv b \pmod{m}$ so $x \in [b]$. Conversely, if $x \in [b]$, then $x \equiv b \pmod{m}$. By symmetry since $a \equiv b \pmod{m}$, $b \equiv a \pmod{m}$, so again by transitivity $x \equiv a \pmod{m}$ and $x \in [a]$. This proves that $[a] = [b]$.

Theorem 1.21.3

For every a there is a unique r such that

$$[a] = [r] \quad \text{and} \quad 0 \leq r < m.$$

Proof

Let $r = a \bmod m$. Then by [Theorem 1.17.2](#) (1) we have $a \equiv r \pmod{m}$. By definition of $a \bmod m$ we have $0 \leq r < m$. Since $a \equiv r \pmod{m}$ by [Theorem 1.21.2](#) $[a] = [r]$. To prove that r is unique, suppose also $[a] = [r']$ where $0 \leq r' < m$. By [Theorem 1.21.2](#) this implies that $a \equiv r' \pmod{m}$. This, together with $0 \leq r' < m$, implies by [Theorem 1.17.4](#) that $r' = a \bmod m = r$.

Theorem 1.21.4

Given $m > 0$, there are exactly m distinct residue classes modulo m , namely,

$$[0], [1], [2], \dots, [m-1].$$

Proof

By [Theorem 1.21.3](#) we know that every residue class $[a]$ is equal to one of the residue classes: $[0], [1], \dots, [m-1]$. So there are no residue classes not in this list. These residue classes are distinct by the uniqueness part of [Theorem 1.21.3](#), namely if $0 \leq r_1 < m$ and $0 \leq r_2 < m$ and $[r_1] = [r_2]$, then by the uniqueness part of [Theorem 1.21.3](#) we must have $r_1 = r_2$.

Definition 1.21.2: Representative

Any element $x \in [a]$ is said to be a *representative* of the residue class $[a]$.

As you are asked to show in [Exercise 1.21.4](#) below, if x is a representative of $[a]$ then $[x] = [a]$, that is, *any element of a residue class may be used to represent it*.

The integers modulo m

Henceforth in this chapter let m be a fixed integer that is greater than 1.

Definition 1.21.3: Ring of Integers Modulo m

We define

$$\mathbb{Z}_m = \{[a] \mid a \in \mathbb{Z}\},$$

that is, $\mathbb{Z}_m$ is the set of all residue classes modulo m . We call $\mathbb{Z}_m$ *the ring of integers modulo m* . In the next chapter we shall show how to add and multiply residue classes. This makes $\mathbb{Z}_m$ into a ring. (See [Appendix C](#) for the definition of *ring*.) Often we drop “the ring of” and just call $\mathbb{Z}_m$ *the integers modulo m* . From [Theorem 1.21.4](#)

$$\mathbb{Z}_m = \{[0], [1], \dots, [m-1]\},$$

and since no two of the residue classes $[0], [1], \dots, [m-1]$ are equal, we see that $\mathbb{Z}_m$ has exactly m elements.¹ By [Exercise 1.21.4](#) if we choose

$$a_0 \in [0], a_1 \in [1], \dots, a_{m-1} \in [m-1]$$

then

$$[a_0] = [0], [a_1] = [1], \dots, [a_{m-1}] = [m-1].$$

So we also have

$$\mathbb{Z}_m = \{[a_0], [a_1], \dots, [a_{m-1}]\}.$$

✓ Example 1.21.2

If $m = 4$ we have, for example,

$$8 \in [0], 5 \in [1], -6 \in [2], 11 \in [3].$$

And hence:

$$\mathbb{Z}_4 = \{[8], [5], [-6], [11]\}.$$

We now show how to define addition and multiplication of residue classes modulo m . It is with respect to these binary operations that $\mathbb{Z}_m$ is a ring (again, see [Appendix C](#)).

Definition 1.21.4

For $[a], [b] \in \mathbb{Z}_m$ we define

$$[a] + [b] = [a + b]$$

and

$$[a][b] = [ab].$$

✓ Example 1.21.3

For $m = 5$ we have

$$[2] + [3] = [5],$$

and

$$[2][3] = [6].$$

Note that since $5 \equiv 0 \pmod{5}$ and $6 \equiv 1 \pmod{5}$ we have $[5] = [0]$ and $[6] = [1]$ so we can also write

$$[2] + [3] = [0]$$

$$[2][3] = [1].$$

Since a residue class can have many representatives, it is important to check that the rules given in [Definition 1.21.4](#) do not depend on the representatives chosen. For example, when $m = 5$ we know that

$$[7] = [2] \text{ and } [11] = [1]$$

so we *should* have

$$[7] + [11] = [2] + [1]$$

and

$$[7][11] = [2][21].$$

In this case we can check that

$$[7] + [11] = [18] \text{ and } [2] + [21] = [23].$$

Now $23 \equiv 18 \pmod{5}$ since $5 \mid (23 - 18)$. Hence $[18] = [23]$, as desired. Also $[7][11] = [77]$ and $[2][21] = [42]$. Then $77 - 42 = 35$ and $5 \mid 35$ so $77 \equiv 42 \pmod{5}$ and hence $[77] = [42]$, as desired.

Theorem 1.21.5

For any modulus $m > 0$ if $[a] = [b]$ and $[c] = [d]$ then

$$[a] + [c] = [b] + [d]$$

and

$$[a][c] = [b][d].$$

Proof

See [Exercise 1.21.8](#) below.

When performing addition and multiplication in $\mathbb{Z}_m$ using the rules in [Definition 1.21.4](#) due to [Theorem 1.21.5](#), we may at any time replace $[a]$ by $[a']$ if $a \equiv a' \pmod{m}$. This will sometimes make calculations easier.

Example 1.21.4

Take $m = 151$. Then $150 \equiv -1 \pmod{151}$ and $149 \equiv -2 \pmod{151}$, so

$$[150][149] = [-1][-2] = [2]$$

and

$$[150] + [149] = [-1] + [-2] = [-3] = [148]$$

since $148 \equiv -3 \pmod{151}$.

When working with $\mathbb{Z}_m$ it is often useful to write each residue class using its name $[a]$, where a is the least nonnegative number in the set. We do this in constructing the following addition and multiplication tables for $\mathbb{Z}_4$. For example, $[2] \cdot [3] = [6]$, but since $[6] = [2]$ and 2 is the smallest nonnegative member of this set, the table records $[2] \cdot [3] = [2]$.

+	[0]	[1]	[2]	[3]
[0]	[0]	[1]	[2]	[3]
[1]	[1]	[2]	[3]	[0]
[2]	[2]	[3]	[0]	[1]
[3]	[3]	[0]	[1]	[2]

·	[0]	[1]	[2]	[3]
[0]	[0]	[0]	[0]	[0]
[1]	[0]	[1]	[2]	[3]
[2]	[0]	[2]	[0]	[2]
[3]	[0]	[3]	[2]	[1]

Recall that by [Theorem 1.17.2](#) (1) we have for all a and $m > 0$

$$a \equiv a \bmod m \pmod{m}.$$

So using residue classes modulo m this gives

$$[a] = [a \bmod m].$$

Hence,

$$\begin{aligned} [a] + [b] &= [(a+b) \bmod m] \\ [a][b] &= [(ab) \bmod m] \end{aligned}$$

So if a and b are in the set $\{0, 1, \dots, m-1\}$, these equations give us a way to obtain representations of the sum and product of $[a]$ and $[b]$ using “names” (representatives) also in $\{0, 1, \dots, m-1\}$. This leads to an alternative way to define $\mathbb{Z}_m$ and addition and multiplication in $\mathbb{Z}_m$. We will use slightly different notation.

Definition 1.21.5

For $m > 0$ define

$$\mathbb{Z}_m = \{0, 1, 2, \dots, m-1\}$$

and for $a, b \in \mathbb{Z}_m$

$$\begin{aligned} a + b &= (a + b) \bmod m \\ ab &= (ab) \bmod m. \end{aligned}$$

where the addition and multiplication *inside the parentheses* are the usual addition and multiplication of integers; what is new in our redefinition is the practice of always reducing the result modulo m .

Remark 1.21.2

The set $\mathbb{Z}_m$ with addition and multiplication as defined is *isomorphic* to $\mathbb{Z}_m$ with addition and multiplication given by [Definition 1.21.4](#) (Students taking a course in elementary abstract algebra will learn a rigorous definition of the term isomorphic. For now, we take “isomorphic” to mean “has the same form.”) The addition and multiplication tables for $\mathbb{Z}_4$ are here:

+	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

✓ Example 1.21.5

Let's solve the congruence

$$272x \equiv 901 \pmod{9}.$$

Using residue classes modulo 9, this congruence is equivalent to

$$[272x] = [901],$$

which is equivalent to

$$[272][x] = [901],$$

which is equivalent to

$$[2][x] = [1].$$

Now we know $[x] \in \{[0], [1], \dots, [8]\}$ so by trial and error we see that $x = 5$ is a solution; actually, every integer in the residue class $[5]$ is a solution.

Exercises

? Exercise 1.21.1

Show that if $m = 2$ then $[1]$ is the set of all odd integers and $[0]$ is the set of all even integers. Show also that $\mathbb{Z} = [0] \cup [1]$ and $[0] \cap [1] = \emptyset$.

? Exercise 1.21.2

Show that if $m = 3$, then $[0]$ is the set of integers divisible by 3, $[1]$ is the set of integers whose remainder when divided by 3 is 1, and $[2]$ is the set of integers whose remainder when divided by 3 is 2. Show also that $\mathbb{Z} = [0] \cup [1] \cup [2]$ and $[0] \cap [1] = [0] \cap [2] = [1] \cap [2] = \emptyset$.

? Exercise 1.21.3

Given the modulus $m > 0$ show that $[a] = [a + m]$ and $[a] = [a - m]$ for all a .

? Exercise 1.21.4

For any $m > 0$, show that if $x \in [a]$ then $[a] = [x]$.

? Exercise 1.21.5

For any $m > 0$, show that if $[a] \cap [b] \neq \emptyset$ then $[a] = [b]$.

? Exercise 1.21.6

For any $m > 0$, show that if $[a] \neq [b]$ then $[a] \cap [b] = \emptyset$.

? Exercise 1.21.7

Let $m = 2$. Show that

$$[0] = [2] = [4] = [32] = [-2] = [-32]$$

and

$$[1] = [3] = [-3] = [31] = [-31].$$

? Exercise 1.21.8

Prove [Theorem 1.21.5](#).

(Hint: Use [Theorems 1.17.3](#) and [1.21.2](#))

? Exercise 1.21.9

Construct addition and multiplication tables for $\mathbb{Z}_5$.

? Exercise 1.21.10

Without doing it, tell how to obtain addition and multiplication tables for $\mathbb{Z}_5$ from the work in [Exercise 1.21.9](#).

? Exercise 1.21.11

- If p is a prime, show that $x^2 \equiv 1 \pmod{p}$ if and only if $x \equiv -1 \pmod{p}$ or $x \equiv 1 \pmod{p}$. (Hint: as part of your answer, explain why $x^2 \equiv 1 \pmod{p}$ implies that $p \mid (x+1)(x-1)$ and apply Euclid's Lemma ([Lemma 1.11.2](#)).
- Conclude that, modulo a prime p , the congruence $[x]^2 = [1]$ has solutions $[x] = [1]$ and $[x] = [-1]$.
- Find an example of a modulus m and a residue class $[a]$ such that $[a]^2 = [1]$ but $[a] \neq [1]$ and $[a] \neq [-1]$ in $\mathbb{Z}_m$.

? Exercise 1.21.12

Solve the congruence $544x \equiv 863 \pmod{7}$.

Footnotes

[1] Each of those elements, like $[0]$ or $[1]$, is itself a set with infinitely many elements, but we will often ignore this. The m sets $[0], [1], \dots, [m-1]$ in $\mathbb{Z}_m$ are its m elements.

This page titled [1.21: Residue Classes and the Integers](#) Modelo m is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Mike Barrus & W. Edwin Clark](#).

1.22: The Groups U_m

Definition 1.22.1

Let $m > 0$. A residue class $[a] \in \mathbb{Z}_m$ is called a *unit* if there is another residue class $[b] \in \mathbb{Z}_m$ such that $[a][b] = [1]$. In this case $[a]$ and $[b]$ are said to be *inverses of each other* in $\mathbb{Z}_m$.

Theorem 1.22.1

Let $m > 0$. A residue class $[a] \in \mathbb{Z}_m$ is a unit if and only if $\gcd(a, m) = 1$.

Proof

Let $[a]$ be a unit. Then there is some $[b]$ such that $[a][b] = [1]$. Hence $[ab] = [1]$ so $ab \equiv 1 \pmod{m}$. So by [Theorem 1.20.2](#), $\gcd(a, m) = 1$.

To prove the converse, let $\gcd(a, m) = 1$. Then by [Theorem 1.20.1](#), there is an integer a^* such that $aa^* \equiv 1 \pmod{m}$. Hence, $[aa^*] = [1]$. So $[a][a^*] = [aa^*] = [1]$, and we can take $b = a^*$.

We see from [Theorem 1.20.6](#) that if $[a] = [b]$ (i.e., $a \equiv b \pmod{m}$) then $\gcd(a, m) = 1 \Leftrightarrow \gcd(b, m) = 1$. So in checking whether or not a residue class is a unit, we can use any representative of the class.

The elements $[1]$ and $[m-1]$ are always units in $\mathbb{Z}_m$ (see [Exercise 1.22.1](#)). The collection all units in $\mathbb{Z}_m$ will be our next focus.

Definition 1.22.2

The set of all units in $\mathbb{Z}_m$ is denoted by U_m and is called the *group of units* of $\mathbb{Z}_m$. See [Appendix C](#) for the definition of a *group*.

Theorem 1.22.2

Let $m > 0$, then

$$U_m = \{[i] \mid 1 \leq i \leq m \text{ and } \gcd(i, m) = 1\}.$$

Proof

We know that if $[a] \in \mathbb{Z}_m$ then $[a] = [i]$ where $0 \leq i \leq m-1$. If $m = 1$ then $\mathbb{Z}_m = \mathbb{Z}_1 = \{[0]\} = \{[1]\}$ and since $[1][1] = [1]$, $[1]$ is a unit, $U_1 = \{[1]\}$ and the theorem holds. If $m \geq 2$, then $\gcd(i, m) = 1$ can only happen if $1 \leq i \leq m-1$, since $\gcd(0, m) = \gcd(m, m) = m \neq 1$. So the theorem follows from [Theorem 1.22.1](#) and the above remarks.

Theorem 1.22.3

(U_m is a group¹ under multiplication.)

1. If $[a], [b] \in U_m$ then $[a][b] \in U_m$.
2. For all $[a], [b], [c]$ in U_m we have $([a][b])[c] = [a]([b][c])$.
3. $[1][a] = [a][1] = [a]$ for all $[a] \in U_m$.
4. For each $[a] \in U_m$ there is a $[b] \in U_m$ such that $[a][b] = [1]$.
5. For all $[a], [b] \in U_m$ we have $[a][b] = [b][a]$.

Proof

See [Exercise 1.22.2](#)

✓ Example 1.22.1

Using [Theorem 1.22.2](#) we see that

$$\begin{aligned} U_{15} &= \{[1], [2], [4], [7], [8], [11], [13], [14]\} \\ &= \{[1], [2], [4], [7], [-7], [-4], [-2], [-1]\}. \end{aligned}$$

Note that using representatives for residue classes modulo 15 with the smallest possible absolute value simplifies multiplication somewhat. (It is easier to multiply by one of $-1, -2, -4$ or -7 , usually, than to multiply by one of 14, 13, 11, or 8.) Rather than write out the entire multiplication table, we just find the inverse of each element of U_{15} :

$$\begin{aligned} [1] &= [1] \\ [2] [-7] &= [2] [8] = [1] \\ [4] [4] &= [1] \\ [7] [-2] &= [7] [13] = [1] \\ [-4] [-4] &= [11] [11] = [1] \\ [-1] [-1] &= [14] [14] = [1]. \end{aligned}$$

Corollary 1.22.1

If $m > 0$,

$$|U_m| = \phi(m),$$

where ϕ denotes Euler's totient function.

Recall that ϕ was introduced in [Section 1.15](#). Observe that

U_1	=	$\{[1]\}$	and	$\phi(1)$	=	1
U_2	=	$\{[1]\}$	and	$\phi(2)$	=	$2 - 1 = 1$
U_3	=	$\{[1], [2]\}$	and	$\phi(3)$	=	$3 - 1 = 2$
U_4	=	$\{[1], [3]\}$	and	$\phi(4)$	=	$2^2 - 2^1 = 2$
U_5	=	$\{[1], [2], [3], [4]\}$	and	$\phi(5)$	=	$5 - 1 = 4$
U_6	=	$\{[1], [5]\}$	and	$\phi(6)$	=	$(2 - 1)(3 - 1) = 2$
U_7	=	$\{[1], [2], [3], [4], [5], [6]\}$	and	$\phi(7)$	=	$7 - 1 = 6$.

Exercises

? Exercise 1.22.1

Given $m \geq 2$, show that $[1]$ and $[m - 1]$ are always units in $\mathbb{Z}_m$.

(Hint: Use the fact that $[m - 1] = [-1]$.)

? Exercise 1.22.2

Prove [Theorem 1.22.3](#).

? Exercise 1.22.3

List the elements of U_7 in at least two different ways (i.e., using two different sets of representatives for the names) and find the inverse of each element, as in [Example 1.22.1](#).

? Exercise 1.22.4

Find the sets U_m , for $8 \leq m \leq 20$. Note that $|U_m| = \phi(m)$. Use [Theorem 1.15.6](#) to calculate $\phi(m)$ and check that you have the right number of elements for each set U_m , $8 \leq m \leq 20$.

? Exercise 1.22.5

Using the fact that $[3]$ and $[19]$ are elements of U_{20} , use addition and multiplication of residue classes (NOT subtraction or division, which we have not defined) to solve the congruences for $[x]$ below. Assume that the modulus is $m = 20$.

- a. $[3][x] + [11] = [4]$
- b. $[19][x] + [2] = [7]$

Footnotes

[1] Actually (1){(4) are all that is required for U_n to be a *group*. Property (5) says that U_n is an *Abelian* group. See [Appendix C](#).

This page titled [1.22: The Groups \$U_m\$](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Mike Barrus & W. Edwin Clark](#).

1.23: Chinese Remainder Theorem

The Chinese Remainder Theorem is an important theorem appearing for perhaps the first time in *Sunzi Suanjing*, a Chinese mathematical text written sometime during the 3rd to 5th centuries AD. We will illustrate its usefulness with an anecdote.

The child of a number theorist is sorting a large pile of pennies (worth less than a dollar) into groups of 3 pennies each. At the end, the child reports that 2 pennies are left over. The child starts over, instead sorting the pennies into groups of 4 and reports that 1 penny is left over. The child starts over again, sorting the pennies into groups of 11 and reports that 7 pennies are left over. The number theorist didn't originally know how many pennies were in the pile, but at this point she speaks up.

What does she say? Did the child make a mistake in sorting the pennies? Or does the number theorist have enough information to tell how many pennies are in the pile?

We will answer this question with the Chinese Remainder Theorem. Here it is:

Theorem 1.23.1: Chinese Remainder Theorem

Let $m_1, m_2, \dots, m_k$ be natural numbers such that each is greater than 1, and every pair of them is relatively prime. Let $M = m_1 m_2 \cdots m_k$, and let $b_1, b_2, \dots, b_k$ be integers. The system of congruences

$$\begin{aligned} x &\equiv b_1 \pmod{m_1}; \\ x &\equiv b_2 \pmod{m_2}; \\ &\vdots \\ x &\equiv b_k \pmod{m_k}; \end{aligned}$$

has a unique solution in the set $\{0, 1, 2, \dots, M-1\}$.

Before we prove the theorem, let's see what the number theorist from our story may have said.

Example 1.23.1: Chinese Remainder Theorem Pennies

Suppose that x is the number of pennies in the child's pile. If we assume for a moment that the child didn't make any mistakes in sorting the pennies into piles, then x satisfies the three congruences

$$x \equiv 2 \pmod{3}; \quad x \equiv 1 \pmod{4}; \quad x \equiv 7 \pmod{11}.$$

At this point, since the moduli 3, 4, and 11 have the property that every two are relatively prime, the Chinese Remainder Theorem states that there is a unique solution to these congruences among the integers between 0 and 131 (here $3 \cdot 4 \cdot 11 = 132$). In fact, we are told that there is a positive number of pennies, but there are fewer than one hundred, so exactly one number in $\{1, 2, \dots, 99\}$ can be our solution to the number of pennies.

This is assuming that the child didn't make any mistakes in counting, but the number theorist trusts her child (the child has been trained from the time he was in diapers to check his work), and the Chinese Remainder Theorem *can't* point out any mistakes in the numbers on the right-hand sides of the congruences (the residues, not the moduli), since every choice of these numbers leads to a solution.

At this point the number theorist knows, at a minimum, that there is a unique solution to the system of congruences. However, if she was at this point able to tell her child how many pennies he has, we still don't know what she said, because the Chinese Remainder Theorem alone does not tell us what the solution *is*—just that it exists. After we prove the Chinese Remainder Theorem, we will discuss a practical way of solving these systems of congruences, of *finding* the solution whose existence has been guaranteed.

Proof of Theorem 1.23.1

Proof

We first prove that there exists a solution in $\{0, \dots, M-1\}$; we will do this algorithmically.

Consider the first two congruences in the system:

$$x \equiv b_1 \pmod{m_1}; \quad x \equiv b_2 \pmod{m_2}.$$

Since m_1 and m_2 are relatively prime, Bezout's Lemma ([Lemma 1.9.1](#)) implies that there exist integers s_1, t_1 such that

$$s_1 m_1 + t_1 m_2 = 1. \tag{1.23.1}$$

We find these integers, perhaps using a technique from [Section 1.10](#). Once we have, we construct a *partial solution* x_2 by

$$x_2 = b_1 t_1 m_2 + b_2 s_1 m_1.$$

Note here that we have attached the summands in equation (1.23.1) to the right-hand sides of the congruences, but in the “opposite” order. We use the subscript 2 and call x_2 a partial solution because it satisfies the first two congruences we began with;

$$\begin{aligned} x_2 &= b_1 t_1 m_2 + b_2 s_1 m_1 \\ &\equiv b_1 t_1 m_2 \pmod{m_1} \\ &\equiv b_1 (1 - s_1 m_1) \pmod{m_1} \\ &\equiv b_1 \pmod{m_1}, \end{aligned}$$

and a similar argument shows that $x_2 \equiv b_2 \pmod{m_2}$. Our next step is to record the information about our partial solution in its own congruence and pair it with the next unused congruence from our system:

$$\begin{aligned} x &\equiv x_2 \pmod{m_1 m_2}; \\ x &\equiv b_3 \pmod{m_3}. \end{aligned}$$

We solve this pair of congruences the same way we did the first pair; since $m_1 m_2$ and m_3 can have no common prime factors, Bezout's Lemma implies the existence of integers s_2, t_2 such that

$$s_2 (m_1 m_2) + t_2 m_3 = 1,$$

and our next partial solution is

$$x_3 = x_2 t_2 m_3 + b_3 s_2 (m_1 m_2).$$

It's possible to check that x_3 satisfies the congruences modulo m_1, m_2 , and m_3 . We once again move on to the next pair of congruences, using our partial solution and the next unused congruence:

$$\begin{aligned} x &\equiv x_3 \pmod{m_1 m_2 m_3}; \\ x &\equiv b_4 \pmod{m_4}. \end{aligned}$$

We continue in this way until we reach the partial solution x_k , which will be a solution to all k of the original k congruences. Thus the system has at least one solution.

We now must prove the uniqueness of the solution. Suppose that x and y both were congruent to b_1 modulo m_1 , congruent to b_2 modulo m_2 , and so on. Then $x \equiv y \pmod{m_i}$ for each i , meaning that each modulus m_i divides $x - y$. Since every two of the moduli are relatively prime, this implies (see [Exercise 1.23.1](#)) that $m_1 m_2 \cdots m_k$ divides $x - y$, which implies that $x \equiv y \pmod{M}$. Thus there can only be one element of $\{0, 1, 2, \dots, M-1\}$ that is a solution to the system of congruences.

The existence part of the proof of the Chinese Remainder Theorem gives us a way to construct solutions. Please carefully read the sections there. We illustrate the technique with a few examples.

✓ Example 1.23.1: Continued

Resuming our earlier example, we solve the system of congruences

$$x \equiv 2 \pmod{3}; \quad x \equiv 1 \pmod{4}; \quad x \equiv 7 \pmod{11}.$$

Beginning with the first two, we find coefficients for Bezout's Lemma for the moduli 3 and 4, obtaining $-1 \cdot 3 + 1 \cdot 4 = 1$. We build our partial solution

$$x_2 = 2 \cdot 4 + 1 \cdot -3 = 5.$$

Thus

$$x \equiv 5 \pmod{12} \quad \text{and} \quad x \equiv 7 \pmod{11}.$$

Now $1 \cdot 12 - 1 \cdot 11 = 1$, so our next partial solution is

$$x_3 = 5 \cdot -11 + 7 \cdot 12 = 29.$$

Since we have used all of the original congruences, we conclude (and can check) that $x = 29$ is a solution to the system of congruences. The number theorist replies to her child that his pile has exactly 29 pennies.

Note that the uniqueness part of the Chinese Remainder Theorem only guarantees uniqueness of a solution among the residue classes modulo M . For example, in our example it was necessary to know that there were fewer than 100 pennies, since $x = 161$ ($= 29 + 132$), $x = 293$ ($= 29 + 2 \cdot 132$), and so on also satisfy all the congruences reported by the child. In general, a system of congruences satisfying the requirements of the Chinese Remainder Theorem is satisfied by an infinite number of integers x , though only one of these belongs to the set $\{0, 1, \dots, M-1\}$.

✓ Example 1.23.2

Let's now describe *all* solutions to the system of congruences

$$\begin{aligned} 2x &\equiv 5 \pmod{7}; \\ 4x &\equiv 5 \pmod{9}; \\ x &\equiv 4 \pmod{11}; \\ x &\equiv 8 \pmod{13}. \end{aligned}$$

This system is not yet ready for our technique from the previous example, since the first two congruences have extra coefficients on the left. We can address this problem by multiplying the first congruence by the inverse of 2 modulo 7 (which is 4) and multiplying the second congruence by the inverse of 4 modulo 9 (which is 7), obtaining

$$\begin{aligned} x &\equiv 6 \pmod{7}; \\ x &\equiv 8 \pmod{9}; \\ x &\equiv 4 \pmod{11}; \\ x &\equiv 8 \pmod{13}. \end{aligned}$$

We begin now with the first two congruences. Since $4 \cdot 7 - 3 \cdot 9 = 1$, our first partial solution is $x_2 = 6 \cdot -27 + 8 \cdot 28 = 62$.

Working now with $x \equiv 62 \pmod{63}$ and $x \equiv 4 \pmod{11}$, using Blankinship's Method or the Extended Euclidean Algorithm yields $-4 \cdot 63 + 23 \cdot 11 = 1$. Our next partial solution is $x_3 = 62 \cdot 253 + 4 \cdot -252 = 14678$. This should make each of the first three congruences true, but it's a rather large number; does it have to be this large?

Notice that $7 \cdot 9 \cdot 11 = 693$, and division yields $14678 = 21 \cdot 693 + 125$. Let $x'_3 = 125$. Let's look at the congruences x_3 satisfies with x'_3 in place of x_3 :

$21 \cdot 693 + 125$	$\equiv$	6	$\pmod{7};$		125	$\equiv$	6	$\pmod{7};$
$21 \cdot 693 + 125$	$\equiv$	8	$\pmod{9};$		125	$\equiv$	8	$\pmod{9};$
$21 \cdot 693 + 125$	$\equiv$	4	$\pmod{11}$		125	$\equiv$	4	$\pmod{11}$
			;					.

Since 693 is divisible by each of 7, 9, 11, our partial solution $x_3 = 14678$ has the same residues modulo these numbers as does $x'_3 = 125$, so henceforth let's replace 14678 by 125.

Moving on, the last pair of congruences we consider are $x \equiv 125 \pmod{693}$ and $x \equiv 8 \pmod{13}$. A little computation yields $-3 \cdot 693 + 160 \cdot 13$, so $x_4 = 125 \cdot 2080 + 8 \cdot -2079 = 243368$. Since $7 \cdot 9 \cdot 11 \cdot 13 = 9009$, we can replace this number by $x'_4 = 243368 \bmod 9009 = 125$.

This last partial solution satisfies all four of our starting congruences, and by the Chinese Remainder Theorem, 125 is the only number in $\{0, 1, \dots, 9008\}$ to do so. However, the question here is to determine *all* integer solutions to the congruences. These are precisely the integers x for which $x \equiv 125 \pmod{9009}$, so the set of all solutions is precisely the set

$$\{125 + 9009n \mid n \in \mathbb{Z}\}.$$

We can now give a proof of the statement from the previous chapter that if a and b are relatively prime positive integers, then $\phi(ab) = \phi(a)\phi(b)$.

Proof of Theorem 1.15.4

Consider the map f from U_{ab} defined by

$$f([x]) = ([x]_a, [x]_b),$$

where $[x]_a$ and $[x]_b$ in the ordered pair indicate the residue classes containing x in U_a and in U_b , respectively. Observe that each image under f is an ordered pair where the first entry is a residue class from U_a (of which there are $\phi(a)$ possibilities) and the second entry is a residue class from U_b (of which there are exactly $\phi(b)$ possibilities). Thus there are $\phi(a)\phi(b)$ possible ordered pairs that *could* appear as the outcomes of the mapping f . Now we may apply the Chinese Remainder Theorem, since a and b are relatively prime. For every ordered pair $([c], [d])$ there exists a unique solution x modulo ab to the congruences $x \equiv c \pmod{a}$ and $x \equiv d \pmod{b}$. Then by our definition of f we conclude that $f([x]) = ([c], [d])$, and $[x]$ is the only element of U_{ab} that gets sent to $([c], [d])$. This is true no matter what $([c], [d])$ is, so there are exactly as many elements in U_{ab} as there are distinct ordered pairs $([c], [d])$; thus $\phi(ab) = \phi(a)\phi(b)$.

Before ending the chapter we mention that the Chinese Remainder Theorem is true in a more general context than that of the integers. For a certain broader class of rings¹ (see [Appendix C](#)), if we use analogous notions of congruence and relative primeness, the Chinese Remainder Theorem still holds true. In particular, the Theorem is true for the Gaussian integers (see [Exercise 1.23.6](#)).

Exercises

Exercise 1.23.1

Prove the following: if n and $m_1, m_2, \dots, m_n$ and N are all integers, and every two of $m_1, m_2, \dots, m_n$ are relatively prime, and $m_1, m_2, \dots, m_n$ each divide N , then $m_1 m_2 \cdots m_n$ divides N .

(Hint: there are multiple proofs that can be given, using the prime factorization or Bezout's Lemma and induction.)

Exercise 1.23.2

Describe all solutions to the following systems of congruences.

- $x \equiv 1 \pmod{8}; x \equiv 2 \pmod{9}$.
- $3x \equiv 8 \pmod{10}; x \equiv 13 \pmod{21}$.

Exercise 1.23.3

Solve this problem attributed to Euler: Find two numbers whose sum is 100 such that one is divisible by seven and the other is divisible by 11.

? Exercise 1.23.4

Give all solutions to the system of congruences:

$$x \equiv 2 \pmod{3}; \quad x \equiv 2 \pmod{5}; \quad x \equiv 5 \pmod{7}.$$

? Exercise 1.23.5

a. Explain why the system of congruences

$$x \equiv 3 \pmod{6}; \quad x \equiv 4 \pmod{8}$$

has no solution. Explain why this does not violate the Chinese Remainder Theorem.

b. Find all numbers in $\{0, 1, 2, \dots, 100\}$ that are solutions to the system of congruences

$$x \equiv 2 \pmod{6}; \quad x \equiv 4 \pmod{8}.$$

Can you describe (and justify) what the set of *all* integer solutions to this system of congruences is?

c. Can you describe a moral to this story? To you, what does this exercise illustrate about systems of congruences and the Chinese Remainder Theorem?

? Exercise 1.23.6

Use ideas from this chapter and from [Section 1.13](#) (including the exercises) to solve the following systems of congruences. Your answers should be Gaussian integers.

a. $x \equiv i \pmod{1+i}; x \equiv 2 \pmod{3i}.$

b. $x \equiv -3+i \pmod{5}; x \equiv -i \pmod{1-i}.$

Footnotes

[1] Namely, the *principal ideal domains*.

This page titled [1.23: Chinese Remainder Theorem](#) is shared under a [not declared](#) license and was authored, remixed, and/or curated by [Mike Barrus & W. Edwin Clark](#).

1.24: Theorems of Wilson, Euler, and Fermat

As the Chinese Remainder Theorem illustrated in the last chapter, some useful and interesting number theoretic results deal with congruences. In this chapter we present a few more well known theorems involving congruences.

Wilson's Theorem

For example, suppose that you took all the nonzero integers modulo m and multiplied them together (modulo m). What would you get? Let's examine a few cases:

In $\mathbb{Z}_2$,	$[1]$	$=$	$[1]$.
In $\mathbb{Z}_3$,	$[1][2]$	$=$	$[2]$.
In $\mathbb{Z}_4$,	$[1][2][3]$	$=$	$[2]$.
In $\mathbb{Z}_5$,	$[1][2][3][4]$	$=$	$[4]$.
In $\mathbb{Z}_6$,	$[1][2][3][4][5]$	$=$	$[0]$.
In $\mathbb{Z}_7$,	$[1][2][3][4][5][6]$	$=$	$[6]$.
In $\mathbb{Z}_8$,	$[1][2][3][4][5][6][7]$	$=$	$[0]$.
In $\mathbb{Z}_9$,	$[1][2][3][4][5][6][7][8]$	$=$	$[0]$.
In $\mathbb{Z}_{10}$,	$[1][2][3][4][5][6][7][8][9]$	$=$	$[0]$.
In $\mathbb{Z}_{11}$,	$[1][2][3][4][5][6][7][8][9][10]$	$=$	$[10]$.
In $\mathbb{Z}_{12}$,	$[1][2][3][4][5][6][7][8][9][10][11]$	$=$	$[0]$.
In $\mathbb{Z}_{13}$,	$[1][2][3][4][5][6][7][8][9][10][11][12]$	$=$	$[12]$.

A pattern seems to be emerging. When m is composite, it seems that

$$[1][2][3] \cdots [m-1] = [0], \quad (1.24.1)$$

except in the case $m = 4$. When m is prime, it looks like

$$[1][2][3] \cdots [m-1] = [m-1]. \quad (1.24.2)$$

Perhaps you can see why at least some of this is true; [Exercise 1.24.1](#) asks you to justify the pattern in equation (1.24.1). The equation (1.24.2) is known as *Wilson's Theorem*,¹ and we present it here in an equivalent form.

Theorem 1.24.1: Wilson's Theorem

If p is a prime, then

$$(p-1)! \equiv -1 \pmod{p}.$$

Proof

We present a sketch of the proof, which will rely on a few statements that are true but will not be proved here.

Starting out, recall that $(p-1)! = 1 \cdot 2 \cdot 3 \cdots (p-1)$, and the residue class modulo p to which this number belongs is the same class that $[1][2] \cdots [p-1]$ equals.

Suppose that p is prime; then the set $\{[1], [2], \dots, [p-1]\}$ of nonzero elements in $\mathbb{Z}_p$ is exactly the set of elements in U_p . By [Theorem 1.22.3](#) (4), each element in U_p has an inverse. A basic result in the theory of groups is that in every group, including U_p , no two distinct elements have the same inverse, and if one group element is the inverse of another, then the latter element is the inverse of the former as well. This means that some of the elements in the product $[1][2] \cdots [p-1]$ can be collected into pairs of elements that are inverses of each other. These pairs will each “cancel,” leaving $[1]$ in their place (which we may ignore in the product, since they don't affect that outcome, by [Theorem 1.22.3](#) (3)). What is left in the product? The only terms

of $[1][2] \cdots [p-1]$ that have not been canceled in this way are those that do not belong to an inverse pair; by [Theorem 1.22.3\(4\)](#) the only way this can happen is when an element is its own inverse. Which elements in U_p are their own inverses? Clearly $[1]$ and $[p-1]$ are their own inverses, so they remain when $[1][2] \cdots [p-1]$ is simplified. However, any element that is its own inverse satisfies $[x]^2 = [1]$, which is equivalent to the statement $p \mid (x^2 - 1) = (x+1)(x-1)$. By Euclid's Lemma ([Lemma 1.11.2](#)) this implies that $p \mid (x+1)$ or $p \mid (x-1)$, which means $x \equiv -1 \pmod{p}$ or $x \equiv 1 \pmod{p}$. Hence there are no other solutions to $[x]^2 = [1]$ besides $[1]$ and $[p-1]$. Hence only $[1]$ and $[p-1]$ are their own inverses, and

$$[1][2] \cdots [p-1] = [1][p-1] = [p-1].$$

Since $-1 \in [p-1]$, our proof is complete.

Let's use an example to illustrate the proof of Wilson's Theorem. Suppose that $p = 11$, and consider the product

$$10! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdot 8 \cdot 9 \cdot 10$$

modulo 11. In U_{11} , we find the pairs of inverse elements, illustrated by

$[1][1] = [1]$	$[2][6] = [1]$	$[3][4] = [1]$
$[6pt][5][9] = [1]$	$[7][8] = [1]$	$[10][10] = [1]$

Now

$$\begin{aligned}
 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdot 8 \cdot 9 \cdot 10 &\equiv 1 \cdot (2 \cdot 6) \cdot (3 \cdot 4) \cdot (5 \cdot 9) \cdot (7 \cdot 8) \cdot 10 \pmod{11} \\
 &\equiv 1 \cdot 1 \cdot 1 \cdot 1 \cdot 1 \cdot 10 \pmod{11} \\
 &\equiv 10 \pmod{11} \\
 &\equiv -1 \pmod{11}.
 \end{aligned}$$

Fermat's Little Theorem and Euler's Theorem

We start this section by introducing *Fermat's Little Theorem*.² This theorem has far reaching consequences for applications to cryptography and secure transmission of data on the Internet.

Theorem 1.24.2: Fermat's Little Theorem

If p is prime and a is relatively prime to p then

$$a^{p-1} \equiv 1 \pmod{p}.$$

Let's look at some examples.

✓ Example 1.24.1

When $p = 2$, Fermat's Little Theorem says that for any integer a relatively prime to 2 is congruent to 1 modulo 2; in other words, a is odd (hardly surprising!).

When $p = 3$, the Theorem says that if a is not a multiple of 3, then a^2 is one more than a multiple of three. We can verify this directly by noting that

$$\begin{aligned}
 (3k+1)^2 &= 9k^2 + 6k + 1 = 3(3k^2 + 2k) + 1 \quad \text{and} \\
 (3k+2)^2 &= 9k^2 + 12k + 4 = 3(3k^2 + 4k + 1) + 1
 \end{aligned}$$

for all integers k .

When $p = 11$, Fermat's Little Theorem says that any number a not divisible by 11 satisfies $a^{10} \equiv 1 \pmod{11}$. For example,

$$\begin{aligned}
 2^{10} &= 1024 = 11 \cdot 93 + 1; \\
 6^{10} &= 60466176 = 11 \cdot 5496925 + 1; \text{ and of course} \\
 10^{10} &= 10000000000 = 999999999 + 1 = 11 \cdot 909090909 + 1.
 \end{aligned}$$

From records available to us, it seems that Fermat's Little Theorem was originally stated (without proof) in a letter Fermat wrote in 1640. A proof was presented by Euler in 1736, and later Euler published a theorem that holds not only for primes but for all positive integer moduli:³

Theorem 1.24.3: Euler's Theorem

If $m > 0$ and a is relatively prime to m then

$$a^{\phi(m)} \equiv 1 \pmod{m}.$$

✓ Example 1.24.2

Take $m = 12$; then

$$\phi(m) = \phi(2^2 \cdot 3) = (2^2 - 2)(3 - 1) = 4.$$

The positive integers $a < m$ with $\gcd(a, m) = 1$ are 1, 5, 7 and 11.

$$\begin{aligned} 1^4 &\equiv 1 \pmod{12} \text{ is clear} \\ 5^2 &\equiv 1 \pmod{12} \text{ since } 12 \mid (25 - 1) \\ \therefore (5^2)^2 &\equiv 1^2 \pmod{12} \\ \therefore 5^4 &\equiv 1 \pmod{12}. \end{aligned}$$

Now $7 \equiv -5 \pmod{12}$, and since 4 is even,

$$\begin{aligned} 7^4 &\equiv (-5)^4 \equiv 5^4 \pmod{12} \\ \therefore 7^4 &\equiv 1 \pmod{12}. \end{aligned}$$

Finally, $11 \equiv -1 \pmod{12}$ and again, since 4 is even, we have

$$11^4 \equiv (-1)^4 \pmod{12}$$

and

$$11^4 \equiv 1 \pmod{12}.$$

This completes the verification of Euler's Theorem for the case $m = 12$.

[Exercise 1.24.7](#) at the end of the chapter asks you to explain why Fermat's Little Theorem can be derived as a consequence of Euler's Theorem. Let's now work towards proving Euler's Theorem.

Definition 1.24.1: Powers of Residue Classes

If $[a] \in U_m$ define $[a]^1 = [a]$ and for $n > 1$, $[a]^n = [a][a] \cdots [a]$ where there are n copies of $[a]$ on the right.

Theorem 1.24.4

If $[a] \in U_m$, then $[a]^n \in U_m$ for $n \geq 1$ and $[a]^n = [a^n]$.

Proof

We prove that $[a]^n = [a^n] \in U_m$ for $n \geq 1$ by induction on n .

If $n = 1$, $[a]^1 = [a] = [a^1]$ and by assumption $[a] \in U_m$. Suppose

$$[a]^k = [a^k] \in U_m$$

for some $k \geq 1$. Then

$$\begin{aligned}
 [a]^{k+1} &= [a]^k [a] \\
 &= [a^k] [a] \quad \text{by the induction hypothesis} \\
 &= [a^k a] \quad \text{by Definition 21.1.4} \\
 &= [a^{k+1}] \quad \text{since } a^k a = a^{k+1}.
 \end{aligned}$$

So by the PMI, the theorem holds for $n \geq 1$.

Note that for fixed $m > 0$ if $\gcd(a, m) = 1$ then $[a] \in U_m$. And using [Theorem 1.24.4](#) we have

$$a^n \equiv 1 \pmod{m} \iff [a^n] = [1] \iff [a]^n = [1].$$

It follows that Euler's Theorem ([Theorem 1.24.3](#)) is equivalent to the following theorem.

Theorem 1.24.5

If $m > 0$ and $[a] \in U_m$ then

$$[a]^{\phi(m)} = [1].$$

A proof of [Theorem 1.24.5](#) is outlined in the [Exercise 1.24.6](#) at the end of the chapter.

Also [Theorem 1.24.5](#) is an easy consequence of *Lagrange's Theorem*, which students who take (or have taken) a course in abstract algebra will learn about (or will already know).

To end this chapter, we note that Fermat's Little Theorem can be used to simplify the computation of $a^n \bmod p$ in the special case that p is prime. Recall that if $a^n \equiv r \pmod{p}$ where $0 \leq r < p$, then $a^n \bmod p = r$. We can do two things to simplify the computation:

1. Replace a by $a \bmod p$.
2. Replace n by $n \bmod (p-1)$.

We illustrate the technique with an example.

Example 1.24.3

Suppose we want to calculate

$$1234^{7865435} \bmod 11.$$

We first replace 1234 by $1234 \bmod 11$. Since the modulus is 11, our job is made easier by [Theorem 1.18.1 \(e\)](#), and we can write $1234 \equiv (4 - 3 + 2 - 1) \pmod{11}$, that is, $1234 \equiv 2 \pmod{11}$. Since $\gcd(2, 11) = 1$ we have $2^{10} \equiv 1 \pmod{11}$. Now $7865435 = (786543) \cdot 10 + 5$ so

$$\begin{aligned}
 2^{7865435} &\equiv 2^{(786543) \cdot 10 + 5} \pmod{11} \\
 &\equiv (2^{10})^{786543} \cdot 2^5 \pmod{11} \\
 &\equiv 1^{786543} \cdot 2^5 \pmod{11} \\
 &\equiv 2^5 \pmod{11},
 \end{aligned}$$

and $2^5 = 32 \equiv 10 \pmod{11}$. Hence,

$$1234^{7865435} \equiv 10 \pmod{11}.$$

It follows that

$$1234^{7865435} \bmod 11 = 10.$$

Exercises

? Exercise 1.24.1

Explain why when m is composite and not equal to 4, the equation (1.24.1) holds, i.e.,

$$[1][2][3] \cdots [m-1] = [0]$$

in $\mathbb{Z}_m$.

(Hint: handle the case where m is of the form p^2 , where p is prime, separately. When you handle this case, think about why $m = 4$ is an exception in the statement above.)

? Exercise 1.24.2

Let p be a prime.

- a. Use Wilson's Theorem to show that if $p \equiv 3 \pmod{4}$, then

$$\left[\left(\frac{p-1}{2} \right)! \right]^2 \equiv 1 \pmod{p}.$$

(Hint: show that the expression on the left is congruent to $-(p-1)!$ modulo p .)

- b. Show that if $p \equiv 1 \pmod{4}$, then

$$\left[\left(\frac{p-1}{2} \right)! \right]^2 \equiv -1 \pmod{p}.$$

? Exercise 1.24.3

Let p be a prime such that $p \equiv 3 \pmod{4}$.

1. Show that

$$(p-1)! \equiv -(1 \cdot 3 \cdot 5 \cdots (p-2))^2 \pmod{p}.$$

(Hint: replace all the even multiplicands in $(p-1)!$ by congruent negative numbers.)

2. Show that

$$(1 \cdot 3 \cdot 5 \cdots (p-2))^2 \equiv 1 \pmod{p}.$$

3. Explain why

$$1 \cdot 3 \cdot 5 \cdots (p-2) \equiv \pm 1 \pmod{p}.$$

(Use an idea from the proof of Wilson's Theorem, or the result of Exercise 1.21.11.)

4. Prove that if $p \equiv 3 \pmod{4}$, then

$$2 \cdot 4 \cdot 6 \cdots (p-1) \equiv \mp 1 \pmod{p}.$$

? Exercise 1.24.4

Let f be the function defined on the positive integers by

$$f(n) = \left\lfloor \frac{n! \bmod (n+1)}{n} \right\rfloor (n-1) + 2.$$

- Compute $f(1), f(2), \dots, f(10)$ and record the values.
- Explain why $f(n)$ is always a prime number, no matter what positive integer n is.
- What do you think? How good is this function as a means for generating prime numbers?

? Exercise 1.24.5

Verify that [Theorem 1.24.2](#) holds if $p = 5$ by direct calculation (as [Theorem 1.24.3](#) was verified for $m = 12$ in [Example 1.24.2](#)).

? Exercise 1.24.6

In this exercise you will prove [Theorem 1.24.5](#). Let $U_m = \{X_1, X_2, \dots, X_{\phi(m)}\}$. Here we write X_i for a residue class in U_m to simplify notation.

1. Show that if $X \in U_m$ then

$$\{XX_1, XX_2, \dots, XX_{\phi(m)}\} = U_m.$$

2. Explain why if $X \in U_m$ then

$$(XX_1)(XX_2) \cdots (XX_{\phi(m)}) = X_1 X_2 \cdots X_{\phi(m)}.$$

3. Let $A = X_1 X_2 \cdots X_{\phi(m)}$. Show that if $X \in U_m$ then $X^{\phi(m)} A = A$.
4. Conclude from the last step that $X^{\phi(m)} = [1]$ and hence [Theorem 1.24.5](#) is true.

? Exercise 1.24.7

Show that Fermat's Little Theorem follows quickly from Euler's Theorem.

? Exercise 1.24.8

Show that if p is prime then $a^p \equiv a \pmod{p}$ for all integers a .

(Hint: Consider two cases: (i) $\gcd(a, p) = 1$ and (ii) $\gcd(a, p) > 1$. Note that in the second case $p \mid a$.)

? Exercise 1.24.9

Let $m > 0$. Let $\gcd(a, m) = 1$. Show that $a^{\phi(m)-1}$ is an inverse for a modulo m . (See [Theorem 1.20.1](#).)

? Exercise 1.24.10

For all $a \in \{1, 2, 3, 4, 5, 6\}$ find an inverse a^* of a modulo 7 by using [Exercise 1.24.9](#). Choose a^* in each case so that $1 \leq a^* \leq 6$.

? Exercise 1.24.11

- a. Use the technique in [Example 1.24.3](#) to calculate

$$28^{1202} \bmod 13$$

(Since the modulus is not 11, you cannot use the same mod 11 trick ([Theorem 1.18.1](#) (e)), of course.)

- b. Can you use Euler's Theorem and reasoning similar to that of [Example 1.24.3](#) to simplify

$$29^{1202} \bmod 12?$$

If so, compute the answer; if not, explain why not.

? Exercise 1.24.12

Wilson's Theorem implies that for every prime p , the number $(p-1)! + 1$ is divisible by p . We call p a *Wilson prime* if even more is true—if the number $(p-1)! + 1$ is divisible by p^2 . Find the smallest two Wilson primes (each is less than 20), and

write a few sentences reporting on what is known and not currently known about Wilson primes, based on what you find through some research online.

Footnotes

[1] After John Wilson (1741{1793), though historians have identified work of the Arab mathematician and scientist Abu Ali al-Hasan ibn al-Haytham (also known as Alhazen, 965-1040) that shows he was aware of and made use of this fact. [7]

[2] Fermat's Little Theorem is not the only theorem named after Fermat. His "big" theorem, or, as it is better known, *Fermat's Last Theorem*, states that $x^n + y^n = z^n$ has no solutions in positive integers x, y, z when $n > 2$. This was proved by Andrew Wiles in 1995 over 350 years after it was first mentioned by Fermat. It's an interesting story that has been the subject of books and a PBS special (*Nova*, "The Proof"), and you are encouraged to learn more about it.

[3] Note the appearance here of *Euler's* totient function ϕ .

This page titled [1.24: Theorems of Wilson, Euler, and Fermat](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Mike Barrus & W. Edwin Clark](#).

1.25: Primality Tests

Two theorems from the previous chapter, Wilson's Theorem and Fermat's Little Theorem, connect prime numbers and congruences in perhaps surprising ways. In this chapter we look at how these theorems can be turned into techniques for *testing* whether a number n is prime. Those tests will differ quite a bit from the first, simplest means of testing primes we describe below, which is based on [Section 1.11](#) and specifically [Theorem 1.11.2](#):

✚ Primality Test #1

Given an integer $n \geq 2$, determine whether n is divisible by any integer in $2, \dots, \lfloor \sqrt{n} \rfloor$ (more specifically, by any prime number in this range). If so, then n is not prime; otherwise, n is prime.

Now we use the theorems from the last chapter.

A Primality Test Using Wilson's Theorem

To introduce our next test, observe the following.

- If n is prime, then $(n-1)! \equiv -1 \pmod{n}$ (this is Wilson's Theorem).
- If $n = 4$, then $(n-1)! \equiv 1 \pmod{n}$; if instead $n \neq 4$ but n is not prime, then $(n-1)! \equiv 0 \pmod{n}$ (see [Exercise 1.24.1](#)).

Hence an integer $n \geq 2$ is prime *if and only if* $(n-1)! \equiv -1 \pmod{n}$. This gives us a test.

✚ Primality Test #2

Given an integer $n \geq 2$, determine whether $(n-1)! \equiv -1 \pmod{n}$. If so, n is prime; if not, then n is composite.

For example, $4! = 24$, which is congruent to -1 modulo 5, so this test guarantees that 5 is prime; meanwhile, $11! = 39916800 \equiv 0 \pmod{12}$, so the test shows that 12 is not a prime number.

One big drawback of Primality Test #2 is that there does not seem to be a generally applicable way to efficiently compute $(n-1)!$ modulo n for large n . For instance $15!$ rounds to 1.3076744×10^{12} , and $20!$ rounds to 2.432902×10^{18} ; the factorial numbers $n!$ are very large even for relatively small n . We *can* compute $(n-1)! \pmod{n}$ by reducing modulo n after every multiplication; for instance, we compute $8! \pmod{9}$ by

$$\begin{aligned} 1 \bmod 9 &= 1; \\ (2 \cdot 1) \bmod 9 &= 2; \\ (3 \cdot 2 \cdot 1) \bmod 9 &= (3 \cdot 2) \bmod 9 = 6; \\ (4 \cdot 3 \cdot 2 \cdot 1) \bmod 9 &= (4 \cdot 6) \bmod 9 = 6; \\ (5 \cdot 4 \cdot 3 \cdot 2 \cdot 1) \bmod 9 &= (5 \cdot 6) \bmod 9 = 3; \\ (6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1) \bmod 9 &= (6 \cdot 3) \bmod 9 = 0; \\ (7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1) \bmod 9 &= (7 \cdot 0) \bmod 9 = 0; \\ (8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1) \bmod 9 &= (8 \cdot 0) \bmod 9 = 0. \end{aligned}$$

However, note that each line of such a calculation requires one more multiplication and one more application of “mod.” Hence testing whether $(n-1)! \equiv -1 \pmod{n}$ in this way takes a number of steps that is proportional to n , so as n gets larger and larger, we would expect this primality test to take longer and longer to perform (whether it is carried out by a human or a computer).

In contrast, note that the number of steps necessary to carry out Primality Test #1 is proportional to $\sqrt{n}$, so Primality Test #1 may take less time to apply than Primality Test #2, especially if n is large. Hence Wilson's Theorem, while it does give us a *new* test for primality, hasn't given us a *great* test for primality.

We wonder—can we find a still better test for primality?

A Primality Test Using Fermat's Little Theorem

According to Fermat's Little Theorem, if p is prime and $1 \leq a \leq p-1$, then

$$a^{p-1} \equiv 1 \pmod{p}.$$

The converse is also true in the following sense:

Theorem 1.25.1: Fermat's Little Theorem Converse

If $m \geq 2$ and for all a such that $1 \leq a \leq m-1$ we have

$$a^{m-1} \equiv 1 \pmod{m}$$

then m must be prime.

Proof

If the hypothesis holds, then for all a with $1 \leq a \leq m-1$, we know that a has an inverse modulo m , namely, a^{m-2} is an inverse for a modulo m . By [Theorem 1.20.2](#), this says that for $1 \leq a \leq m-1$, $\gcd(a, m) = 1$. But if m were not prime, then we would have $m = ab$ with $1 < a < m$, $1 < b < m$. Then $\gcd(a, m) = a > 1$, a contradiction. So m must be prime.

We can turn this into a test for primality.

Primality Test #3

Given an integer $n \geq 2$, for *each* a between 1 and $n-1$, test whether $a^{n-1} \equiv 1 \pmod{n}$. If the congruence is true for every value of a , then n is prime. If the congruence fails for any value of a between 1 and $n-1$, then n is not prime.

As an example, using Primality Test #3 to decide whether 13 is prime, we would verify the following:

$1^{12} = 1 \equiv 1 \pmod{13};$		$7^{12} = 13841287201 \equiv 1 \pmod{13};$
$2^{12} = 4096 \equiv 1 \pmod{13};$		$8^{12} = 68719476736 \equiv 1 \pmod{13};$
$3^{12} = 531441 \equiv 1 \pmod{13};$		$9^{12} = 282429536481 \equiv 1 \pmod{13};$
$4^{12} = 16777216 \equiv 1 \pmod{13};$		$10^{12} = 1000000000000 \equiv 1 \pmod{13};$
$5^{12} = 244140625 \equiv 1 \pmod{13};$		$11^{12} = 3.1384284... \times 10^{12} \equiv 1 \pmod{13};$
$6^{12} = 2176782336 \equiv 1 \pmod{13};$		$12^{12} = 8.9161004... \times 10^{12} \equiv 1 \pmod{13};$

Primality Tests Using Fermat's Little Theorem, Continued

Using Primality Test #3 to check that p is prime, we would have to check that $a^{p-1} \equiv 1 \pmod{p}$ for $a = 1, 2, 3, \dots, p-1$. As the last example showed, this is a lot of work.

What if $n > 2$ and we knew that $2^{n-1} \equiv 1 \pmod{n}$, though we didn't know what a^{n-1} was congruent to modulo n for other values of a ? Could we conclude that n was prime?

Unfortunately, the answer is no. For example, look at $n = 341$. This number is not prime (why not?). Still, it is true that $2^{341-1} \equiv 1 \pmod{341}$ (see [Exercise 1.25.1](#)).

The moral is that even if $2^{n-1} \equiv 1 \pmod{n}$, the number n need not be prime.¹ To be *absolutely certain* that n was prime using Fermat's Little Theorem and [Theorem 1.25.1](#), we would need to check whether $a^{n-1} \equiv 1 \pmod{n}$ for all integers a in $2, 3, \dots, n-1$.

On the other hand, consider the case of $m = 63$. Note that

$$2^6 = 64 \equiv 1 \pmod{63}.$$

Hence, $2^6 \equiv 1 \pmod{63}$. Raising both sides to the 10th power we have

$$2^{60} \equiv 1 \pmod{63}.$$

Then multiplying both sides by 2^2 we get

$$2^{62} \equiv 4 \pmod{63}.$$

Since

$$4 \not\equiv 1 \pmod{63},$$

we have

$$2^{62} \not\equiv 1 \pmod{63}.$$

This tells us that 63 is *not* prime, *without factoring* 63. We emphasize that in general if $2^{m-1} \not\equiv 1 \pmod{m}$ then we can be sure that m is not prime.

Can we turn this into a test (or partial test) for primality? Let's look at some experimental data.

FACT.

There are 455,052,511 odd primes $p \leq 10^{10}$, all of which satisfy $2^{p-1} \equiv 1 \pmod{p}$. There are only 14,884 composite numbers $2 < n \leq 10^{10}$ that satisfy $2^{n-1} \equiv 1 \pmod{n}$. Thus, if $2 < n \leq 10^{10}$ and n satisfies $2^{n-1} \equiv 1 \pmod{n}$, the probability that n is prime is

$$\frac{455,052,511}{455,052,511 + 14,884} \approx .999967292.$$

In other words, if you find that $2^{n-1} \equiv 1 \pmod{n}$, then it is highly likely (but not a certainty) that n is prime, at least when $n \leq 10^{10}$.

Thus, unlike the previous three primality tests in this chapter, our next primality test is a *probabilistic* primality test.

Primality Test #4

Given an integer $n > 2$, test whether $2^{n-1} \equiv 1 \pmod{n}$. If not, then n is definitely not prime. If $2^{n-1} \equiv 1 \pmod{n}$, then n is probably prime.

One might ask what happens if we use 3 instead of 2 in the above probabilistic primality test. Or, better yet, what if we evaluate $a^{n-1} \pmod{n}$ for several different values of a . Consider the following data:

- The number of primes $p \leq 10^6$ is 78,498.
- The number of composite numbers $n \leq 10^6$ such that $2^{n-1} \equiv 1 \pmod{n}$ is 245.
- The number of composite numbers $n \leq 10^6$ such that $2^{n-1} \equiv 1 \pmod{n}$ and $3^{n-1} \equiv 1 \pmod{n}$ is 66.
- The number of composite numbers $n \leq 10^6$ such that $a^{n-1} \equiv 1 \pmod{n}$ for all $a \in \{2, 3, 5, 7, 11, 13, 17, 19, 31, 37, 41\}$ is 0.

Pausing here, we have the another probabilistic primality test, more elaborate than the last:

Primality Test #5

Given an integer $n > 2$, test whether $a^{n-1} \equiv 1 \pmod{n}$ for $a \in \{2, 3, 5, 7, 11, 13, 17, 19, 31, 37, 41\}$ such that $n \nmid a$. If the congruences are not true for all these a , then n is definitely not prime. If the congruences are all true, then n is definitely prime if $n \leq 10^6$ and probably prime (*highly likely*) if $n > 10^6$.

Conclusion

In practice, there are better probabilistic primality tests than the ones mentioned in this chapter, using more efficient approaches than our applications of Wilson's Theorem and Fermat's Little Theorem. For more details see, for example, *Elementary Number Theory*, fourth edition, by Kenneth Rosen.

Computer algebra systems such as Maple run very sophisticated probabilistic primality tests (the computations for numbers $n \leq 10^6$ above were computed using Maple). For example, Maple's command `isprime(n)` returns false if n is not prime and returns true if n is probably prime. (The underlying primality test used by `isprime` uses a somewhat different idea than our

probilistic tests above.) While ideally we would like to eliminate all uncertainty from the results of a test like `isprime(n)`, so far no one has found an integer n for which `isprime(n)` gives the wrong answer.

In this chapter, our first three primality tests gave definite answers but took many steps to complete. The primality tests in the last section gave probable, but not always certain, answers using much less work.² In theory it is possible remove *all* uncertainty from our primality testing using an algorithm requiring a number of steps proportional to the *number of digits* in n (recall from [Exercise 1.6.11](#) that the number of digits in n is proportional to $\log n$). This is by using the *Agrawal–Kayal–Saxena*, or *AKS, primality test*, which was published in 2002 and won its authors multiple prestigious prizes, since it was the first primality test to accomplish all that it does. Note that the AKS test is of great *theoretical* importance, but in practice it is not used much at all—there are much more efficient ways of testing the primality of n .

We end this chapter with an invitation: Do a little research of your own, online or otherwise...What can you find out about the current state of the art in primality testing?

Exercises

? Exercise 1.25.1

Use a computer (or do it via hand and/or calculator) to verify that $2^{340} \equiv 1 \pmod{341}$ and that 341 is not prime. (Hint: for faster calculations, start by determining $2^{10} \bmod 341$.)

? Exercise 1.25.2

Use a computer (or do it via hand and/or calculator) to show that

1. $3^{90} \equiv 1 \pmod{91}$, but 91 is not prime.
2. $2^{n-1} \equiv 1 \pmod{n}$ and $3^{n-1} \equiv 1 \pmod{n}$ for $n = 1105$, but 1105 is not prime.

(Hint: note that $a^k \equiv 1 \pmod{n} \Leftrightarrow a^k \bmod n = 1$.)

? Exercise 1.25.3

Demonstrate the five primality tests from this chapter using the number $n = 11$, showing all computations and the determination of each test. Then repeat these same five tests for $n = 12$.

Footnotes

[1] If n is composite but $a^{n-1} \equiv 1 \pmod{n}$ for some a that is relatively prime to n , we call n a *Fermat pseudoprime*.

[2] One thing we have not commented on is how much work is involved in raising an integer a to a high exponent like $n - 1$, when n is large. This turns out to need less work than you might think. Stay tuned for the next chapter!

This page titled [1.25: Primality Tests](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Mike Barrus & W. Edwin Clark](#).

1.26: Computation of $a^n \bmod m$

Let's first consider the question: *What is the smallest number of multiplications required to compute a^N where N is any positive integer?*

Suppose we want to calculate 2^8 . One way is to perform the following 7 multiplications:

$$\begin{aligned} 2^2 &= 2 \cdot 2 = 4; \\ 2^3 &= 2 \cdot 4 = 8; \\ 2^4 &= 2 \cdot 8 = 16; \\ 2^5 &= 2 \cdot 16 = 32; \\ 2^6 &= 2 \cdot 32 = 64; \\ 2^7 &= 2 \cdot 64 = 128; \\ 2^8 &= 2 \cdot 128 = 256. \end{aligned}$$

But we can do it in only 3 multiplications:

$$\begin{aligned} 2^2 &= 2 \cdot 2 = 4; \\ 2^4 &= (2^2)^2 = 4 \cdot 4 = 16; \\ 2^8 &= (2^4)^2 = 16 \cdot 16 = 256. \end{aligned}$$

In general, using the first method's steps

$$a^2 = a \cdot a, \quad a^3 = a^2 \cdot a, \quad a^4 = a^3 \cdot a, \quad \dots, \quad a^n = a^{n-1} \cdot a$$

requires $n - 1$ multiplications to compute a^n .

On the other hand, if $n = 2^k$ then we can compute a^n by successive squaring with only k multiplications:

$$\begin{aligned} a^2 &= a \cdot a \\ a^{2^2} &= (a^2)^2 = a^2 \cdot a^2 \\ a^{2^3} &= (a^{2^2})^2 = a^{2^2} \cdot a^{2^2} \\ &\vdots \\ a^{2^k} &= (a^{2^{k-1}})^2 = a^{2^{k-1}} \cdot a^{2^{k-1}} \end{aligned}$$

Note that the fact that

$$2^k = (2^{k-1})^2 = 2^{k-1} \cdot 2^{k-1},$$

together with the exponent laws

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

and

$$a^n \cdot a^m = a^{n+m},$$

is what makes this method work. Note that if $n = 2^k$ then k is generally a lot smaller than $n - 1$. For example,

$$1024 = 2^{10}$$

and 10 is quite a bit smaller than 1023.

If n is not a power of 2 we can use the following method to compute a^n .

The Binary Method for Exponentiation

Let n be a positive integer. Let x be any real number. This is a method for computing x^n .

- **Step 1.** Find the binary representation

$$n = [a_r, a_{r-1}, \dots, a_0]_2$$

for n .

- **Step 2.** Compute the powers

$$x^2, x^{2^2}, x^{2^3}, \dots, x^{2^r}$$

by successive squaring as shown above.

- **Step 3.** Compute the product

$$x^n = x^{a_r 2^r} \cdot x^{a_{r-1} 2^{r-1}} \dots x^{a_1 2} \cdot x^{a_0}.$$

[Note each a_i is 0 or 1, so all needed factors were obtained in Step 2.]

✓ Example 1.26.1

Let's compute 3^{15} . Note that $15 = 2^3 + 2^2 + 2 + 1 = [1, 1, 1, 1]_2$. So this takes care of Step 1. For Step 2, we note that

$$\begin{aligned} 3^2 &= 3 \cdot 3 = 9; \\ 3^{2^2} &= 9 \cdot 9 = 81; \\ 3^{2^3} &= 81 \cdot 81 = 6561. \end{aligned}$$

So $3^{15} = 3^{2^3} \cdot 3^{2^2} \cdot 3^2 \cdot 3^1$. For this we need 3 multiplications:

$$\begin{aligned} 3 \cdot 3^2 &= 3 \cdot 9 = 27; \\ (3 \cdot 3^2) \cdot 3^{2^2} &= 27 \cdot 81 = 2187; \\ (3 \cdot 3^2 \cdot 3^{2^2}) \cdot 3^{2^3} &= 2187 \cdot 6561 = 14348907. \end{aligned}$$

So we have

$$3^{15} = 14348907.$$

Note that we have used just 6 multiplications, which is less than the 14 it would take if we used the naive method. Let's not forget that some additional effort was needed to compute the binary representation of 15, but not much.

🔗 Theorem 1.26.1

Computing x^n using the binary method requires $\lfloor \log_2(n) \rfloor$ applications of the Division Algorithm and at most $2\lfloor \log_2(n) \rfloor$ multiplications.

Proof

If $n = [a_r, \dots, a_0]_2$, $a_r = 1$, then $n = 2^r + \dots + a_1 2 + a_0$. Hence

$$2^r \leq n \leq 2^r + 2^{r-1} + \dots + 2 + 1 = 2^{r+1} - 1 < 2^{r+1}.$$

Since $\log_2(2^x) = x$ and when $0 < a < b$ we have $\log_2(a) < \log_2(b)$, the inequalities above yield

$$\log_2(2^r) \leq \log_2(n) < \log_2(2^{r+1})$$

or

$$r \leq \log_2(n) < r + 1.$$

Hence $r = \lfloor \log_2(n) \rfloor$. Note that r is the number of times we need to apply the Division Algorithm to obtain the binary representation $n = [a_r, \dots, a_0]_2$, $a_r = 1$. To compute the powers $x, x^2, x^{2^2}, \dots, x^{2^r}$ by successive squaring requires $r = \lfloor \log_2(n) \rfloor$ multiplications and similarly to compute the product

$$x^{2^r} \cdot x^{a_{r-1} 2^{r-1}} \dots x^{a_1 2} \cdot x^{a_0}$$

requires r multiplications. So after obtaining the binary representation we need at most $2r = 2\lfloor \log_2(n) \rfloor$ multiplications.

Use of a calculator to compute $\log_2(x)$

To find $\log_2(x)$ one may use the formula

$$\log_2(x) = \frac{1}{\ln(2)} \ln(x)$$

or

$$\log_2(x) \approx \frac{1}{0.69314718} \ln(x) \approx 1.442695 \ln(x),$$

where $\ln(x)$ is the natural logarithm of x . For small values of x it is sometimes faster to use the fact that $r = \lfloor \log_2(x) \rfloor$ is equivalent to

$$2^r \leq x < 2^{r+1},$$

that is, r is the largest positive integer such that $2^r \leq x$.

Note that if we count an application of the Division Algorithm and a multiplication as the same, the above tells us that we need at most $3\lfloor \log_2(n) \rfloor$ operations to compute x^n . So, for example, if $n = 10^6$, then it is easy to see that $3\lfloor \log_2(n) \rfloor = 57$. So we may compute $x^{1,000,000}$ with only 57 operations.

Computing $a^n \bmod m$

We use the binary method for exponentiation with the added trick that after every multiplication we reduce modulo m , that is, we divide by m and take the remainder. This keeps the products from getting too big.

✓ Example 1.26.2

We compute $3^{15} \bmod 10$:

$$\begin{aligned} 3^2 &= 3 \cdot 3 = 9 \equiv 9 \pmod{10}; \\ 3^4 &= 9 \cdot 9 = 81 \equiv 1 \pmod{10}; \\ 3^8 &\equiv 1 \cdot 1 \equiv 1 \pmod{10}; \\ \therefore 3^{15} &= 3^8 \cdot 3^4 \cdot 3^2 \cdot 3^1 \equiv 1 \cdot 1 \cdot 9 \cdot 3 = 27 \equiv 7 \pmod{10}. \end{aligned}$$

Note that $3^{15} \equiv 7 \pmod{10}$ implies that $3^{15} \bmod 10 = 7$. (Recall that we calculated that $3^{15} = 14348907$ which is clearly congruent to 7 mod 10, but the multiplications were not so easy.)

✓ Example 1.26.3

Let's find $2^{644} \bmod 645$. It is easy to see that

$$644 = [1, 0, 1, 0, 0, 0, 0, 1, 0, 0]_2.$$

That is, $644 = 2^9 + 2^7 + 2^2 = 512 + 128 + 4$. Now by successive squaring and reducing modulo 645 we get

$$\begin{aligned} 2^2 &= 2 \cdot 2 = 4 \equiv 4 \pmod{645}; \\ 2^4 &\equiv 4 \cdot 4 = 16 \equiv 16 \pmod{645}; \\ 2^8 &\equiv 16 \cdot 16 = 256 \equiv 256 \pmod{645}; \\ 2^{16} &\equiv 256 \cdot 256 = 65,536 \equiv 391 \pmod{645}; \\ 2^{32} &\equiv 391 \cdot 391 = 152,881 \equiv 16 \pmod{645}; \\ 2^{64} &\equiv 16 \cdot 16 = 256 \equiv 256 \pmod{645}; \\ 2^{128} &\equiv 256 \cdot 256 = 65,536 \equiv 391 \pmod{645}; \\ 2^{256} &\equiv 391 \cdot 391 = 152,881 \equiv 16 \pmod{645}; \\ 2^{512} &\equiv 16 \cdot 16 = 256 \equiv 256 \pmod{645}. \end{aligned}$$

Now

$$2^{644} = 2^{512} \cdot 2^{128} \cdot 2^4,$$

hence

$$2^{644} \equiv 256 \cdot 391 \cdot 16 \pmod{645}.$$

So

$$256 \cdot 391 = 100096 \equiv 121 \pmod{645}$$

and

$$121 \cdot 16 = 1936 \equiv 1 \pmod{645},$$

so we have $2^{644} \equiv 1 \pmod{645}$. Hence $2^{644} \bmod 645 = 1$.

Exercises

? Exercise 1.26.1

Calculate $3 \lfloor \log_2(n) \rfloor$ for $n = 2,000,000$.

? Exercise 1.26.2

Use the binary method to compute 2^{25} .

? Exercise 1.26.3

Approximately how many operations would be required to compute 2^n when $n = 10^{100}$? Explain.

? Exercise 1.26.4

Note that 6 multiplications are used to compute 3^{15} using the binary method. Show that one can compute 3^{15} with fewer than 6 multiplications. (You will have to experiment.)

? Exercise 1.26.5

Calculate $2^{513} \bmod 10$.

? Exercise 1.26.6

Calculate $2^{517} \bmod 100$.

? Exercise 1.26.7

If you multiplied out 2^{517} , how many decimal digits would you obtain? (See [Exercise 1.6.9](#).)

? Exercise 1.26.8

Note that we calculated $1234^{7865435} \bmod 11$ with very few multiplications. Why can we not use that method to compute $1234^{7865435} \bmod 12$?

This page titled [1.26: Computation of \$a^n \bmod m\$](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Mike Barrus & W. Edwin Clark](#).

1.27: The RSA Scheme

In this chapter we discuss the basis of the RSA scheme. This is one of the most important examples of a *public key cryptographic scheme*. Put simply, in a public key cryptographic system, information (called the “public key”) is provided to the world by a person or company (let’s call the person “Bob”) who wants to receive messages. Anyone in the world, including a person we’ll call “Alice,” can safely send messages to Bob by disguising, or *encrypting* messages before sending them. Once Alice’s encrypted messages are received by Bob, he uses private information that only he knows to *decrypt* the messages, or change them back into their original state before the decryption. What is important is that the messages Alice sends to Bob be encrypted in such a way that even if these messages were intercepted by a third party (call her “Eve,”¹), they would be unreadable even if Eve knew the information Bob made public that Alice used to encrypt the messages.

In other words, in order for a public key cryptographic scheme to be successful, everyone should be able to follow Bob’s instructions for encrypting messages to be sent to him, but only Bob should know how to decrypt those messages once they are encrypted. A big question immediately arises: is this even possible? If Eve knew how the messages were encrypted, shouldn’t she be able to figure out how to “undo” the encryption and read Alice’s messages that were sent to Bob?

A big step forward in public key cryptography happened in 1977 as R. Rivest, A. Shamir and L. Adelman (note their last initials) discovered and developed a number theoretic technique for encrypting information in such a way that the encryption could not be undone unless the recipient had secret information, or computing power and *lots* of time.²

Let’s play the role of Alice to see how the RSA scheme works. In RSA Bob begins by making public two integers m and e , called the *modulus* and *encryption exponent*, respectively. In our role as Alice, we take a message we wish to send to Bob and convert it to an integer in the set $Z_m = \{0, 1, 2, \dots, m-1\}$ where m is Bob’s modulus (m is usually a very large number, too big for anyone to factor quickly).

Having written our intended message to Bob as an integer (let’s denote this integer by x), we encrypt our message by raising x to the power e and reducing it modulo m . In symbols, we encrypt our message by computing

$$M = x^e \bmod m.$$

The number M resulting from this computation is our encrypted message; it’s M , instead of x , that we send to Bob.

You are invited to think about how an eavesdropper Eve, seeing the encrypted message $M = x^e \bmod m$ and knowing m and e but not x , could determine what x was. You’ll probably come to the conclusion that it’s not an easy task, especially if m is very large.

Now in the RSA scheme Bob receives our encrypted message M and needs to decrypt it to see what our original message x was. Luckily, Bob (and only Bob) knows a special integer d called the *decryption exponent*, which has the property that

$$M^d \bmod m = x.$$

Bob simply raises the encrypted message M to the decryption exponent d , reduces it modulo m , and the original, unencrypted message x pops out.

Let’s see an example (using a much smaller modulus and encryption exponent than would be used in practice).

✓ Example 1.27.1

Suppose that Bob uses modulus $m = 57$ and $e = 25$; he publishes these online so that anyone in the world can use this information to encrypt messages they plan to send to him. Alice wants to communicate the number $x = 23$ to Bob, but she doesn’t want Eve to know what her message is. Alice encrypts the message by computing

$$M = x^e \bmod m = 23^{25} \bmod 57;$$

to do so she might use the binary method from [Section 1.26](#)—she finds that $25 = 16 + 8 + 1$, and

$$\begin{aligned} 23^1 &\equiv 23 \pmod{57}; \\ 23^2 &\equiv 16 \pmod{57}; \\ 23^4 &\equiv 28 \pmod{57}; \\ 23^8 &\equiv 43 \pmod{57}; \\ 23^{16} &\equiv 25 \pmod{57}, \end{aligned}$$

so

$$23^{25} = 23^{16} 23^8 23^1 \equiv 25 \cdot 43 \cdot 23 \equiv 44 \pmod{57}.$$

Thus $M = 44$, and Alice sends the encrypted message 44 to Bob.

Now Eve intercepts the message 44 and knows it has to equal $x^{25} \pmod{57}$, but Eve is unable to determine x directly (she might be able to just experiment by raising guessed numbers to the 25th power and reducing modulo 57 until she finds one that produces 44, but this process would take longer than Eve has patience for).

Bob, on the other hand, receives the message 44 and raises it to his decryption exponent d ; here $d = 13$, but no one knows this but Bob. He computes

$$44^{13} \pmod{57} = 44^8 44^4 44^1 \pmod{57} = (16 \cdot 4 \cdot 44) \pmod{57} = 23$$

and sees that Alice sent him the message 23.

There is a lot that can be said about implementing the RSA scheme in ways to make it more efficient computationally and more secure. Our purpose here, though, is to give the number theoretic underpinning of the basic scheme.

To explain why RSA works, we will require two functions:

$$E : Z_m \rightarrow Z_m \quad (\text{E for encipher})$$

and

$$D : Z_m \rightarrow Z_m \quad (\text{D for decipher}).$$

To be able to use D to decipher what E has enciphered we need to have $D(E(x)) = x$ for all $x \in Z_m$.

Now as described above, we will use

$$E(x) = x^e \pmod{m} \quad \text{and} \quad D(M) = M^d \pmod{m}.$$

To see why E and D are chosen this way, and to suggest how Bob can decide what to use for m , we first prove a lemma:

Lemma 1.27.1

Let p and q be any two distinct primes and let $m = pq$. Let e and d be any two positive integers which are inverses of each other modulo $\phi(m)$. Then

$$(x^e)^d = x^{ed} \equiv x \pmod{m}$$

for all x .

Proof

By [Theorem 1.15.4](#), $\phi(m) = (p-1)(q-1)$. Since $ed \equiv 1 \pmod{\phi(m)}$ we have $ed - 1 = k\phi(m) = k(p-1)(q-1)$ for some k . Note $k > 0$ unless $ed = 1$ in which case the theorem is obvious. So we have

$$ed = k\phi(m) + 1 = k(p-1)(q-1) + 1 \tag{1.27.1}$$

for some $k > 0$.

Now by Fermat's Little Theorem, if $\gcd(x, p) = 1$ we have $x^{p-1} \equiv 1 \pmod{p}$ and raising both sides of the congruence to the power $(q-1)k$ we obtain:

$$x^{(p-1)(q-1)k} \equiv 1 \pmod{p}$$

and multiplying both sides by x we have

$$x^{(p-1)(q-1)k+1} \equiv x \pmod{p}$$

That is, by equation [\(1.27.1\)](#),

$$x^{ed} \equiv x \pmod{p}.$$

Now we proved this last statement when $\gcd(x, p) = 1$, but if $\gcd(x, p) = p$, the same conclusion is true, since then $x \equiv 0 \pmod{p}$. Hence $x^{ed} \equiv x \pmod{p}$ holds in all cases. A similar argument proves that for all x ,

$$x^{ed} \equiv x \pmod{q}.$$

So by [Exercise 1.17.11](#), since $\gcd(p, q) = 1$ and $m = pq$, we have

$$x^{ed} \equiv x \pmod{m}$$

for all x .

The congruence in the statement of this lemma can be rewritten as $D(E(x)) = x$ for all x , meaning that the decryption function D , as we have defined it, manages to undo the encryption function E if the following statements are true:

- The modulus m is the product of two distinct primes p and q .
- The decryption exponent d is the inverse of the encryption exponent e in $Z_{\phi(m)}$. Hence d and e are relatively prime to $\phi(m)$.

In the example above, prior to Alice sending Bob any messages, Bob had created the modulus m by choosing two primes $p = 3$ and $q = 19$ and letting $m = pq$. Since d and e needed to be inverses of each other modulo $\phi(m)$, Bob found

$$\phi(m) = \phi(pq) = \phi(p)\phi(q) = (p-1)(q-1) = (3-1)(19-1) = 36.$$

Hence e must be a number that is relatively prime to 36; Bob chose $e = 25$ and computed 25^* to be 13, which told him d .

Now Bob had done all of this preliminary work in private. When he was ready, he shared none of this information with the world except for the modulus m and the encryption exponent e . Still, m and e are the only things Alice (or anyone else) needs to send Bob encrypted messages.

Attacks on the RSA cryptographic scheme

Since its discovery in the late 1970's, many attempts have been made to attack the RSA scheme, or in other words, find ways for an eavesdropper Eve to take an encrypted message M and determine what the original message x was without knowing the decryption exponent d ahead of time. Though we won't say much about these here, we will encourage the interested reader to do a little library research; there are many articles describing such attacks.³

In practice, every step of Bob's preparation must be carried out carefully to prevent against attacks Eve can carry out. For instance, if Bob's modulus m is too small, then Eve might be able to factor it and deduce what primes p and q were used to form it; from there $\phi(m)$ is easy to compute and so is d , since Eve knows what e is. Hence Bob's primes p and q must both be very large—perhaps 100 digits long each, so that m has 200 digits and no algorithm is likely to stumble on a factorization of m in any reasonable amount of time.

Other quite clever attacks are possible if, for instance, multiple individuals are all known to use the same modulus m , or if Bob's decryption exponent is very small in comparison to m . These attacks often use computing power to test a staggering number of possible cases, but if p, q, d, e are not chosen carefully, the number of cases to check is small enough that the encryption exponent d can somehow be determined by Eve, and all encrypted traffic intended for Bob can be read by Eve as soon as it is intercepted.

Since d may be determined from m and e if m can be factored into p and q , and since one way to improve the security of the encryption is by making sure m is large enough to make factoring it unlikely, an unanswered question suggests itself:

Open Question

Is there a fast algorithm for factoring large integers?

A truly fast algorithm for factoring would have important implications for cryptography and data security. It is interesting (and a little sobering!) to reflect upon the idea that the answers to abstract number theoretic questions have the potential to affect the well-being of humans across the world.

Exercises

? Exercise 1.27.1

Suppose you plan to use RSA to allow people to send you encrypted messages. Your use of RSA employs modulus $m = pq$, encryption exponent e , and decryption exponent d .

- a. Which of the following pieces of information should be made public to people wanting to send you messages?

$$p \quad d \quad m \quad q \quad e \quad \phi(m)$$

- b. For each of the pieces of information in part (a) that should not be made public, explain in detail how a person discovering this information could read the encrypted messages.

? Exercise 1.27.2

- a. Using modulus $m = 851$ and encryption exponent $e = 7$, encrypt the message 28.
 b. The decryption exponent for $e = 7$ is $d = 679$. Show that decrypting your answer to part (a) produces the original message (i.e., 28). It may help to know that $679 = 512 + 128 + 32 + 4 + 2 + 1$.

? Exercise 1.27.3

A message has been previously encrypted as 100 using modulus $m = 851$ and a suitable encryption exponent. If we know that the decryption exponent is 97, then what is the original message?

? Exercise 1.27.4

- a. Suppose that we wish to carry out RSA using $m = 35$. Which numbers in $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ could successfully be used as encryption exponents?
 b. For each potential encryption exponent you found in part (a), determine the corresponding decryption exponent.

In the next two Exercises, use the following table to convert letters to numbers before encryption or after decryption.

A	B	C	D	E	F	G	H	I	J	K	L	M
2	3	4	5	6	7	8	9	10	11	12	13	14
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
15	16	17	18	19	20	21	22	23	24	25	26	27

? Exercise 1.27.5

Using modulus $m = 391$ and encryption exponent 3, and the table above, what are the resulting encrypted numbers produced from the original message 'NUMBER'?

? Exercise 1.27.6

Suppose an individual wants to use modulus $m = 55$ and encryption exponent 3. This is a poor choice, since it will easily allow the encryption to be broken. Let's demonstrate this:

- a. What would the decryption exponent be? (Explain how you arrived at your answer.)
 b. Suppose you intercepted a message that was encrypted using this modulus and encryption exponent, and the encrypted message reads

$$51 \quad 8 \quad 25 \quad 31.$$

What was the message before encryption? Use the table above to recognize the pre-encryption message as a word.

? Exercise 1.27.7

Is it possible to choose a modulus m and an encryption exponent e such that the decryption exponent d equals e ? If not, explain why not. If so, provide an example of such an m and e .

Footnotes

[1] For "eavesdropper" - or perhaps "evil".

[2] A copy of their paper "A Method for Obtaining Digital Signatures and Public-Key Cryptosystems" may be downloaded from <http://citeseer.nj.nec.com/rivest78method.html> .

[3] One well-known survey is "Twenty Years of Attacks on the RSA Cryptosystem" by Dan Boneh of Stanford University, available online at <https://crypto.stanford.edu/dabo/papers/RSA-survey.pdf>.

This page titled [1.27: The RSA Scheme](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Mike Barrus](#) & [W. Edwin Clark](#).

1.28: Sum of Squares

In this chapter we present the answers to the first question asked in the first chapter of this book: which positive integers can be expressed as the sum of two, three, or four perfect squares?

Sums of two squares

Shown here is the list of the integers from 0 to 100 that can be written as a sum of two squares, organized into two sets of four columns, with spaces marking integers that are not equal to the sum of two squares.

		0,	49,	50,	52,
1,	2,	4,	53,		
5,		8,		58,	
9,	10,		61,		64,
13,		16,	65,		68,
17,	18,	20,			72,
			73,	74,	
25,	26,				80,
29,		32,	81,	82,	
	34,	36,	85,		
37,		40,	89,	90,	
41,					
45,			97,	98,	100

Figure 1.28.1

A few patterns seem to jump out immediately:

- We see no number that is congruent to 3 modulo 4.
- In the second column we see mostly numbers that are separated by some multiple of 8—so we see no number that is congruent to 6 modulo 8.

Now numbers that are congruent to 6 modulo 8 are 2 times numbers that are congruent to 3 modulo 4, so perhaps whether a number does or does not appear depends on the prime factorization of the number. Indeed, the third column contains none of the primes 3, 7, 11, 19, 23, 31, 43, 47, 59, 67, 71, 79, or 83, and many—in fact all—of the numbers missing from the list have one of these primes as a prime factor. In fact,

- The only multiples of 3 appearing in the list are 9, 18, 36, 45, 72, 81, 90—the multiples of $9 (= 3^2)$ but not of $27 (= 3^3)$, except for $81 (= 3^4)$.
- The only multiples of 7 appearing in the list are 49 and 98—the multiples of $49 (= 7^2)$.
- No multiple of 11, 19, or any of the other missing primes appears in the list (though clearly if we extended the list, then $121 = 11^2 + 0^2$ would appear on the list).

Examining how these patterns continue among higher numbers, we might arrive at the following statement, which we will partially justify in this section.

Theorem 1.28.1: Sum of 2 Squares

A positive integer n is equal to the sum of two perfect squares if and only if the prime factorization of n contains no odd exponent on any prime that is congruent to 3 modulo 4.

Why might this be true? While a full proof would need more preparation than we will give in this text, We can establish a few helpful lemmas to hint at where this theorem comes from.

Lemma 1.28.1

If a and b can both be written as the sum of two squares, then ab can also be written as the sum of two squares.

Proof

Suppose that $a = w^2 + x^2$ and $b = y^2 + z^2$. We use a factoring trick from the Gaussian integers:

$$\begin{aligned} a &= w^2 - (-x^2) = (w + ix)(w - ix); & b &= y^2 - (-z^2) = (y + iz)(y - iz). \\ ab &= (w + ix)(w - ix)(y + iz)(y - iz) \\ &= [(w + ix)(y + iz)][(w - ix)(y - iz)] \\ &= [(wy - xz) + i(wz + xy)][(wy - xz) - i(wz + xy)] \\ &= (wy - xz)^2 + (wz + xy)^2. \end{aligned}$$

Since $wy - xz$ and $wz + xy$ are integers, ab is a sum of two squares.

Because of this lemma, if we can determine which powers of primes can be written as sums of two squares, then we'll be a lot closer to showing that [Theorem 1.28.1](#) is correct. The easiest statement is this:

Proposition 1.28.1

If p is a prime and n is a nonnegative integer, then $p^{2n} = (p^n)^2 + 0^2$, so any prime raised to an even power can be written as the sum of two squares.

Here's another piece of the puzzle.

Lemma 1.28.2

If $n \equiv 3 \pmod{4}$, then n cannot be written as a sum of two squares.

Proof

Observe that for any integer k ,

$$(2k)^2 = 4k^2 \quad \text{and} \quad (2k+1)^2 = 4k^2 + 4k + 1 = 4(k^2 + k) + 1,$$

so even numbers have squares that are congruent to 0 modulo 4, and odd numbers have squares that are congruent to 1 modulo 4. Hence the sum of two squares can be congruent to either 0 ($= 0 + 0$) or 1 ($= 0 + 1$) or 2 ($= 1 + 1$), but it cannot be congruent to 3 modulo 4.

To complete the proof of [Theorem 1.28.1](#), we would need to show the following:

- Every prime that is congruent to 1 modulo 4 can be written as a sum of two squares. (This statement was made by A. Girard in 1632, and again by Fermat, who claimed in a 1654 letter to Pascal to have a proof, though we don't have a record of it; Euler published a proof in 1758.)
- If n has a prime factorization where some prime that is congruent to 3 modulo 4 is raised to an odd exponent, then n cannot be written as a sum of two squares.

Proofs of both of these statements are made possible through the study of *quadratic residues*. These are simply the numbers that occur as perfect squares with respect to a given modulus; for instance, modulo 10, the quadratic residues are 0, 1, 4, 9, 6, 5 (but not 2, 3, 7, or 8—these never show up as the units digit in a perfect square). There is a wealth of interesting properties and results related to quadratic residues that is *just barely* beyond the scope of this book. You are encouraged to look these up in a bit more advanced of a text and do some reading on your own. In the meantime, we will move on now to the next problem.

Sums of three squares

As we did in the previous section, let's start by listing the nonnegative integers that are at most 100 and able to be written as the sum of three squares. Recalling from [Section 1.1](#) that 7, 15, and 23 could not be written in this way, we put our list in 8 columns.

						0
1,	2,	3,	4,	5,	6,	8,
9,	10,	11,	12,	13,	14,	16,
17,	18,	19,	20,	21,	22,	24,
25,	26,	27,		29,	30,	32,
33,	34,	35,	36,	37,	38,	40,
41,	42,	43,	44,	45,	46,	48,
49,	50,	51,	52,	53,	54,	56,
57,	58,	59,		61,	62,	64,
65,	66,	67,	68,	69,	70,	72,
73,	74,	75,	76,	77,	78,	80,
81,	82,	83,	84,	85,	86,	88,
89,	90,	91,		93,	94,	96,
97,	98,	99,	100			

Figure 1.28.2

Here the patterns are even starker:

- We see no number that is congruent to 7 modulo 8.
- The only other missing numbers belong to the 4th column, where they are $28 (= 4 \cdot 7)$, $60 (= 4 \cdot 15)$, and $92 (= 4 \cdot 23)$; these numbers are each 4 times a number that is congruent to 7 modulo 8.

The last observation suggests that again factorizations of a number may make a difference in expressibility as a sum of squares.

In fact, the correct statement, proved by Legendre in 1797 or 1798, is this:

Theorem 1.28.2: The Sum of 3 Squares

A positive integer n is equal to the sum of three perfect squares if and only if n does not have the form $4^a(8b+7)$.

Like that of [Theorem 1.28.1](#), this proof is beyond our grasp at the moment, but once again we will say what we can. We start with a simple corollary to [Theorem 1.28.1](#).

Proposition 1.28.2

Any number that can be written as the sum of two squares can be written as the sum of three squares, since if $n = a^2 + b^2$ then $n = a^2 + b^2 + 0^2$. Hence we may write as the sum of three squares any n for which the prime factorization of n contains no odd exponent on any prime that is congruent to 3 modulo 4.

We may also write as the sum of three squares any number that is 1 more than a number that is the sum of two squares, since if $n = a^2 + b^2$, then $n + 1 = a^2 + b^2 + 1^2$.

As look at the gaps in our list above, we see that numbers that are congruent to 7 modulo 8 are all congruent to 3 modulo 4 and so are not expressible as the sum of two squares. We also note that each gap in our list corresponds to a gap of at least two consecutive numbers in the list of two-squares numbers in the last section.

We turn to explaining why some numbers cannot be written as the sum of three squares.

Lemma 1.28.3

If $n \equiv 7 \pmod{8}$, then n cannot be written as a sum of three squares.

Proof

See [Exercise 1.28.4](#).

Lemma 1.28.4

No integer of the form $4^a(8b+7)$, where a and b are nonnegative integers, is equal to the sum of three squares.

Proof

We prove the result by induction on a . Observe that if $a = 0$, then $4^a(8b+7)$ cannot equal the sum of three squares by [Lemma 1.28.3](#). Suppose now that no integer of the form $4^s(8m+7)$, where m is an integer, is equal to the sum of three squares for any $s \in \{0, 1, \dots, k\}$, where k is a nonnegative integer. Let $n = 4^{k+1}(8b+7)$, and suppose to the contrary that $n = p^2 + q^2 + r^2$ for integers p, q, r . Now $n \bmod 8 \in \{0, 4\}$, and it can be shown (see [Exercise 1.28.4](#)) that each of p^2, q^2, r^2 is congruent to 0 or 1 or 4 modulo 8. The only way these numbers can sum to 0 or 4 modulo 8 is by having them all be congruent to 0, or two congruent to 0 and one congruent to 4, or all congruent to 4, modulo 8. In each case, each of p^2, q^2, r^2 is divisible by 4, and hence each of p, q, r is divisible by 2. We may thus write $p = 2p', q = 2q',$ and $r = 2r'$, where p', q', r' are integers, and thus

$$n = (2p')^2 + (2q')^2 + (2r')^2 = 4[(p')^2 + (q')^2 + (r')^2],$$

and hence the integer $n/4$ can be written as a sum of three squares:

$$4^k(8b+7) = (p')^2 + (q')^2 + (r')^2.$$

Since the exponent on 4 above makes the left-hand side fit our induction hypothesis, this is a contradiction. Hence by PMI, no integer of the form $4^a(8b+7)$, where a and b are nonnegative integers, is equal to the sum of three squares.

To complete the proof of [Theorem 1.28.2](#), we would need to show that every number n that is not of the form $4^a(8b+7)$ can be written as the sum of three squares. This result is quite a bit more involved than either the two-square theorem of the last section or the four-square theorem in the next section. Its proofs, like the two-square theorem's proof, often make use of properties of quadratic residues (and in particular, a theorem known as the *law of quadratic reciprocity*) that are beyond the scope of this book, but which are worth pursuing, if you are interested.

Sums of four squares

In [Section 1.1](#) we identified no positive integer that could not be written as the sum of four squares. This is no accident.

Theorem 1.28.3: Sum of 4 Squares

Every positive integer n is equal to the sum of four perfect squares.

This statement is believed to have been known (though without proof) by Diophantus, based on examples in the *Arithmetica*; a proof was given by Lagrange in 1770.

Once the three-square theorem ([Theorem 1.28.2](#)) was proved by Legendre in 1797–1798, the four-square theorem became a corollary, as we now show.

Proof of Theorem 1.28.3

Proof

Let n be any positive integer. If $n = a^2 + b^2 + c^2$ for integers a, b, c , then $n = a^2 + b^2 + c^2 + 0^2$. If n cannot be written as the sum of three perfect squares, then n has the form $4^k(8\ell+7)$ for nonnegative integers k, ℓ . The number $4^k(8\ell+6)$ does not have this form (the exponent on 2 in the prime factorization is odd, while the exponent on 2 is even in the prime factorization of $4^k(8\ell+7)$). Thus there exist integers a, b, c such that $4^k(8\ell+6) = a^2 + b^2 + c^2$, and

$$n = 4^k(8\ell+7) = 4^k(8\ell+6) + 4^k = a^2 + b^2 + c^2 + (2^k)^2.$$

Hence n may always be written as the sum of four squares.

However, Lagrange's proof predates Legendre's proof by nearly two decades, and it is much simpler. We will not give the full proof here—once again, it involves facts about quadratic residues that we do not currently have—but we will mention a four-square

analogue of the trick in [Lemma 1.28.1](#) that is useful here as well.

Lemma 1.28.5: Product of 4 Sums

If a and b can both be written as the sum of four squares, then ab can also be written as the sum of four squares.

Proof

Suppose that $a = a_1^2 + a_2^2 + a_3^2 + a_4^2$ and $b = b_1^2 + b_2^2 + b_3^2 + b_4^2$. Tedious though straightforward algebra verifies that

$$\begin{aligned} ab &= (a_1^2 + a_2^2 + a_3^2 + a_4^2)(b_1^2 + b_2^2 + b_3^2 + b_4^2) \\ &= (a_1b_1 - a_2b_2 - a_3b_3 - a_4b_4)^2 \\ &\quad + (a_1b_2 + a_2b_1 + a_3b_4 - a_4b_3)^2 \\ &\quad + (a_1b_3 - a_2b_4 + a_3b_1 + a_4b_2)^2 \\ &\quad + (a_1b_4 + a_2b_3 - a_3b_2 + a_4b_1)^2. \end{aligned}$$

Each of the quantities inside a matched pair of parentheses is an integer, so ab is the sum of four squares.¹

It follows that if we can show that every prime is the sum of four squares, then every positive integer can be written as the sum of four squares. Results from the study of quadratic residues do allow us to prove the four-squares result for primes, and [Theorem 1.28.3](#) follows.

As we conclude this chapter, we mention that the results of this chapter tie in with *Waring's problem*, proposed in 1770 (the same year as Lagrange's four-square theorem proof) by Edward Waring. Waring's problem asks whether every power k has an associated number t_k such that every positive integer can be expressed as the sum of t_k k -th powers. For example, when $k = 2$, we have shown that we can take $t_2 = 4$, since every number can be written as the sum of four perfect squares. Must t_k exist for every k ? The answer is yes, as shown by David Hilbert in 1909.

Other similar results exist. For example, Fermat stated in 1638 (again, without proof) that every positive integer is a sum of at most n different n -gonal numbers, where an n -gonal number or *polygonal number* is a number that counts the number of dots used when the dots are arranged in "shells" shaped like a regular n -gon. The triangular numbers of [Exercise 1.3.14](#) are an example. In Gauss's mathematical diary on July 10, 1796, he is reported to have written, using Greek and some personal symbolic shorthand,

ΕΥΡΕΚΑ! num = $\Delta + \Delta + \Delta$.

This result is accordingly called Gauss's "Eureka Theorem"; in words, it states that every positive integer can be written as the sum of at most three triangular numbers. Of course [Theorem 1.28.3](#) Lagrange's 1770 result, shows that every positive integer can be written as the sum of at most four square numbers. Finally, Fermat's polygonal number claim was proved for all n by Cauchy in 1813.

These theorems highlight a special property of these numbers. By way of contrast, observe that though every positive integer can be written as a sum of integral powers of 2 (or any greater integer base—this was [Theorem 1.6.1](#)), we cannot always do it with three powers of 2, or four, or even ten powers of 2. The number of summands necessary quickly grows larger than any particular finite number—in fact, $2^n - 1$ will always need at least n summands, no matter how big n gets. So square numbers and polygonal numbers somehow possess useful properties that powers of an integer do not. There's probably a lesson there...or maybe even a research question.

Exercises

Exercise 1.28.1

Which of the following may be written as a sum of two squares? For any that cannot, give a detailed justification of why not. For any that can, list two numbers whose squares add up to the given number.

- 315
- 569
- 100,000
- 54,925.

? Exercise 1.28.2

Prove that the norm of a Gaussian integer can never be congruent to 3 modulo 4.

? Exercise 1.28.3

Prove that a prime integer that is congruent to 1 modulo 4 is never a Gaussian prime (as defined in [Section 1.13](#)).

(Hint: as part of your answer, recall how sums of squares can be factored in the Gaussian integers.)

? Exercise 1.28.4

Show that no integer of the form $8b + 7$ is equal to the sum of three squares.

(Hint: Follow the spirit of the proof of [Lemma 1.28.2](#) looking at the possible residues of $a^2 + b^2 + c^2$, modulo 8, over all the possible cases for the remainders when a, b, c are divided by 4.)

? Exercise 1.28.5

Which of the following may be written as a sum of three squares? For any that cannot, give a detailed justification of why not. For any that can, list three numbers whose squares add up to the given number.

- 331
- 156
- 76.

? Exercise 1.28.6

Write each of the numbers n , for $30 \leq n \leq 40$, as a sum of squares, in each case using as few squares as possible.

? Exercise 1.28.7

(For students who have seen matrices and determinants.) In many ways, the matrices of the form $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$, where $a, b \in \mathbb{Z}$, behave like Gaussian integers $a + bi$. (To see this, add and multiply $a + bi$ and $c + di$, and compare the results to the sum and product of $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$ and $\begin{bmatrix} c & d \\ -d & c \end{bmatrix}$.) Now the determinant of $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$ is $a^2 + b^2$, which equals $(a + bi)(\overline{a + bi})$ and also equals a sum of two squares. This gives us another way to derive the identity in [Lemma 1.28.1](#), which you will now carry out:

Compute the determinant of $\begin{bmatrix} a & b \\ -b & a \end{bmatrix} \begin{bmatrix} c & d \\ -c & d \end{bmatrix}$ by multiplying the matrices and afterwards finding the determinant. Then use the property $\det XY = (\det X)(\det Y)$ of matrices to conclude [Lemma 1.28.1](#).

? Exercise 1.28.8

(For students who have seen matrices and determinants.) When we tweak the matrices used in the previous exercise, we get even more interesting results. Matrices of the form $\begin{bmatrix} a + bi & c + di \\ -c - di & a - bi \end{bmatrix}$, where $a, b, c, d \in \mathbb{Z}$, behave like elements of a number system known as the *quaternions*, where numbers have the form $a + bi + cj + dk$ for special symbols i, j, k that, like the imaginary number i , have certain special properties. A quick search of the internet will reveal details of computing with quaternions; do some reading on your own, if you wish. The quaternions can be used to prove [Lemma 1.28.5](#) in the same way the Gaussian integers were used to prove [Lemma 1.28.1](#). However, for simplicity(?), we will work with the matrix analogues of the quaternions, which involve only matrices with Gaussian integer entries.

Now the determinant of $\begin{bmatrix} a+bi & c+di \\ -c-di & a-bi \end{bmatrix}$ is $a^2+b^2+c^2+d^2$, a sum of four squares. We can use this in a similar manner as in the previous exercise.

Compute the determinant of $\begin{bmatrix} a_1+a_2i & a_3+a_4i \\ -a_3+a_4i & a_1-a_2i \end{bmatrix} \begin{bmatrix} b_1+b_2i & b_3+b_4i \\ -b_3+b_4i & b_1-b_2i \end{bmatrix}$ by multiplying the matrices and afterwards finding the determinant. Then use the property $\det XY = (\det X)(\det Y)$ of matrices to conclude [Lemma 1.28.5](#).

Footnotes

[1] This identity was included in a letter that Euler wrote to Goldbach in 1748. Like the proof of [Lemma 1.28.1](#), there is a way to arrive at this identity in an almost straightforward way by using a different set of numbers. Instead of Gaussian integers and complex conjugates, though, the numbers and concept we need here are the *quaternions* and quaternion conjugates. Look these up! Or for an alternative proof, see [Exercise 1.28.8](#).

This page titled [1.28: Sum of Squares](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Mike Barrus & W. Edwin Clark](#).

1.29: Epilogue

With our discussion of sums of squares in the last chapter, at this point we've answered the last of our questions from [Section 1.1](#). In this text we have seen some gems of number theory from throughout ancient and more recent history. Still, please *don't* let this text contain your final interactions with number theory!

Exciting discoveries are being made today that lead to the resolution of old conjectures and spur intriguing new questions of their own. If you wish to see other books, browse the shelves in a university library; many number theory books have a Library of Congress call number of QA 241 or something nearby. One may also obtain much interesting and current information about number theory online. See particularly the websites listed in the Bibliography. The websites by Chris Caldwell [\[2\]](#) and by Eric Weisstein [\[11\]](#) are especially recommended. To see what is going on at the frontier of research on the subject, you may take a look at some recent issues of the *Journal of Number Theory* and other number theory journals, which you may find through a university library.

You don't have to just read other people's work, though. Hopefully by this point you've developed some curiosity of your own about the integers. What questions occur to you that you'd like to see answered? Can you come up with any answers, or at least helpful first steps, on your own?

Where will you go next?

This page titled [1.29: Epilogue](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Mike Barrus & W. Edwin Clark](#).

CHAPTER OVERVIEW

2: Appendices

[2.1: Prime Numbers less than 200](#)

[2.2: Prime Numbers Up to 400](#)

[2.3: A Rings and Groups](#)

[2.4: Bibliography](#)

This page titled [2: Appendices](#) is shared under a [not declared](#) license and was authored, remixed, and/or curated by [Mike Barrus & W. Edwin Clark](#).

2.1: Prime Numbers less than 200

Organized by position

	2	3	5	7	
11		13		17	19
		23			29
31				37	
41		43		47	
		53			59
61				67	
71		73			79
		83			89
				97	
101		103		107	109
		113			
				127	
131				137	139
					149
151				157	
		163		167	
		173			179
181					
191		193		197	199

Figure 2.1.1

Organized by congruence class modulo 6

		2,3		
-1	+1		-1	+1
5	7		107	109
11	13		113	
17	19			
23				127
29	31		131	
	37		137	139
41	43			
47			149	151
53			155	157
59	61			163
	67		167	
71	73		173	
	79		179	181
83				
89			191	193
	97		197	199
101	103			

Figure 2.1.2

This page titled [2.1: Prime Numbers less than 200](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Mike Barrus & W. Edwin Clark](#).

2.2: Prime Numbers Up to 400

Listed

n	n^2	n	n^2
0	0	10	100
1	1	11	121
2	4	12	144
3	9	13	169
4	16	14	196
5	25	15	225
6	36	16	256
7	49	17	289
8	64	18	324
9	81	19	361
		20	400

Figure 2.2.1

Organized by position

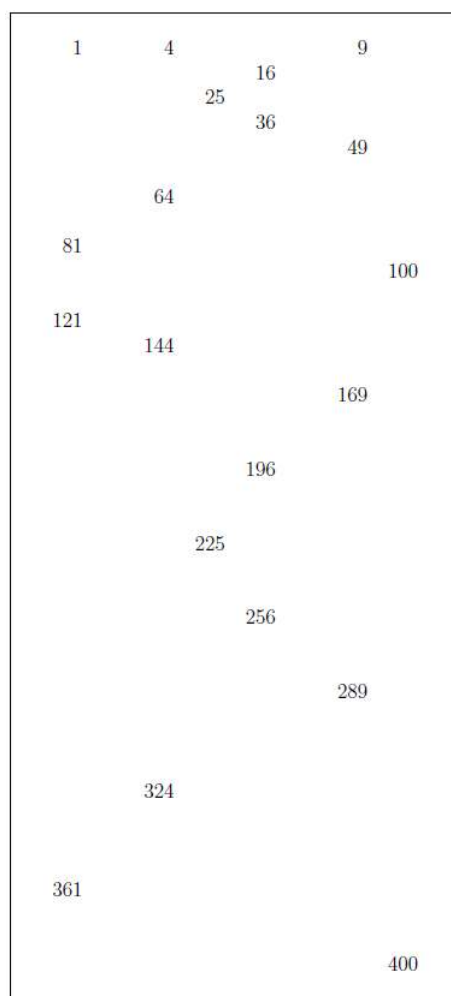


Figure 2.2.2

This page titled [2.2: Prime Numbers Up to 400](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Mike Barrus & W. Edwin Clark](#).

2.3: A Rings and Groups

The material in this appendix is optional reading. However, for the sake of completeness we state here the definition of a *ring* and the definition of a *group*. If you are interested in learning more you might take the course *Elementary Abstract Algebra*. Having had this course should make it a little easier to understand the ideas in abstract algebra and vice versa.

For more details you may download the free book *Elementary Abstract Algebra* from the webpage:

<http://www.math.usf.edu/~eclark>

Alternatively, look in almost any book whose title contains the words *Abstract Algebra* or *Modern Algebra*. Look for one with *Introductory* or *Elementary* in the title.

Definition 2.3.1

A *ring* is an ordered triple $(R, +, \cdot)$ where R is a set and $+$ and $\cdot$ are binary operations on R satisfying the following properties:

A1 $a + (b + c) = (a + b) + c$ for all a, b, c in R .

A2 $a + b = b + a$ for all a, b in R .

A3 There is an element $0 \in R$ satisfying $a + 0 = a$ for all a in R .

A4 For every $a \in R$ there is an element $b \in R$ such that $a + b = 0$.

M1 $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ for all a, b, c in R .

D1 $a \cdot (b + c) = a \cdot b + a \cdot c$ for all a, b, c in R .

D2 $(b + c) \cdot a = b \cdot a + c \cdot a$ for all a, b, c in R .

Thus, to describe a ring one must specify three things:

1. a set,
2. a binary operation on the set called multiplication,
3. a binary operation on the set called addition.

Then, one must verify that the properties above are satisfied.

Example 2.3.1

Here are some examples of rings. The two binary operations $+$ and $\cdot$ are in each case the ones that you are familiar with.

1. $(\mathbb{R}, +, \cdot)$ —the ring of real numbers.
2. $(\mathbb{Q}, +, \cdot)$ —the ring of rational numbers.
3. $(\mathbb{Z}, +, \cdot)$ —the ring of integers.
4. $(\mathbb{Z}_n, +, \cdot)$ —the ring of integers modulo n .
5. $(M_n(\mathbb{R}), +, \cdot)$ —the ring of all $n \times n$ matrices over $\mathbb{R}$.

Definition 2.3.2: Group

A **group** is an ordered pair $(G, *)$ where G is a set and $*$ is a binary operation on G satisfying the following properties

1. $x * (y * z) = (x * y) * z$ for all x, y, z in G .
2. There is an element $e \in G$ satisfying $e * x = x$ and $x * e = x$ for all x in G .
3. For each element x in G there is an element y in G satisfying $x * y = e$ and $y * x = e$.

 Definition 2.3.3

A group $(G, *)$ is said to be *Abelian* if $x * y = y * x$ for all $x, y \in G$.

Thus, to describe a group one must specify two things:

1. a set, and
2. a binary operation on the set.

Then, one must verify that the binary operation is associative, that there is an identity in the set, and that every element in the set has an inverse.

 Example 2.3.2

Here are some examples of groups. The binary operations are in each case the ones that you are familiar with.

1. $(\mathbb{Z}, +)$ is a group with identity 0. The inverse of $x \in \mathbb{Z}$ is $-x$.
2. $(\mathbb{Q}, +)$ is a group with identity 0. The inverse of $x \in \mathbb{Q}$ is $-x$.
3. $(\mathbb{R}, +)$ is a group with identity 0. The inverse of $x \in \mathbb{R}$ is $-x$.
4. $(\mathbb{Q} - \{0\}, \cdot)$ is a group with identity 1. The inverse of $x \in \mathbb{Q} - \{0\}$ is x^{-1} .
5. $(\mathbb{R} - \{0\}, \cdot)$ is a group with identity 1. The inverse of $x \in \mathbb{R} - \{0\}$ is x^{-1} .
6. $(\mathbb{Z}_n, +)$ is a group with identity 0. The inverse of $x \in \mathbb{Z}_n$ is $n - x$ if $x \neq 0$, the inverse of 0 is 0.
7. $(U_n, \cdot)$ is a group with identity $[1]$. The inverse of $[a] \in U_n$ was shown to exist in [Section 1.22](#).
8. $(\mathbb{R}^n, +)$ where $+$ is vector addition. The identity is the zero vector $(0, 0, \dots, 0)$ and the inverse of the vector $\mathbf{x} = (x_1, x_2, \dots, x_n)$ is the vector $-\mathbf{x} = (-x_1, -x_2, \dots, -x_n)$.
9. $(M_n(\mathbb{R}), +)$. This is the group of all $n \times n$ matrices over $\mathbb{R}$ and $+$ is matrix addition.

This page titled [2.3: A Rings and Groups](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Mike Barrus & W. Edwin Clark](#).

2.4: Bibliography

- [1] Tom Apostol, *Introduction to Analytic Number Theory*, Springer-Verlag, New York-Heidelberg, 1976.
 - [2] Chris Caldwell, *The Primes Pages*,
<http://primes.utm.edu>
 - [3] W. Edwin Clark, *Number Theory Links*,
http://www.math.usf.edu/~eclark/numtheory_links.html
 - [4] Earl Fife and Larry Husch, *Number Theory (Mathematics Archives)*,
<http://archives.math.utk.edu/topics/numberTheory.html>
 - [5] Ronald Graham, Donald Knuth, and Oren Patashnik, *Concrete Mathematics*, Addison-Wesley, 1994.
 - [6] Donald Knuth *The Art of Computer Programming*, Vols I and II, Addison-Wesley, 1997.
 - [7] J.J. O'Connor and E.F. Robertson, "Abu Ali al-Hasan ibn al-Haytham," *MacTutor History of Mathematics archive*, last updated November 1999. Entry accessed August 30, 2021 at <http://www-history.mcs.st-andrews.ac.uk/Biographies/Al-Haytham.html> .
 - [8] Oystein Ore, *Number Theory and its History*, Dover Publications, 1988.
 - [9] Carl Pomerance and Richard Crandall, *Prime Numbers – A Computational Perspective*, Springer -Verlag, 2001.
 - [10] Kenneth A. Rosen, *Elementary Number Theory*, (Fourth Edition), Addison-Wesley, 2000.
 - [11] Eric Weisstein, *World of Mathematics –Number Theory Section*,
<http://mathworld.wolfram.com/topics/NumberTheory.html>
-

2.4: Bibliography is shared under a [not declared](#) license and was authored, remixed, and/or curated by LibreTexts.

Index

B

Bézout's lemma

- [1.9: Bezout's Lemma](#)
- [1.10: Computing Coefficients for Bezout's Lemma](#)

Binary Method for Exponentiation

- [1.26: Computation of \$a^n \bmod m\$](#)

C

ceiling function

- [1.2: Basic Axioms for \$\mathbb{Z}\$](#)

Chinese remainder theorem

- [1.23: Chinese Remainder Theorem](#)

composite

- [1.11: Prime Numbers](#)

congruences

- [1.20: More Properties of Congruences](#)

D

division algorithm

- [1.5: The Division Algorithm](#)

E

Euclid's theorem

- [1.11: Prime Numbers](#)

Euclidean Algorithm

- [1.8: The Euclidean Algorithm](#)

Euler's theorems

- [1.24: Theorems of Wilson, Euler, and Fermat](#)

F

Fermat Primes

- [1.14: Fermat Primes and Mersenne Primes](#)

Fermat's Little Theorem

- [1.24: Theorems of Wilson, Euler, and Fermat](#)
- [1.25: Primality Tests](#)

floor

- [1.2: Basic Axioms for \$\mathbb{Z}\$](#)

Fundamental Theorem of Arithmetic

- [1.12: Unique Factorization](#)

G

Gaussian integers

- [1.13: The Gaussian Integers](#)

greatest common divisor

- [1.7: Greatest Common Divisor and Least Common Multiple](#)

I

induction

- [1.3: Proof by Induction](#)

induction hypothesis

- [1.3: Proof by Induction](#)

L

least common multiple

- [1.7: Greatest Common Divisor and Least Common Multiple](#)

M

Mersenne primes

- [1.14: Fermat Primes and Mersenne Primes](#)
- [1.16: Perfect Numbers and Mersenne Primes](#)

P

perfect numbers

- [1.16: Perfect Numbers and Mersenne Primes](#)

primality

- [1.11: Prime Numbers](#)

primality test

- [1.25: Primality Tests](#)

prime number theorem

- [1.11: Prime Numbers](#)

prime numbers

- [1.11: Prime Numbers](#)

proofs by mathematical induction

- [1.3: Proof by Induction](#)

S

sieve of Eratosthenes

- [1.11: Prime Numbers](#)

W

Wilson's theorem

- [1.24: Theorems of Wilson, Euler, and Fermat](#)
- [1.25: Primality Tests](#)

Glossary

Sample Word 1 | Sample Definition 1

Detailed Licensing

Overview

Title: [Elementary Number Theory \(Barrus and Clark\)](#)

Webpages: 46

Applicable Restrictions: Noncommercial

All licenses found:

- [CC BY-NC-SA 4.0](#): 76.1% (35 pages)
- [Undeclared](#): 23.9% (11 pages)

By Page

- [Elementary Number Theory \(Barrus and Clark\) - CC BY-NC-SA 4.0](#)
 - [Front Matter - CC BY-NC-SA 4.0](#)
 - [TitlePage - Undeclared](#)
 - [InfoPage - CC BY-NC-SA 4.0](#)
 - [Table of Contents - Undeclared](#)
 - [Licensing - Undeclared](#)
 - [Preface - CC BY-NC-SA 4.0](#)
 - [1: Chapters - CC BY-NC-SA 4.0](#)
 - [1.1: What Is Number Theory? - Undeclared](#)
 - [1.2: Basic Axioms for \$\mathbb{Z}\$ - CC BY-NC-SA 4.0](#)
 - [1.3: Proof by Induction - CC BY-NC-SA 4.0](#)
 - [1.4: Elementary Divisibility Properties - CC BY-NC-SA 4.0](#)
 - [1.5: The Division Algorithm - CC BY-NC-SA 4.0](#)
 - [1.6: The Base \$b\$ Representation of \$n\$ - CC BY-NC-SA 4.0](#)
 - [1.7: Greatest Common Divisor and Least Common Multiple - CC BY-NC-SA 4.0](#)
 - [1.8: The Euclidean Algorithm - CC BY-NC-SA 4.0](#)
 - [1.9: Bezout's Lemma - Undeclared](#)
 - [1.10: Computing Coefficients for Bezout's Lemma - CC BY-NC-SA 4.0](#)
 - [1.11: Prime Numbers - CC BY-NC-SA 4.0](#)
 - [1.12: Unique Factorization - CC BY-NC-SA 4.0](#)
 - [1.13: The Gaussian Integers - Undeclared](#)
 - [1.14: Fermat Primes and Mersenne Primes - CC BY-NC-SA 4.0](#)
 - [1.15: Number Theoretic Functions - Undeclared](#)
 - [1.16: Perfect Numbers and Mersenne Primes - CC BY-NC-SA 4.0](#)
 - [1.17: Congruences - CC BY-NC-SA 4.0](#)
 - [1.18: Divisibility Tests for 2, 3, 5, 9, 11 - CC BY-NC-SA 4.0](#)
 - [1.19: Divisibility Tests for 7 and 13 - CC BY-NC-SA 4.0](#)
 - [1.20: More Properties of Congruences - CC BY-NC-SA 4.0](#)
 - [1.21: Residue Classes and the Integers Modulo \$m\$ - CC BY-NC-SA 4.0](#)
 - [1.22: The Groups \$U_m\$ - CC BY-NC-SA 4.0](#)
 - [1.23: Chinese Remainder Theorem - Undeclared](#)
 - [1.24: Theorems of Wilson, Euler, and Fermat - CC BY-NC-SA 4.0](#)
 - [1.25: Primality Tests - CC BY-NC-SA 4.0](#)
 - [1.26: Computation of \$a^n \bmod m\$ - CC BY-NC-SA 4.0](#)
 - [1.27: The RSA Scheme - CC BY-NC-SA 4.0](#)
 - [1.28: Sum of Squares - CC BY-NC-SA 4.0](#)
 - [1.29: Epilogue - CC BY-NC-SA 4.0](#)
 - [2: Appendices - Undeclared](#)
 - [2.1: Prime Numbers less than 200 - CC BY-NC-SA 4.0](#)
 - [2.2: Prime Numbers Up to 400 - CC BY-NC-SA 4.0](#)
 - [2.3: A Rings and Groups - CC BY-NC-SA 4.0](#)
 - [2.4: Bibliography - Undeclared](#)
 - [Back Matter - CC BY-NC-SA 4.0](#)
 - [Index - CC BY-NC-SA 4.0](#)
 - [Glossary - CC BY-NC-SA 4.0](#)
 - [Detailed Licensing - Undeclared](#)