



Pure and Applied
UNDERGRADUATE TEXTS

31

An Experimental Introduction to Number Theory

Benjamin Hutz



AMERICAN
MATHEMATICAL
SOCIETY

An
Experimental
Introduction
to Number
Theory



Pure and Applied
UNDERGRADUATE TEXTS • 31

An Experimental Introduction to Number Theory

Benjamin Hutz



AMS
AMERICAN
MATHEMATICAL
SOCIETY

Providence, Rhode Island USA

EDITORIAL COMMITTEE

Gerald B. Folland (Chair) Steven J. Miller
Jamie Pommersheim Serge Tabachnikov

2010 *Mathematics Subject Classification*. Primary 11-01, 11Axx, 11Dxx, 11G05, 11K60, 11T06, 35P05.

For additional information and updates on this book, visit
www.ams.org/bookpages/amstext-31

Library of Congress Cataloging-in-Publication Data

Names: Hutz, Benjamin, 1978– author.

Title: An experimental introduction to number theory / Benjamin Hutz.

Description: Providence, Rhode Island : American Mathematical Society, [2018] | Series: Pure and applied undergraduate texts ; volume 31 | Includes bibliographical references and index.

Identifiers: LCCN 2017036105 | ISBN 9781470430979 (alk. paper)

Subjects: LCSH: Number theory--Textbooks. | AMS: Number theory – Instructional exposition (textbooks, tutorial papers, etc.). msc | Number theory – Elementary number theory – Elementary number theory. msc | Number theory – Diophantine equations – Diophantine equations. msc | Number theory – Arithmetic algebraic geometry (Diophantine geometry) – Elliptic curves over global fields. msc | Number theory – Probabilistic theory: distribution modulo 1; metric theory of algorithms – Diophantine approximation. msc | Number theory – Finite fields and commutative rings (number-theoretic aspects) – Polynomials. msc | Dynamical systems and ergodic theory – Arithmetic and non-Archimedean dynamical systems – Polynomial and rational maps. msc

Classification: LCC QA241 .H88 2018 | DDC 512.7–dc23

LC record available at <https://lcn.loc.gov/2017036105>

Copying and reprinting. Individual readers of this publication, and nonprofit libraries acting for them, are permitted to make fair use of the material, such as to copy select pages for use in teaching or research. Permission is granted to quote brief passages from this publication in reviews, provided the customary acknowledgment of the source is given.

Republication, systematic copying, or multiple reproduction of any material in this publication is permitted only under license from the American Mathematical Society. Permissions to reuse portions of AMS publication content are handled by Copyright Clearance Center's RightsLink® service. For more information, please visit: <http://www.ams.org/rightslink>.

Send requests for translation rights and licensed reprints to reprint-permission@ams.org.

Excluded from these provisions is material for which the author holds copyright. In such cases, requests for permission to reuse or reprint material should be addressed directly to the author(s). Copyright ownership is indicated on the copyright page, or on the lower right-hand corner of the first page of each article within proceedings volumes.

© 2018 by the American Mathematical Society. All rights reserved.

The American Mathematical Society retains all rights
except those granted to the United States Government.

Printed in the United States of America.

∞ The paper used in this book is acid-free and falls within the guidelines
established to ensure permanence and durability.

Visit the AMS home page at <http://www.ams.org/>

10 9 8 7 6 5 4 3 2 1 23 22 21 20 19 18

Contents

Preface	ix
Introduction	1
Chapter 1. Integers	5
1. The Integers and the Well Ordering Property	5
2. Divisors and the Division Algorithm	6
3. Greatest Common Divisor and the Euclidean Algorithm	10
4. Prime Numbers and Unique Factorization	20
Exercises	29
Chapter 2. Modular Arithmetic	39
1. Basic Arithmetic	39
2. Inverses and Fermat's Little Theorem	44
3. Linear Congruences and the Chinese Remainder Theorem	51
Exercises	58
Chapter 3. Quadratic Reciprocity and Primitive Roots	65
1. Quadratic Reciprocity	65
2. Computing m th Roots Modulo n	76
3. Existence of Primitive Roots	81
Exercises	86
Chapter 4. Secrets	91
1. Basic Ciphers	92
2. Symmetric Ciphers	95
3. Diffie–Hellman Key Exchange	97
4. Public Key Cryptography (RSA)	98


5. Hash Functions and Check Digits	101
6. Secret Sharing	104
Exercises	105
Chapter 5. Arithmetic Functions	109
1. Euler Totient Function	109
2. Möbius Function	113
3. Functions on Divisors	121
4. Partitions	130
Exercises	134
Chapter 6. Algebraic Numbers	143
1. Algebraic or Transcendental	143
2. Quadratic Number Fields and Norms	145
3. Integers, Divisibility, Primes, and Irreducibles	148
4. Application: Sums of Two Squares	152
Exercises	154
Chapter 7. Rational and Irrational Numbers	157
1. Diophantine Approximation	157
2. Height of a Rational Number	159
3. Heights and Approximations	162
4. Continued Fractions	166
5. Approximating Irrational Numbers with Convergents	171
Exercises	181
Chapter 8. Diophantine Equations	187
1. Introduction and Examples	187
2. Working Modulo Primes	189
3. Pythagorean Triples	198
4. Fermat's Last Theorem	200
5. Pell's Equation and Fundamental Units	202
6. Waring Problem	208
Exercises	213
Chapter 9. Elliptic Curves	221
1. Introduction	221
2. Addition of Points	224
3. Points of Finite Order	229
4. Integer Points and the Nagel–Lutz Theorem	230
5. Mordell–Weil Group and Points of Infinite Order	236
6. Application: Congruent Numbers	237

Exercises	240
Chapter 10. Dynamical Systems	247
1. Discrete Dynamical Systems	247
2. Dynatomic Polynomials	254
3. Resultant and Reduction Modulo Primes	258
4. Periods Modulo Primes	262
5. Algorithms for Rational Periodic and Preperiodic Points	266
Exercises	269
Chapter 11. Polynomials	275
1. Introduction to Polynomials	275
2. Factorization and the Euclidean Algorithm	278
3. Modular Arithmetic for Polynomials	282
4. Diophantine Equations for Polynomials	288
Exercises	294
Bibliography	299
List of Algorithms	303
List of Notation	305
Index	307

Preface

Note to the Instructor

This book presents material suitable for an undergraduate course in elementary number theory from a computational perspective. It seeks to not only introduce students to the standard topics in elementary number theory, but also to the formulation of conjectures from experimental data. Each topic is motivated by a question to be answered, followed by some experimental data and, eventually, a statement and proof of a theorem. There are numerous opportunities throughout the chapters and exercises for the students to engage in (guided) open-ended exploration. The goal of this exploration is for students to engage with examples, deepen understanding, notice patterns, and formulate conjectures. To be effective as a learning mechanism, it is important that these explorations culminate in clear mathematical statements of what has been discovered. It is my hope that at the end of their course the students will understand how mathematics is developed from asking questions, to gathering data, to formulating and proving theorems.

There is a heavy emphasis on computation throughout the book, including several investigations within each chapter guiding the student through experimental investigations related to the material. At the end of each chapter, the exercises are divided into three sections: computational exercises, theoretical (proof-based) exercises, and explorations. The computational exercises range from simple calculations to more difficult algorithms, with an emphasis on working with explicit examples of the concepts described in the chapter. The ones marked with  are meant to be done with paper and pencil, the rest with a computer algebra system. The theoretical exercises are more like the “standard” exercises you would see in any number theory textbook, where students are asked to prove some additional concepts related to those in the chapter. As with the computational exercises, the difficulty varies, but the emphasis is on developing rigorous proofs of the statements. The explorations are meant to be treated as open-ended projects, where students put into practice the following methodology:

- (a) Ask a question.
- (b) Generate data.
- (c) Formulate conjectures.
- (d) Test your conjectures.
- (e) Try to prove your conjectures.

Many of the exploration topics are areas of current research. However, the student is provided with guiding questions that ensure they will be able to make progress on a subset of the problem. These explorations also serve as mathematical writing exercises, where the student must describe the problem, their steps toward gathering data, and their conclusions in the form of conjectures.

The mathematical prerequisites for this book are few. A basic understanding of functions from first semester calculus is sufficient. However, it is assumed that students have some familiarity with proof techniques. The level of the book is geared toward a junior-level mathematics student, but it could easily be adapted to a lower or higher level. Modest experience with a computer algebra system would be helpful, but a willingness to learn to use one is all that is required. What is most important is to approach this book in the proper frame of mind. This is a subject for exploration and experimentation. Students use the computer algebra system to gather data from which to formulate new conjectures. The definitions and main theorems are presented and proven, but the more the reader wanders down interesting side paths, the more he or she will get out of this book. The exercises present many such side paths, especially the in-chapter investigations and end-of-chapter exploration exercises.

The book is not tied into any one computer algebra system, and there are several freely available. The systems SageMath and PARI/GP are two excellent freely available systems. Similarly, Mathematica, Maple, or other commercial software could also be used. Because there are many excellent tutorials for each of these systems freely available, they will not be presented in this book.

Organization

This book contains more than can be typically covered in a standard semester. The typical material of a first semester number theory course is covered in Chapters 1–8. The remaining chapters, 9–11, represent more specialized topics. Various sections can be omitted from the core chapters to allow more time for other topics. For example, Chapter 4 and Sections 3.2, 3.3, 5.3, and 5.4 are used only lightly, if at all, in later chapters.

Chapters 1 and 2 cover the standard topics in divisibility and modular arithmetic and are prerequisites for every other chapter. Chapter 3 covers Quadratic Reciprocity and primitive roots. Chapter 4 is a brief side journey into cryptography. Chapter 5 investigates a few arithmetic functions, while Chapters 6 and 7 cover height functions, partial fractions, algebraic numbers, and Diophantine approximation. Chapter 8 is an introduction to solving Diophantine equations by examining several standard Diophantine problems. Chapter 9 treats the specific Diophantine equations that are called elliptic curves, with an emphasis on points of finite order.

Chapter 10 examines dynamical systems from a number theoretic perspective by studying rational preperiodic points for iterated systems. Finally, Chapter 11 introduces the notion of number theory with polynomials, that is, an introduction to number theory on function fields.

The following table gives an idea of the dependencies.

Chapter/Section	Dependencies
Chapter 1	–
Chapter 2	Chapter 1
Chapter 3	Chapters 1 and 2
Chapter 4	Chapters 1 and 2
Chapter 5	Chapter 1
Chapter 6	Chapter 1
Section 6.4	Additionally Chapter 2 and Section 3.1
Chapter 7	Chapters 1 and 6
Chapter 8	Chapters 1, 2, and Section 3.1
Section 8.5	Additionally Section 7.4
Chapter 9	Chapters 1 and 2
Section 9.6	Additionally Section 8.3
Chapter 10	Chapters 1, 2, Section 5.2
Chapter 11	Chapters 1, 2, Section 5.1, Chapter 6
Section 11.4	Additionally Chapter 8

Acknowledgments

As with any work of this sort, the author owes a great debt to those who came before. The written sources consulted can be found in the references. Additionally, there are many people over the course of many years that fostered my interest in number theory and computation, from my first number theory course as a freshman at Duke University taught by William Pardon, to my PhD advisor Joseph Silverman at Brown University, and all the professors and textbook authors in between. Particular thanks to Joe for always offering astute advice on everything from mathematics to publishers and for being a great role model by writing such excellent books.

There are many people who read early versions of the text and provided feedback. Stephen Kennedy provided detailed feedback on an early version and several helpful discussions about publishing in general. Steven J. Miller and his number theory class at Williams College provided helpful feedback and identified several errors in the text. The spring 2013 number theory class at the Florida Institute of Technology and the spring 2017 number theory class at Saint Louis University also showed remarkable patience using an early version of this text and providing useful feedback.

Special thanks goes to my mom Linda Pesante for bravely reading every page despite having to treat much of the technical material as a foreign language. Without her, grammatical errors and simply bad writing would be much more prevalent in text.

Benjamin Hutz
July 2017

Introduction

Number theory at its most basic is the study of properties of numbers. These properties can be as simple as even and odd numbers or the more complicated amicable pairs or algebraic integers. The following types of numbers represent a small fraction of those studied in number theory. Each of these types of numbers is studied because they satisfy some kind of special condition.

Even Numbers: $\dots, -4, -2, 0, 2, 4, 6, 8, \dots$

Prime Numbers: $2, 3, 5, 7, 11, \dots$

Square Numbers: $1, 4, 9, 16, 25, \dots$

Fibonacci Numbers: $0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, \dots$

Perfect Numbers: $6, 28, 496, 8128, 33550336, 8589869056, \dots$

Abundant Numbers: $12, 18, 20, 24, 30, 36, 40, 42, \dots$

Triangular Numbers: $1, 3, 6, 10, 15, 21, 28, 36, 45, 55, 66, 78, 91, 105, \dots$

Carmichael Numbers: $561, 1105, 1729, 2465, 2821, 6601, 8911, \dots$

Several early civilizations (~ 2000 B.C.E.) had elements of number theory in their mathematics. From the early Babylonians survives a list of Pythagorean triples. From the ancient Egyptians survive lists of problems, mostly computational. From the ancient Chinese mathematicians survive a great deal of arithmetic, including decimals and negative numbers, and many problems in Diophantine equations.

One of the perpetual allures of number theory is that it takes little more than basic arithmetic to understand and state interesting mathematical questions, but often those interesting questions require sophisticated mathematics to resolve. A famous example is Fermat's Last Theorem, conjectured by Pierre de Fermat¹ in 1637.

¹Pierre de Fermat (1601–1665) was a French lawyer and mathematician.

Theorem (Fermat’s Last Theorem). *For $n \geq 3$, there are no positive integers x, y, z that satisfy*

$$x^n + y^n = z^n.$$

This conjecture was found written in the margin of Fermat’s copy of Diophantus’s *Arithmetica*. It was not until 1995, 358 years later, that Andrew Wiles² gave a proof, a *very* complicated proof.³

Someone new to mathematics may be mystified as to how Fermat may have come up with such a statement. While one must talk to Fermat to know exactly why he was considering this problem, we can speculate. If you study right triangles, you notice that it is possible that the lengths of all three sides are integers, such as (3, 4, 5) and (5, 12, 13). Writing down a few such integer-sided right triangles, you notice that it seems that the side lengths (x, y, z) always satisfy $x^2 + y^2 = z^2$. This fact is precisely the Pythagorean theorem. Now, you can ask yourself, Are there other such relations? Two ways to generalize the Pythagorean theorem is either to increase the number of variables

$$x_1^2 + x_2^2 + \cdots + x_n^2 = z^2$$

or to increase the exponent

$$x^n + y^n = z^n.$$

Increasing the number of variables is a special case of the Waring problem, which we will discuss in Chapter 8. Increasing the exponent gives the problem studied by Fermat in Fermat’s Last Theorem. Having formulated the new question, you could now try several things. One is to try to find a solution for some particular values of n . If that works, then you know there are solutions. If that does not work for some n , then you may try to prove that there are no solutions for that value of n . Fermat was able to show there are no solutions for $n = 4$, whose proof we will see in Chapter 8. Sophie Germain⁴ proved there are no solutions when n is a prime less than 100, and Ernst Kummer⁵ proved it for a certain infinite class of primes. Eventually, Andrew Wiles was able to prove the theorem for all n . This developing process will occur often throughout this book:

- (a) experimentation to find a pattern or conjecture,
- (b) precise mathematical statement of the conjecture,
- (c) proving special cases of the conjecture,
- (d) full proof of the conjecture,
- (e) further generalization.

This process is not usually evident to a student since mathematics is most often presented as well-formed definition, theorem, and proof. However, it is rare that this fully developed mathematics jumps straight onto the page out of a mathematician’s head. Throughout this book the emphasis is on the earlier stages: the computation of examples and data leading to formulation of precise conjectures. While it is

²Sir Andrew John Wiles (1953–) is a British mathematician.

³There are many excellent references on the development and eventual proof of Fermat’s Last Theorem; see [51] for a general overview or [53] for a more in-depth treatment of the mathematics.

⁴Marie-Sophie Germain (1776–1831) was a French mathematician.

⁵Ernst Eduard Kummer (1810–1893) was a German mathematician.

important to be able to construct rigorous proofs of mathematical facts (and we will do plenty of that), it is also important to be able to develop your own mathematical questions, theories, and conjectures. Accordingly, there is a heavy emphasis on computation of both examples and experimental data in this book.

To solve a computational problem, we determine an *algorithm*. An algorithm is a finite list of performable steps that is guaranteed to produce an answer. For example, Algorithm 0.1 is an algorithm to find all the divisors of a given positive integer n .

Algorithm 0.1. Trial Division

Input: a positive integer $n \geq 1$

Output: a list of divisors

Algorithm:

1: Set $d = 1$.

2: Repeat until $d > \sqrt{n}$.

 a: If d divides n , append d to the list of divisors.

 b: Increase d by one.

3: Return the list of divisors.

This is called “trial division”, and we are attempting to divide n by the numbers $\{1, 2, 3, \dots, \sqrt{n}\}$. If any such division succeeds, we have found a divisor. This is an algorithm because it requires a finite number of steps and conclusively answers the question. If n is very large, say 10^{40} , then even though the number of steps is finite, the algorithm will not terminate in a *reasonable* amount of time. For example, if you could try $1,000,000,000 = 10^9$ divisions per second, then it would take more than 3,000 years to try all values up to $\sqrt{n} = 10^{20}$. So while this is an algorithm, that does not necessarily mean that solving the problem is feasible. This distinction between the existence of an algorithm and the existence of a *feasible* algorithm is extremely important. For example, for certain classes of cryptographic systems (such as public key cryptography) the security is based on the fact that the algorithms for factoring large integers are infeasible. The development of efficient algorithms often requires sophisticated mathematics and clever ingenuity.

We will spend very little time in this book examining the feasibility of algorithms, but it will often come up in practice. For example, if you are trying to generate data on a particular phenomenon, the more data you can gather, the better. Thus, the more efficient and sophisticated your algorithms are, the better data you will be able to gather. You would be surprised at how often conjectures are found to be false by carrying the computations a little further. Fermat numbers, discussed in Exploration Exercise 1.35, are a well-known example in which Fermat conjectured that certain numbers were prime based on minimal data, but he turned out to be wrong.

It may seem surprising, but there are many instances for which we can prove an algorithm does not exist to solve the problem. A famous example is Hilbert’s Tenth Problem, one of the 23 problems the mathematician David Hilbert⁶ stated in

⁶David Hilbert (1862–1943) was a German mathematician.

his famous 1900 lecture at the International Congress of Mathematicians in Paris.⁷ The following is an English translation of the tenth problem.

Problem. Given a Diophantine equation with any number of unknown quantities and with rational integral numerical coefficients, devise a process according to which it can be determined by a finite number of operations whether the equation is solvable in rational integers.

In more modern terminology, Hilbert asks for a way to determine if a polynomial equation with integer coefficients has an integer solution. In other words, given a polynomial $f(x_1, x_2, \dots, x_n)$ with integer coefficients, can you determine if there exists (a_1, \dots, a_n) with all the a_i integers such that

$$f(a_1, \dots, a_n) = 0?$$

The proof that such an algorithm does not exist was completed in 1970, building on the work of many mathematicians; see [31].

The idea of gathering data to provide insight into difficult problems is often called *experimental mathematics*. While it often lacks the rigor and elegance of pure mathematical theorems, it is nevertheless an interesting and exciting field. There is no doubt that working through some explicit examples is a great way to understand abstract theorems; experimental mathematics goes even further. In some sense you could say you are working to try to find the theorems. With the ever-increasing power of modern computers, there are some fairly amazing “theorems” that resulted from experimental computation. A good example is an infinite sum formula for π discovered in 1995 by Bailey, Borwein, and Plouffe through the PSQ algorithm, which can find linear relations between real numbers. Especially interesting is that the formula can be used to calculate digits of π from any starting point

$$\pi = \sum_{k=0}^{\infty} \frac{1}{16^k} \left(\frac{4}{8k-1} - \frac{2}{8k+4} - \frac{1}{8k+5} - \frac{1}{8k+6} \right).$$

Having “discovered” the formula through computation, they were then able to prove its validity.

In summary, we mention just a few ways experimental mathematics can be used:

- providing insight into hard problems through computational examples;
- discovering new insights by
 - quickly exploring many dead ends,
 - visualizing patterns through graphs and tables,
 - generating a **lot** of data quickly;
- determining what is true and what merits proof, including
 - Gödel famously showed not everything that is true can be proven,
 - my PhD advisor once advised me to prove only the things that are true;
- finding counterexamples and disproving conjectures
 - a well-known adage in mathematics is to try to prove your conjectures by day and search for counterexamples by night.

⁷See German original [20] or English translation [21].

Integers

1. The Integers and the Well Ordering Property

Consider the set of positive integers known as the *natural numbers* $\mathbb{N} = \{1, 2, 3, \dots\}$. These numbers represent the fundamental building blocks of arithmetic and are the numbers you encounter when you count. If you add or multiply two natural numbers, you again have a natural number. However, subtraction may result in negative values; if you have \$3 and spent \$5, how much money do you have? The answer is that you still owe \$2. To represent the concept of “owing”, we introduce negative numbers. Then we can represent owing \$2 as having $-\$2$. We denote the set of all integers as $\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$, which additionally includes both 0 and negative integers. Extending the natural numbers to the integers allows for addition, multiplication, and subtraction. We will take as axioms many of the basic properties of arithmetic, such as

- (a) *Commutativity*: $a + b = b + a$ and $ab = ba$;
- (b) *Associativity*: $a + (b + c) = (a + b) + c$ and $(ab)c = a(bc)$;
- (c) *Distribution*: $a(b + c) = ab + ac$.

We have not discussed the operation of division. Some fractions, such as $\frac{4}{2}$, are actually integers, but many fractions, such as $\frac{1}{2}$, are not integers. Consequently, we must be very careful with the operation of division for integers. In the next section we will carefully define the notion of divisibility. However, we first discuss one more property of the integers called the well ordering property.

Well Ordering Property. Every nonempty set of nonnegative integers has a least element.

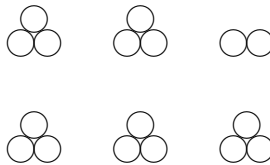
Example 1.1.

- The finite set $\{2, 5, 7, 11\}$ has least element 2.
- The infinite set $\{3n : n \in \mathbb{N}\} = \{3, 6, 9, \dots\}$ has least element 3.

The well ordering property may seem obvious, but it represents a powerful tool for proving theorems in number theory.

2. Divisors and the Division Algorithm

We approach the notion of divisibility with the concrete notion of splitting up a quantity of objects into equal sized groups. When we can split a quantity of objects into two equal halves, we say that the quantity is divisible by 2. This is only possible if we start with an even number of objects. If we started with an odd number of objects, we could split them into two equal groups with 1 left over. In fact, we can always divide a quantity into equal sized groups with some number left over. For example, we can divide 17 objects into five groups of 3 with 2 left over



or two groups of 8 with 1 left over



Intuitively, if we divide the quantity into n equal size groups, then we can have at most $n - 1$ left over; otherwise, we would be able to add one more to each group. We will formalize this idea with the *division algorithm*, but first we define the notion of divisibility.

Definition 1.2. Let n be an integer. We say that d *divides* n or d is a *divisor* of n if there exists an integer m such that $n = dm$. When d divides n , we denote this as $d \mid n$. If d does not divide n , we denote this as $d \nmid n$.

To say that $d \mid n$, is to say that we can divide n into equal groups of size d with none left over; in other words, $\frac{n}{d}$ is an integer.

Notice that with this definition, 1 divides every integer.

Example 1.3. We say that 4 divides 12 since we can take $m = 3$ in Definition 1.2,

$$12 = 4 \cdot 3.$$

However, 3 does not divide 11 since 11 is not a multiple of 3.

Investigation 1.4 (Number of divisors). An interesting question we may ask is how many positive divisors a given integer has. For example, 2 has two divisors $\{1, 2\}$ and 4 has three divisors $\{1, 2, 4\}$.

- (a) Consider the number of divisors of the integers between 1 and 30. Do any of these values stand out?

- (b) Partition the integers 1 to 30 by their number of divisors. Can you say anything about the form of the numbers which have the same number of divisors? For example, which numbers have exactly 2 divisors? 3 divisors? etc.

You may have noticed in the previous investigation that there are certain integers that have very few divisors. Every integer is divisible by 1 and itself, so every integer (greater than 1) has at least two divisors. The integers with exactly two divisors we call *prime numbers*.

Definition 1.5. A natural number $n > 1$ is *prime* if its only (positive) divisors are 1 and n . Otherwise, we say that n is *composite*.

Example 1.6. We see that 6 is composite since we can write $1 \cdot 6 = 2 \cdot 3 = 6$ so that 1, 2, 3, and 6 all divide 6. However, 5 is prime since only 1 and 5 divide 5. The first few prime numbers are 2, 3, 5, 7, 11, 13, 17, and 19.

Prime numbers play an essential role in number theory, and they will occur throughout the book. We will return to prime numbers later in this chapter, but now we return to the notion of divisibility.

Proposition 1.7. Let a , b , and n be integers.

- (a) If n divides a and b , then n divides $a + b$.
- (b) If $n \mid a$, then $n \mid ab$.
- (c) If $n \mid a$ and $a \mid b$, then $n \mid b$.
- (d) If $a \mid b$ and $b \mid a$, then $a = \pm b$.

If you are new to mathematical proofs, take a moment to read the following proofs carefully. Notice that what we are doing is using Definition 1.2 to prove divisibility. In other words, we need to show the existence of an integer m with the right properties.

Proof.

- (a) We assume that $n \mid a$ so, by the definition of divisibility, there exists an integer c such that $a = nc$. We also assume that $n \mid b$, so that there exists an integer d such that $b = nd$. So we can write

$$a + b = nc + nd = n(c + d).$$

Thus, we have found an integer $m = c + d$ such that $a + b = nm$. This satisfies the definition of divisibility, so we conclude that $n \mid a + b$.

- (b) We assume $n \mid a$, so there exists an integer c such that $a = nc$. So we may write

$$ab = (nc)b = n(cb).$$

Since cb is an integer, $n \mid ab$.

- (c) We assume $n \mid a$, so there exists an integer c such that $a = nc$. We also assume that $a \mid b$, so there exists an integer d such that $b = ad$. So we may write

$$b = ad = (nc)d = n(cd).$$

Since cd is an integer, $n \mid b$.

- (d) By assumption there exist integers k and ℓ such that $a = kb$ and $b = \ell a$. We have

$$a = k(\ell a) = k\ell a.$$

So we must have $k\ell = 1$; k and ℓ are either both 1 or both -1 . \square

Remark. Note that part (a) of Proposition 1.7 can be used in reverse as follows. If $n \mid a$ and $n \nmid b$, then $n \nmid a + b$. For example, $2 \mid 4$ and $2 \nmid 3$, thus, $2 \nmid 7$ since $7 = 4 + 3$.

Going back to dividing a collection of objects into equal groups, the division algorithm gives the formal statement that the number of groups and the number left over are unique.

Theorem 1.8 (Division algorithm). *If d, n are positive integers, then there are unique integers q, r such that $n = qd + r$ with $0 \leq r < d$.*

Notice that there are actually two statements in this theorem: first there exist integers q and r with the desired properties; second, they are unique.

Example 1.9. When dividing 7 by 2, we see

$$7 = 3 \cdot 2 + 1.$$

In other words, if we have seven objects ($n = 7$) and we divide them into groups of size 2 ($d = 2$), then we have three groups ($q = 3$) with 1 remaining ($r = 1$).

Proof of Theorem 1.8. We first prove the existence of q and r . If $d > n$, then $q = 0$ and $r = n$ satisfy the desired properties. If $d \leq n$, then consider the numbers

$$\{n, n - d, n - 2d, n - 3d, \dots\}.$$

This is a decreasing sequence of integers. Since there are only finitely many non-negative integers at most n , this list has a minimal nonnegative element (the well ordering property). Let $n - qd$ be the minimal nonnegative element, and truncate the list to include only the nonnegative values

$$(1) \quad \{n, n - d, n - 2d, n - 3d, \dots, n - qd\}.$$

Now, set r to be the minimal nonnegative element $r = n - qd$, and we will show that $0 \leq r < d$. It is clear that r is at least 0 because we chose it as the minimal nonnegative element of the list (1). Now, if $r \geq d$, then $r - d > 0$ and the number

$$r - d = (n - qd) - d = n - (q + 1)d$$

must also be in list (1). But r was chosen to be the smallest nonnegative value, so we cannot have $r - d$ in the list. Thus, we must have $r - d < 0$, which is the same as $r < d$.

Now we show q and r are unique. Assume there are two distinct pairs (q, r) and (q', r') , both of which satisfy the necessary properties. Then we have

$$\begin{aligned} n &= qd + r, \\ n &= q'd + r'. \end{aligned}$$

By subtracting the two previous equations, we have

$$0 = (q - q')d + (r - r').$$

In particular,

$$(2) \quad (r' - r) = d(q - q'),$$

and so $d \mid (r' - r)$. But since $0 \leq r, r' < d$ we must also have

$$(3) \quad -d < r' - r < d.$$

Since the inequalities in equation (3) are strict, the only way we have d dividing $r' - r$ is if $r' - r = 0$. Therefore, by equation (2) we also have $q - q' = 0$. The equalities

$$r' - r = 0 \quad \text{and} \quad q - q' = 0$$

mean that the two pairs (q, r) and (q', r') are the same. \square

Definition 1.10. The integer q is called the *quotient*, and the integer r is called the *remainder*.

Remark. The remainder is 0 if and only if $d \mid n$.

The proof of Theorem 1.8 is a constructive proof. Not only does it prove the existence of a unique q and r , but it does so by finding their values. Therefore, it is not too hard to turn the proof into an algorithm. Algorithm 1.1 describes an algorithm to determine the unique q and r in the division algorithm. Algorithm 1.2 uses the division algorithm as a test for divisibility.

Algorithm 1.1. Division Algorithm

Input: two nonnegative numbers m, n

Output: the quotient and remainder q, r

Algorithm:

- 1: If $m \leq 0$, return error.
 - 2: If $m > n$.
 - a: Return $q = 0$ and $r = n$.
 - 3: Set $q = 1$.
 - 4: While $n - qm \geq m$.
 - a: $q = q + 1$.
 - 5: Return $(q, n - qm)$.
-

Algorithm 1.2. Divisibility Test

Input: two nonnegative numbers n, d

Output: True if $d \mid n$, False if $d \nmid n$

Algorithm:

- 1: If $d = 0$, return error.
 - 2: Find (q, r) from the division algorithm applied to (n, d) .
 - 3: If $r = 0$, return True. Otherwise, return False.
-

3. Greatest Common Divisor and the Euclidean Algorithm

We saw in the previous section what it means for an integer d to divide an integer n and how you could implement a test for divisibility. We start this section with a question that at first glance seems unrelated.

Question 1.11. You are given two jars, one that holds 3 liters and one that holds 5 liters. Is it possible to measure out 1 liter of water? More generally, what quantities is it possible to measure with these two jars?

We could answer the first question by trial and error and eventually see that if we add water with the 3-liter jar twice and remove water with the 5-liter jar once, we get 1 liter,

$$(4) \qquad 3 \cdot 2 - 5 = 1.$$

However, the second question is not about a single specific quantity, so it requires more than just experimentation. Let's think about what we can do with the two jars. We can fill either jar with water and add it to the amount, or we can remove the amount of either jar. The final result is some combination of 3's and 5's of the form

$$3a + 5b$$

for some integers a and b . The integer a represents how many times the quantity of the 3-liter jar was added and b the 5-liter jar, where negative represents removing that amount. So by taking multiples of equation (4), we see that we can get any integer number of liters

$$n(3 \cdot 2 - 5) = 3 \cdot 2n - 5 \cdot n = n.$$

In other words, by adding water from $2n$ 3-liter jars and removing water with n 5-liter jars, we end up with n liters of water.

If we replace 3 and 5 with variables, we have the following more general question.

Question 1.12. Given integers x and y , determine all possible values of the form $ax + by$ for integers a and b . In mathematical notation, determine the set

$$\{ax + by : a, b \in \mathbb{Z}\}.$$

Example 1.13. Consider two jars of sizes 4 liters and 10 liters. Recall that we are looking at combinations of the form

$$4a + 10b$$

representing how many times each jar is added or subtracted from the total. It is clear that whatever the result is, it must be even. Thus, we already know that we cannot get any number of liters because we cannot get an odd number of liters. It remains to be seen if we can get any *even* number of liters.

We can again take multiples, so if we can get 2 liters, then we can get any even number of liters. We prove this statement formally.

Proof. Assume there exists integers a and b such that

$$4a + 10b = 2.$$

We can also write any even number as $2m$, where m is an integer. Then to get $2m$ liters, we simply take

$$m(4a + 10b) = 4(am) + 10(bm) = 2m. \quad \square$$

So, if we can find a way to get 2 liters, we can get any even number of liters. We see that

$$10 - 2 \cdot 4 = 2,$$

so we can get any even number of liters.

Investigation 1.14. Since the key to knowing all possible linear combinations is knowing the smallest positive combination, consider the following problems.

- (a) Determine the smallest positive quantity that can be measured with a 2-liter and 4-liter jar.
- (b) Determine the smallest positive quantity that can be measured with a 12-liter and 30-liter jar.
- (c) Determine the smallest positive quantity that can be measured with a 234-liter and 1926-liter jar.

Do you see a relationship between the size of the jars and the smallest positive quantity they can measure?

We have now gathered some data on Question 1.12 and conjectured a relationship in Investigation 1.14. The next step is to state and prove the general mathematical solution to Question 1.12. It turns out the answer depends on the divisibility properties of x and y . We make some general definitions.

Definition 1.15. Let x, y be integers. A number of the form $ax + by$ for integers a and b is called a *linear combination* of x and y .

Definition 1.16. Given integers x and y that are not both 0, define the *greatest common divisor* of x and y , notated $\gcd(x, y)$, to be the largest positive integer that divides both x and y . If $\gcd(x, y) = 1$, then x and y are said to be *relatively prime*.

Example 1.17.

- We compute $\gcd(12, 15) = 3$. In other words, 3 divides both 12 and 15, but no larger integer divides both 12 and 15.
- We compute $\gcd(1462, 408) = 34$.
- The numbers $\{1, 2, 4, 5, 7, 8\}$ are all relatively prime to 9.

Definition 1.18. We define the *least common multiple* of x and y , notated $\text{lcm}(x, y)$, to be the smallest positive integer n that is divisible by both x and y .

Example 1.19. We compute $\text{lcm}(9, 12) = 36$. In other words, 36 is the smallest integer that is divisible by both 9 and 12.

We can extend these definitions to any number of integers and denote them by $\text{gcd}(a_1, \dots, a_n)$ and $\text{lcm}(a_1, \dots, a_n)$. Notice that in both definitions we say “the”, meaning that, unlike divisors, there is a unique greatest common divisor and a unique least common multiple. We next give a more general characterization of $\text{lcm}(x, y)$.

Proposition 1.20. *Given integers x, y , and n , if $x \mid n$ and $y \mid n$, then $\text{lcm}(x, y) \mid n$.*

Example 1.21. We have $6 \mid 60$ and $5 \mid 60$ so the conclusion is $\text{lcm}(5, 6) = 30$ also divides 60.

Proof. Let n be any integer such that $x \mid n$ and $y \mid n$. Then there exists an integer k such that $n = xk$. Since x also divides $\text{lcm}(x, y)$, there is another integer k' such that $\text{lcm}(x, y) = k'x$. Now apply the division algorithm to n with $d = \text{lcm}(x, y)$ to write

$$n = q \text{lcm}(x, y) + r, \quad 0 \leq r < \text{lcm}(x, y).$$

Now we have

$$xk = qxk' + r,$$

so

$$x(k - qk') = r.$$

Thus $x \mid r$.

We can apply the same argument with y instead of x to see that $y \mid r$.

If $r > 0$, r is a (positive) integer divisible by both x and y which, by the division algorithm, is smaller than $\text{lcm}(x, y)$, contradicting the minimality of $\text{lcm}(x, y)$. So we must have $r = 0$, which is equivalent to $\text{lcm}(x, y) \mid n$. \square

We answer Question 1.12 with the following theorem and corollary.

Theorem 1.22. *The greatest common divisor of two integers x and y not both 0 is the smallest positive integer that is a linear combination of x and y .*

Proof. Let d be the smallest positive integer that is a linear combination of x and y . Write $d = ax + by$ for integers a and b . We need to see that $d \mid x$ and $d \mid y$.

Using the division algorithm (Theorem 1.8) we write $x = qd + r$ for $0 \leq r < d$. Then we have

$$r = x - qd = x - q(ax + by) = (1 - qa)x - qby.$$

Thus, r is a linear combination of x and y . Since d is the smallest positive linear combination and $0 \leq r < d$, we must have $r = 0$. Since r is the remainder of x after division by d , $r = 0$ implies that $d \mid x$. Similarly, we can show $d \mid y$.

Therefore, $d \mid \text{gcd}(x, y)$. \square

Corollary 1.23. *The set of linear combinations of x and y are the multiples of $\text{gcd}(x, y)$.*

$$\{ax + by : a, b \in \mathbb{Z}\} = \{m \text{gcd}(x, y) : m \in \mathbb{Z}\}.$$

Proof. Let $d = \gcd(x, y)$. Since $d \mid x$ and $d \mid y$, by Proposition 1.7, $d \mid ax + by$ for any integers a and b . The definition of divisibility gives the existence of an integer m such that

$$ax + by = dm.$$

In other words, every linear combination is a multiple of $\gcd(x, y)$.

To show that we obtain every multiple, let a and b be such that

$$\gcd(x, y) = ax + by.$$

Then for any integer m , am , and bm satisfy

$$m \cdot \gcd(x, y) = amx + bmy.$$

Therefore, we can represent every multiple of the greatest common divisor as a linear combination. \square

Example 1.24. Consider $x = 9$ and $y = 6$. Then $\gcd(x, y) = 3$, and we compute a few linear combinations to see they are all multiples of 3.

$$\begin{aligned} x + y &= 15 = 5 \cdot 3, \\ x - y &= 3 = 1 \cdot 3, \\ x + 2y &= 21 = 7 \cdot 3, \\ -x + y &= -3 = -1 \cdot 3, \\ 7x - 3y &= 45 = 15 \cdot 3, \\ 12x - 5y &= 78 = 26 \cdot 3. \end{aligned}$$

As we have already seen, the answer to Question 1.11 is yes, we can measure 1 liter given 3-liter and 5-liter jars. Now we know the reason this is possible is that $\gcd(3, 5) = 1$.

We also get another characterization of the greatest common divisor.

Corollary 1.25. *If $d \mid x$ and $d \mid y$, then $d \mid \gcd(x, y)$.*

Proof. From Theorem 1.22 we can write the $\gcd(x, y)$ as a linear combination $ax + by$ with $a, b \in \mathbb{Z}$. If $d \mid x$ and $d \mid y$, by Proposition 1.7 we have $d \mid ax + by$. \square

Now we turn to the practical (computational) question of computing the gcd.

Question 1.26. Given two integers x and y , how do we compute $\gcd(x, y)$ and find the pair of integers (a, b) such that

$$ax + by = \gcd(x, y)?$$

To compute the greatest common divisor, we could simply test the divisibility of every positive integer at most $\min(|x|, |y|)$ and find the largest that divides both x and y . While this takes a finite number of steps and is guaranteed to produce the greatest common divisor, we hope for a more efficient solution. Additionally, this method gives us no information about (a, b) . We describe the method of Euclid,¹

¹Euclid, also known as Euclid of Alexandria, was a Greek mathematician living around 300 B.C.E.

called the *Euclidean algorithm*, that solves both of these problems at once in a very efficient manner. It relies on a connection between the greatest common divisor of two numbers and the remainder from the division algorithm.

Lemma 1.27. *Let n and d be integers with $d > 0$. Let $q, r \in \mathbb{Z}$ be the quotient and remainder from the division algorithm ($n = qd + r$), then*

$$\gcd(n, d) = \gcd(d, r).$$

Proof. Let $a = \gcd(n, d)$ and $b = \gcd(d, r)$. We will show that $a \mid b$ and that $b \mid a$, proving that $a = b$ (Proposition 1.7).

First we show $a \mid b$. From Theorem 1.22 we can find integers x and y such that

$$(5) \quad b = xd + yr.$$

We also know that

$$n = qd + r \quad \text{which is equivalent to} \quad r = n - qd.$$

Substituting this expression for r into equation (5), we see that

$$b = xd + y(n - qd) = (x - qy)d + yn.$$

Since $a \mid d$ and $a \mid n$, we have $a \mid b$ by Proposition 1.7.

Now we show $b \mid a$. We can find integers x and y such that

$$a = xn + yd$$

and

$$n = qd + r.$$

Combining yields

$$a = x(qd + r) + yd = (xq + y)d + xr.$$

Since $b \mid d$ and $b \mid r$, we have $b \mid a$ by Proposition 1.7.

We conclude that since $a \mid b$ and $b \mid a$ with both positive, then $a = b$. \square

The next example illustrates how Lemma 1.27 can be used to find the greatest common divisor of two numbers by repeating, making the problem smaller.

Example 1.28. We want to find $\gcd(36, 28)$. We use the division algorithm

$$36 = 1 \cdot 28 + 8$$

and Lemma 1.27 to see that

$$\gcd(36, 28) = \gcd(28, 8).$$

This has made the original problem simpler since 8 is smaller than 36. We do it again.

$$28 = 3 \cdot 8 + 4$$

so that

$$\gcd(36, 28) = \gcd(28, 8) = \gcd(8, 4).$$

We could now see that the answer is 4 or carry it one step further.

$$8 = 2 \cdot 4 + 0$$

so that

$$\gcd(36, 28) = \gcd(28, 8) = \gcd(8, 4) = \gcd(4, 0) = 4.$$

We formalize the procedure in Example 1.28 as the Euclidean algorithm.

Theorem 1.29 (Euclidean algorithm). *Given integers n, d with $n \geq d > 0$, let $r_0 = n$ and $r_1 = d$. We construct a sequence of integers $\{r_0, r_1, r_2, \dots\}$ using the division algorithm*

$$r_i = q_{i+1}r_{i+1} + r_{i+2}.$$

This sequence terminates when $r_m = 0$ for some m . Then $r_{m-1} = \gcd(n, d)$.

In following the proof, it may help to keep Example 1.28 in mind. In the notation of Theorem 1.29, we have $n = 36, d = 28$, and we constructed the sequence of remainders

$$r_0 = 36, r_1 = 28, r_2 = 8, r_3 = 4, r_4 = 0.$$

This is a decreasing sequence and the last nonzero value is 4, the greatest common divisor.

Proof. We need to show that the algorithm terminates and that the value is the greatest common divisor.

For termination, recall that by the division algorithm $r_i > r_{i+1} \geq 0$ for all i . Thus, we have a decreasing sequence of nonnegative integers, which can only be finite in length. As a consequence, we must eventually get to an m such that $r_m = 0$.

By Lemma 1.27 we attain the greatest common divisor since

$$\begin{aligned} \gcd(n, d) &= \gcd(r_0, r_1) = \gcd(r_1, r_2) = \dots \\ &= \gcd(r_{m-1}, r_m) = \gcd(r_{m-1}, 0) = r_{m-1}. \end{aligned} \quad \square$$

Example 1.30. Compute $\gcd(126, 34)$.

$$\begin{aligned} 126 &= 3 \cdot 34 + 24, \\ 34 &= 1 \cdot 24 + 10, \\ 24 &= 2 \cdot 10 + 4, \\ 10 &= 2 \cdot 4 + \boxed{2}, \\ 4 &= 2 \cdot 2 + 0. \end{aligned}$$

Thus,

$$\begin{aligned} \gcd(126, 34) &= \gcd(34, 24) = \gcd(24, 10) = \gcd(10, 4) \\ &= \gcd(4, 2) = \gcd(2, 0) = 2. \end{aligned}$$

The Euclidean algorithm solves the first half of Question 1.26, computing the greatest common divisor. But we still want to find the coefficients a, b in the linear combination that produce the gcd:

$$ax + by = \gcd(x, y).$$

We can find a, b by working backward through the Euclidean algorithm steps. For example,

$$\begin{array}{ll} 2 &= 10 - 2 \cdot 4 && \text{now use } 4 = 24 - 2 \cdot 10, \\ &= 10 - 2(24 - 2 \cdot 10) = 5 \cdot 10 - 2 \cdot 24 && \text{now use } 10 = 34 - 1 \cdot 24, \\ &= 5(34 - 24) - 2 \cdot 24 = 5 \cdot 34 - 7 \cdot 24 && \text{now use } 24 = 126 - 3 \cdot 34, \\ &= 5 \cdot 34 - 7(126 - 3 \cdot 34) = 26 \cdot 34 - 7 \cdot 126. \end{array}$$

In particular,

$$2 = 26 \cdot 34 - 7 \cdot 126.$$

The Euclidean algorithm turns out to be a very fast way to compute greatest common divisors. In fact, it is much faster to compute the greatest common divisor of two numbers than it is to factor a number!

Investigation 1.31. How many steps does it take to complete the Euclidean algorithm? Find pairs (x, y) which take the most number of steps relative to their size to complete the Euclidean algorithm.

- Find the pair (x, y) with $x, y < 7$ that takes the most steps.
- Find the pair (x, y) with $x, y < 10$ that takes the most steps.
- Find the pair (x, y) with $x, y < 15$ that takes the most steps.
- Find the pair (x, y) with $x, y < 25$ that takes the most steps.
- Can you find the pattern in these numbers?

Algorithm 1.3 describes the algorithm for finding the greatest common divisor using the Euclidean algorithm.

Algorithm 1.3. Euclidean Algorithm

Input: two positive integers n, d

Output: $\gcd(n, d)$

Algorithm:

- 1: If $d \leq 0$ or $n \leq 0$, return error.
 - 2: Let $r_0 = \max(n, d)$, let $r_1 = \min(n, d)$.
 - 3: While $r_i \neq 0$.
 - a: Let r_i be the remainder from the division algorithm of (r_{i-1}, r_{i-2}) .
 - 4: Return r_{i-1} .
-

To get the coefficients a and b , we have already seen a method in Example 1.30 of working backward through the steps of the Euclidean algorithm. If we wish to write a program to do this, it would be better if we could work forward since we do not know ahead of time when the algorithm will terminate. (For the more programming advanced, recursion would allow us to work backward, but it has the penalty of using more memory.) Let's see how this would work in an example.

Example 1.32. We compute $\gcd(186, 144)$. We first use the division algorithm to get

$$186 = 1 \cdot 144 + 42,$$

which we rewrite as

$$(6) \quad 42 = 186 - 1 \cdot 144,$$

which expresses 42 as a linear combination of $(186, 144)$. Now we continue with the next step in the Euclidean algorithm

$$144 = 3 \cdot 42 + 18$$

or

$$18 = 144 - 3 \cdot 42.$$

We substitute equation (6) in for 42 to get

$$(7) \quad 18 = 144 - 3(186 - 1 \cdot 144) = 4 \cdot 144 - 3 \cdot 186,$$

which expresses 18 as a linear combination of (186, 144). Continuing, we have

$$42 = 2 \cdot 18 + 6$$

or

$$6 = 42 - 2 \cdot 18.$$

We substitute equation (6) in for 42 and equation (7) in for 18 to have

$$\begin{aligned} 6 &= (186 - 1 \cdot 144) - 2(4 \cdot 144 - 3 \cdot 186) \\ &= 7 \cdot 186 - 9 \cdot 144, \end{aligned}$$

which expresses 6 as a linear combination of (186, 144). The last step is

$$18 = 3 \cdot 6 + 0$$

so that

$$\gcd(186, 144) = 6 = 7 \cdot 186 - 9 \cdot 144.$$

We can work out the process of Example 1.32 in a general way to provide what is known as the *Extended Euclidean Algorithm*.

Given two integers r_0, r_1 , we can write

$$\begin{aligned} r_0 &= a_0 r_0 + b_0 r_1 \quad \text{with} \quad (a_0, b_0) = (1, 0), \quad \text{i.e., } r_0 = 1 \cdot r_0 + 0 \cdot r_1, \\ r_1 &= a_1 r_0 + b_1 r_1 \quad \text{with} \quad (a_1, b_1) = (0, 1), \quad \text{i.e., } r_1 = 0 \cdot r_0 + 1 \cdot r_1. \end{aligned}$$

Now consider

$$r_0 = q_1 r_1 + r_2 \quad \text{which is the same as} \quad r_2 = r_0 - q_1 r_1.$$

Substituting the expressions for r_0 and r_1 , we get

$$\begin{aligned} r_2 &= (a_0 r_0 + b_0 r_1) - q_1 (a_1 r_0 + b_1 r_1) \\ &= (a_0 - q_1 a_1) r_0 + (b_0 - q_1 b_1) r_1. \end{aligned}$$

Repeating this process with $r_{i+1} = r_{i-1} - q_i r_i$, we get

$$r_{i+1} = (a_{i-1} - q_i a_i) r_0 + (b_{i-1} - q_i b_i) r_1.$$

Since the greatest common divisor of n and d is the last nonzero r_i , we are “adjusting” the coefficients a_i, b_i at each step of the algorithm as

$$\begin{aligned} a_{i+1} &= a_{i-1} - q_i a_i, \\ b_{i+1} &= b_{i-1} - q_i b_i. \end{aligned}$$

Thus, we can implement the extended algorithm as Algorithm 1.4.

Algorithm 1.4. Extended Euclidean Algorithm

Input: two nonnegative numbers n, d

Output: $\gcd(n, d)$ and a, b such that $an + bd = \gcd(n, d)$

Algorithm:

- 1: If $d \leq 0$ or $n \leq 0$, return error.
 - 2: Let $r_0 = \max(n, d)$, let $r_1 = \min(n, d)$.
 - 3: Let $a_0 = 1$ and $b_0 = 0$.
 - 4: Let $a_1 = 0$ and $b_1 = 1$.
 - 5: While $r_i \neq 0$.
 - a: Use the division algorithm to determine $r_{i-1} = q_i r_i + r_{i+1}$.
 - b: Let $a_{i+1} = a_{i-1} - q_i a_i$ and $b_{i+1} = b_{i-1} - q_i b_i$.
 - 6: Return $(r_{i-1}, a_{i-1}, b_{i-1})$.
-

3.1. Efficiency of the Euclidean Algorithm. We stated above that the Euclidean algorithm is very efficient for finding the greatest common divisor of two integers, and Investigation 1.31 looked empirically at how many steps it can take. We now prove a bound on the number of steps the Euclidean algorithm takes by examining the worst possible case. By worst case, we mean given a number of steps k , we want the smallest integers m and n such that computing $\gcd(m, n)$ takes k steps with the Euclidean algorithm.

Question 1.33. How does the number of steps required in the Euclidean algorithm to compute $\gcd(m, n)$ depend on the inputs m and n ?

The idea to construct the worst possible inputs is actually quite simple, we want the remainder r in each step to be as large as possible compared to n . The largest remainder occurs when $q = 1$ in the division algorithm. In particular, we have a series of steps that look like

$$\begin{aligned} n &= m + r_1, \\ m &= r_1 + r_2, \\ r_1 &= r_2 + r_3, \\ &\vdots \end{aligned}$$

Notice that each number is the sum of the previous two. This gives a recursively defined sequence of numbers that depends only on the first two numbers in the sequence. Taking the initial two numbers as 0 and 1 defines the Fibonacci numbers.

Definition 1.34. Let $F_0 = 0$ and $F_1 = 1$, and for $n \geq 2$ define

$$F_n = F_{n-1} + F_{n-2}.$$

The elements of this sequence of positive integers are called the *Fibonacci numbers*.

We prove that the Fibonacci numbers represent the worst possible inputs for the Euclidean algorithm and use a formula for their size to get an approximate worst-case running time for the Euclidean algorithm.

Theorem 1.35 (Lamé's Theorem). *If the Euclidean algorithm applied to $n > m \geq 1$ requires k steps, then $n \geq F_{k+2}$ and $m \geq F_{k+1}$.*

Proof. We proceed by induction on k . For $k = 1$, since $n > m \geq 1$, we must have $n \geq 2 = F_3$ and $m \geq 1 = F_2$.

Now consider $k > 1$ steps. We can write

$$(8) \quad n = qm + r$$

so that the Euclidean algorithm applied to (m, r) takes $(k-1)$ -steps. By induction, we know that $m \geq F_{k+1}$ and $r \geq F_k$. Since $m \geq 1$ in (8) we have $n \geq F_{k+1} + F_k = F_{k+2}$. \square

We can now find an upper bound on the number of steps.

Lemma 1.36. *Let $\phi = \frac{1+\sqrt{5}}{2}$ be the golden ratio. For $n \geq 1$, the Fibonacci numbers satisfy*

$$F_{n+1} < \phi^n < F_{n+2}.$$

Proof. We first compute directly that

$$(9) \quad \phi^2 = \frac{3 + \sqrt{5}}{2} = \phi + 1.$$

Then we can multiply each side of equation (9) by any power of ϕ and still have an equality so that, for all $k > 1$,

$$\phi^k = \phi^{k-1} + \phi^{k-2}.$$

Now we proceed by induction. We check that

$$F_2 = 1 < \phi < 2 = F_3.$$

Now we assume that

$$F_{k+1} < \phi^k < F_{k+2} \quad \text{and} \quad F_{k+2} < \phi^{k+1} < F_{k+3}.$$

Adding the two inequalities together yields

$$F_{k+1} + F_{k+2} < \phi^k + \phi^{k+1} < F_{k+2} + F_{k+3},$$

which gives

$$F_{k+3} < \phi^{k+2} < F_{k+4},$$

completing the induction. \square

Theorem 1.37. *For integers $n > m \geq 1$, the Euclidean algorithm takes at most $(5 \log_{10}(m) + 1)$ steps.*

Proof. Using Theorem 1.35, we need to find the largest Fibonacci number less than m . Combining this with Lemma 1.36, to take k steps, we need $m \geq F_{k+1} > \phi^{k-1}$. We compute

$$\log_{\phi}(m) > \log_{\phi}(\phi^{k-1}) = (k-1)$$

so that

$$\log_{\phi}(m) + 1 > k.$$

Consequently,

$$k < \log_{\phi}(m) + 1 = \frac{\log_{10}(m)}{\log_{10}(\phi)} + 1 < 5 \log_{10}(m) + 1. \quad \square$$

In particular, Theorem 1.37 says that the Euclidean algorithm takes at most 5 times the number of decimal digits in the smaller of the two inputs. For example, we can find the greatest common divisor of two 2^{2048} -bit numbers (the size of a strong modulus for public key cryptography) in at most about 7100 steps; in other words, very quickly.

There are many other fascinating properties of the Fibonacci numbers which can be explored in Exploration Exercise 1.31.

4. Prime Numbers and Unique Factorization

Recall that we defined a prime number as an integer greater than one whose only positive divisors are 1 and itself. In this section we show that primes are the fundamental building blocks of numbers (*unique factorization*) and examine how many prime numbers there are. Before we state the theorem of unique factorization, we state a useful property of prime numbers that we will need for the proof.

Lemma 1.38. *Let a , b , and p be integers with p prime. If $p \mid ab$, then $p \mid a$ or $p \mid b$.*

Proof. If $p \mid a$, then the conclusion is satisfied, so assume $p \nmid a$. Since p is prime and the $\gcd(p, a)$ must divide p , we know that $\gcd(p, a) = 1$ or p . Since $p \nmid a$ and $\gcd(p, a)$ must also divide a , then $\gcd(p, a) = 1$. From Theorem 1.22 we can find integers m, n such that $np + ma = 1$. Multiplying both sides by b , we have

$$b(np + ma) = b.$$

Since $p \mid p$ and $p \mid ab$, then p divides the left-hand side. Thus, p also divides the right-hand side, i.e., $p \mid b$. \square

Example 1.39. We have

$$5 \mid 20, \quad \text{and} \quad 20 = 2 \cdot 10$$

and

$$5 \mid 10.$$

Example 1.40. Lemma 1.38 can be false when p is a composite number. For example,

$$4 \mid 20$$

but $4 \nmid 2$ and $4 \nmid 10$.

Theorem 1.41 (Unique factorization of integers). *Every integer $n \geq 2$ can be written as a product of prime numbers in exactly one way (up to reordering)*

$$n = p_1 \cdots p_r,$$

where p_1, \dots, p_r are prime numbers.

Example 1.42. We can factor any integer into a product of prime numbers in only one way:

$$\begin{aligned} 70 &= 2 \cdot 5 \cdot 7, \\ 228 &= 2 \cdot 2 \cdot 3 \cdot 19. \end{aligned}$$

Proof. There are two statements: first, n can be written as a product of primes; and second, factorization is unique. We prove the first by (strong) induction on n .

- Base case $n = 2$: 2 is prime, so there is nothing to do.
- Induction step: Assume that every positive integer up to $n - 1$ can be written as a product of primes. Consider n . If n is prime, then we are done. Otherwise, n is composite, and we can write $n = ab$ for two integers $1 < a, b < n$. By the induction assumption, we can write

$$\begin{aligned} a &= p_1 \cdots p_r, \\ b &= q_1 \cdots q_s \end{aligned}$$

for prime numbers $p_1, \dots, p_r, q_1, \dots, q_s$. Thus, we can write n as a product of prime numbers as

$$n = p_1 \cdots p_r \cdot q_1 \cdots q_s.$$

Now for the uniqueness. Suppose that we may factor n as the product of prime numbers in two different ways, $n = p_1 \cdots p_r$ and $n = q_1 \cdots q_s$. Consider p_1 . We know that $p_1 \mid n$ and thus $p_1 \mid q_1 \cdots q_s$. By Lemma 1.38 we must then have $p_1 \mid q_i$ for some $1 \leq i \leq s$ so that $p_1 = q_i$. We may as well rename q_i to q_1 so that $q_1 = p_1$ and

$$n = p_1 p_2 \cdots p_r = p_1 q_2 \cdots q_s.$$

We cancel p_1 from both sides and consider the equality

$$p_2 \cdots p_r = q_2 \cdots q_s.$$

We repeat the above process to see that $p_2 = q_i$ for some i , etc. In particular, every prime p_j is equal to exactly one q_i . Thus, after renaming, we see that we must have $p_i = q_i$ for $1 \leq i \leq r$ and $r = s$, so the two factorizations are identical. \square

While the theorem says that there exists a unique factorization, it is not effective in the sense that it does not give a way to actually find the factorization of a given integer.

Question 1.43. Given an integer n , can we find its prime factorization?

This is actually quite a hard question, and there are many algorithms (both deterministic and probabilistic) for factoring, but none are efficient for large integers. In particular, the notion that factoring is hard is the basic assumption for the security of many public key cryptographic systems. We will not discuss factorization algorithms in detail. (See also Exploration Exercise 1.39).

Unique factorization also leads to the question about the number of prime numbers. We now know that all (positive) integers are made up of products of prime numbers and there are infinitely many (positive) integers. Does that mean there are infinitely many prime numbers, or do we eventually reach a point where every integer is a product from the same set of smaller integers?

Question 1.44. How large is the set of prime numbers?

Investigation 1.45.

- (a) Find all the prime numbers less than 100.
- (b) Find all the prime numbers less than 1000.
- (c) Consider the following limit on the proportion of prime numbers:

$$\lim_{N \rightarrow \infty} \frac{\{p : p \text{ is prime, } p \leq N\}}{N}.$$

What is the limiting value? What does this say about the number of prime numbers?

The ancient Greeks knew that there were infinitely many prime numbers. We consider the proof given by Euclid. The proof is by contradiction. Euclid showed that given *any* finite list of prime numbers, there is at least one prime number not in this list. Thus, there are infinitely many prime numbers. The key idea is that no prime can divide both n and $n + 1$ since it must also divide their difference. Thus, we would have the contradiction $p \mid 1$.

Theorem 1.46 (Euclid). *The set of prime numbers is infinite.*

Proof. Assume there are finitely many prime numbers, and list them as $\{p_1, \dots, p_n\}$. Consider the integer

$$p = p_1 \cdot p_2 \cdots p_n + 1.$$

Let p_i be any prime in $\{p_1, \dots, p_n\}$. Then p_i divides p only if it also divides 1 (Proposition 1.7). Since no prime divides 1, every prime in the list does not divide p . However, p may be factored into a product of prime numbers by the unique factorization of integers theorem (Theorem 1.41). Thus, there exists some prime number $q \mid p$ and $q \notin \{p_1, \dots, p_n\}$. \square

Investigation 1.47 (New prime numbers). Euclid’s proof of the infinitude of prime numbers relies on n and $n + 1$ having distinct prime factors. Let’s take a moment and explore this way of constructing new primes.

- (a) Which pairs of prime number p_1, p_2 have $p_1 p_2 + 1$ prime?
- (b) Start with a list of two primes $\{p_1, p_2\}$. Add all the prime factors of $p_1 p_2 + 1$ to your list. Next add all the prime factors of $1 + \prod p_i$ to your list. Repeat. What proportion of primes end up in your list?

While Theorem 1.46 gives a partial answer to Question 1.44, unfortunately “infinitely many” is not very descriptive. For example, there are infinitely many even integers and infinitely many multiples of 5, but our intuition tells us that there are somehow “more” even numbers than multiples of 5. We would like to get a better sense of the number of prime numbers. We can address this issue satisfactorily with a counting function.

Definition 1.48. We want to determine how many integers satisfy some property P . We can define a *counting function* which counts the number of numbers that satisfy P up to some bound

$$c_P(N) = \#\{n \in \mathbb{N}: n \text{ satisfies } P \text{ and } n \leq N\}.$$

From this count, we can compute the *density* of positive integers that satisfy property P as

$$\lim_{N \rightarrow \infty} \frac{c_P(N)}{N}.$$

In some sense, the density gives an average number of integers with some property.

Example 1.49. For a positive integer N let $c_2(N)$ be the number of even numbers at most N , and let $c_5(N)$ be the number of multiples of 5 at most N . Then we have

$$\begin{aligned} c_2(10) &= \#\{2, 4, 6, 8, 10\} = 5, \\ c_5(10) &= \#\{5, 10\} = 2. \end{aligned}$$

To quantify this in a more meaningful way, we take the ratio $\frac{c(N)}{N}$. We compute a few ratios below.

N	$\frac{c_2(N)}{N}$	$\frac{c_5(N)}{N}$
10	$\frac{5}{10} = 0.5$	$\frac{2}{10} = 0.2$
11	$\frac{5}{11} = 0.454545 \dots$	$\frac{2}{11} = 0.181818 \dots$
12	$\frac{6}{12} = 0.5$	$\frac{2}{12} = 0.16666 \dots$
13	$\frac{6}{13} = 0.461538 \dots$	$\frac{2}{13} = 0.153846 \dots$
14	$\frac{7}{14} = 0.5$	$\frac{2}{14} = 0.142857 \dots$
15	$\frac{7}{15} = 0.46666 \dots$	$\frac{3}{15} = 0.2$
100	$\frac{50}{100} = 0.5$	$\frac{20}{100} = 0.2$
101	$\frac{50}{101} = 0.495049 \dots$	$\frac{20}{101} = 0.1980 \dots$
1000	$\frac{500}{1000} = 0.5$	$\frac{200}{1000} = 0.2$
1001	$\frac{500}{1001} = 0.499500 \dots$	$\frac{200}{1001} = 0.199800 \dots$

Just by cursory inspection, it appears that the variation around $\frac{1}{2}$ and $\frac{1}{5}$ decreases as N increases. In fact we could prove that

$$\lim_{N \rightarrow \infty} \frac{c_2(N)}{N} = \frac{1}{2},$$

and we would say that multiples of 2 have density $\frac{1}{2}$. Similarly for multiples of 5, we could prove that

$$\lim_{N \rightarrow \infty} \frac{c_5(N)}{N} = \frac{1}{5}$$

to see that multiples of 5 have density $\frac{1}{5}$.

This gives us a satisfactory answer to our original question of how to quantify that there are more even numbers than multiples of 5, even though there are infinitely many of both. While one out of every two numbers (on average) is even, only one out of every five numbers (on average) is a multiple of 5.

Now back to our question about prime numbers. We define a prime counting function.

Definition 1.50. Let N be an integer, and define $\pi(N)$ to be the number of primes $\leq N$.

Question 1.51. Determine the function $\pi(N)$.

To generate data about Question 1.51, we have already encountered a problem: we must be able to find all the prime numbers up to some bound.

Remark. A related question, although not exactly the same, is the following.

Question 1.52. Given a positive integer n , how do you determine efficiently if n is prime?

While this is similar to Question 1.43, which asks for the factorization of a given number, it is not exactly the same. If we know the factorization of a number, then we certainly know if it is prime; but it is possible to determine if a number is prime or composite without actually finding a factorization. Such methods are called *primality tests* (see Exploration Exercise 2.31).

The Sieve of Eratosthenes² (Algorithm 1.5) is one way to determine all primes up to some bound. Essentially, we start with a list of all numbers and remove all the multiples of the prime numbers. At the end of each step, the smallest number not crossed out is prime since it has no smaller divisors. To find all the primes at most N , we stop when we are larger than \sqrt{N} .

²Eratosthenes (276–194 B.C.E) born in Cyrene, a Greek colony west of Egypt.

Algorithm 1.5. Sieve of Eratosthenes

Input: a positive integer N **Output:** a list of prime numbers $\leq N$ **Algorithm:**

- 1: Start with an ordered list of all numbers $2, \dots, N$.
 - 2: Repeat the following until you have examined every number in the list.
 - a: Let p be the smallest number in the list not already examined. Note that p is a prime since nothing smaller can divide it.
 - b: Remove all multiples of p from the list.
 - 3: Return the remaining elements in the list.
-

We give a graphical representation of Algorithm 1.5 for $N = 100$. The numbers in bold are the prime numbers.

	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

We see that

$$\pi(100) = 28 \quad \text{and} \quad \frac{\pi(100)}{100} = \frac{28}{100} = 0.28.$$

We could continue to compute with larger bounds to see the following.

N	$\pi(N)$	$\frac{\pi(N)}{N}$
100	28	0.28
1,000	168	0.168
10,000	1,220	0.1220
10^5	9,592	0.09592
10^6	78,498	0.078498
10^7	664,579	0.0664579
10^8	5,761,455	0.05761455
10^9	50,847,534	0.050847534

The density appears to be heading toward 0, and we must refine our question yet again because we do not know how fast it is approaching 0. For example, $f(x) = x$ and $g(x) = x^{10}$ both approach 0 as $x \rightarrow 0$, but $g(x)$ approaches 0 much more quickly:

$$f(0.1) = 0.1 \quad \text{whereas} \quad g(0.1) = 0.0000000001.$$

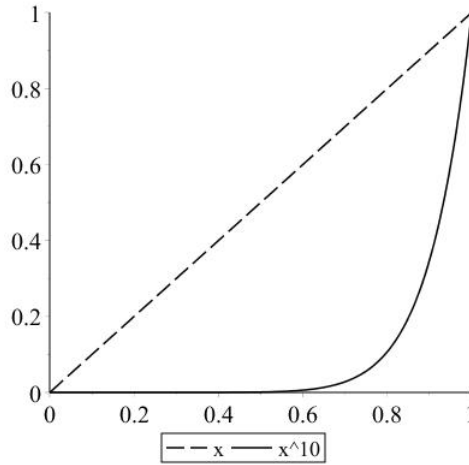


Figure 1.1

We can also see this difference graphically (Figure 1.1). Notice that the solid line representing x^{10} is decreasing more quickly as x decreases (i.e., moving right to left) than the dashed line representing x .

Question 1.53. Can you find a function that has similar values to $\pi(N)$ for each positive integer N ?

If we think of $\pi(x)$ as a function of a variable x , Question 1.51 asks for a function that takes on similar values to $\pi(x)$ for each positive integer x .

Investigation 1.54. To try to answer Question 1.53, you could compare the values of $\pi(x)$ with other functions $f(x)$. You might consider comparing the ratio of the functions instead of the function values themselves.

- (a) Compare the function $\pi(x)$ to various powers of x , i.e., functions $f(x) = x^n$ for integers n .
- (b) Try again, but this time allow fractional powers.
- (c) If you still have not found a good match for $\pi(x)$, perhaps try combining one of the power functions with $\ln(x)$ or e^x .

While Question 1.51 asks for a function with similar values to $\pi(x)$, if it is possible to find such a function, it is exceedingly difficult. Instead, we are looking for functions which are *asymptotic* to $\pi(x)$.

Definition 1.55. We say that $f(x)$ is *asymptotic* to $g(x)$ if

$$\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = 1,$$

and denote this as $f \sim g$.

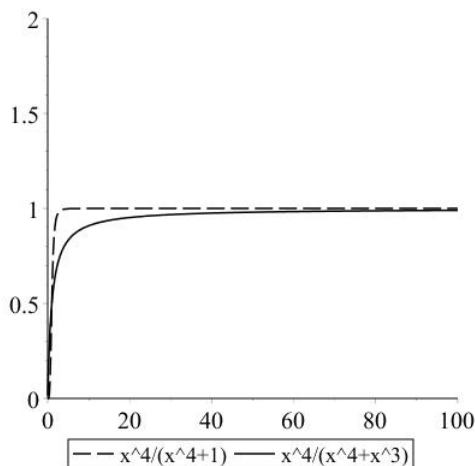


Figure 1.2

Example 1.56. The functions $f(x) = x^4$, $g(x) = x^4 + 1$, and $h(x) = x^4 + x^3$ are all asymptotic:

$$\lim_{x \rightarrow \infty} \frac{x^4}{x^4 + 1} = \lim_{x \rightarrow \infty} \frac{x^4}{x^4 + x^3} = 1.$$

We can see this graphically (Figure 1.2) as the graphs of their ratios have a horizontal asymptote of 1. However, $F(x) = x^4$ and $G(x) = x^3 + 1$ are not asymptotic:

$$\lim_{x \rightarrow \infty} \frac{x^4}{x^3 + 1} = \infty.$$

We can see this graphically (Figure 1.3) as the graph of their ratio does not approach a horizontal asymptote as x increases. Essentially, we are saying that the values of

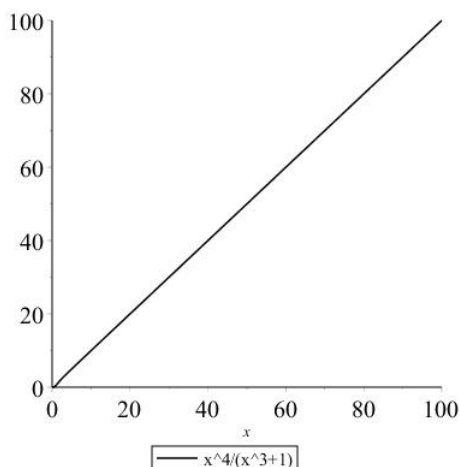


Figure 1.3

f , g , and h are getting (relatively) closer together as x goes to infinity, but that the values of F and G are getting (relatively) farther apart.

Finally, we have a precise mathematical question about the number of prime numbers.

Question 1.57. Can you find a function $f(x)$ such that $\pi(x)$ is asymptotic to $f(x)$?

There is an asymptotic approximation of $\pi(x)$ called the *Prime Number Theorem*. We state it now, but the proof is beyond our scope.

Theorem 1.58 (Prime Number Theorem). $\pi(x) \sim \frac{x}{\ln x}$.

Figure 1.4 shows the ratio $\frac{\pi(x)}{x/\ln(x)}$ with the x -axis plotted logarithmically (\log_{10}). While the ratio is asymptotic to 1, it is approaching 1 quite slowly.

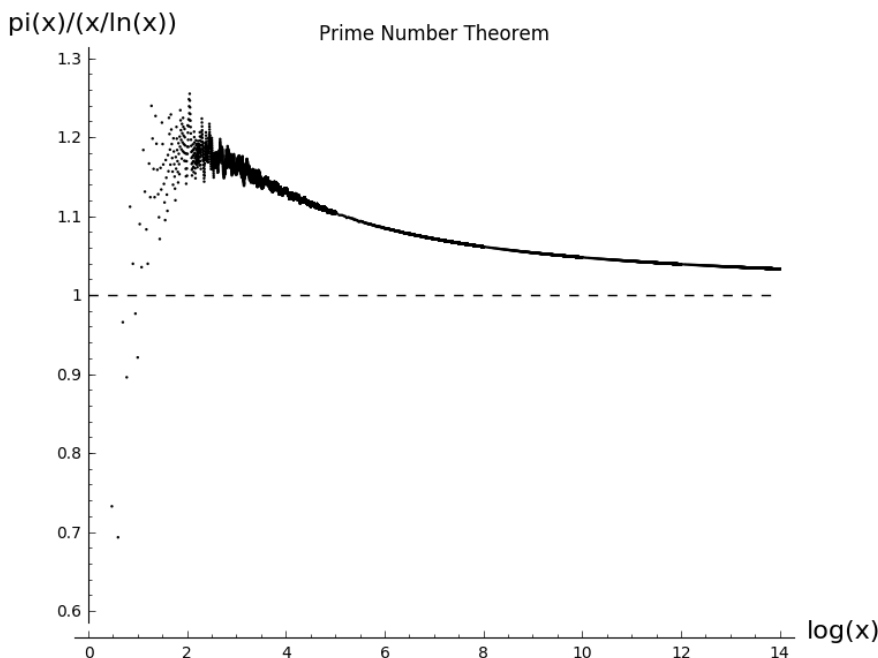


Figure 1.4

This theorem is the pinnacle of counting: it gives an explicit function that is asymptotic to the counting function. In the exercises you will investigate many other types of integers and should strive to find such a counting formula for them even if you are unable to prove it.

COMPUTATIONAL EXERCISES**1.1.**


- a. Write a program to perform trial division on a given number n .
- b. Find the smallest prime number larger than 10010.

1.2.


- a. Write a program to perform the division algorithm. Compute q, r for $(n, d) = (103, 11)$.
- b. Write a program to perform the Euclidean algorithm. Compute

$$\gcd(891555, 191415).$$


- c. Write a program to perform the extended Euclidean algorithm. Find a, b such that $ax + by = \gcd(x, y)$ for
 1. $(x, y) = (23, 18)$.
 2. $(x, y) = (2^{12} - 1, 2^{15} - 1)$.

-  **1.3.** Use the Euclidean algorithm to compute the following greatest common divisors.

- a. $\gcd(154, 147)$
- b. $\gcd(9108, 510)$

-  **1.4.** Use the extended Euclidean algorithm to find integers x and y that satisfy the following.

- a. $28x + 13y = 1$
- b. $42x + 15y = 6$

-  **1.5.** Is it possible to measure one pint of water if you have only an 8-pint and an 11-pint container? If yes, describe the procedure.

- 1.6.** Find the number of ordered triples (a, b, c) of positive integers for which $\text{lcm}(a, b) = 1000$, $\text{lcm}(b, c) = 2000$, and $\text{lcm}(c, a) = 2000$.

1.7.

- a. Write a program to perform the Sieve of Eratosthenes up to some bound n .
- b. Determine the number of prime numbers at most 1,000,000.

- 1.8.** Find the first five prime numbers larger than 1000 that are palindromes, i.e., they are the same written backward and forward (e.g., $p = 131$).

- 1.9.** The Bang–Zsigmondy theorem states that for all $n \geq 2$ with $n \neq 6$, there is a prime p such that $p \mid 2^n - 1$ and $p \nmid 2^d - 1$ for all $1 \leq d < n$.

- a. Find p for $n = 4$.
- b. Write a function that takes as input n and outputs all such p .
- c. Find p for $n = 100$.

1.10.

- a. Write a program that takes as input integers a, b, N and computes the sequence $a_n = \gcd(a^{2^n} - 1, b^{2^n} - 1)$ for $0 \leq n \leq N$.
- b. Find a pair (a, b) for which the sequence (a_n) converges.
- c. Find a pair (a, b) for which the sequence (a_n) diverges.

1.11. Find a sequence of at least three consecutive numbers the sum of whose squares is a square.

1.12. For $n \in \mathbb{N}$ consider $T_n = 2^{2^n} - 1$.

- a. Factor T_n for $n = 1, \dots, 5$.
- b. Prove that T_n has at least n prime divisors.

1.13. Consider $a_n = 3^{2^n} - 1$ and $b_n = 2^{n+3}$.

- a. Determine the remainder when a_n is divided by b_n for $n = 1, \dots, 4$.
- b. Conjecture a general formula for the remainder.

1.14. Find all integers 2^n with the property that after deleting the first digit of its decimal representation, we again get a power of 2. For a nonexample $2^7 = 128$, removing the first digit gives us 28, which is not a power of 2.

THEORETICAL EXERCISES

1.15. Prove that any finite set of integers has a maximal and a minimal element.

1.16. Use induction to prove that $5 \mid n^5 - n$ for any positive integer n .

1.17. For positive integers a, b , and n , prove that if $a \mid n$ and $b \mid n$ with $\gcd(a, b) = 1$, then $ab \mid n$.

1.18. Prove that the product of three consecutive integers is divisible by 6.

1.19. Let p be a prime number, and let a, n be positive integers.

- a. Prove that if $p \mid a^n$, then $p \mid a$.
- b. Use the previous part to prove that for positive integers a and b , if $\gcd(a, b) = d$, then $\gcd(a^n, b^n) = d^n$.

1.20. Let a, b , and n be positive integers. Prove that if $a^n \mid b^n$, then $a \mid b$.

1.21. For two integers a and b , with $\gcd(a, b) = 1$, prove that $\gcd(a + b, a - b)$ is either 1 or 2.

1.22. Let n be an integer. Prove that $\gcd(2n + 5, n + 2) = 1$.

1.23. Let a and b be positive integers. Prove that $ab = \gcd(a, b) \operatorname{lcm}(a, b)$.

1.24. Let a, b, m , and n be integers. Prove the following properties.

- a. If $\gcd(a, b) = d$, then $\gcd(a/d, b/d) = 1$.
- b. If $\gcd(m, n) = 1$, then $\gcd(mn, a) = \gcd(m, a) \cdot \gcd(n, a)$.
- c. $\gcd(ma, mb) = m \gcd(a, b)$.

1.25. In this exercise, we generalize Question 1.12. Given positive integers x_1, \dots, x_n , prove that

$$\{a_1x_1 + a_2x_2 + \dots + a_nx_n : a_i \in \mathbb{Z}\} = \{m \gcd(x_1, \dots, x_n) : m \in \mathbb{Z}\}.$$

1.26. Let a and b be positive integers such that $\text{lcm}(a, b) + \gcd(a, b) = a + b$. Prove that $a \mid b$ or $b \mid a$.

1.27. Let $\phi = \frac{1+\sqrt{5}}{2}$ be the golden ratio. Prove that for $n \geq 1$, the n th Fibonacci number is given by

$$F_n = \frac{\phi^n - (1 - \phi)^n}{\sqrt{5}}.$$

1.28. Prove that there are infinitely many primes with remainder 3 when divided by 4.

1.29. For a positive integer n , prove that the density of multiples of n is $\frac{1}{n}$.

1.30. Prove that for a positive integer s

$$\sum_{n=1}^{\infty} n^{-s} = \prod_{p \text{ prime}} \frac{1}{1 - p^{-s}}.$$

EXPLORATION EXERCISES

1.31 (Fibonacci numbers).

Definition 1.59. Let $F_0 = 0$ and $F_1 = 1$, and for $n \geq 2$ define

$$F_n = F_{n-1} + F_{n-2}.$$

The elements of this sequence of positive integers are called the *Fibonacci numbers*.

- Compute the first few Fibonacci numbers.
- Can you find a relationship between $\gcd(F_n, F_m)$ and m, n ?
- Can you find a relationship between F_n^2 and F_{n-1}, F_{n+1} ?
- Sum formulas:
 - Find a formula for $\sum_{i=0}^n F_i$.
 - Find a formula for $\sum_{i=0}^n F_i^2$.

You can also ask questions about primes.

- How many of the Fibonacci numbers are prime?
- Given a prime p , what is the smallest n such that $p \mid F_n$?
 - Is there an n for every p ?
 - What properties does the sequence of n satisfy?
- For each n is there a prime which divides F_n and which does not divide F_0, \dots, F_{n-1} ? (We would call such a prime a *primitive prime divisor*).
- Generalizations.* We may take any starting values F_0 and F_1 and ask the same questions. For example, Lucas numbers are defined as taking $F_0 = 2$ and $F_1 = 1$. Can you find any particularly interesting starting values?

1.32 (Arithmetic progressions).

Definition 1.60. An *arithmetic progression* is a sequence of integers (a_1, a_2, \dots) such that $a_n - a_{n-1} = c$ for all n for some integer constant c . We often write arithmetic progressions as a sequence $\{an + b : n \in \mathbb{N}\}$.

Example 1.61. $\{2, 7, 12, 17, 22, \dots\}$ is an arithmetic progression.

- a. Given a sequence of integers $\{a_n\}$, prove that if $a_n - a_{n-1} = c$ for some integer constant c and for all n , then there exist integers a, b such that

$$a_n = an + b.$$

Theorem 1.62 (Dirichlet's theorem on arithmetic progressions, 1837). Given two integers a, b with $\gcd(a, b) = 1$, there are infinitely many primes in the sequence

$$\{a + b, 2a + b, 3a + b, \dots\}.$$

- b. Consider $(a, b) = (4, 1)$. Determine an asymptotic counting function for the number of primes in the progression $\{an + b : n \in \mathbb{N}\}$. Can you characterize all the primes in this progression?
- c. Consider $(a, b) = (4, 3)$. Determine an asymptotic counting function for the number of primes. Can you characterize all the primes in this progression?
- d. Can a nonconstant arithmetic progression contain terms that are all primes?
- e. Which sequences have the most/least number of primes?
- f. For which progression is there a prime that divides a_n but does not divide a_1, \dots, a_{n-1} for every n ? (We call such a prime a *primitive prime divisor*.)

1.33 (Prime generating polynomials). In this exploration we look at primes generated by polynomials of degree at least 2. Consider the sequence of numbers of the form $n^2 + 1$ for $n \in \mathbb{N}$. The first ten terms of the sequence are

$$2, 5, 10, 17, 26, 37, 50, 65, 82, 101.$$

Notice that $\{2, 5, 17, 37, 101\}$ are prime, which is exactly half of the terms.

Conjecture 1.63. *There are infinitely many primes of the form $n^2 + 1$.*

- a. Conjecture an asymptote for the counting function.
- b. Can you describe which primes are in this sequence?
- c. Can you find other polynomials that take on infinitely many prime values? What are their counting functions?
- d. Can you find polynomials that do not take on infinitely many prime values?

What about consecutive distinct primes? Euler³ in 1772 gave the example $n^2 - n + 41$, which yields primes for the consecutive values $n = 0, \dots, 40$.

- e. For which primes p does the polynomial $n^2 - n + p$ generate prime numbers for $n = 0, \dots, p - 1$?

³Leonhard Euler (1707–1783) was a Swiss mathematician.

- f. For each degree d find a polynomial that has the most consecutive distinct primes. In other words, $f(x)$ is a different prime for $x = 0, 1, \dots, n$ for the largest n .
- g. Can you find a polynomial which takes only prime values on the positive integers or prove that you cannot?

1.34 (Prime gaps).

Definition 1.64. Given two prime numbers p and q with $p < q$ such that there are no primes ℓ such that $p < \ell < q$, we say that p and q are *consecutive primes*. The difference between two consecutive primes is called a *prime gap*.

Example 1.65. The numbers 3 and 5 are consecutive primes with prime gap 2.

Definition 1.66. Two consecutive primes p and q are called *twin primes* if their prime gap is 2.

Conjecture 1.67 (Twin primes). *There are infinitely many twin primes.*

Conjecture 1.68 (Polignac's Conjecture). *For every positive integer n , $2n$ occurs infinitely often as a prime gap.*

- a. Compute the prime gaps for the first 100 primes.
- b. What integers can occur as prime gaps? How frequently does each prime gap occur?
- c. Is there a largest or smallest prime gap? In particular, determine the growth rate of the largest prime gap. In other words, find the largest gap g for primes up to size B , and consider g/B .
- d. Determine the average prime gap over a certain range. (Compare this to the Prime Number Theorem (Theorem 1.58).)
- e. Can you find a function f such that every interval (for n large enough) of the form $[n, f(n)]$ contains a prime?

Write the prime gaps as a sequence (g_n) . For example, the first few primes are $\{2, 3, 5, 7, 11, \dots\}$. Thus, the prime gap sequence begins $(1, 2, 2, 4, \dots)$.

- f. What sequences of integers can occur as a sequence of prime gaps? For example, can $(2, 2, 2)$ occur as a sequence of prime gaps?
- g. Given a function f , how many prime gaps satisfy $g_n = f(m)$ for some $m \in \mathbb{N}$? For example, $f(x) = x^2$ means, How many prime gaps are perfect squares?

1.35 (Fermat numbers).

Definition 1.69. We define a sequence of numbers called *Fermat numbers* as

$$F_n = 2^{2^n} + 1 \text{ for } n \in \mathbb{N}.$$

If F_n is prime, then it is called a *Fermat prime*. Fermat conjectured that F_n is prime for all $n \in \mathbb{N}$.

Example 1.70. We compute $F_1 = 2^2 + 1 = 5$ and $F_2 = 2^{2^2} + 1 = 17$ are both prime.

- a. Determine the factorization of the first few Fermat numbers. Do you believe Fermat's conjecture?
- b. Find a recurrence relation that describes the Fermat numbers.
- c. Find the greatest common divisor of several pairs of Fermat numbers. Conjecture a general statement.
- d. What can be said about the sum of the reciprocals of the Fermat numbers?
 1. Does it converge? If yes, is the value a rational number?
- e. Which Fermat numbers can be written as the sum of two primes?
- f. Which Fermat numbers can be written uniquely as the sum of two nonzero squares?
- g. What can be said about the number of prime factors of Fermat numbers?
- h. What can be said about the size of the largest prime factor of a Fermat number?
- i. What about more general Fermat numbers such as

$$2^n + 1 \text{ for } n \in \mathbb{N}?$$

1. For what n do you get a prime number?
 2. Can you find a general form for n for which you always get a prime number?
- j. Even more generally, for integers a, b with $a > 1$, what about

$$a^n + b^n?$$

1. Can you find any (a, b, n) that result in a prime?
2. Can you find a fixed (a, b) such that successive values of n result in primes?

1.36 (Mersenne numbers).

Definition 1.71. A *Mersenne⁴ number* is a number of the form

$$M_n = 2^n - 1 \text{ for } n \in \mathbb{N}.$$

A *Mersenne prime⁵* is a Mersenne number that is prime.

Example 1.72. We compute $M_1 = 2 - 1 = 1$ and $M_2 = 2^2 - 1 = 3$.

- a. Compute the first few Mersenne numbers.
- b. Which of the first N Mersenne numbers are prime?

Mersenne conjecture: Mersenne in his *Cogitata Physica-Mathematica* (1644; see Dickson 1919) conjectured that the numbers $2^n - 1$ were prime for

$$n = 2, 3, 5, 7, 13, 17, 19, 31, 67, 127, \text{ and } 257$$

and were composite for all other positive integers $n < 257$.

Was he right?

⁴Marin Mersenne (1588–1648) was a French theologian, philosopher, and mathematician.

⁵The first four Mersenne primes, $M_2 = 3$, $M_3 = 7$, $M_5 = 31$, and $M_7 = 127$, were known in antiquity. The fifth, $M_{13} = 8191$, was discovered anonymously before 1461; the next two (M_{17} and M_{19}) were found by Cataldi in 1588. After nearly two centuries, M_{31} was verified to be prime by Euler in 1772.

- c. How many Mersenne primes are there?
 - 1. Finite or infinite?
 - 2. If infinite, find a counting function for Mersenne primes.
- d. If M_n is prime, what can be said about n ? (Is it prime?)
- e. Generalization: For which a can $a^n - 1$ be prime?
- f. The New Mersenne conjecture of Bateman, Selfridge, and Wagstaff states that for any odd natural number p , if any two of the following conditions hold, then so does the third:
 - Condition I:** $p = 2^k \pm 1$ or $p = 4^k \pm 3$ for some natural number k .
 - Condition II:** $2^p - 1$ is prime (a Mersenne prime).
 - Condition III:** $(2^p + 1)/3$ is prime (a Wagstaff prime).
 - 1. Show that if p is an odd composite number, then $2^p - 1$ and $(2^p + 1)/3$ are both composite.
 - 2. Determine all numbers that satisfy these three conditions up to some bound.

1.37 (Almost primes).

Definition 1.73. A number n is called a k -almost prime if it has exactly k prime factors.

Example 1.74. 6 is a 2-almost prime. 30 is a 3-almost prime. 5 is a 1-almost prime (i.e., a prime).

- a. Find the first few k -almost primes for $1 \leq k \leq 1000$. Can you say anything about the numbers that appear in each class?
- b. Prove that (p, q) is a twin prime pair if and only if $p(p+2)$ is a 2-almost prime.
- c. How does the smallest k -almost prime grow with k ?
- d. Determine a counting function for k -almost primes.

1.38 (Smooth numbers).

Definition 1.75. An integer is called N -smooth if all its prime factors are less than N .

Example 1.76. 10 is 5-smooth (and 6-smooth and 7-smooth, etc.)

Smooth does not mean small; it means small prime factors. For example, 1,073,741,824 is 2-smooth.

Smooth numbers are important in cryptographic applications which rely on factoring.

- a. Find the first few k -smooth numbers in numerical order for various k .
- b. Find a counting function for the number of k -smooth numbers.
- c. For a prime p , determine k such that $p - 1$ is k -smooth. Is there a relationship between k and p ?

- d. For an integer n , determine k such that $n^2 + 1$ is k -smooth. Is there a relationship between k and n ?
- e. Can you find a function that generates numbers that are all k -smooth? (Or are all not k -smooth?)

1.39 (Factorization algorithms). There are many factoring algorithms. Here we explore just a few that use only what we have learned so far.

- Trial division

Input: positive integer n
Output: a factor of n
Algorithm:
 1: Let $a = 2$.
 2: While $a \nmid n$ and $a \leq \sqrt{n}$.
 a: $a = a + 1$.
 3: If $a > \sqrt{n}$ return n , else return a .

- Fermat factorization (for odd integers)

The basic idea is that if we can write $n = x^2 - y^2$, then we can write $n = (x + y)(x - y)$. If $x - y \neq 1$ and $x + y \neq 1$, then we have found a factor. Note that we can always do this since if $n = ab$, then $n = \left(\frac{a+b}{2}\right)^2 - \left(\frac{a-b}{2}\right)^2$.

Input: positive odd integer n
Output: a factor of n
Algorithm:
 1: Let a be the smallest integer larger than \sqrt{n} .
 2: Check if $a^2 - n$ is a square ($\neq 1$). If it is, then we are done. If not, then we change a .
 3: While $(a^2 - n)$ is not a square).
 a: $a = a + 1$.
 4: Return $a - \sqrt{a^2 - n}$.

- Euler factorization

This method relies on writing n as the sum of squares in two different ways. Say $n = a^2 + b^2 = c^2 + d^2$. Let $x = \gcd(a - c, d - b)$ and $y = \gcd(a + c, b + d)$. Then we can find $a - c = xr$ and $a + c = ys$. Then we may write

$$n = \frac{1}{4}(x^2 + y^2)(r^2 + s^2)$$

or

$$n = \left(\left(\frac{x}{2} \right)^2 + \left(\frac{y}{2} \right)^2 \right) (r^2 + s^2).$$

The problem of writing a number as the sum of two squares will be more fully explored in section 6.4. For now we adopt an algorithm similar to the search in Fermat factorization.

Input: positive odd integer n
Output: a factorization of n
Algorithm:
 1: Let $a = 1$.
 2: Repeat until we have found 2 different representations of n as the sum of two squares or $a > \sqrt{n}$.
 a: Check if $n - a^2$ is a square ($\neq 1$). If it is, then store $(a, \sqrt{n - a^2})$. If not, then we change a .
 b: While $(n - a^2$ is not a square).
 i: $a = a + 1$.
 3: Return $(\frac{1}{4}(x^2 + y^2), r^2 + s^2)$.

- Pollard Rho

The idea is that if we apply a suitably generic function, the remainder after division (see Chapter 2) will be suitably random. Thus, Pollard Rho uses a function to generate a finite-length pseudo-random sequence to test for divisibility.

Input: n
Output: a factor of n
Algorithm:
 1: Choose a function $f(x)$ that is not too simple.
 2: Let $x = 2, y = 2, d = 1$.
 3: While $(d = 1)$.
 a: Let $t = f(x)$.
 b: Let $s = f(f(y))$.
 c: Compute $d = \gcd(|t - s|, n)$.
 4: If $d = n$ return failure, else return d .

Note that this can return failure for composite n . In that case, simply try a different function $f(x)$.

- Lucas–Lehmer test for Mersenne numbers
 Determine if $2^p - 1$ is prime.

Input: p
Output: prime or composite
Algorithm:
 1: $s = 4, n = 2^p - 1$.
 2: Repeat $p - 2$ times.
 a: Let $t = s^2 - 2$.
 b: Define s as the remainder after t is divided by n .
 3: If $s = 0$, return prime, else return composite.

Questions:

- a. Implement each of the above algorithms.
- b. Can you find a number for each algorithm above which is factored faster with that method than the others?
- c. Can you find a number for each algorithm which factors in a reasonable amount of time, but does not for the other algorithms?
- d. The algorithms above are simple examples of these methods. Can you improve the implementation?
- e. Find the largest prime you can.
- f. If the computer system you are using has a built-in factoring function, determine if this algorithm is better than your functions. Why do you think that is?

Modular Arithmetic

1. Basic Arithmetic

We saw in Chapter 1 that given any two positive integers n and m , we can always split a collection of n objects into groups of m with possibly some number left over. We expressed this mathematically as the division algorithm (Theorem 1.8),

$$n = qm + r, \quad 0 \leq r < m.$$

We call this unique r the *remainder after division by m* . In this chapter we explore the arithmetic of these remainders. For example, consider two people each with some number of apples. One person has one apple left over after dividing his apples into groups of five and the other person has two apples left over when dividing her apples into groups of five. If these two people combine their apples, the resulting total number of apples has three left over when divided into groups of five. Notice that the total remainder does not depend how many groups of five each person has, only the remainder left over. We have *added* their remainders together.



Another familiar example is a 12-hour clock, which divides the time into groups of twelve, and the current time is the remainder. For example, if it is currently 11 a.m., what time is 4 hours later? The answer is of course 3 p.m. The mathematics behind this answer is to take $11 + 4 = 15$ and consider the remainder after division by 12:

$$15 = 1 \cdot 12 + 3.$$

To make a formal definition we use the language of divisibility.

Definition 2.1. Let a , r , and n be integers with $n > 0$. We say that a is *congruent to r modulo n* if n divides $(a - r)$, i.e., $n \mid (a - r)$, and we write $a \equiv r \pmod{n}$. The number n is called the *modulus*.

Example 2.2. Our clock example becomes

$$15 \equiv 3 \pmod{12} \quad \text{since} \quad 12 \mid (15 - 3).$$

Keep in mind that the previous definition is really just remainder after division as in the division algorithm. If $a = qn + r$, then $a - r = qn$, so $n \mid (a - r)$. Our new definition states this as $a \equiv r \pmod{n}$; in other words, a number is congruent to its remainder after division by the modulus. However, unlike the remainder, congruence is not unique. There are infinitely many numbers that have a remainder of 2 after division by 7 and we say they are all congruent. For example,

$$-5 \equiv 2 \equiv 9 \equiv 16 \equiv 23 \pmod{7}.$$

Definition 2.3. We call the set of numbers all congruent to the same value modulo n a *residue class modulo n* . For example, $\{\dots, -12, -5, 2, 9, 16, \dots\}$ is a residue class modulo 7. We denote a residue class with a bar

$$\bar{2} = \{\dots, -12, -5, 2, 9, 16, \dots\},$$

meaning every integer congruent to 2 modulo 7.

Investigation 2.4. We noted above that the remainder of a sum is the sum of the remainders. Written in residue notation, $\overline{a + b} = \bar{a} + \bar{b}$.

- (a) Verify with a few examples, that if you add $\bar{2} + \bar{3}$ modulo 7, then it does not matter which element of the residue classes you choose.
- (b) What about subtraction? Is it true that $\overline{a - b} = \bar{a} - \bar{b}$?
- (c) What about multiplication? Is it true that $\overline{ab} = \bar{a}\bar{b}$?
- (d) What about division? How might you define $\frac{\bar{a}}{\bar{b}}$?

We know that there are infinitely many integers in each residue class, so we might ask the following question.

Question 2.5. Given a positive integer n , how many different residue classes are there modulo n ?

The division algorithm provides the answer. Every number is congruent to exactly one number from the set of possible remainders $\{0, 1, \dots, n - 1\}$. Because the difference between any two remainders is not divisible by n , no two remainders are congruent. So there are n different residue classes modulo n . We could take a slightly different set, such as $\{1, \dots, n - 1, n\}$ or $\{2, 3, \dots, n, n + 1\}$, and still have every number congruent to exactly one element of the set. This is because we are choosing exactly one representative from each residue class.

Definition 2.6. A *complete system of residues modulo n* is a set of integers such that every integer is congruent to exactly one element of the set.

Example 2.7. Consider $n = 3$. Then the following sets are all complete systems of residues modulo 3.

- $\{0, 1, 2\}$
- $\{3, 7, 11\}$
- $\{-3, -2, -1\}$
- $\{51, -38, 104\}$

Since the definition of congruence is based on divisibility, it is not surprising that we can use congruences to test divisibility. Thinking of congruences as remainders after division, this is even less surprising. We have divisibility when we have remainder 0.

Theorem 2.8. *Let a, n be integers with $n > 0$. Then*

$$a \equiv 0 \pmod{n}$$

if and only if n divides a .

Proof. If $a \equiv 0 \pmod{n}$, the definition of congruence says there is an integer k such that

$$a = kn$$

so that n divides a .

If n divides a , by the definition of divisibility there is an integer k such that

$$a = kn$$

so that $a \equiv 0 \pmod{n}$. □

We saw in the previous chapter that if $n \mid a$ and $n \mid b$, then $n \mid a + b$. Written in the language of congruences, this implication is the same as

$$\text{if } a \equiv 0 \pmod{n} \text{ and } b \equiv 0 \pmod{n}, \text{ then } a + b \equiv 0 \pmod{n}.$$

Investigation 2.9. This notion of using remainder 0 for divisibility is a powerful proof technique which occurs frequently in number theory.

- (a) Use this fact to prove that a number m is divisible by 3 if and only if the sum of its digits is divisible by 3.
- (b) Can you find a divisibility test for 9 similar to the test for divisibility by 3?

We stated without proof that the addition of integers always leads to addition of residues. The following theorem proves this fact for addition as well as describing multiplication and “division” of residue classes. We call these induced actions on residue classes *modular arithmetic*.

Theorem 2.10. *Let a, b, c, d be integers, and let n be a positive integer. Assume that $a \equiv b \pmod{n}$ and $c \equiv d \pmod{n}$.*

(a) $a + c \equiv b + d \pmod{n}$.

(b) $ac \equiv bd \pmod{n}$.

Let $g = \gcd(c, n)$.

(c) *If $ac \equiv bc \pmod{n}$, then $a \equiv b \pmod{n/g}$.*

Proof. We are assuming that $a \equiv c \pmod{n}$ and $b \equiv d \pmod{n}$. Using the definition of congruence, there are integers q_1, q_2, q_3, q_4, r_1 , and r_2 such that

$$a = q_1n + r_1,$$

$$b = q_2n + r_1,$$

$$c = q_3n + r_2,$$

$$d = q_4n + r_2.$$

Notice that the remainder after division by n is r_1 for both a and b and is r_2 for both c and d . The first two statements in the theorem come from basic arithmetic.

(a) We write

$$a + c = (q_1n + r_1) + (q_2n + r_2) = (q_1 + q_2)n + (r_1 + r_2),$$

$$b + d = (q_3n + r_1) + (q_4n + r_2) = (q_3 + q_4)n + (r_1 + r_2).$$

Note the remainders after division by n are the same for both sums, $r_1 + r_2$, so we have verified the congruence.

(b) We write

$$ac = (q_1n + r_1)(q_2n + r_2) = q_1q_2n^2 + (q_1r_2 + q_2r_1)n + r_1r_2,$$

$$bd = (q_3n + r_1)(q_4n + r_2) = q_3q_4n^2 + (q_3r_2 + q_4r_1)n + r_1r_2.$$

Since we can write the n^2 terms as an integer times n ,

$$q_1q_2n^2 + (q_1r_2 + q_2r_1)n = (q_1q_2n + q_1r_2 + q_2r_1)n,$$

$$q_3q_4n^2 + (q_3r_2 + q_4r_1)n = (q_3q_4n + q_3r_2 + q_4r_1)n,$$

the remainders after division by n are the same, r_1r_2 .

For the last statement we translate the assumption

$$ac \equiv bc \pmod{n}$$

to mean

$$ac = bc + kn$$

for some integer k . Now we divide through by c , to get

$$(10) \quad a = b + \frac{kn}{c}.$$

Since the left-hand side is an integer and b is an integer, we know $\frac{kn}{c}$ must be an integer. In other words, c divides kn . It may be that some factors of c divide n and

some divide k . The largest factor of c that divides n is $g = \gcd(c, n)$. If we write $c = c'g$, then we have

$$\frac{kn}{c} = \frac{k}{c'} \frac{n}{g}.$$

Since $\frac{kn}{c}$ and $\frac{n}{g}$ are integers, then $\frac{k}{c'}$ is an integer, call it k' . So we have from (10)

$$a = b + k' \frac{n}{g}.$$

Since k' is an integer, the integer $\frac{n}{g}$ divides $a - b$, the definition of congruence. \square

Example 2.11. Consider $n = 4$. Then we know that $6 \equiv 10 \pmod{4}$ and $11 \equiv 23 \pmod{4}$. We illustrate the three properties.

(a) We compute

$$\begin{aligned} 6 + 11 &= 17 = 4 \cdot 4 + 1 \equiv 1 \pmod{4}, \\ 10 + 23 &= 33 = 4 \cdot 8 + 1 \equiv 1 \pmod{4}. \end{aligned}$$

(b) We compute

$$\begin{aligned} 6 \cdot 11 &= 66 = 4 \cdot 16 + 2 \equiv 2 \pmod{4}, \\ 10 \cdot 23 &= 230 = 4 \cdot 57 + 2 \equiv 2 \pmod{4}. \end{aligned}$$

(c) Consider $a = 9, b = 7, c = 6$ modulo 4. We compute

$$\begin{aligned} 9 \cdot 6 &= 54 = 4 \cdot 13 + 2 \equiv 2 \pmod{4}, \\ 7 \cdot 6 &= 42 = 4 \cdot 10 + 2 \equiv 2 \pmod{4}. \end{aligned}$$

Since $\gcd(6, 4) = 2$, the conclusion is that

$$\begin{aligned} 9 \cdot 6 &\equiv 7 \cdot 6 \pmod{4} \\ \Downarrow \\ 9 &\equiv 7 \pmod{2}. \end{aligned}$$

Note that the conclusion is not true modulo 4, i.e., we have

$$9 \not\equiv 7 \pmod{4}.$$

In other words, given

$$ac \equiv bc \pmod{n},$$

we cannot simply cancel c from both sides. Instead, we must take into consideration the $\gcd(c, n)$.

If we take the special case of Theorem 2.10 where $d = c$, we may state the following simple corollary that looks more like what you are familiar with from integer arithmetic: you can add or multiply the same value to both sides of a congruence.

Corollary 2.12. *Let n be a positive integer, and assume that $a \equiv b \pmod{n}$. Then for any integer c ,*

- (a) $a + c \equiv b + c \pmod{n}$.
- (b) $ac \equiv bc \pmod{n}$.

Example 2.13. Consider $2 \equiv 7 \pmod{5}$. If we add 11 to both sides, we get

$$\begin{aligned} 2 + 11 &= 13 \equiv 3 \pmod{5}, \\ 7 + 11 &= 18 \equiv 3 \pmod{5}. \end{aligned}$$

In particular,

$$13 \equiv 18 \pmod{5}.$$

If we multiply both sides by -2 , we get

$$\begin{aligned} 2(-2) &= -4 \equiv 1 \pmod{5}, \\ 7(-2) &= -14 \equiv 1 \pmod{5}. \end{aligned}$$

In particular,

$$-4 \equiv -14 \pmod{5}.$$

We have seen that modular arithmetic follows basically the same rules as integer arithmetic, except you have to be careful about “division”. Let’s now clarify the notion of division with the more general notion of inverse.

2. Inverses and Fermat’s Little Theorem

An important use of division is to solve linear equations of the form

$$ax = b.$$

Unless a is zero, dividing both sides by a allows us to solve for x . Notice that dividing by a is the same as multiplying both sides by the inverse of a , $\frac{1}{a}$. When trying to generalize to modular arithmetic, the key property to keep in mind is that

$$a \cdot \frac{1}{a} = 1.$$

Thus, division by a number is the same as multiplying by its inverse. Even though we do not have fractions in modular arithmetic, we sometimes have inverses.

Definition 2.14. We say that a is the *inverse* of b modulo m if

$$a \cdot b \equiv 1 \pmod{m}.$$

Example 2.15. Working modulo 15 we see that

$$2 \cdot 8 \equiv 1 \pmod{15},$$

so 2 is the inverse of 8 modulo 15. We could also turn this around and say that 8 is the inverse of 2 modulo 15.

As another example

$$4 \cdot 4 \equiv 1 \pmod{15},$$

so 4 is its own inverse modulo 15.

However, inverses modulo m do not always exist.

Example 2.16. We can check that 2 does not have an inverse modulo 4 by trying every possible residue class $\{\overline{0}, \overline{1}, \overline{2}, \overline{3}\}$. Since modular arithmetic depends only on the residue class, we need to consider just one representative from each class. So we can easily check this claim by checking whether any of $\{0, 1, 2, 3\}$ are inverses of 2:

$$2 \cdot 0 \equiv 0 \pmod{4},$$

$$2 \cdot 1 \equiv 2 \pmod{4},$$

$$2 \cdot 2 \equiv 0 \pmod{4},$$

$$2 \cdot 3 \equiv 2 \pmod{4}.$$

Since none of the multiplications results in 1 and we have tried every possible residue class, 2 does not have an inverse modulo 4.

Remark. Example 2.16 illustrates a very useful feature of modular arithmetic for computing purposes. Because there are only finitely many residue classes, you can solve many problems simply by checking every residue class!

Question 2.17. Given an integer m , which residue classes have an inverse modulo n ?

Investigation 2.18.

- (a) Choose some values for m and determine which residue classes have inverses.
- (b) Is there any pattern to which numbers do or do not have inverses?
- (c) Given a specific m , can you find a formula for the number of residue classes that have inverses?

The following table compiles some computational data on the existence of inverses modulo various n . Can you find a pattern?

n	Has inverse	No inverse
2	$\{1\}$	$\{0\}$
3	$\{1, 2\}$	$\{0\}$
4	$\{1, 3\}$	$\{0, 2\}$
5	$\{1, 2, 3, 4\}$	$\{0\}$
6	$\{1, 5\}$	$\{0, 2, 3, 4\}$
7	$\{1, 2, 3, 4, 5, 6\}$	$\{0\}$
8	$\{1, 3, 5, 7\}$	$\{0, 2, 4, 6\}$

The following theorem proves that numbers relatively prime to the modulus have inverses.

Theorem 2.19. *Given a positive integer n and an integer a relatively prime to n (i.e., $\gcd(a, n) = 1$), there is a unique residue class b modulo n such that $ab \equiv 1 \pmod{n}$.*

Proof. We know from Theorem 1.22 that the greatest common divisor of two integers can be written as a linear combination. Thus, if $\gcd(a, n) = 1$, there are integers s and t such that

$$sa + tn = 1.$$

Since $tn \equiv 0 \pmod{n}$, we have

$$sa \equiv 1 \pmod{n},$$

and s is the inverse of a modulo n .

To prove uniqueness, assume that x and y are both inverses of a modulo n . Then we have

$$ax \equiv ay \equiv 1 \pmod{n}.$$

Since $\gcd(a, n) = 1$, we have from Theorem 2.10(c)

$$x \equiv y \pmod{n}. \quad \square$$

Note that the previous theorem states the existence of an inverse but does not give a value. However, from the proof, we see that the extended Euclidean algorithm can be used to find the inverse. We describe the process in Algorithm 2.1.

Algorithm 2.1. Inverses from the Euclidean Algorithm

Input: two positive integers (n, a) that are relatively prime

Output: the inverse of a modulo n

Algorithm:

1: Find x, y from the Euclidean algorithm such that

$$xa + yn = 1.$$

2: Return x .

Example 2.20. Consider $n = 14$ and $a = 9$. We compute

$$2 \cdot 14 - 3 \cdot 9 = 1$$

so that the inverse of 9 modulo 14 is -3 and

$$-3 \equiv 11 \pmod{14}.$$

You may have remarked from your data on inverses, or as an easy consequence of Theorem 2.19, that every nonzero residue class has an inverse when the modulus is prime. This is because a prime number is relatively prime to every integer not a multiple of itself. In particular, when the modulus is a prime p , there are $p - 1$ residue classes relatively prime to p . Fermat capitalized on this fact in the next theorem.

Theorem 2.21 (Fermat's Little Theorem). *If p is a prime and a an integer not divisible by p , then*

$$a^{p-1} \equiv 1 \pmod{p}.$$

We need a lemma to help prove this theorem.

Lemma 2.22. Let p be a prime number. For an integer a with $\gcd(a, p) = 1$, the set of numbers $\{a, 2a, 3a, \dots, (p-1)a\}$ is a complete set of (nonzero) residues. (i.e., it is equivalent modulo p to the set $\{1, 2, \dots, p-1\}$).

Proof. Since a is relatively prime to p , all the numbers $\{a, 2a, 3a, \dots, (p-1)a\}$ are relatively prime to p , in particular, none of them are congruent to 0 modulo p . Thus, each number ka is equivalent to some number in the set $\{1, 2, \dots, p-1\}$. We just have to show there are no repeats.

Assume that $ka \equiv ja \pmod{p}$. Since $\gcd(a, p) = 1$, we can apply Theorem 2.10(c) to have

$$k \equiv j \pmod{p}.$$

Thus, each of the numbers in the set $\{a, 2a, 3a, \dots, (p-1)a\}$ is distinct. \square

We will need the notion of a factorial in the proof of Fermat's Little Theorem.

Definition 2.23. For a positive integer n , we define the *factorial* of n as the product of all positive integers less than or equal to n :

$$n! = n \cdot (n-1) \cdot (n-2) \cdots 2 \cdot 1.$$

For the ease of writing formulas, we also define

$$0! = 1.$$

Example 2.24. $5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$.

Proof of Theorem 2.21. From Lemma 2.22 we know that the set $\{a, 2a, \dots, (p-1)a\}$ is a complete set of (nonzero) residues. The set $\{1, 2, \dots, p-1\}$ is also a complete set of (nonzero) residues. Since modular arithmetic does not depend on the choice of representative in the residue class, we can equate the products of the two sets of residues

$$\prod_{k=1}^{p-1} ka \equiv (p-1)! \pmod{p}.$$

In particular,

$$a^{p-1}(p-1)! \equiv (p-1)! \pmod{p}.$$

Since each number $1, \dots, (p-1)$ is relatively prime to p , we cancel the factorial term from both sides with Theorem 2.10(c) to get

$$a^{p-1} \equiv 1 \pmod{p}. \quad \square$$

Investigation 2.25. Does p have to be prime for Fermat's Little Theorem to be true? Here are some related questions to think about.

- (a) Is there a composite m such that $a^{m-1} \equiv 1 \pmod{m}$ for all a relatively prime to m ?
- (b) If there is an a relatively prime to m such that $a^{m-1} \not\equiv 1 \pmod{m}$, is m necessarily composite?
- (c) Can you determine the correct generalization for Fermat's Little Theorem to composite moduli?

Primes are useful building blocks for numbers and are used in certain kinds of cryptographic systems. In Question 1.52 we asked about determining whether a given number is prime, a primality test. Fermat's Little Theorem gives a test for composite numbers, that is, a negative test for primality. Note that even if we determine that n is composite with this test, we still do not know any of its factors.

Corollary 2.26. *Let p be a positive integer. If there exists an integer a with $\gcd(a, p) = 1$ such that*

$$a^p \not\equiv a \pmod{p},$$

then p is not prime.

Unfortunately, this statement is not an “if and only if” statement. There can be composite numbers n such that there are integers a with $\gcd(a, n) = 1$ and

$$a^n \equiv a \pmod{n}.$$

The composite values that satisfy Fermat's Little Theorem are called *Fermat pseudoprimes base a* .

Example 2.27. For $a = 5$ and $p = 124$, we have

$$5^{123} \equiv 1 \pmod{124},$$

but $124 = 4 \cdot 31$ is composite. We can find a different a value that does not satisfy Fermat's Little Theorem. For example, $a = 3$ gives

$$3^{123} \equiv 27 \pmod{341}.$$

However, there are composite numbers n which pass Fermat's Little Theorem test for every choice of a relatively prime to n . In other words, for every value of a with $\gcd(a, n) = 1$,

$$a^{n-1} \equiv 1 \pmod{n}.$$

Such n are called *Carmichael numbers*. The smallest is 561.

Question 2.28. Can we generalize Fermat's Little Theorem? In other words, given a positive integer n , with n not necessarily prime, can we find a positive integer e such that for all integers a with $\gcd(a, n) = 1$ we have

$$a^e \equiv 1 \pmod{n}?$$

Let's generate some data.

n	a	e	n	a	e
4	1	1	10	1	1
	3	2		3	4
6	1	1		7	4
	5	2		9	2
8	1	1	12	1	1
	3	2		5	2
	5	2		7	2
	7	2		11	2
9	1	1	14	1	1
	2	6		3	6
	4	3		5	6
	5	6		9	6
	7	3		11	6
	8	2		13	6

It seems clear that there is a particular value of e that works for all a for each n . For $n = 6$, we have $e = 2$; for $n = 9$, we have $e = 6$; etc. However, identifying how e relates to n is not as clear. You may notice that in five out of the seven examples, the e that works is the same as the number of rows, that is, the number of a values relatively prime to n . In the other two examples, it is exactly half that value.

Definition 2.29. We define the *Euler totient function* $\varphi(n)$, also called the *Euler phi function*, for a positive integer n as the number of integers between 1 and n which are relatively prime to n .

Example 2.30. We compute a few values of $\varphi(n)$.

n	relatively prime numbers	$\varphi(n)$
2	{1}	1
3	{1, 2}	2
4	{1, 3}	2
5	{1, 2, 3, 4}	4
6	{1, 5}	2
7	{1, 2, 3, 4, 5, 6}	6
8	{1, 3, 5, 7}	4
9	{1, 2, 4, 5, 7, 8}	6
10	{1, 3, 7, 9}	4

Rephrasing our observation above, the values of $\varphi(n)$ correspond to the e values in five out of seven examples. In the other two cases the experimental e value divides $\varphi(n)$. For example, $n = 8$ has experimentally $e = 2$ and $\varphi(8) = 4$.

We can now generalize Fermat's Little Theorem.

Theorem 2.31 (Euler's formula). *Let n be a positive integer, and let a be any integer. If $\gcd(a, n) = 1$, then $a^{\varphi(n)} \equiv 1 \pmod{n}$.*

Proof. This proof proceeds similarly to the proof of Fermat's Little Theorem (Theorem 2.21).

Let $\{m_1, \dots, m_{\varphi(n)}\}$ be the set of numbers relatively prime to n . Since a is relatively prime to n , multiplication by a permutes this set. (Check that $\gcd(am_i, n) = 1$ for all i and that $am_i \not\equiv am_j \pmod{n}$ for $m_i \not\equiv m_j$.) In other words, the sets $\{am_1, \dots, am_{\varphi(n)}\}$ and $\{m_1, \dots, m_{\varphi(n)}\}$ consist of the same residue classes modulo n . Then the product of those two sets must be the same:

$$\prod_{i=1}^{\varphi(n)} m_i a \equiv \prod_{i=1}^{\varphi(n)} m_i \pmod{n}.$$

Since each number m_i is relatively prime to n , their product is also relatively prime to n , and we can apply Theorem 2.10(c) to arrive at

$$a^{\varphi(n)} \equiv 1 \pmod{n}. \quad \square$$

We will discuss the Euler totient function in more detail in Chapter 5.

2.1. Application: Linear Congruences.

Definition 2.32. A *linear congruence equation* is an equation of the form

$$ax \equiv b \pmod{n},$$

where a , b , and n are fixed integers and x is the unknown variable.

When $\gcd(a, n) = 1$, we can solve a linear congruence by multiplying both sides of the equation by the inverse of a .

Example 2.33. Consider

$$3x \equiv 5 \pmod{17}.$$

The inverse of 3 modulo 17 is 6. We multiply both sides of the equation by 6 to get

$$18x \equiv 30 \pmod{17}.$$

Since $18 \equiv 1 \pmod{17}$ and $30 \equiv 13 \pmod{17}$, we have

$$x \equiv 13 \pmod{17}.$$

If $\gcd(a, n) \neq 1$, then a does not have an inverse and this procedure does not work. However, that does not mean there are no solutions.

Example 2.34. For the equation

$$2x \equiv 4 \pmod{6},$$

we can try representatives from all possible residue classes $\{0, 1, 2, 3, 4, 5\}$ to see that

$$x \equiv 2 \pmod{6} \quad \text{and} \quad x \equiv 5 \pmod{6}$$

are both solutions to this equation.

Example 2.34 is interesting for two reasons. First, there is no inverse of 2 modulo 6, so we cannot solve this equation in the “normal” way, by multiplying both sides by the inverse of a . Second, not only is there a solution, but there are multiple distinct solutions. This is in direct contrast to regular polynomial equations. The Fundamental Theorem of Algebra says that a polynomial of degree d has at most d roots. Here, we have a polynomial of degree 1 (linear), which has two roots!

Question 2.35. For which integers a , b , and n is it possible to solve the congruence equation

$$ax \equiv b \pmod{n}?$$

When it is solvable, how many solutions does it have?

Investigation 2.36. Gather numerical data on Question 2.35.

- Find an equation with exactly one solution. What conditions on a , b , and n guarantee exactly one solution?
- Find an equation with 2, 3, 4, or more solutions. What conditions on a , b , and n determine the number of solutions?
- Find an equation with no solutions. What conditions on a , b , and n guarantee there are no solutions?

3. Linear Congruences and the Chinese Remainder Theorem

As a first step to resolving Question 2.35, we can try to find a linear equation that does not have any solutions. If there are no solutions, we must have $\gcd(a, n) \neq 1$; otherwise, a has an inverse modulo n , and we can solve the congruence. Choosing a and n with $\gcd(a, n) \neq 1$, we look for linear equations with no solutions. Since we hope to uncover some kind of pattern, we probably should not consider random equations; instead, we should examine a well-defined set of equations. For example, we could fix a and n that have a nontrivial greatest common divisor and try to solve

$$ax \equiv b \pmod{n}$$

for all possible residue classes b . The following data is for $a = 3, n = 12$, and $b \in \{0, 1, \dots, 11\}$.

Solutions	No Solutions
$3x \equiv 0 \pmod{12}$	$3x \equiv 1 \pmod{12}$
$3x \equiv 3 \pmod{12}$	$3x \equiv 2 \pmod{12}$
$3x \equiv 6 \pmod{12}$	$3x \equiv 4 \pmod{12}$
$3x \equiv 9 \pmod{12}$	$3x \equiv 5 \pmod{12}$
	$3x \equiv 7 \pmod{12}$
	$3x \equiv 8 \pmod{12}$
	$3x \equiv 11 \pmod{12}$

We see that there are definitely equations which do not have solutions, and they seem to all have b relatively prime to a . The next example analyzes the details of one of the equations with solutions from this set of data.

Example 2.37. If we recall the properties of “division” from Theorem 2.10(c), we know that

$$3x \equiv 3 \pmod{12}$$

implies that

$$x \equiv 1 \pmod{4}.$$

The residues classes $\{1, 5, 9\}$ modulo 12 are all equivalent to 1 modulo 4, and we can check that $x \in \{1, 5, 9\}$ are all the solutions to

$$3x \equiv 3 \pmod{12}.$$

By looking at this particular example in detail, we are able to not only have an idea as to which equations have solutions, but also why those particular equations have solutions. You will see that the proof of the general theorem below is based exactly on what we did to find the solutions to

$$3x \equiv 3 \pmod{12}.$$

Theorem 2.38. *Given a positive integer n and two integers a and b with $\gcd(a, n) = d$, the linear congruence $ax \equiv b \pmod{n}$ has a solution if and only if d divides b . Moreover, there are exactly d residue classes of solutions. Given one solution x_0 , then the solutions are*

$$x_0 + t \frac{n}{d}, \quad 0 \leq t < d.$$

Proof. Assume first that $ax \equiv b \pmod{n}$ has a solution x_0 . Then we know that there is an integer y such that

$$ax_0 + ny = b.$$

From Theorem 1.22 we know that $\gcd(a, n)$ divides b . This proves one direction of the “if and only if” statement.

Assume now that the $\gcd(a, n) = d$ divides b . Then we can reduce the equation to

$$\frac{a}{d}x \equiv \frac{b}{d} \pmod{\frac{n}{d}}.$$

Since $\frac{a}{d}$ is an integer relatively prime to the integer $\frac{n}{d}$, we can solve this equation by finding the inverse of $\frac{a}{d}$ modulo $\frac{n}{d}$. Call this solution x_0 . Notice that x_0 is also a solution modulo n since

$$\frac{a}{d}x_0 - \frac{b}{d} = k \frac{n}{d}$$

for some integer k implies

$$ax_0 - b = kn.$$

Then for any integer t

$$x = x_0 + t \frac{n}{d}$$

is also a solution modulo $\frac{n}{d}$ (and modulo n). Furthermore, if x is any solution modulo n , then

$$ax \equiv ax_0 \equiv b \pmod{n},$$

$$a(x - x_0) \equiv 0 \pmod{n},$$

$$\frac{a}{d}(x - x_0) \equiv 0 \pmod{\frac{n}{d}}.$$

Since $\frac{a}{d}$ has an inverse modulo $\frac{n}{d}$, then

$$x - x_0 \equiv 0 \pmod{\frac{n}{d}}.$$

In other words, there is an integer t such that

$$x - x_0 = t \frac{n}{d}.$$

We have shown that every solution is of the form $x_0 + t \frac{n}{d}$ and every integer of the form $x_0 + t \frac{n}{d}$ is a solution. Now we just have to count how many of the $x_0 + t \frac{n}{d}$ are distinct modulo n . Let s and t be integers. Then we have

$$x_0 + s \frac{n}{d} \equiv b \pmod{n},$$

$$x_0 + t \frac{n}{d} \equiv b \pmod{n},$$

and taking their difference yields

$$(t - s) \frac{n}{d} \equiv 0 \pmod{n}.$$

This equation is the same as

$$(t - s) \frac{n}{d} = kn$$

for some integer k , which is the same as

$$t - s = kd,$$

which is the same as

$$t \equiv s \pmod{d}.$$

Thus, the distinct residue classes modulo d give the distinct solutions. \square

Example 2.39. Solve

$$21x \equiv 9 \pmod{30}.$$

We compute $\gcd(21, 30) = 3$ and that $3 \mid 9$, so there are three solutions. To find one solution, we consider

$$7x \equiv 3 \pmod{10}.$$

We solve this by computing the inverse of 7 modulo 10, which is 3. So we have

$$x \equiv 9 \pmod{10}.$$

Then the rest of the solutions are

$$\left\{ 9, 9 + 1 \frac{30}{3}, 9 + 2 \frac{30}{3} \right\} = \{9, 19, 29\}.$$

We can verify

$$21(9) \equiv 189 \equiv 9 \pmod{30},$$

$$21(19) \equiv 399 \equiv 9 \pmod{30},$$

$$21(29) \equiv 609 \equiv 9 \pmod{30}.$$

We have completely answered Question 2.35.

Investigation 2.40. We mentioned earlier that a linear congruence having multiple solutions contrasts with the Fundamental Theorem of Algebra, which says that a degree n polynomial equation can have at most n solutions. Make a conjecture about the number of solutions of higher degree polynomial congruences. As a place to start you could consider the following polynomials for various choices of n :

- (a) $x^2 + x + 1 \equiv 0 \pmod{n}$,
- (b) $x^2 + x \equiv 0 \pmod{n}$,
- (c) $x^3 + x \equiv 0 \pmod{n}$.

We next consider the problem of solving several linear congruences simultaneously. We begin with a motivating problem found in the *Sun Tzu Suan Ching*, an early Chinese textbook on arithmetic dating from about the third century C.E.

Problem 2.41. We have a number of things, but we do not know exactly how many. If we count them by threes, we have two left over. If we count them by fives, we have three left over. If we count them by sevens, we have two left over. How many things are there?

We can translate the problem into three linear equations:

$$\begin{aligned}x &\equiv 2 \pmod{3}, \\x &\equiv 3 \pmod{5}, \\x &\equiv 2 \pmod{7}.\end{aligned}$$

We can check that any solution x must satisfy

$$x \equiv 23 \pmod{105}.$$

In general, we are considering systems of linear congruences with different moduli. For example, for $m_1 \neq m_2$, consider a set of equations of the form

$$\begin{aligned}a_1x &\equiv b_1 \pmod{m_1}, \\a_2x &\equiv b_2 \pmod{m_2}.\end{aligned}$$

We ask the following general question.

Question 2.42. Which systems of linear congruences are solvable? When the system is solvable, can you find a solution?

Investigation 2.43.

- (a) Find a system of linear congruences that has a solution.
- (b) Find a system of linear congruences that does not have a solution.
- (c) Can you determine a criterion to tell if a linear system of congruences has a solution?

The author Sun Tzu gave a general method for solving many such problems, and we call it the Chinese Remainder Theorem.

Theorem 2.44 (Chinese Remainder Theorem). *Let n_1, \dots, n_r be positive integers such that any pair is relatively prime. Then the system of linear congruences*

$$x \equiv a_1 \pmod{n_1}$$

$$\vdots$$

$$x \equiv a_k \pmod{n_k}$$

has a unique solution modulo $N = \prod_{i=1}^k n_i$. The solution is given by

$$x = a_1 \frac{N}{n_1} y_1 + \dots + a_k \frac{N}{n_k} y_k,$$

where y_i is the inverse of $\frac{N}{n_i}$ modulo n_i .

Before proving the theorem, let's see how to solve Problem 2.41.

Example 2.45. We are trying to solve

$$x \equiv 2 \pmod{3},$$

$$x \equiv 3 \pmod{5},$$

$$x \equiv 2 \pmod{7}.$$

The three moduli are pairwise relatively prime, that is,

$$\gcd(3, 5) = \gcd(3, 7) = \gcd(5, 7) = 1,$$

so the Chinese Remainder Theorem gives a solution modulo $N = 3 \cdot 5 \cdot 7 = 105$.

Using the notation from the theorem, we have

$$a_1 = 2, \quad n_1 = 3,$$

$$a_2 = 3, \quad n_2 = 5,$$

$$a_3 = 2, \quad n_3 = 7,$$

and

$$N = 3 \cdot 5 \cdot 7 = 105,$$

$$m_1 = \frac{N}{n_1} = 5 \cdot 7 = 35,$$

$$m_2 = \frac{N}{n_2} = 3 \cdot 7 = 21,$$

$$m_3 = \frac{N}{n_3} = 3 \cdot 5 = 15.$$

We need the inverses of the m_i modulo n_i , which we call y_i :

$$2 \cdot 35 \equiv 1 \pmod{3} \Rightarrow y_1 = 2,$$

$$1 \cdot 21 \equiv 1 \pmod{5} \Rightarrow y_2 = 1,$$

$$1 \cdot 15 \equiv 1 \pmod{7} \Rightarrow y_3 = 1.$$

So we have

$$x \equiv a_1 m_1 y_1 + a_2 m_2 y_2 + a_3 m_3 y_3,$$

$$x \equiv 2 \cdot 35 \cdot 2 + 3 \cdot 21 \cdot 1 + 2 \cdot 15 \cdot 1 \equiv 233 \equiv 23 \pmod{105}.$$

We now proceed to the proof.

Lemma 2.46. *Let a and b be two integers, and let n_1, \dots, n_k be pairwise relatively prime positive integers. If $a \equiv b \pmod{n_i}$ for $1 \leq i \leq k$, then $a \equiv b \pmod{\prod_{i=1}^k n_i}$.*

Proof. First since $a \equiv b \pmod{n_i}$, then $a - b$ is a multiple n_i for each i . Thus, $a - b$ is a multiple of $\text{lcm}(n_1, \dots, n_k)$.

We now have to show that

$$\text{lcm}(n_1, \dots, n_k) = \prod_{i=1}^k n_i.$$

We work pairwise using the formula that for any two integers x, y , we have

$$x \cdot y = \text{lcm}(x, y) \text{gcd}(x, y).$$

Since $\text{gcd}(n_1, n_2) = 1$, we have

$$n_1 \cdot n_2 = \text{lcm}(n_1, n_2).$$

Then we have

$$(n_1 \cdot n_2) \cdot n_3 = \text{lcm}(\text{lcm}(n_1, n_2), n_3) = \text{lcm}(n_1, n_2, n_3).$$

We could proceed inductively at this point or continue in this manner through the finite number of n_i to see that

$$\text{lcm}(n_1, \dots, n_k) = \prod_{i=1}^k n_i.$$

Putting this all together, we have that $a - b$ is a multiple of $n_1 \cdots n_k$, which is the same as saying

$$a - b \equiv 0 \pmod{\prod_{i=1}^k n_i}$$

or

$$a \equiv b \pmod{\prod_{i=1}^k n_i}. \quad \square$$

Proof of Theorem 2.44. Define $N = n_1 \cdots n_k$ and $m_i = \frac{N}{n_i}$. For each $1 \leq i \leq k$,

$$\text{gcd}(m_i, n_i) = 1$$

so that there are integers y_i, t_i such that

$$y_i m_i + t_i n_i = 1.$$

Let

$$x = a_1 m_1 y_1 + \cdots + a_k m_k y_k.$$

Then for each i we have

$$x \equiv a_i \pmod{n_i}$$

since $n_i \mid m_j$ for $j \neq i$ and $y_i m_i \equiv 1 \pmod{n_i}$.

We have now shown that x is a solution. We must show that it is unique modulo $N = n_1 \cdots n_k$. Assume that z is another solution, in other words, that for each $1 \leq i \leq k$,

$$x \equiv z \equiv a_i \pmod{n_i}.$$

In particular, $x - z \equiv 0 \pmod{n_i}$ and $n_i \mid (x - z)$. Since the n_i are all relatively prime, then by Lemma 2.46 we also have that

$$N \mid (x - z),$$

which shows that

$$x \equiv z \pmod{N}. \quad \square$$

We formalize the steps of solving systems of congruences with the Chinese Remainder Theorem in Algorithm 2.2.

Algorithm 2.2. Chinese Remainder Theorem

Input: pairwise relatively prime (positive) integers $\{n_1, \dots, n_k\}$ and integers $\{a_1, \dots, a_k\}$

Output: an integer x such that x solves the system of linear equations

$$x \equiv a_i \pmod{n_i}$$

Algorithm:

1: Compute $N = \prod_{i=1}^k n_i$.

2: Compute y_i , the inverse of N/n_i modulo n_i .

3: Compute

$$x = a_1 \frac{N}{n_1} y_1 + \cdots + a_k \frac{N}{n_k} y_k.$$

4: Return x .

We conclude with an example where the moduli are not prime, just relatively prime.

Example 2.47. Solve the system of linear congruences

$$x \equiv 5 \pmod{9},$$

$$x \equiv 1 \pmod{8}.$$

In the notation from the theorem, we have

$$a_1 = 5, \quad n_1 = 9,$$

$$a_2 = 1, \quad n_2 = 8,$$

and

$$N = 9 \cdot 8 = 72,$$

$$m_1 = \frac{N}{n_1} = 8,$$

$$m_2 = \frac{N}{n_2} = 9.$$

What we need are the inverses of the m_i modulo n_i , which we call y_i .

$$\begin{aligned} 8 \cdot 8 &\equiv 1 \pmod{9} &\Rightarrow y_1 &= 8, \\ 1 \cdot 9 &\equiv 1 \pmod{8} &\Rightarrow y_2 &= 1. \end{aligned}$$

So we have

$$\begin{aligned} x &\equiv a_1 m_1 y_1 + a_2 m_2 y_2, \\ x &\equiv 5 \cdot 8 \cdot 8 + 1 \cdot 9 \cdot 1 \equiv 329 \equiv 41 \pmod{72}. \end{aligned}$$

COMPUTATIONAL EXERCISES

2.1.

- a. Determine the last digit of 932^7 .
- b. Determine the last two digits of 2103^5 .

2.2. Find the smallest positive residue of $(n-1)!$ modulo n for integers $2 \leq n \leq 50$. Notice any patterns?

2.3. Given integers a and m with $\gcd(a, m) = 1$ and $m > 0$, write a function to determine the inverse of a modulo m .

2.4. A composite number n is called a *Fermat pseudoprime base a* if $\gcd(a, n) = 1$ and

$$a^{n-1} \equiv 1 \pmod{n}.$$

- a. Find the first n which is a Fermat pseudoprime base 2.
- b. Find the first n which is a Fermat pseudoprime base 3.

2.5. Find all the composite numbers less than 100 that are Fermat pseudoprimes for at least one base.

2.6. Find all the Carmichael numbers (page 48) less than 10^4 .

2.7. Compute the following values.

- a. $3^{113} \pmod{29}$
- b. $11^{11^{11}} \pmod{6}$

2.8. Recall that the Fibonacci numbers (Definition 1.34) are defined as $F_0 = 0$, $F_1 = 1$ and

$$F_n = F_{n-1} + F_{n-2} \quad \text{for } n \geq 2.$$


- a. Compute the sequence $\{F_n \bmod p : 1 \leq n \leq 20\}$ for $p = 3$. What does this say about the Fibonacci numbers divisible by 3?
- b. Repeat the previous part for $p = 5$.

2.9. Shiu proved that if a and q are relatively prime integers, then there exist arbitrarily long strings of consecutive prime numbers that are all congruent to a modulo q .


- a. For $(a, q) = (3, 11)$, find the longest such string of consecutive primes p , with $p \leq 10,000$.
- b. Determine the pair (a, q) with $0 \leq a, q \leq 10$ that has the longest possible such string of consecutive primes p with $p \leq 10,000$.

2.10. Let $p = 23$. Find all the primes $q \equiv 1 \pmod{p}$ less than 10,000 such that

$$2^{\frac{q-1}{p}} \equiv 1 \pmod{q}.$$

 **2.11.** Find all solutions to the congruence

$$18x \equiv 6 \pmod{42}.$$

 **2.12.** Use the Chinese Remainder Theorem to solve the following system of congruences:

$$\begin{aligned} x &\equiv 3 \pmod{7}, \\ x &\equiv 5 \pmod{13}. \end{aligned}$$

2.13. A truck full of eggs crashes, but the number of eggs is not known. The driver recalls that when counted by 2 there were none left over, when counted by 331 there were 2 left over, and when counted by 521 there were none left over. What is the smallest number of eggs that the truck could contain?

2.14.

- Find a system of linear congruences whose moduli are not relatively prime that has a solution.
- Find a system of linear congruences whose moduli are not relatively prime that does not have a solution.

THEORETICAL EXERCISES

2.15. Prove that the last digit of a perfect square is never 2, 3, 7, or 8.

2.16. Let m and n be positive integers. Prove that if m divides n and $a \equiv b \pmod{n}$, then $a \equiv b \pmod{m}$.

2.17. Let n be a positive integer, and let a and b be any integers. Prove that if $a \equiv b \pmod{n}$, then $\gcd(a, n) = \gcd(b, n)$.

2.18. Let n be an odd integer. Prove that $n^2 \equiv 1 \pmod{8}$.

2.19.

- Prove that an integer n is divisible by 11 if and only if the alternating sum of its digits is divisible by 11.
- Use the previous part to show that there are no 4-digit prime palindromes.

2.20. Prove the inverse of Theorem 2.10(c). In other words, given positive integers a , b , c , and n , prove if

$$a \equiv b \pmod{n},$$

then

$$ac \equiv bc \pmod{nc}.$$

2.21. Let n be a positive integer, and let a be any integer. Prove that if $\gcd(a, n) \neq 1$, then a does not have an inverse modulo n .

2.22. Let n be a positive integer. The Carmichael function $\lambda(n)$ is the smallest positive integer such that

$$a^{\lambda(n)} \equiv 1 \pmod{n}$$

for all a with $\gcd(a, n) = 1$.

a. Find an example where $\lambda(n) < \varphi(n)$.

b. Prove that $\lambda(n) \mid \varphi(n)$.

2.23. Let n be a positive integer. Let $\{r_1, \dots, r_m\}$ be a complete set of residues modulo n . Let a and b be any two integers satisfying $\gcd(a, n) = 1$. Prove that $\{ar_1 + b, ar_2 + b, \dots, ar_m + b\}$ is also a complete set of residues modulo n .

2.24. Let n be a positive integer. Prove that if $n \equiv 3 \pmod{4}$, then n cannot be written as the sum of two integer squares ($a^2 + b^2 = n$).

2.25.

a. Let p be a prime. Prove that $x^{p-1} - 1 \equiv (x-1) \cdots (x-(p-1)) \pmod{p}$ for an indeterminate x .

b. Prove Wilson's Theorem: n is prime if and only if

$$(n-1)! \equiv -1 \pmod{n}.$$

2.26. Use the Chinese Remainder Theorem (Theorem 2.44) to prove that Euler's totient function $\varphi(n)$ is multiplicative.

2.27. Let p and q be distinct primes. Prove that $p^{q-1} + q^{p-1} \equiv 1 \pmod{pq}$.

2.28. For two integers a, b and positive integers m_1, \dots, m_r , prove that if $a \equiv b \pmod{m_i}$ for $1 \leq i \leq r$, then $a \equiv b \pmod{\text{lcm}(m_1, \dots, m_r)}$.

2.29. Prove that every integer can be written as the sum of two squares modulo a prime p . In other words, every circle $x^2 + y^2 = a$ has at least one point modulo p .

2.30. Let p be a positive integer. Prove that p is prime if and only if

$$x^p - a \equiv (x-a)^p \pmod{p}$$

for every $0 \leq a < p$ where x is an indeterminate.

EXPLORATION EXERCISES

2.31 (Primality testing). The interest in large prime numbers has become a practical matter since the invention of computers and the use of prime numbers in encryption algorithms, such as public key cryptography. Consequently, there is a strong need to be able to determine if a given (large) integer n is prime.

A *primality test* is an algorithm to determine whether a given number is prime. Some tests conclude that n is prime, some conclude that n is composite, and some conclude that n is prime with some probability.

You will explore a few different methods and their efficiency. It turns out that determining the efficiency of a primality test is more difficult than it seems. Depending on the form of n (i.e., number and size of its prime factors), different primality

tests are more or less efficient. For the probabilistic tests, there may be numbers that pass the test but that are not prime. These are called *pseudoprimes*. Can you find pseudoprimes for all the probabilistic tests? How many pseudoprimes does each test have? Are pseudoprimes for one test also pseudoprimes for other tests?

- a. Trial division (Deterministic).
 1. Try dividing n by all positive integers $\leq \sqrt{n}$.
 2. Try dividing n by all prime numbers $\leq \sqrt{n}$.
- b. Wilson's Theorem (Deterministic).

Theorem (Theoretical Exercise 2.25). *A positive integer n is prime if and only if $(n-1)! \equiv -1 \pmod{n}$.*

- c. Fermat primality test (Probabilistic).

Theorem (Theorem 2.21). *If n is prime, then for all a relatively prime to n , $a^{n-1} \equiv 1 \pmod{n}$.*

Notice that this theorem is not “if and only if”. In fact, there are numbers which pass this test for every choice of a , called *Carmichael numbers*.

1. Find the first few Carmichael numbers.
- d. Miller–Rabin primality test (Probabilistic). Choose some integer $1 < a < n$. Let $2^s d = n-1$, where d is odd. If $a^d \not\equiv 1 \pmod{n}$ and $a^{2^r d} \not\equiv -1 \pmod{n}$ for all $0 \leq r \leq s-1$, then n is composite.
 1. Can you find any pseudoprimes?
 2. Determine a probability of primality based on randomly generated a 's.
- e. Lucas test (Deterministic). If there is an integer $1 < a < n-1$ such that $a^{n-1} \equiv 1 \pmod{n}$ and for every prime factor $q \mid (n-1)$ and we have $a^{(n-1)/q} \not\equiv 1 \pmod{n}$, then n is prime. If no such a exists, n is composite.
- f. Agrawal–Kayal–Saxena (AKS) primality test.

Theorem (Theoretical Exercise 2.30). *A positive integer p is prime if and only if*

$$x^p - a \equiv (x - a)^p \pmod{p}.$$

This test provides an “if and only if” criteria, but the congruence criteria takes exponential time to check. The actual AKS test is a polynomial time algorithm and uses a slightly different congruence.

2.32 (Prime divisors of Fibonacci numbers). Recall that the Fibonacci numbers (Definition 1.34) are defined as $F_0 = 0$, $F_1 = 1$, and

$$F_n = F_{n-1} + F_{n-2} \quad \text{for all } n \geq 2.$$

- a. Consider the Fibonacci numbers modulo various primes p .
- b. If 0 occurs in the sequence, that means p divides the corresponding Fibonacci number. Fix a prime p and determine which Fibonacci numbers are divisible by p .
- c. How many Fibonacci numbers are odd versus even?
- d. Notice that the Fibonacci numbers repeat themselves modulo primes. How long or short can this sequence be for different primes?

- e. What about the Fibonacci numbers modulo a composite number?
- f. What about the prime divisors of more general sequences?

2.33 (Carmichael numbers). Recall that we have defined Carmichael numbers as numbers that are Fermat pseudoprimes for every base (page 48).

- a. Compute the first few Carmichael numbers. Notice that they are all odd.
- b. Can you find a base a for which every even composite number n satisfies $a^{n-1} \not\equiv 1 \pmod{n}$?
- c. Factor each of the Carmichael numbers. What do you notice about the number of prime divisors?
- d. Compare the prime divisors of each Carmichael number n to the divisors of $n - 1$. Can you use this to conclude Carmichael numbers must be odd?

2.34 (Multivariable linear congruences).

- a. Let n be a positive integer. For which integer triples (a, b, c) can you solve

$$ax + by \equiv c \pmod{n}?$$

- b. How many solutions are there?
- c. For which integer triples (a, b, c) can you not solve

$$ax + by \equiv c \pmod{n}?$$

- d. What about more variables?

$$a_1x_1 + \cdots + a_mx_m \equiv b \pmod{n}$$

2.35 (Nonlinear congruences).

- a. Pick a random polynomial of degree 2 with integer coefficients

$$ax^2 + bx + c, \quad a \neq 0.$$

What can you say about the number of zeros modulo p for primes p ? You want to consider as separate cases primes where $p \mid a$ and primes where $p \nmid a$.

- b. What about degree 3? or higher degree?
- c. Can you find an irreducible polynomial that has a root modulo every prime?
- d. Consider the same questions, but for two variable polynomials. Start by letting $f(x)$ be a single variable polynomial, and count the number of solutions to the following congruences modulo primes.
 1. $y = f(x)$
 2. $y^2 = f(x)$
 3. $y^n = f(x)$
- e. What about arbitrary two-variable polynomials $f(x, y)$?

2.36 (Carmichael function). Let n be a positive integer. We define the *Carmichael function* $\lambda(n)$ as the smallest positive integer such that

$$a^{\lambda(n)} \equiv 1 \pmod{n}$$

for all a with $\gcd(a, n) = 1$.

- a. Relate the values of the Carmichael function to the Euler totient function. When are they equal? When are they different?
- b. Given two positive integers m and n with m dividing n , what can you say about $\lambda(m)$ and $\lambda(n)$?
- c. Given two positive integers m and n , what can you say about $\lambda(mn)$ in terms of $\lambda(m)$ and $\lambda(n)$? Perhaps start with the case $\gcd(m, n) = 1$.
- d. How different can the Carmichael function be than the totient function? In particular, how large can $\frac{\varphi(n)}{\lambda(n)}$ be?

You could also examine the values of the Carmichael function.

- e. Which integers k satisfy $\lambda(n) = k$ for some n ? How many n exist for each k ?
- f. Can you find a way to compute $\lambda(n)$ based on the prime factorization of n ?
- g. For which n is $\lambda(n)$ prime?
- h. For which n is $\lambda(n)$ a perfect square, cube, etc.?

Quadratic Reciprocity and Primitive Roots

In this chapter, we consider congruence equations of degree 2 or higher. This topic leads to the Theorem of Quadratic Reciprocity (Theorem 3.14) and to studying primitive roots.

1. Quadratic Reciprocity

Consider the simplest quadratic equation $x^2 = a$ for an integer a . This has an integer solution exactly when a is a perfect square (this is the definition of a perfect square). There are exactly two solutions $\pm\sqrt{a}$, unless $a = 0$, then there is one solution. What happens in modular arithmetic?

Question 3.1. Given a positive integer n and an integer a , when can you solve

$$x^2 \equiv a \pmod{n}?$$

How many solutions does it have?

Example 3.2. With a few simple examples, we can encounter a wide range of behaviors.

- $x^2 \equiv 2 \pmod{3}$ has no solutions.
- $x^2 \equiv 2 \pmod{7}$ has two solutions.
- $x^2 \equiv 1 \pmod{5}$ has two solutions.
- $x^2 \equiv 1 \pmod{12}$ has four solutions.

One of the most interesting of these is the first example, which shows that square roots do not always exist in modular arithmetic!

We give a name to the numbers that have a square root.

Definition 3.3. Given a positive integer n and an integer a relatively prime to n , we say that a is a *quadratic residue modulo n* if the equation $x^2 \equiv a \pmod{n}$ has a solution. If the equation $x^2 \equiv a \pmod{n}$ does not have a solution, we say that a is a *quadratic nonresidue modulo n* .

Investigation 3.4. Given a modulus n , determine how many residue classes are quadratic residues modulo n . Do you see a pattern? Start by considering the case when n is a prime number.

We will address the following question.

Question 3.5. Let p be a prime number. How many of the complete set of residues $\{0, 1, \dots, p-1\}$ modulo p are quadratic residues?

1.1. Euler's Criterion and the Legendre Symbol.

Example 3.6. We take the first few primes and list the nonzero a values for which we can solve $x^2 \equiv a \pmod{p}$ and those for which we cannot.

p	$x^2 \equiv a \pmod{p}$ for some x	$x^2 \not\equiv a \pmod{p}$ for some x
2	1	
3	1	2
5	1, 4	2, 3
7	1, 2, 4	3, 5, 6
11	1, 3, 4, 5, 9	2, 6, 7, 8, 10
13	1, 3, 4, 9, 10, 12	2, 5, 6, 7, 8, 11
17	1, 2, 4, 8, 9, 13, 15, 16	3, 5, 6, 7, 10, 11, 12, 14

It seems that about half of all residue classes are quadratic residues. For $p = 2$ all nonzero residue classes are quadratic residues, so our theorem must exclude 2.

Theorem 3.7. *If $p > 2$ is a prime and $a \neq 0$, then $x^2 \equiv a \pmod{p}$ has either zero or two solutions.*

Proof. Assume that b is a solution, then $-b$ is also a solution. For $p > 2$, since b is nonzero and $-b \not\equiv b \pmod{p}$, if there is one solution, then there are at least two solutions. Now assume that c is any solution. Being a solution implies

$$c^2 \equiv b^2 \equiv a \pmod{p},$$

which is the same as

$$c^2 - b^2 \equiv 0 \pmod{p}.$$

Then

$$p \mid b^2 - c^2 = (b - c)(b + c).$$

Since p is prime, $p \mid b + c$ or $p \mid b - c$. In other words, $b \equiv c \pmod{p}$ or $-b \equiv c \pmod{p}$. Therefore, c is not a new solution. \square

Corollary 3.8. *If $p > 2$ is a prime, then there are exactly $\frac{p-1}{2}$ nonzero quadratic residues modulo p (and $\frac{p-1}{2}$ quadratic nonresidues).*

Proof. Consider the complete system of nonzero residues $\{1, \dots, p-1\}$. Let $\{a_1, \dots, a_{p-1}\}$ be their squares modulo p ; in other words,

$$\begin{aligned} 1^2 &\equiv a_1 \pmod{p} \\ 2^2 &\equiv a_2 \pmod{p} \\ &\vdots \\ (p-1)^2 &\equiv a_{p-1} \pmod{p}. \end{aligned}$$

Since $x^2 \equiv a \pmod{p}$ has either zero or two solutions for $a \neq 0$, each a_i occurs exactly twice in the list. Thus, there are $\frac{p-1}{2}$ nonzero quadratic residues modulo p . \square

We now know how often we can solve $x^2 \equiv a \pmod{p}$ and how many solutions can exist. However, given a particular a and p , we still have no way to determine whether a solution exists other than simply checking all possible residues. To help with this problem, we define the Legendre¹ symbol.

Definition 3.9. Let $p > 2$ be a prime, and let a be an integer not divisible by p . Then we define the *Legendre symbol* as

$$\left(\frac{a}{p}\right) = \begin{cases} 1 & \text{if } a \text{ is a quadratic residue modulo } p, \\ -1 & \text{if } a \text{ is a quadratic nonresidue modulo } p. \end{cases}$$

Notice that the Legendre symbol depends only on the residue class of a ; that is, if $a \equiv b \pmod{p}$, then we have

$$\left(\frac{a}{p}\right) = \left(\frac{b}{p}\right).$$

In 1748 Euler proved a way to compute Legendre symbols and, thus, determine whether a is a quadratic residue modulo p .

Theorem 3.10 (Euler's criterion). *Let $p > 2$ be a prime, and let a be an integer not divisible by p . Then*

$$\left(\frac{a}{p}\right) = a^{\frac{p-1}{2}} \pmod{p}.$$

Proof. We start with Fermat's Little Theorem (Theorem 2.21),

$$a^{p-1} - 1 \equiv 0 \pmod{p}.$$

We can factor this as

$$\left(a^{\frac{p-1}{2}} - 1\right) \left(a^{\frac{p-1}{2}} + 1\right) \equiv 0 \pmod{p}.$$

Since p is prime, by Lemma 1.38 we must have

$$a^{\frac{p-1}{2}} \equiv 1 \pmod{p} \quad \text{or} \quad a^{\frac{p-1}{2}} \equiv -1 \pmod{p}.$$

¹Adrien-Marie Legendre (1752–1833) was a French mathematician.

If a is a quadratic residue, then there is some b such that

$$b^2 \equiv a \pmod{p}.$$

Then we have

$$(b^2)^{\frac{p-1}{2}} \equiv b^{p-1} \equiv 1 \pmod{p}.$$

Thus, every quadratic residue satisfies

$$a^{\frac{p-1}{2}} \equiv 1 \pmod{p}.$$

Now assume that a is a quadratic nonresidue. Then we may partition the nonzero residues $\{1, \dots, p-1\}$ into pairs of distinct numbers (x, y) such that

$$xy \equiv a \pmod{p}.$$

In particular, given y , there is a unique x that solves the linear equation

$$(11) \quad xy \equiv a \pmod{p}.$$

We can always solve equation (11) because every nonzero residue y is relatively prime to p . There are $\frac{p-1}{2}$ such pairs, so we have that

$$a^{\frac{p-1}{2}} = \prod_{\text{pairs } (x,y)} xy = (p-1)!.$$

Since $x^2 \equiv 1 \pmod{p}$ has exactly two solutions, ± 1 , each number $\{1, \dots, p-1\}$ has an inverse modulo p and, except for ± 1 , the inverse is different from the number. So we have

$$a^{\frac{p-1}{2}} \equiv (p-1)! \equiv -1 \pmod{p}. \quad \square$$

We apply Euler's criterion to show that the Legendre symbol is multiplicative.

Corollary 3.11. *Let $p > 2$ be a prime, and let a and b be integers not divisible by p . Then*

$$\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right) \left(\frac{b}{p}\right).$$

Proof. We compute

$$\left(\frac{ab}{p}\right) = (ab)^{\frac{p-1}{2}} \pmod{p} = a^{\frac{p-1}{2}} b^{\frac{p-1}{2}} \pmod{p} = \left(\frac{a}{p}\right) \left(\frac{b}{p}\right). \quad \square$$

Because we know the Legendre symbol is multiplicative, we have reduced its computation to the two situations of $\left(\frac{2}{p}\right)$ and $\left(\frac{q}{p}\right)$, where p and q are distinct (odd) primes.

Investigation 3.12.

- (a) Given two distinct odd primes p and q , can you determine when p is a quadratic residue modulo q ?
- (b) If p is a quadratic residue modulo q , does that tell you anything about whether q is a quadratic residue modulo p ?

It may be helpful to separate the cases of p and q congruent to 1 or 3 modulo 4.

1.2. Law of Quadratic Reciprocity. The prime 2, as the only even prime, often must be considered separately. We consider the primes p for which 2 is a quadratic residue. Let's gather some data and see what patterns arise.

primes for which 2 is a quadratic residue
7, 17, 23, 31, 41, 47, 71, 73, 79, 89, 97

primes for which 2 is a quadratic nonresidue
3, 5, 11, 13, 19, 29, 37, 43, 53, 59, 61, 67, 83

Since the primes p are all odd, it makes sense to examine their residue class modulo powers of 2. After some experimentation we see that the residue class modulo 8 is the determining factor. For the quadratic residues we have

$$\begin{aligned} 17, 41, 73, 89, 97 &\equiv 1 \pmod{8}, \\ 7, 23, 31, 47, 71, 79 &\equiv 7 \pmod{8}. \end{aligned}$$

For the quadratic nonresidues we have

$$\begin{aligned} 3, 11, 19, 43, 59, 67, 83 &\equiv 3 \pmod{8}, \\ 5, 13, 29, 37, 53, 61 &\equiv 5 \pmod{8}. \end{aligned}$$

In summary, for primes congruent to 1 or 7 modulo 8, 2 is a quadratic residue. For primes that are congruent to 3 or 5 modulo 8, 2 is a quadratic nonresidue. We state this as the supplemental Law of Quadratic Reciprocity.

Theorem 3.13 (Supplemental law). *Let $p > 2$ be a prime number. Then*

$$\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}}.$$

There are elementary proofs using messy congruences, but instead let's use the roots of unity for a more conceptually pleasing proof.

Proof. Let $\zeta = \sqrt{i}$, where $i^2 = -1$. The number ζ is called an *8th root of unity* since $\zeta^8 = 1$. Let ζ^{-1} be the inverse of ζ , and let $w = \zeta + \zeta^{-1}$. Then we can verify that $w = \sqrt{2}$ as

$$w^2 = (\zeta + \zeta^{-1})^2 = \zeta^2 + 2 + \zeta^{-2} = i + 2 - i = 2.$$

Note that we used the fact that $\frac{1}{i} = -i$. Now we apply Euler's criterion to $\left(\frac{2}{p}\right)$:

$$\begin{aligned} 2^{\frac{p-1}{2}} &= (\sqrt{2})^{p-1} = w^{p-1} = w^p w^{-1} = (\zeta + \zeta^{-1})^p w^{-1} \\ &\equiv (\zeta^p + \zeta^{-p}) w^{-1} \pmod{p}. \end{aligned}$$

The last equivalence follows from the binomial expansion of $(\zeta + \zeta^{-1})^p$ as

$$(\zeta + \zeta^{-1})^p \equiv \zeta^p + \zeta^{-p} \pmod{p}.$$

However, we can compute $\zeta^p + \zeta^{-p}$ using the fact that ζ is an 8th root of unity. In particular, ζ^p can be computed just from the remainder of p after division by 8. For example

$$\zeta^{12} = \zeta^8 \zeta^4 = 1 \cdot \zeta^4 = \zeta^4$$

or

$$\zeta^{27} = \zeta^8 \zeta^8 \zeta^8 \zeta^3 = 1 \cdot 1 \cdot 1 \cdot \zeta^3 = \zeta^3.$$

We compute

$$\zeta^p + \zeta^{-p} \equiv \begin{cases} \zeta + \zeta^{-1} \equiv w \pmod{p} & p \equiv \pm 1 \pmod{8}, \\ \zeta^3 + \zeta^{-3} \equiv -w \pmod{p} & p \equiv \pm 3 \pmod{8}. \end{cases}$$

So we have

$$\begin{aligned} 2^{\frac{p-1}{2}} &\equiv (\zeta^p + \zeta^{-p})w^{-1} \pmod{p} \\ &\equiv \begin{cases} 1 \pmod{p} & p \equiv \pm 1 \pmod{8}, \\ -1 \pmod{p} & p \equiv \pm 3 \pmod{8}. \end{cases} \quad \square \end{aligned}$$

With $\left(\frac{2}{p}\right)$ solved, we now turn to the situation of Legendre symbols for odd primes, $\left(\frac{q}{p}\right)$. First let's gather some data where we abbreviate R for quadratic residue and N for quadratic nonresidue.

	p											
q		3	5	7	11	13	17	19	23	29	31	37
	3		N	N	R	R	N	N	R	N	N	R
	5	N		N	R	N	N	R	N	R	R	N
	7	R	N		N	N	N	R	N	R	R	R
	11	N	R	R		N	N	R	N	N	N	R
	13	R	N	N	N		R	N	R	R	N	N
	17	N	N	N	N	R		R	N	N	N	N
	19	R	R	N	N	N	R		N	N	R	N
	23	N	N	R	R	R	N	R		R	N	N
	29	N	R	R	N	R	N	N	R		N	N
	31	R	R	N	R	N	N	N	R	N		N
	37	R	N	R	R	N	N	N	N	N	N	

Unfortunately, there does not seem to be much of a pattern here. However, if we consider only the cases where $p \equiv 1 \pmod{4}$ or $q \equiv 1 \pmod{4}$, there is a striking symmetry around the diagonal.

	p											
q		3	5	7	11	13	17	19	23	29	31	37
	3		N			R	N			N		R
	5	N		N	R	N	N	R	N	R	R	N
	7		N			N	N			R		R
	11		R			N	N			N		R
	13	R	N	N	N		R	N	R	R	N	N
	17	N	N	N	N	R		R	N	N	N	N
	19		R			N	R			N		N
	23		N			R	N			R		N
	29	N	R	R	N	R	N	N	R		N	N
	31		R			N	N			N		N
	37	R	N	R	R	N	N	N	N	N	N	

Legendre and Euler both conjectured the following theorem, called the Law of Quadratic Reciprocity, but it was first proven by Gauss² in 1801. In fact, Gauss gave six different proofs. To date, there are well over 200 distinct proofs of quadratic reciprocity.

Theorem 3.14 (Quadratic Reciprocity). *Let $p, q > 2$ be distinct prime numbers. Then*

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{(p-1)(q-1)}{4}}.$$

We need two easy facts about the Legendre symbol for the proof.

Lemma 3.15. *Let $p > 2$ be a prime. For every integer a not divisible by p , the equation $x^2 \equiv a \pmod{p}$ has*

$$1 + \left(\frac{a}{p}\right)$$

solutions.

Proof. If there is a solution, then by Theorem 3.7 there are two solutions. Similarly, if there is a solution, then a is a quadratic residue, and thus $\left(\frac{a}{p}\right) = 1$.

If there are no solutions, then a is a quadratic nonresidue and $\left(\frac{a}{p}\right) = -1$. \square

Lemma 3.16. *Let $p > 2$ be a prime. Then*

$$\sum_{a=1}^{p-1} \left(\frac{a}{p}\right) = 0.$$

Proof. Since exactly half of the a are quadratic residues and half are quadratic nonresidues, the sum of their Legendre symbols is 0. \square

Now we are ready to prove the Law of Quadratic Reciprocity.

Proof of Theorem 3.14. The proof proceeds by counting the number of solutions to

$$(12) \quad x_1^2 - x_2^2 + x_3^2 - \cdots + x_p^2 \equiv 1 \pmod{q}$$

in two different ways.

Define N_p to be the number of solutions to equation (12). First substitute $x_1 = x_1 + x_2$ to get

$$x_1^2 + x_3^2 - x_4^2 + \cdots + x_p^2 \equiv -2x_1x_2 \pmod{p}.$$

For each $x_1 \neq 0$ and any choice of (x_3, \dots, x_p) , there is a unique x_2 value that solves the resulting (linear) equation. So there are $q^{p-2}(q-1)$ solutions of this form: $(q-1)$ choices for x_1 , and q choices for each of $\{x_3, \dots, x_p\}$. When $x_1 = 0$, the solutions satisfy

$$x_3^2 - x_4^2 + \cdots + x_p^2 \equiv 1 \pmod{q}$$

for any value of x_2 .

²Carl Friedrich Gauss (1777–1855) was a German mathematician.

By renaming the variables, this is the same as

$$y_1^2 - y_2^2 + \cdots + y_{p-2}^2 \equiv 1 \pmod{q},$$

which has N_{p-2} solutions. These solutions are valid for any choice of x_2 for a total of qN_{p-2} more solutions. Thus,

$$\begin{aligned}
 N_p &= q^{p-2}(q-1) + qN_{p-2} \\
 &\text{now substitute } N_{p-2} = q^{-4}(q-1)qN_{p-4} \text{ to get} \\
 &= q^{p-2}(q-1) + q(q^{p-4}(q-1) + qN_{p-4}) \\
 &= q^{p-1} - q^{p-2} + q^{p-2} - q^{p-3} + q^2N_{p-4} \\
 &= q^{p-1} - q^{p-3} + q^2N_{p-4} \\
 &\text{now substitute for } N_{p-4} \text{ to get} \\
 &= q^{p-1} - q^{p-3} + q^2(q^{p-6}(q-1) + qN_{p-6}) \\
 &= q^{p-1} - q^{p-3} + q^{p-3} - q^{p-4} + q^3N_{p-6} \\
 &= q^{p-1} - q^{p-4} + q^3N_{p-6} \\
 &\vdots \\
 &= q^{p-1} - q^{\frac{p-1}{2}} + q^{\frac{p-1}{2}}N_1 \\
 &= q^{p-1} - q^{\frac{p-1}{2}} + 2q^{\frac{p-1}{2}} \\
 &= q^{p-1} + q^{\frac{p-1}{2}} \\
 &\equiv 1 + \left(\frac{q}{p}\right) \pmod{p}.
 \end{aligned}
 \tag{13}$$

Now we count solutions to equation (12) another way. Let $N(x^2 \equiv a \pmod{q})$ be the number of solutions to the quadratic equation $x^2 \equiv a \pmod{q}$. Then we can also determine N_p as

$$\begin{aligned}
 N_p &= \sum_{t_1 + \cdots + t_p \equiv 1 \pmod{q}} N(x_1^2 \equiv t_1 \pmod{q}) \cdot N(x_1^2 \equiv -t_2 \pmod{q}) \\
 &\quad \cdot N(x_3^2 \equiv t_3 \pmod{q}) \cdots N(x_p^2 \equiv t_p \pmod{q}).
 \end{aligned}$$

First note that in the set of possible tuples (t_1, \dots, t_p) , each t_i takes on each possible residue class modulo q a power of q times. Consequently, by Lemma 3.16 we have

$$\sum_{t_i} \left(\frac{t_i}{q}\right) = 0.$$

We expand this product and apply Lemma 3.16 to get

$$\begin{aligned}
 N_p &= \sum_{t_1+\dots+t_p \equiv 1 \pmod{q}} 1 + \sum_{i=1}^p \left(\frac{(-1)^{i+1} t_i}{q} \right) + \sum_{i,j=1}^p (-1)^{i+j} \left(\frac{t_i}{q} \right) \left(\frac{t_j}{q} \right) \\
 &\quad + \dots + (-1)^{(p-1)/2} \prod_{i=1}^p \left(\frac{t_i}{q} \right) \\
 &= \sum_{t_1+\dots+t_p \equiv 1 \pmod{q}} 1 + \sum_{i=1}^p \left(\frac{-1}{q} \right)^{i+1} \left(\frac{t_i}{q} \right) + \dots + \left(\frac{-1}{q} \right)^{\frac{p-1}{2}} \left(\frac{t_1 \dots t_p}{q} \right) \\
 &= \sum_{t_1+\dots+t_p=1} 1 + 0 + \dots + 0 + \left(\frac{-1}{q} \right)^{\frac{p-1}{2}} \left(\frac{t_1 t_2 \dots t_p}{q} \right) \\
 &= q^{p-1} + \left(\frac{(-1)^{\frac{p-1}{2}}}{q} \right) \sum_{t_1+\dots+t_p \equiv 1 \pmod{q}} \left(\frac{t_1 t_2 \dots t_p}{q} \right).
 \end{aligned}$$

Now we examine this expression for N_p modulo p , as in the previous part of the proof. The only terms of the sum that are nonzero modulo p are those where $t_1 \equiv t_2 \equiv \dots \equiv t_p \equiv p^{-1} \pmod{q}$ since, by symmetry, the other terms can be collected into groups of size p . For example, if $t_1 \equiv 2^{-1} \pmod{q}$, then we have terms of the form $\sum_{t_2+\dots+t_p \equiv 2^{-1} \pmod{q}} \left(\frac{2^{-1} t_2 \dots t_p}{q} \right)$. But there are p such sets of terms where each of the p different t_i satisfies $t_i \equiv 2^{-1} \pmod{q}$. Thus, we can group these terms together to get $p \sum_{t_2+\dots+t_p \equiv 2^{-1} \pmod{q}} \left(\frac{2^{-1} t_2 \dots t_p}{q} \right) \equiv 0 \pmod{p}$. The only time we do not get such symmetry is when all the t_i are equal: $t_1 \equiv t_2 \equiv \dots \equiv t_p \equiv p^{-1} \pmod{q}$.

Thus, we have

$$\begin{aligned}
 (14) \quad N_p &\equiv 1 + \left(\frac{(-1)^{\frac{p-1}{2}}}{q} \right) \left(\frac{p^{-p}}{q} \right) \pmod{p} \\
 &\equiv 1 + (-1)^{\frac{p-1}{2} \frac{q-1}{2}} \left(\frac{p}{q} \right) \pmod{p}.
 \end{aligned}$$

The last equality relies on the observation that p is a square modulo q if and only if p^{-p} is a square modulo q (Theoretical Exercise 3.17). The reciprocity law follows by comparing the two equations (13) and (14). \square

The multiplicativity of the Legendre symbol and the Law of Quadratic Reciprocity combine to give a very efficient method of determining whether a is a quadratic residue or nonresidue modulo a prime. In particular, given a quadratic congruence

$$x^2 \equiv a \pmod{p},$$

the Legendre symbol and the Law of Quadratic Reciprocity give an efficient way of determining whether there is a solution. The algorithm is similar to the Euclidean algorithm, as demonstrated in Example 3.17.

Example 3.17. We first use multiplicativity to reduce to the case of primes:

$$\left(\frac{124}{17}\right) = \left(\frac{2^2 31}{17}\right) = \left(\frac{2}{17}\right)^2 \left(\frac{31}{17}\right).$$

We can then simplify powers since $(-1)^n$ depends only on the parity of n . In other words, all Legendre symbols to even powers are just 1 and Legendre symbols to odd powers are the same as the Legendre symbol to the first power. So we have

$$\left(\frac{2}{17}\right)^2 \left(\frac{31}{17}\right) = \left(\frac{31}{17}\right).$$

Next we can use the fact that if $a \equiv b \pmod{p}$, we have

$$\left(\frac{a}{p}\right) = \left(\frac{b}{p}\right)$$

to get

$$\left(\frac{31}{17}\right) = \left(\frac{14}{17}\right).$$

Now we again factor

$$\left(\frac{14}{17}\right) = \left(\frac{2 \cdot 7}{17}\right) = \left(\frac{2}{17}\right) \left(\frac{7}{17}\right).$$

We compute $\left(\frac{2}{17}\right)$ by looking the residue class of 17 modulo 8 (Theorem 3.13), which is 1, so that 2 is a quadratic residue modulo 17 and

$$\left(\frac{2}{17}\right) \left(\frac{7}{17}\right) = \left(\frac{7}{17}\right).$$

Since we are down to the case of two odd primes, we apply quadratic reciprocity (Theorem 3.14) to flip the Legendre symbol:

$$\left(\frac{7}{17}\right) = (-1)^{\frac{7-1}{2} \frac{17-1}{2}} \left(\frac{17}{7}\right) = -\left(\frac{17}{7}\right).$$

We again take a different representative in the residue class:

$$-\left(\frac{17}{7}\right) = -\left(\frac{3}{7}\right).$$

Now we again apply quadratic reciprocity to flip the Legendre symbol:

$$-\left(\frac{3}{7}\right) = -(-1)^{\frac{3-1}{2} \frac{7-1}{2}} \left(\frac{7}{3}\right) = -\left(\frac{7}{3}\right) = -\left(\frac{1}{3}\right) = -1.$$

Each time we flip the Legendre symbol, we are taking the remainder after division and reducing the size of the numbers we are working with, similar to the Euclidean algorithm.

This example may have seemed long, but in practice this is quite fast. For large numbers, the hardest step is the factoring step, which is a very difficult problem for large numbers.

1.3. Jacobi Symbol. Now we turn our attention to composite moduli.

Question 3.18. When does $x^2 \equiv a \pmod{n}$ have solutions when n is composite?

Investigation 3.19. As a first step toward Question 3.18, see if you can generalize the Legendre symbol to a composite modulus.

- (a) Choose a composite modulus $n = pq$ which is a product of two primes.
- (b) Determine which residue classes are quadratic residues and nonresidues modulo n .
- (c) Determine whether each residue class is a quadratic residue modulo each prime p and q .
- (d) Does this give you any ideas as to how you might generalize the Legendre symbol?

We follow Jacobi³ and define a generalization of the Legendre symbol.

Definition 3.20. Let n be a positive integer, and factor n into a product of distinct prime numbers as $n = p_1^{e_1} \cdots p_r^{e_r}$, where the e_i are positive integers. We can define the *Jacobi symbol* for an integer a relatively prime to n as

$$\left(\frac{a}{n}\right) = \left(\frac{a}{p_1}\right)^{e_1} \cdots \left(\frac{a}{p_r}\right)^{e_r}.$$

But be careful; $\left(\frac{a}{n}\right) = -1$ only if a is a quadratic nonresidue, but it is also possible that $\left(\frac{a}{n}\right) = 1$ when a is a quadratic nonresidue.

Example 3.21. Recall that 2 is a quadratic nonresidue for both 3 and 5. Then we have

$$\left(\frac{2}{15}\right) = \left(\frac{2}{3}\right) \left(\frac{2}{5}\right) = (-1)(-1) = 1,$$

but 2 is a quadratic nonresidue modulo 15.

Theorem 3.22. Let a , b , and n be integers with $n > 0$. If $\gcd(ab, n) = 1$, we have the following properties for the Jacobi symbol:

- (a) If $a \equiv b \pmod{n}$, then $\left(\frac{a}{n}\right) = \left(\frac{b}{n}\right)$.
- (b) $\left(\frac{ab}{n}\right) = \left(\frac{a}{n}\right) \left(\frac{b}{n}\right)$.
- (c) $\left(\frac{-1}{n}\right) = (-1)^{\frac{n-1}{2}}$.
- (d) $\left(\frac{2}{n}\right) = (-1)^{\frac{n^2-1}{8}}$.
- (e) If $\gcd(n, m) = 1$, then

$$\left(\frac{n}{m}\right) \left(\frac{m}{n}\right) = (-1)^{\frac{m-1}{2} \frac{n-1}{2}}.$$

We leave the proof as Theoretical Exercise 3.20.

³Carl Gustav Jacob Jacobi (1804–1851) was a German mathematician.

2. Computing m th Roots Modulo n

We next turn to the question of higher degree equations.

Question 3.23. When can we solve the congruence

$$x^m \equiv a \pmod{n}?$$

How many solutions does it have?

Definition 3.24. Let n and m be positive integers. We say that an integer a relatively prime to n is an m th power residue modulo n if the equation

$$x^m \equiv a \pmod{n}$$

has a solution.

Example 3.25. We generate a few examples to see the different behavior.

- $x^3 \equiv 3 \pmod{7}$ has no solutions.
- $x^3 \equiv 1 \pmod{7}$ has three solutions.
- $x^4 \equiv 9 \pmod{12}$ has two solutions.
- $x^4 \equiv 1 \pmod{12}$ has four solutions.
- $x^5 \equiv 0 \pmod{16}$ has eight solutions.

The key points to notice from these few examples are that it is possible to have no solutions and it is possible to have less than, equal to, or more than m solutions for $x^m \equiv a \pmod{n}$. That is quite a wide range of behaviors.

To answer Question 3.23, we need the notions of multiplicative order and primitive roots.

Definition 3.26. Let n be a positive integer. We say that an integer a has (multiplicative) order d modulo n if

$$a^d \equiv 1 \pmod{n}$$

and

$$a^t \not\equiv 1 \pmod{n} \quad \text{for all } 0 < t < d.$$

Recall from Euler's formula (Theorem 2.31) that for any (positive) modulus n , every integer a with $\gcd(a, n) = 1$ has multiplicative order dividing $\varphi(n)$. Just knowing that the order divides $\varphi(n)$ is not very specific; it could be as large as $\varphi(n)$ and as small as 1. We know that $a = 1$ always has multiplicative order 1, so the smallest order is possible. What about large orders?

Question 3.27. For each positive integer n , is there an a whose multiplicative order modulo n is $\varphi(n)$?

Investigation 3.28. Choose a positive integer n .

- (a) Compute the multiplicative order for all residue classes relatively prime to n . Try several different values of n .
- (b) Can you say anything about the resulting set of multiplicative orders and/or answer Question 3.27?

Example 3.29. We compute a few orders modulo 36 for $\gcd(a, 36) = 1$. Note that $\varphi(36) = 12$.

a	order
5	6
7	6
11	6
13	3
17	2
19	2
23	6
25	3
29	6
31	6
35	2

All have order dividing 12, but their orders are all less than 12.

So the answer to Question 3.27 is no. However, the next example shows that it is possible to have order $\varphi(n)$.

Example 3.30. We compute orders modulo 18 for $\gcd(a, n) = 1$. Note that $\varphi(18) = 6$.

a	order
5	6
7	3
11	6
13	3
17	2

They all have order dividing 6, and both 5 and 11 have order 6.

We modify question Question 3.27 to reflect our new information.

Question 3.31. For which positive integers n is there an a whose multiplicative order modulo n is $\varphi(n)$?

Definition 3.32. We say that a is a *primitive root modulo n* if $\varphi(n)$ is the order of a modulo n .

Investigation 3.33.

- (a) Find some positive integers n that have a primitive root.
- (b) Can you find a restriction on n that guarantees it has a primitive root?

It is interesting to note that $\varphi(n)$ is the number of positive integers less than n and relatively prime to n . An integer a having multiplicative order $\varphi(n)$ means that the $\varphi(n)$ powers $\{a, a^2, a^3, \dots, a^{\varphi(n)}\}$ are all distinct. Since $\gcd(a, n) = 1$, each of the powers a^m , $1 \leq m \leq \varphi(n)$ is relatively prime to n . So the set of residue classes relatively prime to n is the same as the set of powers of a , i.e., the set $\{a, a^2, a^3, \dots, a^{\varphi(n)}\}$ is exactly the set of residue classes modulo n that are relatively prime to n . In particular, every integer relatively prime to the modulus can be written as a power of a primitive root.

Example 3.34. From Example 3.30 we see that 5 and 11 are both primitive roots modulo 18. In particular, we generate all the residue classes relatively prime to 18 as

$$\{5, 5^2, 5^3, 5^4, 5^5, 5^6\} \equiv \{5, 7, 17, 13, 11, 1\} \pmod{18}.$$

Proposition 3.35. *Let n be a positive integer. If n has a primitive root g , then for any integer a with $\gcd(a, n) = 1$ there exists an integer k such that*

$$g^k \equiv a \pmod{n}.$$

Proof. Assume that g is a primitive root of n and that $\gcd(a, n) = 1$. The elements of the set $\{g, g^2, \dots, g^{\varphi(n)}\}$ are all distinct since a primitive root g has multiplicative order $\varphi(n)$. Also, for any k

$$g^k g^{\varphi(n)-k} \equiv 1 \pmod{n},$$

so each number in the set $\{g, g^2, \dots, g^{\varphi(n)}\}$ has an inverse modulo n . By Theorem 2.19, there are $\varphi(n)$ residue classes that have inverses modulo n . Thus, $\{g, g^2, \dots, g^{\varphi(n)}\}$ are all the residue classes that have inverses. Since a is relatively prime to n , it has an inverse (Theorem 2.19) and, thus, $a \in \{g, g^2, \dots, g^{\varphi(n)}\}$. \square

For moduli with primitive roots, we now know we can write a (relatively prime to n) as a power of the primitive root and reduce Question 3.23 to a linear congruence of the exponents. We can then use Theorem 2.38 to solve the linear congruence.

Let's take a simple example involving square roots.

Example 3.36. Solve $x^2 \equiv 7 \pmod{18}$. We know from Example 3.30 that 5 is a primitive root modulo 18, so every residue class relatively prime to 18 can be

written as a power of 5:

$$\begin{aligned} 5^1 &\equiv 5 \pmod{18}, \\ 5^2 &\equiv 7 \pmod{18}, \\ 5^3 &\equiv 17 \pmod{18}, \\ 5^4 &\equiv 13 \pmod{18}, \\ 5^5 &\equiv 11 \pmod{18}, \\ 5^6 &\equiv 1 \pmod{18}. \end{aligned}$$

Then we can replace the variable x by 5^y for some new variable y to have the equation

$$5^{2y} \equiv 7 \pmod{18}.$$

We can also replace 7 by the power 5^2 to have the equation

$$(15) \quad 5^{2y} \equiv 5^2 \pmod{18}.$$

Since 5 is a primitive root, it has order $\varphi(18) = 6$, which means that for any integer m ,

$$5^m \equiv 5^{m+6} \equiv 5^{m+12} \equiv \dots \pmod{18}.$$

So we compare exponents in equation (15) modulo $\varphi(18) = 6$ to get the linear equation

$$2y \equiv 2 \pmod{6}.$$

It is now easy to see that $y = 1$ is a solution, which is

$$x \equiv 5^y \equiv 5^1 \equiv 5 \pmod{18}.$$

We now give the general solution for m th power residues.

Theorem 3.37. *Let n and m be positive integers, and let a be an integer. If n has a primitive root and $\gcd(a, n) = 1$, then a is an m th power residue if and only if $a^{\varphi(n)/d} \equiv 1 \pmod{n}$ for $d = \gcd(m, \varphi(n))$. If there is an m th power residue, then there are d of them.*

Proof. Let g be a primitive root modulo n and $a = g^b$ and $x = g^y$ for some integers b and y . Then the congruence

$$x^m \equiv a \pmod{n}$$

is equivalent to

$$g^{my} \equiv g^b \pmod{n},$$

which is in turn equivalent to

$$my \equiv b \pmod{\varphi(n)}.$$

The last equation is solvable if and only if $\gcd(m, \varphi(n)) = d \mid b$, and if there is one solution, then there are exactly d solutions (Theorem 2.38).

We now need to show that $d \mid b$ is equivalent to $a^{\varphi(n)/d} \equiv 1 \pmod{n}$.

If $d \mid b$, then

$$a^{\varphi(n)/d} \equiv g^{b\varphi(n)/d} \equiv (g^{\varphi(n)})^{b/d} \equiv 1 \pmod{n}.$$

Conversely, if $a^{\varphi(n)/d} \equiv 1 \pmod{n}$, then

$$g^{k\varphi(n)/d} \equiv 1 \pmod{n},$$

which implies $\varphi(n)$ divides $\frac{k\varphi(n)}{d}$, which implies $d \mid k$. \square

Remark. Notice that the condition for the existence of a solution does not require knowing the primitive root. So we can determine existence of a solution even if we cannot find the solution.

Example 3.38. Working modulo 14, we can try to solve

$$x^3 \equiv 11 \pmod{14}.$$

We compute $\varphi(n) = 6$ and

$$d = \gcd(m, \varphi(n)) = \gcd(3, 6) = 3.$$

We compute

$$11^{\varphi(n)/d} \equiv 11^2 \equiv 9 \pmod{14},$$

and we see that there is no solution. We can explicitly compute all the third powers modulo 14 to check.

a	$a^3 \pmod{14}$
1	$1^3 \equiv 1$
3	$3^3 \equiv 13$
5	$5^3 \equiv 13$
9	$9^3 \equiv 1$
11	$11^3 \equiv 1$
13	$13^3 \equiv 13$

Now consider $a = 13$,

$$x^3 \equiv 13 \pmod{14}.$$

We compute

$$a^{\varphi(n)/d} \equiv 13^2 \equiv 1 \pmod{14},$$

so we expect a solution, in fact, three solutions. Using the primitive root 3 and writing

$$x \equiv 3^y \pmod{14} \quad \text{and} \quad 13 \equiv 3^3 \pmod{14}$$

for some integer y , we have

$$3^{3y} \equiv 3^3 \pmod{14}.$$

The resulting linear equation is

$$3y \equiv 3 \pmod{6},$$

and we may choose $y = 1$ to get $x = 3$ as a solution. From Theorem 2.38, the other two solutions are $y = 3$ and $y = 5$, which correspond to $x = 13$ and $x = 5$, respectively. The table of third powers confirms these solutions.

We use Theorem 3.37 to construct Algorithm 3.1 for computing m th roots.

Algorithm 3.1. m th Roots**Input:** positive integers m, n and an integer a such that $\gcd(a, n) = 1$ **Output:** an m th root of a modulo n **Algorithm:**

- 1: Compute $\varphi(n)$.
- 2: Define $d = \gcd(m, \varphi(n))$.
- 3: Find positive integers x, y such that $xm - y\varphi(n) = d$.
- 4: Compute $a^{x/d}$.

Investigation 3.39. The fraction x/d in the exponent of Algorithm 3.1 step (4) is worrisome since unless $d \mid x$, this is a fractional power!

- (a) Apply Algorithm 3.1 to

$$x^3 \equiv 13 \pmod{14}.$$

Was there an issue with step (4)? Did you need to know a primitive root?

- (b) Apply Algorithm 3.1 to

$$x^6 \equiv 6 \pmod{25}.$$

Was there an issue with step (4)? How is this issue resolved by knowing a primitive root?

In the special case $d = 1$, we do not have to worry about the possible divisibility for the exponent in step (3) and, hence, do not need to know primitive roots, and we can use Algorithm 3.1 without concern.

3. Existence of Primitive Roots

We now return to Question 3.31 about the existence of primitive roots. In particular, we are looking for positive integers n for which there is a residue class a whose multiplicative order is $\varphi(n)$. We know that it is possible for primitive roots to exist, and we know there are moduli that do not have primitive roots.

Example 3.40.

- 2 and 3 both have multiplicative order 4 modulo 5, so they are primitive roots.
- There is no primitive root modulo 15.

Investigation 3.41. See if you can conjecture the correct statement for which moduli n have primitive roots.

- (a) First consider $n = p$ a prime.
- (b) Consider $n = p^k$ with $k > 1$, a power of a single prime.
- (c) Consider $n = pq$ the product of two distinct primes.

We start with two lemmas that we will need later in this section.

Lemma 3.42. *Let n be a positive integer, and let a and b be integers relatively prime to n . If the multiplicative order of a modulo n is x and the multiplicative order of b modulo n is y with $\gcd(x, y) = 1$, then the multiplicative order of ab modulo n is xy .*

Proof. Let r be the multiplicative order of ab modulo n , that is, $(ab)^r \equiv 1 \pmod{n}$. We will show that $r = xy$ by seeing that $r \mid xy$ and $xy \mid r$.

First we compute

$$(ab)^{xy} \equiv a^{xy}b^{xy} \equiv (a^x)^y(b^y)^x \equiv 1 \cdot 1 \equiv 1 \pmod{n}.$$

So we must have $xy \mid r$.

For the other direction, consider the following two computations:

$$a^{ry} \equiv a^{ry} \cdot 1 \equiv a^{ry}(b^y)^r \equiv (ab)^{ry} \equiv 1 \pmod{n},$$

$$b^{rx} \equiv b^{rx} \cdot 1 \equiv b^{rx}(a^x)^r \equiv (ab)^{rx} \equiv 1 \pmod{n}.$$

So $x \mid ry$ and $y \mid rx$. Since $\gcd(x, y) = 1$, these become $x \mid r$ and $y \mid r$. Written as congruences, we have

$$r \equiv 0 \pmod{x},$$

$$r \equiv 0 \pmod{y}$$

with $\gcd(x, y) = 1$. We can apply the Chinese Remainder Theorem (Theorem 2.44) to see that

$$r \equiv 0 \pmod{xy}.$$

This congruence is the same as $xy \mid r$. □

Example 3.43. We illustrate Lemma 3.42 with an example. Consider $n = 7$. Then for $a = 2$,

$$a \equiv 2 \pmod{7},$$

$$a^2 \equiv 4 \pmod{7},$$

$$a^3 \equiv 8 \equiv 1 \pmod{7}$$

so that the multiplicative order of a is 3. For $b = 6$ we have

$$b \equiv 6 \pmod{7},$$

$$b^2 \equiv 36 \equiv 1 \pmod{7}$$

so that the multiplicative order of b is 2. Then $2 \cdot 6 \equiv 12 \equiv 5 \pmod{7}$, which has multiplicative order $6 = 2 \cdot 3$. We check this as

$$5 \equiv 5 \pmod{7},$$

$$5^2 \equiv 25 \equiv 4 \pmod{7},$$

$$5^3 \equiv 20 \equiv 6 \pmod{7},$$

$$5^4 \equiv 30 \equiv 2 \pmod{7},$$

$$5^5 \equiv 10 \equiv 3 \pmod{7},$$

$$5^6 \equiv 15 \equiv 1 \pmod{7}.$$

Our next lemma relates the primitive roots modulo n to the primitive roots modulo $2n$.

Lemma 3.44. *Let n be an odd positive integer. There is a primitive root modulo n if and only if there is a primitive root modulo $2n$.*

Proof. First observe that for an odd number n , $\varphi(2n) = \varphi(n)$. Next note that a primitive root g (if it exists) modulo $2n$ must be odd, and so for any positive integer k

$$g^k \equiv 1 \pmod{2}.$$

Since $\gcd(2, n) = 1$, the Chinese Remainder Theorem (Theorem 2.44) says that the system of equations

$$\begin{aligned} g^k &\equiv 1 \pmod{2}, \\ g^k &\equiv 1 \pmod{n} \end{aligned}$$

is equivalent to

$$g^k \equiv 1 \pmod{2n}.$$

In particular, since k is the smallest power such that $g^k \equiv 1 \pmod{2n}$, then k is the smallest power such that $g^k \equiv 1 \pmod{n}$, so that g is also a primitive root modulo n .

Now assume we have a primitive root g modulo n . If g is even, since n is odd, the element $g + p$ is odd and is also a primitive root since it is in the same residue class as g . With g odd, we can apply the above argument in reverse to see that g is also a primitive root modulo $2n$. \square

Example 3.45. Let $n = 11$, then the primitive roots are $\{2, 6, 7, 8\}$. For $2n$ the primitive roots are $\{7, 13, 17, 19\}$. The proof of Lemma 3.44 says that we have a correspondence between these sets of primitive roots. In particular, given a primitive root g modulo 11, then g or $g+11$ is a primitive root modulo 22. We see the following correspondence:

$$\begin{aligned} 2 &\mapsto 2 + 11 = 13, \\ 6 &\mapsto 6 + 11 = 17, \\ 7 &\mapsto 7, \\ 8 &\mapsto 8 + 11 = 19. \end{aligned}$$

The last ingredient we need in order to determine which moduli have primitive roots is a statement about roots of unity modulo primes.

Definition 3.46. We say that x is an n th root of unity if $x^n = 1$. Equivalently, x is an n th root of unity modulo n if $x^n \equiv 1 \pmod{n}$.

Example 3.47. The only roots of unity which are integers are 1 and -1

Roots of unity will be discussed in more detail in Chapter 5, section 2, as an application of the Möbius function.

Lemma 3.48. *Let p be a prime number. If $d \mid p - 1$, then*

$$x^d \equiv 1 \pmod{p}$$

has exactly d solutions.

Proof. Write $p - 1 = md$ for some integer m , and let

$$f(x) = 1 + x^d + (x^d)^2 + \cdots + (x^d)^{m-1}.$$

Then we have

$$x^{p-1} - 1 = (x^d - 1)f(x),$$

and so

$$x^{p-1} - 1 \equiv (x^d - 1)f(x) \pmod{p}.$$

Fermat's Little Theorem (Theorem 2.21) says that the left-hand side is zero for exactly $p - 1 = md$ distinct x values. Since p is prime, each such x value must be a zero of either $(x^d - 1)$ or $f(x)$. Each is a polynomial and of degree d and $dm - d$, respectively. Lagrange's Theorem (Theoretical Exercise 3.27) says that a polynomial of degree d can have at most d roots modulo a prime. So we must have that $x^d - 1$ has exactly d 0's modulo p , since to have fewer would contradict that the total number must be $p - 1$. \square

Investigation 3.49. Lemma 3.48 assumes the modulus is prime. However, there still exist roots of unity for composite moduli.

- (a) What is the appropriate generalization of the condition $d \mid p - 1$ for a composite modulus?
- (b) Are there the “correct” number of roots of unity?
- (c) Conjecture a generalization of Lemma 3.48 for composite moduli.

First we prove the existence of primitive roots for powers of odd primes.

Proposition 3.50. *If p is an odd prime, then there is a primitive root modulo p^k for all integers $k \geq 1$.*

Proof. We first consider $k = 1$. Let q be a prime that divides $p - 1$. Let $m \geq 1$ be the maximal power so that $q^m \mid (p - 1)$. An element of order q^m is a solution to

$$a^{q^m} \equiv 1 \pmod{p},$$

and by Lemma 3.48 there are exactly q^m of them. Each such a has multiplicative order dividing q^m . If the order is less than m , then a is also a solution to

$$a^{q^j} \equiv 1 \pmod{p}$$

for some $1 \leq j < m$. There are exactly q^j such solutions for each j . Since

$$q^m > q^{m-1} + q^{m-2} + \cdots + q + 1,$$

there is at least one a that has multiplicative order q^m modulo p , call it a_q . We can do this for each prime divisor q of $p - 1$. Since each prime is distinct by Lemma 3.42, the product

$$\prod_{\substack{q \mid (p-1) \\ \text{distinct}}} a_q$$

has multiplicative order

$$\prod_{\substack{q \mid (p-1) \\ \text{distinct}}} q^m = p - 1.$$

Now we need to consider powers p^k for $k \geq 2$. Assume that g is a primitive root modulo p . We will show that g or $g + p$ is a primitive root modulo p^k .

We proceed by induction: $k = 1$ is true since g is a primitive root modulo p , so assume that $g^{\varphi(p^k)} \equiv 1 \pmod{p^k}$. Let m be the multiplicative order of g modulo p^{k+1} , so by Euler's formula (Theorem 2.31) we must have $m \mid \varphi(p^{k+1}) = p^k(p-1)$. We then have two possibilities $m = \varphi(p^{k+1})$ and we are done, or $m = \varphi(p^k)$ and we are not done. We break the proof into two cases. First assume that

$$g^{p-1} \not\equiv 1 \pmod{p^2}.$$

Then, we can show by induction (Theoretical Exercise 3.28) that

$$g^{\varphi(p^k)} \not\equiv 1 \pmod{p^{k+1}}.$$

Consequently, $m = \varphi(p^{k+1})$.

Now assume

$$g^{p-1} \equiv 1 \pmod{p^2}.$$

Consider $g + p$, which is in the same residue class as g modulo p , so it is still a primitive root modulo p . We compute

$$(g + p)^{p-1} \equiv g^{p-1} + (p-1)pg^{p-2} \not\equiv 1 \pmod{p^2}.$$

In particular, the multiplicative order of $(g + p)$ must be $\varphi(p^{k+1})$. □

Notice that the proof is not constructive in the sense that given a prime p , it does not produce a primitive root. We are reduced to trying every possible residue class to find a primitive root. However, given a prime power p^k , if we can find a primitive root g modulo p , then at least one of g or $g + p$ is primitive root modulo p^k for each $k \geq 2$. (See also Theoretical Exercise 3.28.)

Example 3.51.

- We compute that 6 is a primitive root modulo 13. Checking the condition in the proof, we see that $6^{12} \equiv 144 \not\equiv 1 \pmod{13^2}$, and we can check that 6 is a primitive root modulo 13^k for all $k \geq 1$.
- For $p = 37$ we have 18 as a primitive root. This primitive root satisfies $18^{36} \equiv 1 \pmod{37^2}$, so that 18 is not a primitive root modulo 37^k for $k \geq 2$, but that $18 + 37 = 55$ is a primitive root modulo 37^k for $k \geq 2$.

We can now bring all our work in this section together to state exactly which moduli have a primitive root.

Theorem 3.52 (Primitive root theorem). *A positive integer n has a primitive root if and only if $n = 2, 4, p^k$, or $2p^k$ for an odd prime p and a positive integer k .*

Proof. For $n = 2$ or 4 we can simply observe that 1 is a primitive root modulo 2 and 3 is a primitive root modulo 4. For $n = 2^3 = 8$ there are no primitive roots. Now we proceed by induction on the power 2^k for $k \geq 3$ to show there are no primitive roots by showing that every odd integer has multiplicative order dividing $2^{k-2} < \varphi(2^k) = 2^{k-1}$. For the base case $k = 3$, we check that every odd residue class modulo 8 has multiplicative order dividing 2. For the induction assumption,

we assume that every odd residue class modulo 2^k has multiplicative order dividing 2^{k-2} . In particular, for each odd a ,

$$a^{2^{k-2}} = 1 + m2^k$$

for some integer m . Squaring both sides, we see that

$$a^{2^{k-1}} = 1 + m2^{k+1} + m^2 2^{2k} \equiv 1 \pmod{2^{k+1}}.$$

This shows that any odd integer a has multiplicative order modulo $2^{k-1} < 2^k$ so is not a primitive root. Then, by induction, $n = 2^k$ for $k \geq 3$ does not have a primitive root.

Proposition 3.50 states that powers of odd primes have primitive roots, so we are left to consider the case where n is divisible by two distinct odd primes. In this case we can write $n = mp^k$ for an odd prime p and an integer $m \geq 3$ with $\gcd(m, p) = 1$. The Euler totient function is multiplicative (Lemma 5.7), so we compute

$$\varphi(n) = \varphi(m)\varphi(p^k),$$

where both $\varphi(m)$ and $\varphi(p^k)$ are even. In particular by Euler's formula (Theorem 2.31) we have the following two congruences for any a relatively prime to n :

$$a^{\varphi(n)/2} \equiv (a^{\varphi(m)})^{\varphi(p^k)/2} \equiv 1 \pmod{m},$$

$$a^{\varphi(n)/2} \equiv (a^{\varphi(p^k)})^{\varphi(m)/2} \equiv 1 \pmod{p^k}.$$

By the Chinese remainder theorem (Theorem 2.44), these congruences imply that

$$a^{\varphi(n)/2} \equiv 1 \pmod{n},$$

and, thus, a is not a primitive root. □

COMPUTATIONAL EXERCISES


3.1. Find all the residue classes which are quadratic residues modulo 61.

3.2. Find all the primes $p < 1000$ for which 17 is a quadratic residue.

 **3.3.** Compute the following Legendre symbols.

a. $\left(\frac{328}{13}\right)$


b. $\left(\frac{420}{17}\right)$

 **3.4.** Compute the following Jacobi symbols $\left(\frac{a}{n}\right)$. Determine whether a is quadratic residue or nonresidue modulo n .

a. $\left(\frac{42}{15}\right)$

b. $\left(\frac{117}{35}\right)$

3.5. Determine all the quadratic residues modulo 1624.

 **3.6.** Find all solutions to

$$x^2 \equiv 1 \pmod{35}.$$

3.7. We say that the set of points (x_1, \dots, x_n) satisfying

$$x_1^2 + \dots + x_n^2 = 1$$


is a hypersphere of dimension n and radius 1. Count the number of points on the hypersphere modulo p for $n = 1, 2, 3, 4$ and $p = 3, 5, 7, 11$.

3.8. Count 8th roots of unity.

- a. Find all the solutions to $x^8 \equiv 1 \pmod{31}$.
- b. Find all the primes p less than 100 for which the equation

$$x^8 \equiv 1 \pmod{p}$$


has eight solutions.

 **3.9.** Assume that 5 is a primitive root modulo 23. Find all solutions to


$$x^4 \equiv 2 \pmod{23}.$$

3.10. Find all the solutions to $x^{12} \equiv 87 \pmod{101}$.

3.11. Determine the m th power residues modulo 11 for $m = 2, 3, 4, 5, 6$. In other words, determine all the a such that $x^m \equiv a \pmod{11}$.

 **3.12.** Determine the list of moduli n with $2 \leq n \leq 30$ that do not have a primitive root.

3.13. Determine all the primitive roots modulo 38.

 **3.14.** Determine a primitive root modulo 5^{10} .

THEORETICAL EXERCISES

3.15. Let $p > 2$ be a prime. Prove that $\left(\frac{-1}{p}\right) = 1$ if and only if $p \equiv 1 \pmod{4}$.

3.16. Prove that 7 is a quadratic residue modulo a prime $p > 2$ if and only if $p \equiv \pm 1, \pm 3, \text{ or } \pm 9 \pmod{28}$.

3.17. For distinct odd primes p and q , prove that p is a quadratic residue modulo q if and only if p^{-p} is a quadratic residue modulo q .

3.18.

- a. Let $p > 2$ be a prime. Prove that $\left(\frac{3}{p}\right) = 1$ if and only if $p \equiv 1 \pmod{3}$.
- b. Prove that $\left(\frac{-3}{p}\right) = 1$ if and only if $p \equiv 1 \pmod{3}$. *Hint:* Let g be a primitive root for p , and consider the solutions to

$$x^3 - 1 \equiv (x - 1)(x^2 + x + 1) \equiv 0 \pmod{p}$$

in terms of g .

Conclude that $\left(\frac{-3}{p}\right) = \left(\frac{p}{3}\right)$.

3.19.

- a. Let $n = pq$ for distinct odd primes p and q . Prove that a is a quadratic residue modulo n if and only if a is a quadratic residue modulo both p and q . Conclude that the number of quadratic residues modulo n is $\frac{(p-1)(q-1)}{4}$.
- b. Let $n = \prod_{i=1}^m p_i$ be a product of distinct odd primes. Prove that a is a quadratic residue modulo n if and only if it is a quadratic residue modulo each p_i , $1 \leq i \leq m$.

3.20. Prove Theorem 3.22. Let n and m be positive integers. Let a and b be integers with $\gcd(ab, n) = 1$.

- a. If $a \equiv b \pmod{n}$, then $\left(\frac{a}{n}\right) = \left(\frac{b}{n}\right)$.
- b. $\left(\frac{ab}{n}\right) = \left(\frac{a}{n}\right) \left(\frac{b}{n}\right)$.
- c. $\left(\frac{-1}{n}\right) = (-1)^{\frac{n-1}{2}}$.
- d. $\left(\frac{2}{n}\right) = (-1)^{\frac{n^2-1}{8}}$.
- e. If $\gcd(n, m) = 1$, then

$$\left(\frac{n}{m}\right) \left(\frac{m}{n}\right) = (-1)^{\frac{m-1}{2} \frac{n-1}{2}}.$$

3.21. Let n and m be positive integers. Let x and a be integers such that

$$x^m \equiv a \pmod{n}.$$

Prove that $\gcd(x, n) = 1$ if and only if $\gcd(a, n) = 1$.

3.22. Let n be a positive integer. We define the *Carmichael function* $\lambda(n)$ as the smallest positive integer such that

$$a^{\lambda(n)} \equiv 1 \pmod{n}$$

for all a with $\gcd(a, n) = 1$.

Prove that n has a primitive root if and only if $\lambda(n) = \varphi(n)$.

3.23. Use primitive roots to prove Euler's criterion (Theorem 3.10).

3.24. Let n be a positive integer. Prove that if there is one, then there are $\varphi(\varphi(n))$ primitive roots modulo n .

3.25. Prove that for a prime p and a positive integer n , if g is a primitive root modulo p^n , then g is a primitive root modulo p .

3.26. Let n be a positive integer that has a primitive root g . Prove that g^m is a primitive root modulo n if and only if m is relatively prime to $\varphi(n)$.

3.27. Lagrange's Theorem: Let p be a prime. Let $f(x)$ be a polynomial of degree d with integer coefficients and with at least one coefficient not divisible by p . Prove that

$$f(x) \equiv 0 \pmod{p}$$

has at most d solutions. *Hint:* Try using induction on the degree of f .

3.28. Let p be an odd prime with primitive root g . Prove that if $g^{p-1} \not\equiv 1 \pmod{p^2}$, then $g^{\varphi(p^k)} \not\equiv 1 \pmod{p^{k+1}}$ for all $k \geq 1$.

EXPLORATION EXERCISES

3.29 (Quadratic residues). Consider the following inverse problem for quadratic residues. Given a nonsquare integer a , determine the set of moduli n such that a is a quadratic residue modulo n .

- a. Assume that $p = 2^k$ for a positive integer k .
- b. Assume that p is an odd prime power.
- c. Assume that $n = pq$ is the product of distinct primes.
- d. Consider any n .

3.30 (Higher order reciprocity).

- a. Cubic reciprocity: $x^3 \equiv a \pmod{p}$.
 1. Define an equivalent of the Legendre symbol. Can you find a criterion like Euler's criterion (Theorem 3.10)?
 2. State a version of cubic reciprocity (Theorem 3.13 and Theorem 3.14).
 3. What about cubic Jacobi symbols?
- b. Higher order reciprocities: $x^n \equiv a \pmod{p}$ for $n > 3$. Consider the same questions as for cubic reciprocity.

3.31 (Roots of unity modulo n). We say that a is an m th root of unity modulo n if

$$a^m \equiv 1 \pmod{n}.$$

We will be counting roots of unity modulo various n , so define $C(m, n)$ to be the number of m th roots of unity modulo n . In particular, we are counting the number of solutions to

$$x^m - 1 \equiv 0 \pmod{n}.$$

- a. Consider the values $C(\lambda(n), n)$, where $\lambda(n)$ is the Carmichael function (see Theoretical Exercise 3.22).
- b. Given positive integers k, m, n with $m \mid n$, what can you say about $C(k, m)$ and $C(k, n)$?
- c. Given positive integers k, m, n , what can you say about $C(k, mn)$?
- d. Given positive integers k, m, n , what can you say about $C(km, n)$ in terms of $C(k, n)$ and $C(m, n)$? Perhaps consider first the case where $\gcd(k, m) = 1$.
- e. Let n be a positive integer, and let p be a prime. What can you say about $C(p^k, n)$ for $k \geq 1$?

If m is the smallest positive integer such that $a^m \equiv 1 \pmod{n}$, we say that a is a *primitive m th root of unity*.

- f. Can you say anything about the number of primitive m th roots modulo n ?
- g. For which integers m are there primitive m th roots of unity modulo a prime p ?

- h.** For which integers m are there primitive m th roots of unity modulo a composite n . What if you assume there are primitive roots modulo n ? What if there are not?

3.32 (m th roots). Number of m th roots.

- a.** For prime moduli, half of all (nonzero) residue classes are squares. How many are cubes? Fourth powers? m th powers?
- b.** What about composite moduli?
- c.** How many solutions can $x^m \equiv a \pmod{p}$ have?
- d.** What about composite moduli?

Secrets

An important application of number theory is the protection of information, for example through cryptography. With so much of our financial and personal information available on-line in today's digital world, the protection of information is of paramount importance to nearly everyone. The idea of using mathematics to protect information is not new, and it was even used by Julius Caesar¹ in the Roman Empire. The basic premise is that some piece of information (a message) needs to be sent from one person to another, but if someone other than the intended recipient intercepts the message, it should be unintelligible. In the days of Julius Caesar, the “interception” was physical interception of a physically written message. For modern communication, the interception could be someone “listening in” as the message travels over a network.

Question 4.1. How do you communicate a message so that only the intended recipient can read it?

Julius Caesar implemented a simple system where a strip of paper was wrapped around a cylinder of some specified diameter. Then the message was written across the cylinder, so that the unwrapped strip was a column of jumbled letters. Anyone intercepting the message would simply have a long strip of paper containing jumbled letters. This jumbled strip would be the *encrypted* message. The intended recipient would have a cylinder of the same diameter with which to read the message. As you can imagine, with a little trial and error it would be simple to “decrypt” the message by trying different size cylinders until you found one that produced a readable message. Since most combinations of letters do not produce language, you could guess that this readable message was, in fact, the correct message.

However, interception is not the only problem a message may face. For example, what if military messages were intercepted by the enemy and replaced with different messages intended to disrupt? Or what if a message were sent to your

¹Gaius Julius Caesar (100–44 B.C.E.) was a Roman general and emperor.

bank to transfer funds from one account to another? What if the tampering simply changed the destination account of a valid bank transfer you requested? This type of tampering, changing only part of the message, is subtle and can be difficult to detect because the rest of the message will be authentic. The counterfeit bank message brings up yet another problem: How does the bank know that the message to transfer funds actually came from the owner of the account? The concept of marking messages with a unique identifier of the sender, such as a wax imprint of a personal seal, is intended both to prevent tampering and to verify the identity of the sender. We now state the two problems of tampering and sender verification as formal questions.

Question 4.2. How do you detect if a received message has been tampered with?

Question 4.3. How do you verify the identity of the sender of a message?

We will discuss solutions to all three questions stated so far in this chapter. However, we will discuss only the mathematics of the solutions. Implementing the solutions in a real system encounters a host of other problems. The methods discussed here represent only a small fraction of topics in cryptography and associated problems.

Investigation 4.4. Determine situations in your life where the protection of information plays an important role.

- (a) In what situations do you send information that it is important that only the designated recipient can read?
- (b) In what situations do you send information that it is important that arrives unchanged?
- (c) In what situations is it important the sent information can be verified as coming from you?
- (d) What about situations where you are the one receiving the information?

1. Basic Ciphers

A function used to encrypt information is called a *cipher*. We apply mathematics to create a cipher by considering the alphabet as an ordered set of 26 letters corresponding to the numbers 0 through 25; in particular, $a \mapsto 0, b \mapsto 1$, etc. So we can associate lists of numbers to words:

$$2, 0, 19 \mapsto \text{cat}.$$

In this way, if we apply any invertible operation (a cipher) to that list of numbers, we can encrypt and decrypt the word.

Example 4.5. Consider taking the operation

$$x \mapsto x + 1 \pmod{26}.$$

Then we would have

$$2, 0, 19 \mapsto 3, 1, 20.$$

Thus, we have encrypted “cat” as “dbu”. To read the message, we apply the inverse operation

$$x \mapsto x - 1 \pmod{26}$$

so that

$$3, 1, 20 \mapsto 2, 0, 19.$$

Thus, we recover “cat” from “dbu”.

This type of cipher is called a *shift cipher*. We have *shifted* each letter of the alphabet to a different letter through modular arithmetic, and we can decrypt by shifting back. To communicate with this method, two people need only to agree on the amount to shift. In the previous example, the shift amount was 1. The shift amount is the *shared secret* used in this encryption scheme. Anyone who knows this secret can read the encrypted messages.

Definition 4.6. More generally, we can define an *affine cipher* as the operation

$$x \mapsto ax + b \pmod{n}$$

for integers a, b with $\gcd(a, n) = 1$.

Let’s see why we need $\gcd(a, n) = 1$. To encrypt, we compute

$$ax + b \equiv y \pmod{n}$$

and use y as the encrypted value of x . To decrypt, we must compute x given

$$y \equiv ax + b \pmod{n}.$$

This is the linear congruence equation we studied in Chapter 3. We know exactly when we can solve such equations and how to find the solutions. We go through the process again here for

$$y \equiv ax + b \pmod{n}.$$

We first subtract b from both sides to get

$$y - b \equiv ax \pmod{n}.$$

Now we must multiply both sides by the inverse of a modulo n to have

$$x \equiv a^{-1}(y - b) \pmod{n}.$$

But not all numbers have inverses in modular arithmetic. In fact, we saw in Theorem 2.19 that for a to have an inverse modulo n it must be relatively prime to n , i.e., $\gcd(a, n) = 1$. Otherwise, the operation is not invertible, and we would not be able to decrypt messages.

For an affine cipher, the pair (a, b) is the shared secret. Encryption through an affine cipher is described in Algorithm 4.1.

Unfortunately, affine ciphers are very easy to defeat. It is no surprise that language is not a random collection of characters. In particular, certain letters occur with higher frequency than others. Figure 4.1 shows the frequency at which individual letters occur in English.

Algorithm 4.1. Encryption with Affine Ciphers

Input: a list of characters M and an affine cipher $f(x) = ax + b$

Output: a list of characters E

Algorithm:

- 1: Convert the list of characters M to a list of integers m .
 - 2: Apply f to each entry of m to get a list of integers e .
 - 3: Convert the list of integers e to a list of characters E .
 - 4: Return E .
-

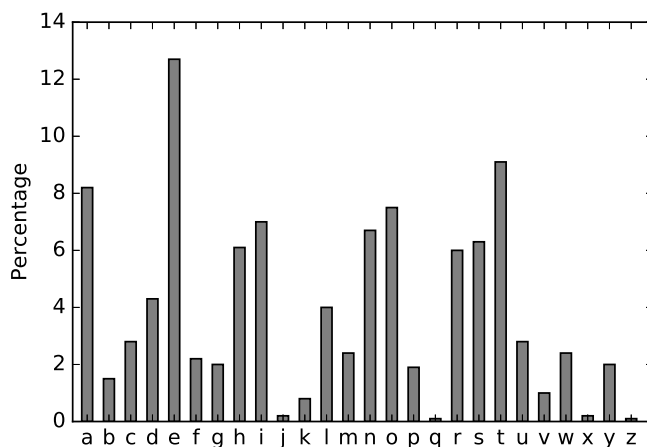


Figure 4.1. English character frequencies

If you analyze an encrypted message for character frequency, you can make an educated guess as to which letter encrypts to which letter. Then you can determine the shared secret (a, b) of the affine cipher using those character pairs to write a system of equations. Unless the message is very short (not enough characters to get a good frequency count), this method will easily defeat an affine cipher.

Investigation 4.7. For this investigation, you will need to work with a partner.

- (a) Agree on a shift cipher and send an encrypted message to your partner. Read each other's messages.
- (b) Send an encrypted message to your partner, but do not tell them what shift you used. Try to decrypt your partner's message.
- (c) Agree on an affine cipher and send an encrypted message to your partner. Read each other's messages.
- (d) Send an encrypted message to your partner, but do not tell them what affine cipher you used. Try to decrypt your partner's message.

Remark. In our simple example (Example 4.5) we chose $n = 26$ using a basic representation of the alphabet, but we could include any number of other characters as long as their representations are agreed upon. Perhaps you have heard the words ASCII or Unicode, which correspond to two common character-to-number representations that computers use.

2. Symmetric Ciphers

The affine cipher can be improved by introducing a *key*. For example, we can encrypt different letters with different affine ciphers based on their location within the message. The correspondence between letter location within the message and which affine cipher to use is the “key.” This method is often called a *Vigenère cipher*. For simplicity, we will only use shift ciphers for our affine ciphers in the following examples.

Example 4.8. Assume we wish to encrypt the message

“The grass is green.”

We can first reduce the message to contain no capitals, spaces, or punctuation.

“thegrassisgreen”.

We can choose a key, say “blue”. The letters of the key “(b,l,u,e)” correspond to the values (1, 11, 20, 4), which correspond to four different shift ciphers:

$$\begin{aligned} f_1(x) &\equiv x + 1 \pmod{26}, \\ f_2(x) &\equiv x + 11 \pmod{26}, \\ f_3(x) &\equiv x + 20 \pmod{26}, \\ f_4(x) &\equiv x + 4 \pmod{26}. \end{aligned}$$

We apply each cipher to one letter of the message until we have encrypted the whole message. In particular, we repeat the key until the entire message is covered:

“theg rass isgr een”,

“blue blue blue blu”.

Then we encrypt each letter of the message with the corresponding cipher to get

“theg rass isgr een”
 “blue blue blue blu”
 ↓ ↓ ↓ ↓
 “usyk slmw jdav fph”.

So the encrypted message is

“usykslmwjdavfph”.

Notice that the first “h” and “r” both encrypted to “s” disrupting a character frequency analysis.

To decrypt, we use the same key but the inverse ciphers, which correspond to the functions

$$\begin{aligned}g_1(x) &\equiv x - 1 \pmod{26}, \\g_2(x) &\equiv x - 11 \pmod{26}, \\g_3(x) &\equiv x - 20 \pmod{26}, \\g_4(x) &\equiv x - 4 \pmod{26}.\end{aligned}$$

Definition 4.9. A *symmetric cipher* is one where the same key is used to both encrypt and decrypt.

The use of a symmetric key makes the cipher more difficult to break with character frequency analysis. However, just as individual letters occur with nonrandom frequency in language (see Figure 4.1), so do pairs or triples of letters. However, you typically need a longer message to get enough statistics for this analysis to work. In fact, the longer your key is, the more data is needed since the frequency statistics of groups of letters is valid only over very large data sets. Taken to the extreme, if your key is the same length as your message, then you cannot break the cipher by analyzing the character frequency in any single message.

Example 4.10. We will use the nursery rhyme “Twinkle, Twinkle, Little Star” to encrypt a message.

“Dear parents, please send money.”

As usual we strip out all capitals, spaces, and punctuation. Then we encrypt with the key of “Twinkle, Twinkle, Little Star. How...” by matching each letter in sequence and shifting.

<i>d</i>	<i>ea</i>	<i>rp</i>	<i>ar</i>	<i>en</i>	<i>ts</i>	<i>pl</i>	<i>ea</i>	<i>se</i>	<i>se</i>	<i>nd</i>	<i>mo</i>	<i>ne</i>	<i>y</i>
<i>t</i>	<i>wi</i>	<i>nk</i>	<i>le</i>	<i>tw</i>	<i>in</i>	<i>kl</i>	<i>el</i>	<i>it</i>	<i>tl</i>	<i>es</i>	<i>ta</i>	<i>rh</i>	<i>o</i>
<i>w</i>	<i>ai</i>	<i>ez</i>	<i>lv</i>	<i>xj</i>	<i>bf</i>	<i>zw</i>	<i>il</i>	<i>ax</i>	<i>lp</i>	<i>rv</i>	<i>fo</i>	<i>el</i>	<i>m</i>

So the encrypted message is

“waiezlvxjbfzwlaxlprvfoelm”.

Since every letter of the message was encrypted with a different shift cipher, it is very difficult to recover the message without knowing the key.

Investigation 4.11. Let’s take a closer look at character frequency analysis.

- Pick a block of text at least six sentences long (the longer the better).
- Determine the frequency at which each character appears in the text. How close are they to the percentages in Figure 4.1?
- Apply a shift cipher to the text, and analyze the character frequencies of the encrypted text. Recover the shift from that analysis.
- Pick a second text the same length at the first. Use this text as an encryption key as in Example 4.8. Do a character analysis of the encrypted text. Can you determine the key?

Several modern cryptographic algorithms in use today are based on symmetric ciphers. For example, the U.S. government approved algorithms AES and outdated DES are both symmetric key ciphers, albeit with much more sophisticated algorithms that just shift ciphers.

3. Diffie–Hellman Key Exchange

For symmetric ciphers, the sender and receiver have to know the same secret. In particular, they need to have some way to ensure one secure communication, such as meeting in person, where they could agree upon the secret. Then, they could start communicating securely over public channels. This problem of an initial agreed-upon secret is a major problem for cryptographic systems. For example, would it really be feasible to meet in person to agree upon a secret key with every person or company you communicate with electronically every time you wanted to initiate secure communication? What we need is a way to choose a secret between two people over public channels so that only they know the secret regardless of who is listening.

Question 4.12. How can two people agree upon a secret over public channels after which they are the only two who can determine the secret?

In 1976 Whitfield Diffie² and Martin Hellman³ solved this initialization problem by determining a way to agree upon a secret using public communications in such a way that an eavesdropper would be unable to determine the secret. Algorithm 4.2 describes their solution.

Algorithm 4.2. Diffie–Hellman Key Exchange

Input: public integers (p, g) where p is a prime and g a primitive root modulo p

Output: a shared secret

Algorithm:

- 1: Alice and Bob exchange integers (p, g) (publicly).
- 2: Alice chooses a secret integer a and sends (publicly) $g^a \bmod p$ to Bob.
- 3: Bob chooses a secret integer b and sends (publicly) $g^b \bmod p$ to Alice.
- 4: They both compute the secret

$$g^{ab} \bmod p = (g^a)^b \bmod p = (g^b)^a \bmod p.$$

The security of Algorithm 4.2 relies on something called the discrete log problem.

Question 4.13 (Discrete log problem). Let p be a prime, and let g, h be integers such that $g^a \equiv h \pmod{p}$ for some positive integer a . Can you find a efficiently knowing g, h , and p ?

²Bailey Whitfield Diffie (1944–) is an American cryptographer.

³Martin Edward Hellman (1945–) is an American cryptographer.

There is no known *efficient* way to solve the discrete log problem. Applying brute force, you could simply try all possible $a < p$, but if p is large enough, this could take many years. In other words, knowing p, g , and $g^a \bmod p$, it is difficult to determine a . In particular, if two people each privately choose an integer, a and b , and agree publicly upon a prime p and primitive root g , then they can create the shared key

$$g^{ab} \bmod p.$$

This works because each can send $g^a \bmod p$ and $g^b \bmod p$ without a listener determining a or b . Then, since

$$(g^a)^b = (g^b)^a,$$

they can both compute the same key.

Example 4.14. Alice and Bob perform the following steps.

- (a) Alice and Bob publicly agree to $(p, g) = (3493417471, 6)$.
- (b) Alice chooses $a = 15234$ and sends to Bob $g^a \bmod p = 599428873$.
- (c) Bob chooses $b = 78695943$ and sends to Alice $g^b \bmod p = 2354817634$.

The secret is therefore

$$g^{ab} \bmod p = 2046929873.$$

Investigation 4.15. For this investigation, you'll need to work with a partner.

- (a) Perform the Diffie–Hellman key exchange with a partner.
- (b) Can you determine the private integer chosen by your partner?

4. Public Key Cryptography (RSA)

Another way to circumvent the problem of agreeing upon keys is to use an asymmetric cipher.

Definition 4.16. A cipher is called *asymmetric* if the encryption and decryption keys are different.

In the asymmetric case, you could make your encryption key public and keep your decryption key private. Then anyone wishing to send you a secure message simply encrypts the message with your public (encryption) key and sends it to you. Since you are the only person who knows the private (decryption) key, you are the only person who can read it.

Rivest,⁴ Shamir,⁵ and Adleman⁶ devised such a scheme, known as RSA, given in Algorithm 4.3.

The mathematics of RSA relies on Euler's formula (Theorem 2.31), which states that for any a relatively prime to n

$$a^{\varphi(n)} \equiv 1 \pmod{n}.$$

⁴Ronald Linn Rivest (1947–) is an American cryptographer.

⁵Adi Shamir (1952–) is an Israeli cryptographer.

⁶Leonard Adleman (1945–) is an American computer scientist.

Algorithm 4.3. RSA Public Key

Set-up: each person performs the following setup

- 1: Generate two prime numbers p, q . Compute $n = pq$.
- 2: Compute the Euler totient $\varphi(n) = (p-1)(q-1)$.
- 3: Choose an integer e such that $\gcd(\varphi(n), e) = 1$.
- 4: Determine d , the inverse of e modulo $\varphi(n)$.
- 5: Publish n, e publicly; d is secret.

Encryption: To send a message m to Alice with her public key (n, e)

- 1: Send $m^e \pmod n$ to Alice.
 - 2: Alice reads it with her private key d as $(m^e)^d \equiv m \pmod n$.
-

This implies that for any a with $\gcd(a, n) = 1$, we have

$$a^k \equiv a \pmod n$$

for any $k \equiv 1 \pmod{\varphi(n)}$. Thus, we can choose public/private key pairs (e, d) that are inverses modulo $\varphi(n)$. Then encryption is performed as

$$a^e \pmod n$$

and decryption as

$$(a^e)^d \pmod n = a^{ed} \pmod n = a.$$

Example 4.17. We setup a public/private key pair following Algorithm 4.3.

- (a) Let $p = 2349827353$ and $q = 5344562761$. Then

$$n = pq = 12558799765623001633.$$

- (b) We compute

$$\varphi(n) = (p-1)(q-1) = 12558799757928611520.$$

- (c) For $e = 1255879975792345523$, we have $\gcd(e, \varphi(n)) = 1$.

- (d) We compute $d = 6225187690547444987$ with the Euclidean algorithm (Algorithm 2.1) and verify that

$$d \cdot e \equiv 1 \pmod{\varphi(n)}.$$

- (e) The public information is (n, e) , and d is the decryption key, which is kept private.

That is great, but asymmetric encryption adds an additional problem. In symmetric encryption, you can verify who sent the message since only the proper sender knows the key. But with asymmetric encryption, everyone knows the public key. For example, if you want to send a secure message to your bank telling them to transfer all your money to a different account, it is important that the bank can verify that it is really you who sent the message. We need the digital equivalent of a signature (or an imprinted wax seal).

Question 4.18. With an asymmetric cipher, how do you verify the sender of a message?

Definition 4.19. A *digital signature* is a piece of information attached to a message that securely identifies the sender.

It is possible to digitally sign messages with RSA. While there are two keys, the order in which they are applied does not actually matter. In other words,

$$m^{de} \equiv (m^d)^e \equiv (m^e)^d \equiv m^{ed} \pmod{n}.$$

Thus, to digitally sign something, you encrypt a piece of information (your signature) with your private key and attach it to the message. Then the receiver can verify the information by decrypting the signature with your public key. Since you are the only one who knows your private key, only signatures created by you will decrypt with your public key. This allows the receiver to verify that you sent the message.

Algorithm 4.4. Digital Signature and Verification with RSA

Input: the sender's RSA data (d, e, n)

Output: True or False

- 1: Compute a piece of information with the sender's identity, h .
- 2: Encrypt h with the private key d as $H = h^d \pmod{n}$.
- 3: Attach H to the message.
- 4: The receiver can verify the sender's identity by computing

$$H^e \equiv (h^d)^e \equiv h \pmod{n}.$$

The security of these systems depends on the inability to compute d knowing (n, e) . However, if we know the factorization of n (i.e., p and q), we can compute $\varphi(n)$ and use the extended Euclidean algorithm (Algorithm 2.1) to compute the inverse. Fortunately, as we have discussed in Chapter 1, factoring large numbers remains an infeasible problem. We explored several different factoring algorithms in Exploration Exercise 1.39. However, the notion of *large* depends on current computing technology and algorithms, so it is continually changing. The National Institute of Standards and Technology (NIST) gives recommended key sizes for both short-term and long-term security; see Table 4.1.

Table 4.1. NIST Recommendations 2012

Algorithm	Recommended key size	Long term security
Asymmetric (RSA)	2^{2048}	2^{3072}
Symmetric (AES/3DES)	2^{112}	2^{128}

Investigation 4.20 (Public key and digital signatures). For this investigation, you will need to work with a partner.

- (a) Choose (n, d, e) and tell your partner (n, e) .
- (b) Using your partner's (n, e) , encrypt a short message (Algorithm 4.3).
- (c) Decrypt the message from your partner with your private key d .
- (d) Choose a personal identifier (perhaps your last name converted to a number). Send your partner a message signed with your identifier (Algorithm 4.4).
- (e) Verify the signature on your partner's message.

5. Hash Functions and Check Digits

Digital signatures are great for verifying the sender of the message, but they provide no assurance that the message contents are valid. In the bank transfer scenario, it is important both that the message to transfer is sent by the account holder and that the transfer details received are the same as the transfer details sent.

Question 4.21. How do you verify that the content of a received message is the same as the content that was sent?

One solution to this problem comes from what is called a *hash function* or a *one-way function*.

Definition 4.22. A *one-way function* is a function for which the input cannot be determined from the output.

Definition 4.23. A *hash function* is a one-way function whose output has a fixed length.

Example 4.24. We can take the hash function that simply uses the last two digits of any number

$$h(n) = n \mod 100.$$

For example,

$$h(238746283764) = 64.$$

This function is clearly one-way since we cannot recover 238746283764 from just 64. However, since many numbers have the same output hash value, this is not a good hash function for cryptographic purposes.

Definition 4.25. We say m, n *collide* for h if $h(n) = h(m)$.

A common hash function is MD5, which produces a 32-digit number associated to any piece of information. However, there are now several known collision attacks against MD5, meaning that it is no longer secure. The National Security Agency (NSA) developed SHA-1 standing for “secure hash algorithm,” but SHA-1 was also broken at about the same time as MD5, in 2005. The NSA developed SHA-2 to replace it. SHA-2 is similar in design to SHA-1 and, as of this writing, has not yet to our knowledge been broken. To develop a hash function based on a different underlying algorithm, NIST held a design competition and selected SHA-3 in 2012.

As you can see, developing a secure hash function is a difficult and ongoing, but feasible, process.

To use a hash function to prevent tampering, we can hash the contents of the message and attach it as part of the digital signature. This way, if the contents are changed, the hash value of the modified message will not match the hash value in the signature and the receiver knows the message has been tampered with. Since the hash value is in the digital signature, it is protected by the sender's RSA key and cannot be (feasibly) altered to correspond to the altered message. Algorithm 4.5 describes the procedure to verify the contents of a message from a hash included in the digital signature.

Algorithm 4.5. Digital Signature and Verification with Hash Function

Input: the sender's RSA data (d, e, n)

Output: True or False

Algorithm:

- 1: Compute the hash value v of the original message.
 - 2: Attach $v^d \bmod n$ to the message.
 - 3: The receiver computes $(v^d)^e \equiv v \pmod{n}$.
 - 4: The receiver computes the hash of the received message and compares it to v .
-

Hash functions used for cryptography should satisfy the following two properties.

- (a) Given a hash value v , it should be difficult to find any message m such that $v = h(m)$ (one-way).
- (b) Given an input m_1 , it should be difficult to find another input m_2 , where $m_1 \neq m_2$ such that $h(m_1) = h(m_2)$ (few collisions).

We next discuss a few simple hash functions.

5.1. Check digits.

5.1.1. Universal Product Code (UPC). An example is the 12-digit UPC found on items for sale. It has 11 digits and a *check digit*. To compute the check digit we perform the following steps on the first 11 digits (numbered left to right).

- (a) Add the digits in the odd-numbered positions together and multiply by 3.
- (b) Add the digits in the even-numbered positions to the result.
- (c) Take the result modulo 10 and subtract this from 10 to get the 12th digit.

Example 4.26. Consider the UPC 123923081939. We perform these steps.

- (a) $3 \cdot (1 + 3 + 2 + 0 + 1 + 3) = 30$.
- (b) $30 + (2 + 9 + 3 + 8 + 9) = 61 \equiv 1 \pmod{10}$.
- (c) Last digit $10 - 1 = 9$.

So, if you tamper with one of the first 11 digits of a UPC number, the last digit will often be wrong—but not always!

Investigation 4.27.

- (a) Find two distinct 11-digit numbers that have the same UPC check digit.
- (b) What percentage of 11-digit numbers have the same UPC check digit?

5.1.2. International Standard Book Number (ISBN). The final character of a 10-digit International Standard Book Number is a check digit. The digit is chosen so that multiplying each digit by its position in the number (counting from the right) and taking the sum of these products modulo 11 is 0. The digit farthest to the right (which is multiplied by 1) is the check digit. It may need to have the value 10, which is represented as the letter X.

Example 4.28. Consider the first nine digits of the ISBN 0375873503; we verify the last digit by solving the following equation for a :

$$a + 2 \cdot 0 + 3 \cdot 5 + 4 \cdot 3 + 5 \cdot 7 + 6 \cdot 8 + 7 \cdot 5 + 8 \cdot 7 + 9 \cdot 3 + 10 \cdot 0 = a + 228 \equiv a + 8 \pmod{11},$$

so we need $a = 3$.

Investigation 4.29.

- (a) Find two distinct ISBN numbers that have the same check digit.
- (b) What percentage of ISBN numbers have the same check digit?

5.2. CvHP hash function. The CvHP hash function is named after its developers Chaum,⁷ van Heijst,⁸ and Pfitzmann.⁹

Let p be a prime such that $q = \frac{p-1}{2}$ is also prime. Let a and b be two primitive roots modulo p .

Define a hash function as

$$h(m) = a^{x_0} b^{x_1} \pmod{p},$$

where $m \equiv x_0 + x_1 q \pmod{q^2}$.

Example 4.30. Let $p = 59$ with primitive roots $a = 6$ and $b = 13$; then $q = 29$ is also prime. Let $m = 1347$, then we have

$$1347 \equiv 13 + 17 \cdot 29 \pmod{29^2}.$$

We compute

$$h(1347) = 6^{13} 13^{17} \pmod{59} = 49.$$

Let s be such that $a^s \equiv b \pmod{p}$. Then (x, y) and $(x + s, y - 1)$ hash to the same value. So finding a collision (and, thus, being able to reverse the hash function in this case) is as hard as the discrete log problem (Question 4.13).

Investigation 4.31. Find p, q, a, b and two distinct numbers that hash to the same value with the CvHP hash function.

⁷David Lee Chaum (1955–) is an American cryptographer.

⁸Eugène van Heijst is a Dutch cryptographer.

⁹Birgit Pfitzmann is a German computer scientist.

6. Secret Sharing

Another common problem with information security is sharing secrets. What if you have a group of five people who have money in a common account? You want to ensure that only by a majority agreement can the money be used. In particular, you want a secret that any three of the people can determine but one or two cannot.

Question 4.32. How can a secret be split among n people so that any m of them can determine the secret, but no fewer?

We examine Mignotte's¹⁰ solution, which uses the Chinese Remainder Theorem (Theorem 2.44).

Definition 4.33. We call a sequence of n positive integers a (k, n) -Mignotte sequence if they are pairwise relatively prime, and the product of the smallest k of them is larger than the product of the largest $k - 1$ of them.

Example 4.34. The sequence $(5, 7, 11)$ is a $(2, 3)$ -Mignotte sequence since the product of any two of them is larger than any one of them:

$$\begin{aligned} 5 \cdot 7 &> 11, \\ 5 \cdot 11 &> 7, \\ 7 \cdot 11 &> 5. \end{aligned}$$

But the sequence $(5, 7, 37)$ is not a $(2, 3)$ -Mignotte sequence since

$$5 \cdot 7 < 37.$$

The sharing scheme is detailed in Algorithm 4.6.

Algorithm 4.6. Secret Sharing with Mignotte Sequences

Input: positive integers k and n

Output: n private keys

Algorithm:

- 1: Choose a (k, n) -Mignotte sequence $p_1 < p_2 < \cdots < p_n$.
 - 2: Let $A = \prod_{i=0}^{k-2} p_{n-i}$ and $B = \prod_{i=1}^k p_i$.
 - 3: Choose any integer S (the secret) with $A < S < B$.
 - 4: The i th person receives the number $(S \bmod p_i)$ for $1 \leq i \leq n$.
-

This scheme works since the Chinese Remainder Theorem for the system of congruences

$$\begin{aligned} x &\equiv a_1 \pmod{p_1} \\ &\vdots \\ x &\equiv a_m \pmod{p_k} \end{aligned}$$

¹⁰Maurice Mignotte is a French mathematician.

gives a unique solution

$$x \pmod{p_1 \cdots p_k}.$$

In order to determine the secret S , the product of the p_i 's used must be larger than S . Thus, enough people must contribute their piece of information so that the modulus of the resulting Chinese Remainder Theorem problem is large enough. In particular, since $A < S$, $k - 1$ people cannot determine S . But since $S < B$, then k (or more) people can determine S .

Example 4.35. We construct an example that has three people, of which at least two are needed to determine the secret. We need a length (2,3)-Mignotte sequence. One such sequence is

$$\{5, 7, 11\}.$$

We compute

$$A = 11, \quad B = 35.$$

Let our secret S be

$$S = 29.$$

Then each person receives their individual secrets, the pairs $(S \pmod{p_i}, p_i)$, which in this case are

$$(4, 5), (1, 7), \text{ and } (7, 11).$$

Recall that these secrets correspond to a congruence equation

$$(4, 5) \rightarrow S \equiv 4 \pmod{5}.$$

We can perform the Chinese Remainder Theorem with any two of them to get

$$\begin{aligned} \{(4, 5), (1, 7)\} &\rightarrow S \equiv 29 \pmod{35}, \\ \{(4, 5), (7, 11)\} &\rightarrow S \equiv 29 \pmod{55}, \\ \{(1, 7), (7, 11)\} &\rightarrow S \equiv 29 \pmod{77}. \end{aligned}$$

Investigation 4.36. For a group with $n = k + 1$ people, have each person perform the following steps:

- (a) Create a (k, n) -Mignotte sequence.
- (b) Choose a secret S and determine the n key pairs $(S \pmod{p_i}, p_i)$ for $1 \leq i \leq n$.
- (c) Have the other k members of the group determine the secret S .

COMPUTATIONAL EXERCISES

4.1. Decrypt the following text by trying all possible affine ciphers:

a.

*RDXUJHNFQUQJFXZWJNSRFYMJRFYNHXWJXYJIUF
WYNHZQFWQDTSNYXUZWJQDXUJHZQFYNAJUFWYG
GTQEFST*

b.


EW H H Z Z H A X H F M V N E W H V N E B X Z B Z B E Z M U H H C F V R
X N A E F U

4.2. Decrypt the following text by frequency analysis. The message was encrypted with an affine cipher.

VYVTHPHYYYBOGZORZUTVWBOHMKBOBLOBYYQHYGBOPHYN
 QZRDKYVIHYBWYQBYQBZOXZGUDPMBOTNVYQTZPDRQTDR
 RBTTWVWUZYKBHIBDTNVYQYQBEOZZGTZGYQBYQBZOBPT
 QBWVTRZIBOBWVUYODYQBDKBOHUWKHLOHULBNQZQHIBU
 ZYWVTWHVUBWYQVTFVUWZGOBTBHORQQHIBEOZIBWPZTY
 ZGYQBTBYQBZOBPTHUWQHIBIBUTDMTYVYDYBWBSYBUT
 VIBYQBZOVBTGZOYQBVTZKHBYWEOZEZTVYVZUTZGGBOP
 HYMDYYQBOBHOBTBIBOHKEOZZGTNQVRQQHIBOBTVTYBW
 YQBVBOGGZOYTHKBLBUWOB

4.3. Encrypt the following messages using the given keys and (Vigenère) shift ciphers.

- a. “The understanding of mathematics is necessary for a sound grasp of ethics.”
Key = “SOCRATES”.
- b. “If only I had the theorems! Then I should find the proofs easily enough.”
Key = “RIEMANN”.

 **4.4.** Given the following affine cipher, determine its inverse cipher:

$$3x + 7 \pmod{26}.$$

4.5. Solve the following discrete logarithms for x :

- a. $2^x \equiv 73 \pmod{523}$.
- b. $2107^x \equiv 48125 \pmod{52321}$.

4.6. You eavesdrop on the following Diffie–Hellman key exchange:

- a. $(p, g) = (5227, 1352)$.
- b. Alice to Bob: 5212.
- c. Bob to Alice: 1453.


What is the secret?

4.7. Given the following public key and modulus for an RSA public key system, find the private key:

modulus: 448832842251643, public key: 298374833.

 **4.8.** Which of these is a valid UPC?

- a. 121004779706
- b. 301222330111

 **4.9.** Compute the check digit for the following first nine digits of an ISBN:

- a. 345720004
- b. 019201001

4.10. Compute the following CvHP hash values for $(p, q) = (167, 83)$ and $(a, b) = (37, 70)$:

- a. $m = 587$.
- b. $m = 7934$.

4.11.

- a. Find (p, q) and (a, b) for a CvHP hash function.
- b. Find two distinct m that hash to the same value.

4.12. Construct a Mignotte sequence of the following types:

- a. $(k, n) = (2, 5)$.
- b. $(k, n) = (3, 5)$.
- c. $(k, n) = (4, 5)$.

EXPLORATION EXERCISES

4.13 (Rabin¹¹ coin flipping). For two people:

Step 1: Alice chooses two primes p and q , computes $N = pq$, and sends N to Bob.

Step 2: Bob chooses a number x and sends $x^2 \bmod N$ to Alice.

Step 3: Alice finds the four possible square roots of x^2 as $\{r_1, r_2, r_3, r_4\}$. Alice picks one of these and sends it to Bob.

Half of Alice's choices allow Bob to find a factor of N , and half of Alice's choices do not. If Bob can find a factor of N , he wins; if he cannot, Alice wins.

- a. Perform a coin flipping. Take turns being Alice and Bob. Does it appear to be an equal chance of winning for each player?
- b. Do your best to cheat.

Modified Rabin (Blum¹²):

Step 1: Alice chooses $N = pq$ and sends N to Bob.

Step 2: Bob chooses $x < \frac{N}{2}$ and sends $x^2 \bmod N$ to Alice.

Step 3: Alice replies with a guess of “larger” or “smaller”.

Step 4: Bob sends x to Alice.

Step 5: Alice sends (p, q) to Bob.

¹¹Michael Oser Rabin (1931–) is an Israeli computer scientist.

¹²Manuel Blum (1938–) is a Venezuelan computer scientist.

The outcome is whether Alice guessed correctly if x is the larger/smaller of the two square roots that are less than $N/2$.

- c. Perform a coin flipping. Take turns being Alice and Bob. Does it appear to be an equal chance of winning for each player?
- d. Do your best to cheat.

4.14 (RSA public key). For at least two people:

Step 1: Each person chooses a modulus and public/private key (n, d, e) .

Step 2: Share public keys with each other.

Step 3: Send each other an encrypted message.

Step 4: Send each other a signed message using a hash of the message.

- a. Try to find each other's private keys.
- b. Tamper with a signed message and see if the other person detects the tampering.
- c. Try to forge a message.

The following are some scenarios to explore while trying to break your partner's encryption scheme.

- d. Are there "good" or "bad" choices for n ?
 - 1. What if n is prime?
 - 2. What if you can factor n ?
 - 3. What if you choose a completely random n ?
- e. Are there "good" or "bad" choices for (d, e) ?
- f. Besides the parameters (n, d, e) , what are some other possible points of attack in a real-world situation?

4.15 (Secret sharing). For three or more people:

- a. Set up a shared secret system based on the number of people you have.
- b. Determine the secret with different combinations of the people.
- c. Try to determine the secret with fewer people.
- d. Encrypt a message with the secret.
- e. Try to decrypt the message with fewer people.
- f. Can you set up a sharing scheme where one person carries twice the weight of the other people? For example, can you set up a scheme where you need four people to determine the secret without person A , but only three people if one of them is person A ?

How secure is your sharing scheme?

- g. Are certain Mignotte sequences "more secure" than others?
- h. What characteristics does a "good" sequence have?
- i. What characteristics does a "bad" sequence have?

Arithmetic Functions

Functions whose domains (possible inputs) are the positive integers are called *arithmetic functions*. Arithmetic functions are somewhat peculiar and require a little getting used to. In calculus, functions are typically something like $f(x) = x^2$. In particular they are defined by equations. For $f(x) = x^2$, we can let x be any number and compute its square as $f(x)$. We could restrict x to only (positive) integers to make it satisfy the definition of being an arithmetic function. However, this restriction does not have any intrinsic number theoretic information. We are interested in arithmetic functions whose values give information on the arithmetic nature of the input such as information about their factorization. We have already seen one such function in the Euler totient function $\varphi(n)$, which computes the number of positive integers less than n and relatively prime to n . In this chapter, we introduce several arithmetic functions and explore some of their properties, starting with the totient function.

1. Euler Totient Function

We introduced the Euler totient function in Chapter 2 and applied it to modular arithmetic. Now we examine properties of the function itself.

Definition 5.1. For a positive integer n , we define the function $\varphi(n)$ as the number of positive integers less than n that are relatively prime to n :

$$\varphi(n) = \#\{m \in \mathbb{N} : m < n \text{ and } \gcd(m, n) = 1\}.$$

The function φ is called the *Euler totient function* or the *Euler phi function*.

Example 5.2.

- $\varphi(3) = \#\{1, 2\} = 2$.
- $\varphi(4) = \#\{1, 3\} = 2$. We exclude 2 since $\gcd(2, 4) = 2$.
- $\varphi(5) = \#\{1, 2, 3, 4\} = 4$.
- $\varphi(6) = \#\{1, 5\} = 2$. We exclude $\{2, 3, 4\}$.

Question 5.3. Can you find a formula for $\varphi(n)$ in terms of n ? In other words, compute $\varphi(n)$ without checking to see if $\gcd(m, n) = 1$ for each $1 \leq m < n$.

One way to approach Question 5.3 is to look at classes of numbers for which you know the answer.

Investigation 5.4.

- (a) Compute $\varphi(n)$ when $n = p$, a prime number. Determine a formula for $\varphi(p)$ in terms of p .
- (b) Compute $\varphi(n)$ when $n = p^e$, a power of a prime number. Determine a formula for $\varphi(p^e)$ in terms of p and e .
- (c) Compute $\varphi(n)$ when $n = pq$, a product of distinct prime numbers. Determine a formula for $\varphi(pq)$ in terms of p and q .

Now that you have made some conjectures as to what the formulas for $\varphi(n)$ should be in Investigation 5.4, we need to prove these formulas. However, it is often a good idea to generate some additional data based on your conjectures before trying to prove them. By testing the conjectures on data not used to make those conjectures, you can often identify errors. For example, looking at n a power of 2 we see the following:

- $\varphi(2) = 1$.
- $\varphi(4) = 2$.
- $\varphi(8) = 4$.
- $\varphi(16) = 8$.

It seems as if $\varphi(2^e) = 2^{e-1}$. To check, let's look at powers of a different prime:

- $\varphi(3) = 2$.
- $\varphi(9) = 6 = 2 \cdot 3$.
- $\varphi(27) = 18 = 2 \cdot 9$.
- $\varphi(81) = 54 = 2 \cdot 27$.

So we see that $\varphi(3^e) \neq 3^{e-1}$ and our initial guess was wrong. If you have not already done so in Investigation 5.4, try to correct the formula for $\varphi(p^e)$ before continuing.

We first show that $\varphi(n)$ is multiplicative, which proves that determining $\varphi(n)$ for n a prime power is enough to determine $\varphi(n)$ for any n .

Definition 5.5. An arithmetic function f is *multiplicative* if for relatively prime positive integers a and b we have $f(ab) = f(a)f(b)$.

We say f is *completely multiplicative* if $f(ab) = f(a)f(b)$ for all positive integers a and b .

Example 5.6.

- The function $f(n) = n$ is completely multiplicative.
- The function $f(n) = \gcd(n, k)$ for a fixed positive integer k is multiplicative (Theoretical Exercise 5.25).
- The function $f(n) = n + 1$ is not multiplicative.

Lemma 5.7. *Let m, n be positive integers such that $\gcd(m, n) = 1$. Then*

$$\varphi(mn) = \varphi(m)\varphi(n).$$

Proof. We consider two sets:

$$A = \{a : \gcd(a, mn) = 1, 1 \leq a < mn\},$$

$$B = \{(b, c) : \gcd(b, m) = 1, \gcd(c, n) = 1, 1 \leq b < m, 1 \leq c < n\}.$$

In particular $\#A = \varphi(mn)$ and $\#B = \varphi(m)\varphi(n)$. We need to show there is a bijection between A and B .

Since m and n are relatively prime, the Chinese Remainder Theorem (Theorem 2.44) provides a bijection between solutions to the system

$$x \equiv b \pmod{m},$$

$$x \equiv c \pmod{n}$$

and residue classes $x \equiv a \pmod{mn}$. First assume that $\gcd(a, mn) = 1$. Then a must be relatively prime to both m and n , so the corresponding b and c are relatively prime to m and n , respectively. If b and c are relatively prime to m and n , respectively, then x must be relatively prime to m and n , so c must also be relatively prime to mn .

Thus, we have a bijection between the sets, so they have the same cardinality. \square

We are ready to prove the general formula for $\varphi(n)$.

Theorem 5.8.

(a) *Let p be a prime, and let e be a positive integer. Then*

$$\varphi(p^e) = (p-1)p^{e-1} = p^e - p^{e-1} = p^e \left(1 - \frac{1}{p}\right).$$

(b) *Let $p_1 \cdots p_r$ be the distinct prime factors of n , then*

$$\varphi(n) = n \left(1 - \frac{1}{p_1}\right) \cdots \left(1 - \frac{1}{p_r}\right).$$

Proof.

(a) From the set $\{1, 2, \dots, p^e\}$ the numbers not relatively prime to p^e are exactly those divisible by p , that is, the set $\{p, 2p, 3p, \dots, p^e\}$. There are p^{e-1} of those, so the total number of relatively prime numbers is

$$p^e - p^{e-1} = (p-1)p^{e-1} = p^e \left(1 - \frac{1}{p}\right).$$

- (b) For this part, we use part (a) and Lemma 5.7. Given $n = p_1^{e_1} \cdots p_r^{e_r}$, we compute

$$\begin{aligned}
 \varphi(n) &= \varphi(p_1^{e_1} \cdots p_r^{e_r}) \\
 &= \varphi(p_1^{e_1}) \cdots \varphi(p_r^{e_r}) \\
 &= p_1^{e_1} \left(1 - \frac{1}{p_1}\right) \cdots p_r^{e_r} \left(1 - \frac{1}{p_r}\right) \\
 &= n \left(1 - \frac{1}{p_1}\right) \cdots \left(1 - \frac{1}{p_r}\right). \quad \square
 \end{aligned}$$

We can now compute the Euler totient function for any integer we can factor. Note that this means that computing $\varphi(n)$ in this manner is as hard as factoring, which is very hard! Even so, this is much faster than checking the greatest common divisor for every positive integer less than n .

Example 5.9.

- (a) For $n = 81$, we compute

$$\varphi(81) = \varphi(3^4) = 3^3(3 - 1) = 54.$$

- (b) For $n = 12$, we compute

$$\begin{aligned}
 \varphi(12) &= 12 \left(1 - \frac{1}{2}\right) \left(1 - \frac{1}{3}\right) \\
 &= 12 \cdot \frac{1}{2} \cdot \frac{2}{3} = 4.
 \end{aligned}$$

A remarkable fact is that the sum over divisors of the Euler totient function is remarkably simple.

Theorem 5.10 (Gauss). *Let n be a positive integer.*

$$\sum_{d|n} \varphi(d) = n.$$

Proof. First consider the case that n is a prime power. We have

$$\begin{aligned}
 \sum_{d|p^e} \varphi(d) &= \varphi(p^e) + \varphi(p^{e-1}) + \cdots + \varphi(p) + \varphi(1) \\
 &= p^{e-1}(p-1) + p^{e-2}(p-1) + \cdots + (p-1) + 1 \\
 &= p^e - p^{e-1} + p^{e-1} - p^{e-2} + \cdots + p^2 - p + p - 1 + 1 \\
 &= p^e,
 \end{aligned}$$

where the last equality comes from the fact that the sum is telescoping.

For the general case write $n = p_1^{e_1} \cdots p_r^{e_r}$, for distinct primes p_1, \dots, p_r , and positive integers e_1, \dots, e_r . Let $n' = p_2^{e_2} \cdots p_r^{e_r} = \frac{n}{p_1^{e_1}}$. Then we can compute

$$\begin{aligned} \sum_{d|n} \varphi(d) &= \sum_{d|n'} \sum_{d_1|p_1^{e_1}} \varphi(dd_1) \\ &= \sum_{d|n'} \sum_{d_1|p_1^{e_1}} \varphi(d)\varphi(d_1) \\ &= \sum_{d|n'} \varphi(d) \sum_{d_1|p_1^{e_1}} \varphi(d_1) \\ &= \sum_{d|n'} \varphi(d) p_1^{e_1} \\ &= p_1^{e_1} \sum_{d|n'} \varphi(d). \end{aligned}$$

Repeating this process for each of the $p_i^{e_i}$, we conclude that

$$\sum_{d|n} \varphi(d) = p_1^{e_1} \cdot p_r^{e_r} = n. \quad \square$$

We end our discussion of the Euler totient function with a question due to Lehmer.¹

Question 5.11. Does there exist a positive composite integer n such that $\varphi(n)$ divides $(n - 1)$?

Clearly, this is true if n is prime, so we have restricted the question to composite n . Lehmer was able to show that any such n must be odd, squarefree, and divisible by at least seven distinct primes. This was improved in 1980 by Cohen and Hagis to $n > 10^{20}$ with at least 14 distinct prime factors. However, the question remains an unsolved problem.

2. Möbius Function

Gauss first studied the Möbius function in 1801 when he showed that sum of the primitive roots modulo p takes the value $\mu(p - 1)$. However, the function is named after August Möbius² who systematically studied its properties in 1832. Our main application of the Möbius function is the inclusion-exclusion formula known as the Möbius inversion formula (Theorem 5.20).

Definition 5.12. Let n be a positive integer. If there is a prime p such that p^2 divides n , then we say that n is *not squarefree*. If there is no such p , we say that n is *squarefree*.

Equivalently, we could look at the prime factorization of n as $n = p_1^{e_1} \cdots p_r^{e_r}$ for distinct primes p_i and say that n is *squarefree* if and only if all the exponents e_i are 1.

¹Derrick Henry Lehmer (1905–1991) was an American mathematician.

²August Ferdinand Möbius (1790–1868) was a German mathematician.

Example 5.13.

- 2 is squarefree, whereas $4 = 2^2$ is not.
- $14 = 2 \cdot 7$ is squarefree, but $28 = 2^2 \cdot 7$ and $56 = 2^3 \cdot 7$ are both divisible by $4 = 2^2$, so they are not.

Definition 5.14. We define the *Möbius function* on a positive integer n as

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{if } n \text{ is not squarefree,} \\ (-1)^r & \text{where } r \text{ is the number of distinct prime factors of } n, \text{ if } n \text{ is squarefree.} \end{cases}$$

Example 5.15.

- $\mu(2) = -1$.
- $\mu(4) = 0$ and $\mu(28) = 0$.
- $\mu(14) = (-1)^2 = 1$.
- $\mu(30) = (-1)^3 = -1$.

Investigation 5.16. The Möbius function has many curious properties. See what patterns you can find.

- Compute $\mu(n)$ for some consecutive values of n . Can you anything about the number of 0's or 1's or -1 's?
- Does the Möbius function have an average value? In other words, does the following limit exist?

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \mu(k)$$

- What about nonconsecutive values?

The first property we prove is multiplicativity.

Proposition 5.17. *The Möbius function is multiplicative.*

Proof. Let a and b be positive integers with $\gcd(a, b) = 1$. If at least one of a or b is 1, then

$$\mu(ab) = \mu(a \cdot 1) = \mu(a) = \mu(a)\mu(1) = \mu(a)\mu(b)$$

or

$$\mu(ab) = \mu(1 \cdot b) = \mu(b) = \mu(1)\mu(b) = \mu(a)\mu(b).$$

Now assume that both a and b are not 1. Then we may factor a and b into products of distinct prime powers as $a = p_1^{e_1} \cdots p_r^{e_r}$ and $b = q_1^{d_1} \cdots q_s^{d_s}$. Since $\gcd(a, b) = 1$, then a and b share no common factors. In particular, the set of primes $\{p_1, \dots, p_r\}$ is disjoint from the set of primes $\{q_1, \dots, q_s\}$. If either a or b is not squarefree, then neither is their product, so $\mu(ab) = \mu(a)\mu(b) = 0$.

Assume that both a and b are squarefree, that is, $e_i = d_j = 1$ for all i, j . Then, using the fact that the primes are all distinct,

$$\mu(ab) = \mu(p_1 \cdots p_r q_1 \cdots q_s) = (-1)^{r+s} = (-1)^r (-1)^s = \mu(a)\mu(b). \quad \square$$

It is not too surprising that a function based on the number of distinct factors is multiplicative. However, the next property of the Möbius function is less intuitive. It says that the sum of the Möbius function applied to all the divisors of a fixed integer n is 0, except for the trivial case $n = 1$.

Example 5.18. Consider $n = 735 = 3 \cdot 5 \cdot 7^2$. We compute

$$\begin{aligned} \sum_{d|735} \mu(d) &= \mu(1) + \mu(3) + \mu(5) + \mu(7) + \mu(15) + \mu(21) + \mu(35) + \mu(49) \\ &\quad + \mu(105) + \mu(147) + \mu(245) + \mu(735) \\ &= \mu(1) + [\mu(3) + \mu(5) + \mu(7)] + [\mu(3 \cdot 5) + \mu(3 \cdot 7) + \mu(5 \cdot 7) + \mu(7^2)] \\ &\quad + [\mu(3 \cdot 5 \cdot 7) + \mu(3 \cdot 7^2) + \mu(5 \cdot 7^2)] + \mu(3 \cdot 5 \cdot 7^2) \\ &= 1 + [(-1) + (-1) + (-1)] + [1 + 1 + 1 + 0] + [(-1) + 0 + 0] + 0 \\ &= 0. \end{aligned}$$

Notice that the terms in the sum are grouped by their number of prime divisors.

Proposition 5.19. *Let n be a positive integer. Then*

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{if } n > 1. \end{cases}$$

Notice that we are not restricting n to be squarefree, so some of the terms in the sum are going to be 0 and some are going to be ± 1 . We need to show in the proof that the number of $+1$'s and the number of -1 's exactly balance one another. We proceed in the same manner as in the example, by grouping the divisors by their number of prime divisors.

Proof. For $n = 1$,

$$\sum_{d|1} \mu(d) = \mu(1) = 1.$$

Let $n = p_1^{e_1} \cdots p_r^{e_r} > 1$ be a product of distinct prime powers. Then the divisors d of n for which $\mu(d) \neq 0$ are precisely the divisors of n that are products of distinct primes. We group these d by the number of distinct prime divisors. So we have

$$\begin{aligned} \sum_{d|n} \mu(d) &= 1 + [\mu(p_1) + \cdots + \mu(p_r)] + [\mu(p_1 p_2) + \mu(p_1 p_3) + \cdots + \mu(p_{r-1} p_r)] \\ &\quad + \cdots + [\mu(p_1 \cdots p_r)]. \end{aligned}$$

Now we need to count the number of pairs of distinct primes, the number of triples of distinct primes, etc. This is a purely combinatorial problem and involves the

binomial coefficient $\binom{a}{b}$, which is the number of different ways to choose b objects from a set of a objects.

$$\begin{aligned}\sum_{d|n} \mu(d) &= 1 + \binom{r}{1}(-1) + \binom{r}{2}(1) + \cdots + \binom{r}{r}(-1)^r \\ &= (1 - 1)^r \quad (\text{binomial theorem}) \\ &= 0.\end{aligned}$$

□

An even more startling property of the Möbius function is the inversion formula, which plays a key role in defining cyclotomic polynomials below.

Theorem 5.20 (Möbius inversion formula). *Given two (arithmetic) functions f and g such that*

$$g(n) = \sum_{d|n} f(d)$$

for every positive integer n , then

$$(16) \quad f(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) g(d).$$

Proof. First note that

$$(17) \quad \sum_{d|n} d = \sum_{d|n} \frac{n}{d}$$

since we are just rearranging the order of the sum and not the actual values we are summing.

We substitute the definition of g into equation (16) as

$$f(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) g(d) = \sum_{d|n} \mu\left(\frac{n}{d}\right) \sum_{a|d} f(a) = \sum_{d|n} \mu\left(\frac{n}{d}\right) \sum_{a|\frac{n}{d}} f(a),$$

where the last equality is from equation (17). Now we make the key rearrangement where we group the terms of the inner sum not in terms of the Möbius values, but in terms of the function values,

$$\sum_{d|n} \mu\left(\frac{n}{d}\right) \sum_{a|\frac{n}{d}} f(a) = \sum_{d|n} \sum_{k|d} \mu(k) f\left(\frac{n}{d}\right).$$

Since $f\left(\frac{n}{d}\right)$ does not depend on k , we can move it outside of the inner sum to have

$$\sum_{d|n} \sum_{k|d} \mu(k) f\left(\frac{n}{d}\right) = \sum_{d|n} f\left(\frac{n}{d}\right) \sum_{k|d} \mu(k).$$

Finally, we use Proposition 5.19 to see that $\sum_{k|d} \mu(k) = 0$, unless $d = 1$, to have

$$\sum_{d|n} f\left(\frac{n}{d}\right) \sum_{k|d} \mu(k) = f(n).$$

□

Remark. The key rearrangement step in the proof is only a rearrangement of the terms, but it is best illustrated by an example. Let $n = 6$. Then we have

$$\begin{aligned} \sum_{d|n} \mu\left(\frac{n}{d}\right) \sum_{a|\frac{n}{d}} f(a) &= \mu(6)f(1) + \mu(3)[f(1) + f(2)] \\ &\quad + \mu(2)[f(1) + f(3)] + \mu(1)[f(1) + f(2) + f(3) + f(6)]. \end{aligned}$$

Rearranging, we group by the values of f instead of the values of μ to have

$$\begin{aligned} \sum_{d|n} \sum_{k|d} \mu(k) f\left(\frac{n}{d}\right) &= f(1)[\mu(1) + \mu(2) + \mu(3) + \mu(6)] \\ &\quad + f(2)[\mu(1) + \mu(3)] + f(3)[\mu(1) + \mu(2)] + f(6)\mu(1). \end{aligned}$$

Corollary 5.21. Let n be a positive integer.

$$\varphi(n) = \sum_{d|n} \mu(d) \frac{n}{d}.$$

Proof. From Theorem 5.10 we know that

$$\sum_{d|n} \varphi(d) = n.$$

So we apply the Möbius inversion formula to $f(n) = \varphi(n)$ and the identity function $g(n) = n$. Then we have

$$\varphi(n) = \sum_{d|n} \mu(d) g\left(\frac{n}{d}\right) = \sum_{d|n} \mu(d) \frac{n}{d}. \quad \square$$

2.1. Cyclotomic Polynomials. Now that we have several properties of the Möbius function, let's apply this knowledge to investigating roots of unity and cyclotomic polynomials. Recall that the solutions to $x^n = 1$ are called the n th roots of unity (Definition 3.46).

Definition 5.22. Let n be a positive integer. The complex numbers x such that $x^n = 1$ and $x^m \neq 1$ for any $m < n$ are called *primitive n th roots of unity*.

Example 5.23. We have that 1 is a root of unity for all positive integers n since $1^n = 1$. We have -1 is a 2nd root of unity since $(-1)^2 = 1$. Moreover, -1 is a *primitive* 2nd root of unity since

$$(-1)^2 = 1 \quad \text{and} \quad (-1)^1 \neq 1.$$

Each n th root of unity is also a root of unity for any multiple of n , but it is primitive for only one value of n . For example, -1 is a primitive 2nd root of unity and is also a 4th root of unity since

$$(-1)^4 = 1.$$

In fact, -1 is an n th root of unity for any even n .

We want to answer the following question. Recall that a *root* of a polynomial $f(x)$ is a number a such that $f(a) = 0$.

Question 5.24. For every positive integer n , is there a polynomial whose roots are exactly the primitive n th roots of unity?

Your first guess might be the polynomial

$$x^n - 1.$$

While, yes, the primitive n th roots of unity are roots of that polynomial, so are the primitive m th roots of any m dividing n . For example, for $n = 4$ the polynomial

$$x^4 - 1$$

has (-1) as a root, which is a primitive 2nd root of unity, whereas we want only the primitive 4th roots of unity.

Investigation 5.25. Factor the first few polynomials $x^n - 1$. Do you see any pattern to their factors?

Let us give a name to the polynomial that answers Question 5.24.

Definition 5.26. Let n be a positive integer. We define the *n th cyclotomic polynomial* as the polynomial whose roots are exactly the primitive n th roots of unity.

We have already seen the polynomial whose roots are the n th roots of unity,

$$x^n - 1.$$

What we want to do is somehow exclude the nonprimitive n th roots. The method will be through division. If you recall, the fundamental theorem of algebra says that if a is a root of a polynomial $f(x)$, then we can factor $f(x)$ as $f(x) = (x - a)f_1(x)$ for some polynomial $f_1(x)$. So if a is a root of both $f(x)$ and $g(x)$, we have cancellation through division,

$$\frac{f(x)}{g(x)} = \frac{(x - a)f_1(x)}{(x - a)g_1(x)} = \frac{f_1(x)}{g_1(x)}.$$

For this cancellation to result in a polynomial, we need to have at least as many factors in the numerator as there are in the denominator. In particular, we want to end up with only the “correct” roots left in the numerator. This is where the Möbius function will be used. First, let’s look at a few specific examples to see how this cancellation works and what difficulties might arise.

Example 5.27. Suppose we want to find the primitive 3rd roots of unity. Since 1st roots of unity are also 3rd roots of unity, we need to take all the 3rd roots of unity and remove those that are also 1st roots of unity. We can simply take the polynomial defining all the 3rd roots of unity and divide by the polynomial defining all the 1st roots of unity, in effect, cancelling the 1st roots of unity from the numerator:

$$\frac{x^3 - 1}{x - 1} = 1 + x + x^2.$$

Note that we do end up with a polynomial.

If n is not a prime number, as in the next example, there can be more factors involved.

Example 5.28. For the primitive 4th roots of unity, we need to be a little more careful since both 1st and 2nd roots of unity are also 4th roots of unity. We need to take all the 4th roots of unity and remove the 2nd and 1st roots of unity. If we try

$$\frac{x^4 - 1}{(x^2 - 1)(x - 1)} = \frac{x^2 + 1}{x - 1},$$

we do not get a polynomial. The reason is that the 1st roots of unity are also 2nd roots of unity, so we have in essence tried to remove them twice, which failed. If we instead try

$$\frac{x^4 - 1}{x^2 - 1} = x^2 + 1,$$

we get exactly the primitive 2nd roots of unity.

As you can probably guess, the more divisors that n has, the more complicated this procedure becomes. What we are looking for is a general procedure that works for all n . Let's look at one more example before stating the general theorem.

Example 5.29. For the primitive 6th roots of unity, we need to remove the 3rd, 2nd, and 1st roots of unity. We know the 1st roots of unity are both 2nd and 3rd roots of unity, so if we try

$$\frac{x^6 - 1}{(x^3 - 1)(x^2 - 1)},$$

we remove the erroneous 2nd and 3rd roots of unity. But we tried to remove the 1st roots of unity twice, once in $x^3 - 1$ and once in $x^2 - 1$! To get the correct polynomial, we are forced to add back in the 1st roots of unity as

$$\frac{(x^6 - 1)(x - 1)}{(x^3 - 1)(x^2 - 1)} = x^2 - x + 1.$$

The Möbius function allows us to state a general procedure.

Theorem 5.30. *Let n be a positive integer. The n th cyclotomic polynomial is given by*

$$\Phi_n(x) = \prod_{d|n} (x^d - 1)^{\mu(n/d)}.$$

So when $\mu(n/d)$ is negative, we are “subtracting” those roots; when $\mu(n/d)$ is positive, we are “adding” those roots back in; and when $\mu(n/d) = 0$, we are doing nothing.

Example 5.31. For $n = 6$, we have

$$\mu(1) = \mu(6) = 1 \quad \text{and} \quad \mu(2) = \mu(3) = -1$$

to have

$$\Phi_6(x) = \frac{(x^6 - 1)(x - 1)}{(x^3 - 1)(x^2 - 1)}.$$

This is exactly what we came up with in Example 5.29.

As the examples indicate, it is possible to prove this theorem by carefully multiplying and dividing by factors in a way that preserves only primitive roots. But, as the examples also indicate, such a proof is tricky to construct. It is an illustrating endeavor to attempt such a proof, and you should make such an effort. Fortunately, there is a clever proof that utilizes the power of the Möbius inversion formula (Theorem 5.20).

Proof of Theorem 5.30. Since $x^n - 1$ is all the n th roots of unity, we can split it up into a product of the primitive roots of unity for $d \mid n$

$$x^n - 1 = \prod_{d \mid n} \Phi_d(x).$$

To use the Möbius inversion formula, we need to change the product to a sum using natural logs. We take the natural log of both sides

$$\ln(x^n - 1) = \ln \left(\prod_{d \mid n} \Phi_d(x) \right) = \sum_{d \mid n} \ln(\Phi_d(x)).$$

Now we apply the Möbius inversion formula (Theorem 5.20) to get

$$\ln(\Phi_n(x)) = \sum_{d \mid n} \mu \left(\frac{n}{d} \right) \ln(x^d - 1).$$

To return back to a product, we exponentiate both sides to get

$$\Phi_n(x) = \prod_{d \mid n} (x^d - 1)^{\mu(n/d)}.$$

□

Example 5.32. We compute the first few cyclotomic polynomials.

$$\begin{aligned}
\Phi_1 &= x - 1, \\
\Phi_2 &= \frac{x^2 - 1}{x - 1} = x + 1, \\
\Phi_3 &= \frac{x^3 - 1}{x - 1} = x^2 + x + 1, \\
\Phi_4 &= \frac{x^4 - 1}{x^2 - 1} = x^2 + 1, \\
\Phi_5 &= \frac{x^5 - 1}{x - 1} = x^4 + x^3 + x^2 + x + 1, \\
\Phi_6 &= \frac{(x^6 - 1)(x - 1)}{(x^3 - 1)(x^2 - 1)} = x^2 - x + 1, \\
\Phi_7 &= \frac{x^7 - 1}{x - 1} = x^6 + x^5 + x^4 + x^3 + x^2 + x + 1, \\
\Phi_8 &= \frac{(x^8 - 1)(x - 1)}{x^4 - 1} = x^4 + 1, \\
\Phi_9 &= \frac{x^9 - 1}{x^3 - 1} = x^6 + x^3 + 1, \\
\Phi_{10} &= \frac{(x^{10} - 1)(x - 1)}{(x^5 - 1)(x^2 - 1)} = x^4 - x^3 + x^2 - x + 1.
\end{aligned}$$

Notice that these are all polynomials with integer coefficients and are irreducible over \mathbb{Q} .

There are many questions one may ask about cyclotomic polynomials; we end our discussion with just two.

Question 5.33. Is $\Phi_n(x)$ irreducible over \mathbb{Q} for all n ?

Question 5.34. Are the coefficients of $\Phi_n(x)$ always ± 1 or 0?

3. Functions on Divisors

We have studied a few properties of divisors, especially the idea of unique factorization into primes. We now look at arithmetic functions that give us more information about the divisors of a given integer.

Definition 5.35. We define a *sum of divisors function*. For an integer $k \geq 0$ and positive integer n , we define $\sigma_k(n)$ by

$$\sigma_k(n) = \sum_{d|n} d^k.$$

Note that $\sigma_0(n)$ gives the number of divisors of n . Recall the Möbius function relied on the number of distinct *prime* factors of an integer. Consequently, we define a function that counts the number of distinct prime divisors of n .

Definition 5.36. For a positive integer n , define $\omega(n)$ as the number of distinct prime divisors of n .

Example 5.37. We compute

$$\sigma_1(6) = 1^1 + 2^1 + 3^1 + 6^1 = 12,$$

$$\sigma_0(6) = 1^0 + 2^0 + 3^0 + 6^0 = 1 + 1 + 1 + 1 = 4,$$

$$\omega(6) = 2 \text{ (the primes 2 and 3),}$$

and

$$\sigma_2(9) = 1^2 + 3^2 + 9^2 = 91,$$

$$\sigma_0(9) = 3,$$

$$\omega(9) = 1 \text{ (the prime 3).}$$

Investigation 5.38. The following questions could apply to any of the functions $\sigma_k(n)$ or $\omega(n)$. We use $\sigma_1(n)$ as an example.

- (a) Are there choices of n for which $\sigma_1(n)$ is particularly large or small?
- (b) For which integers m are there integers n such that $\sigma_1(n) = m$? Are there multiple solutions for a given m ?
- (c) What is its average value of $\sigma_1(n)$ over a range of n values?

As with $\varphi(n)$ we would like a simple formula for computing $\sigma_k(n)$.

Question 5.39. Is there a formula for $\sigma_k(n)$?

For n prime, $\sigma_k(n)$ is easy to compute since the only divisors of a prime are 1 and itself; we have $\sigma_k(p) = 1 + p^k$. Similarly for prime powers, the divisors are easy to enumerate; they are all powers of the prime.

Lemma 5.40. Let p be a prime, and let m be a positive integer.

$$\sigma_k(p^m) = \sum_{i=0}^m p^{ki} = \begin{cases} \frac{p^{k(m+1)} - 1}{p^k - 1} & k \geq 1, \\ m + 1 & k = 0. \end{cases}$$

Proof. The only divisors of p^m are the powers of p , $\{1, p, p^2, p^3, \dots, p^{m-1}, p^m\}$. Thus,

$$\sigma_k(p^m) = 1 + p^k + p^{2k} + \dots + p^{km} = \frac{p^{k(m+1)} - 1}{p^k - 1}.$$

A simple way to see that the fraction is correct is to consider p^k as a variable x . Then we divide

$$\frac{x^{m+1} - 1}{x - 1} = 1 + x + x^2 + \dots + x^m,$$

giving the stated formula. □

Now we examine products of primes. We first show that σ_k is multiplicative.

Lemma 5.41. *Let m and n be positive integers. If $\gcd(m, n) = 1$, then $\sigma_k(mn) = \sigma_k(m)\sigma_k(n)$ for all integers $k \geq 0$.*

Proof. Let $\{d_1, \dots, d_r\}$ be the divisors of m , and let $\{e_1, \dots, e_s\}$ be the divisors of n . Since $\gcd(m, n) = 1$, the intersection of the two sets only contains 1:

$$\{d_1, \dots, d_r\} \cap \{e_1, \dots, e_s\} = \{1\}.$$

Then, the divisors of mn are products ab , where $a \mid m$ and $b \mid n$. So the set of divisors of mn is the set

$$\{d_i e_j : 1 \leq i \leq r, 1 \leq j \leq s\}.$$

Computing σ_k , we have

$$\begin{aligned} \sigma_k(mn) &= \sum_{i,j} (d_i e_j)^k \\ &= \sum_{i,j} d_i^k e_j^k \\ &= (d_1^k + \dots + d_r^k)(e_1^k + \dots + e_s^k) \\ &= \sigma_k(m)\sigma_k(n). \end{aligned} \quad \square$$

Example 5.42. We demonstrate the multiplicativity of $\sigma_k(n)$. Let $n = 6$, $m = 25$, and $k = 2$. We compute

$$\begin{aligned} \sigma_2(6) &= 1 + 2^2 + 3^2 + 6^2 = 50, \\ \sigma_2(25) &= 1 + 5^2 + 25^2 = 651, \\ \sigma_2(6 \cdot 25) &= \sigma_2(150) = 1 + 2^2 + 3^2 + 5^2 + 6^2 + 10^2 + 15^2 + 25^2 \\ &\quad + 30^2 + 50^2 + 75^2 + 150^2 \\ &= 32550 = 50 \cdot 651. \end{aligned}$$

Multiplicativity allows us to compute $\sigma_k(n)$ for any n we can factor.

Theorem 5.43. *Let n be a positive integer that factors into a product of distinct prime powers as $n = p_1^{e_1} \cdots p_r^{e_r}$. Then*

$$\sigma_k(p_1^{e_1} \cdots p_r^{e_r}) = \begin{cases} \frac{p_1^{k(e_1+1)} - 1}{p_1^k - 1} \cdots \frac{p_r^{k(e_r+1)} - 1}{p_r^k - 1} & k \geq 1, \\ (1 + e_1) \cdots (1 + e_r) & k = 0. \end{cases}$$

Proof. We combine Lemmas 5.40 and 5.41. □

Example 5.44.

- For $n = 24$ and $k = 4$,

$$\begin{aligned} \sigma_4(24) &= \sigma_4(2^3 \cdot 3) = \sigma_4(2^3)\sigma_4(3) \\ &= \frac{2^{4(3+1)} - 1}{2^4 - 1} \cdot \frac{3^{4(1+1)} - 1}{3^4 - 1} = \frac{65535}{15} \cdot \frac{6560}{80} = 358258. \end{aligned}$$

- For $n = 4725$ and $k = 0$,

$$\sigma_0(4725) = \sigma_0(3^2 \cdot 5^2 \cdot 7) = (1+2)(1+2)(1+1) = 18.$$

Finally, we give an alternative way to compute the values of σ_k . In particular, we give an infinite series (for each k) whose coefficients are the values $\sigma_k(n)$. Surprisingly, this occurs fairly often in number theory, where the values of arithmetic functions occur as the coefficients of power series. This method will be particularly effective when we consider partitions in Section 5.4.

Theorem 5.45. *For every integer $k \geq 0$,*

$$\sum_{n=1}^{\infty} n^k \frac{x^n}{1-x^n} = \sum_{n=1}^{\infty} \sigma_k(n) x^n.$$

Let's look at an example before proving the theorem.

Example 5.46. We compute $\sigma_3(12)$. First we use Theorem 5.43:

$$\sigma_3(12) = \sigma_3(2^2 \cdot 3) = \frac{2^{3(3)} - 1}{2^3 - 1} \cdot \frac{3^{3(2)} - 1}{3^3 - 1} = (73)(28) = 2044.$$

Now we compute the first few terms of the series. Recall that the power series expansion of $\frac{1}{1-x^m} = 1 + x^m + x^{2m} + x^{3m} + \dots$.

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{n^3 x^n}{1-x^n} &= \frac{x}{1-x} + \frac{8x^2}{1-x^2} + \frac{27x^3}{1-x^2} + \dots + \frac{12^3 x^{12}}{1-x^{12}} + \dots \\ &= x(1 + x + x^2 + x^3 + \dots) + 8x^2(1 + x^2 + x^4 + x^6 + \dots) \\ &\quad + 27x^3(1 + x^3 + x^6 + x^9 + \dots) + \dots \\ &\quad + 12^3 x^{12}(1 + x^{12} + x^{24} + x^{36} + \dots) + \dots \\ &= x + 9x^2 + 28x^3 + 73x^4 + 126x^5 + 252x^6 + 344x^7 \\ &\quad + 585x^8 + 757x^9 + 1134x^{10} + 1332x^{11} + 2044x^{12} + O(x^{13}). \end{aligned}$$

The value of $\sigma_3(12)$ is the coefficient of x^{12} ,

$$\sigma_3(12) = 2044.$$

Notice that since the power in the numerator continues to grow, after summing the first m terms, we know the values of the coefficients for $\{x, x^2, \dots, x^m\}$. Hence in our example, we know the first 12 coefficients, so we have computed $\sigma_3(n)$ for $1 \leq n \leq 12$ at the same time!

Proof of Theorem 5.45. Using geometric series, we expand each term of the sum. We write them in such a way that the same powers of x are in separate columns; that way, when we sum to get the coefficients, we are summing each column.

$$\begin{array}{rcl}
\frac{x}{1-x} & = & x + x^2 + x^3 + x^4 + x^5 + x^6 + x^7 + x^8 + x^9 + x^{10} + \dots \\
\frac{2^k x^2}{1-x^2} & = & 2^k x^2 + 2^k x^4 + 2^k x^6 + 2^k x^8 + 2^k x^{10} + \dots \\
\frac{3^k x^3}{1-x^3} & = & 3^k x^3 + 3^k x^6 + 3^k x^9 + \dots \\
\frac{4^k x^4}{1-x^4} & = & 4^k x^4 + 4^k x^8 + \dots \\
\frac{5^k x^5}{1-x^5} & = & 5^k x^5 + 5^k x^{10} + \dots \\
\frac{6^k x^6}{1-x^6} & = & 6^k x^6 + \dots \\
& \vdots &
\end{array}$$

Summing the first few columns, we see

$$\begin{aligned}
\sum_{n=1}^{\infty} n^k \frac{x^n}{1-x^n} &= 1 \cdot x + (1 + 2^k)x^2 + (1 + 3^k)x^3 + (1 + 2^k + 4^k)x^4 \\
&\quad + (1 + 5^k)x^5 + (1 + 2^k + 3^k + 6^k)x^6 + \dots
\end{aligned}$$

In each column we get coefficients that are sums of the k th powers of the divisors of the exponent. The reason is the terms in the power series expansion of $\frac{1}{1-x^m}$ have exponents that are multiples of m starting with 1, i.e.,

$$\frac{1}{1-x^m} = 1 + x^m + x^{2m} + x^{3m} + \dots$$

Multiplying this by $m^k x^m$ gives the multiples starting with m instead of 1, and weights it with the coefficient m^k . Thus, the x^n column has only coefficients m^k where n is a multiple of m , i.e., m is a divisor of n . Thus, summing the columns gives the sum of the powers of the divisors. \square

So we have, in fact, answered the original question (to find a formula for $\sigma_k(n)$) in two different ways. The first way requires factoring the number. The second requires computing with infinite series. Both operations are time consuming.

3.1. Perfect Numbers. Let's examine the values of $\sigma_1(n)$ for the first few $n > 1$. Since we know that $\sigma_1(n)$ gets larger as n gets larger, let's also compute the ratio $\frac{\sigma_1(n)}{n}$.

n	$\sigma_1(n)$	$\frac{\sigma_1(n)}{n}$
2	3	1.5
3	4	1.33...
4	7	1.75
5	6	1.2
6	12	2
7	8	1.14...
8	15	1.875
9	13	1.44...
10	18	1.8

From this data there are many interesting questions that we could ask.

Question 5.47.

- (a) What is the largest and smallest possible value of $\frac{\sigma_1(n)}{n}$?
- (b) For which $a \in \mathbb{Q}$ does there exist an n such that $\sigma_1(n) = a$?
- (c) Are there infinitely many n such that $\sigma_1(n) = a$ for each possible a ?

You could also ask all of these questions for $\sigma_k(n)$ for other values of k . We will focus on just one situation, $\frac{\sigma_1(n)}{n} = 2$ or, equivalently, $\sigma_1(n) = 2n$. The other parts of Question 5.47 are considered in Exploration Exercise 5.50.

Definition 5.48. Let n be a positive integer. We say that d is a *proper* divisor of n if $d < n$ and d divides n . Define the *aliquot sum* $s(n)$ to be the sum of the proper divisors of n , that is, $s(n) = \sigma_1(n) - n$.

Example 5.49. The proper divisors of 12 are $\{1, 2, 3, 4, 6\}$ and

$$s(12) = 16.$$

Definition 5.50. We say a number is *perfect* if $s(n) = n$, in other words, if $\frac{\sigma_1(n)}{n} = 2$.

Example 5.51.

- 6 is a perfect number

$$1 + 2 + 3 = 6.$$

- 28 is a perfect number

$$1 + 2 + 4 + 7 + 14 = 28.$$

Question 5.52. How many perfect numbers are there?

Euclid proved that $2^{p-1}(2^p - 1)$ is an even perfect number whenever $2^p - 1$ is prime. When $2^p - 1$ is prime it is called a *Mersenne prime* (see Exploration Exercise 1.36). He also conjectured that all the even perfect numbers are of this form. This was proven by Euler.

Theorem 5.53 (Euler). *A number n is an even perfect number if and only if*

$$n = 2^{p-1}(2^p - 1)$$

for a Mersenne prime $2^p - 1$.

Proof. Assume that $2^p - 1$ is prime, and let $n = 2^{p-1}(2^p - 1)$. We need to compute $\sigma_1(n)$. First note that $\gcd(2^{p-1}, 2^p - 1) = 1$ since $2^p - 1$ is prime (by assumption) and odd and 2^{p-1} is a power of 2. For relatively prime numbers, σ_1 is multiplicative (Lemma 5.41), so we have

$$\begin{aligned}\sigma_1(n) &= \sigma_1(2^{p-1})\sigma_1(2^p - 1) \\ &= \frac{2^p - 1}{2 - 1}((2^p - 1) + 1) \\ &= (2^p - 1)2^p = 2n.\end{aligned}$$

In particular, n is perfect.

Now assume that n is an even perfect number, so we may write $n = 2^{k-1}m$ for some odd integer m and $k \geq 2$. We again use the fact that σ_1 is multiplicative to compute

$$\sigma_1(n) = \sigma_1(2^{k-1})\sigma_1(m) = (2^k - 1)\sigma_1(m).$$

We know that n is perfect, so we must have

$$(18) \quad \begin{aligned}\sigma_1(n) &= 2n, \quad \text{which is} \\ (2^k - 1)\sigma_1(m) &= 2^k m.\end{aligned}$$

In particular, since $\sigma_1(m)$ is an integer, we must have $(2^k - 1)$ divides $2^k m$. Since $(2^k - 1)$ is odd and 2^k is even, we must have $(2^k - 1)$ divides m . Let M be such that $(2^k - 1)M = m$. Then from (18) we have that

$$(2^k - 1)\sigma_1(m) = 2^k(2^k - 1)M$$

so that

$$\sigma_1(m) = 2^k M.$$

Furthermore,

$$\sigma_1(m) \geq m + M$$

since both m and M divide m . So we have

$$2^k M = \sigma_1(m) \geq m + M = (2^k - 1)M + M = 2^k M.$$

Since the left-most and right-most terms are the same, the inequality must actually be an equality so that

$$\sigma_1(m) = m + M.$$

Thus, m has only two divisors $\{m, M\}$, so we must have $m = 2^k - 1$ is prime and $M = 1$. □

If we knew there were infinitely many Mersenne primes, then we could conclude that there are infinitely many (even) perfect numbers; but, unfortunately, we do not know if there are infinitely many Mersenne primes. Additionally, this construction produces only even perfect numbers.

Question 5.54. Are there any odd perfect numbers?

The existence of odd perfect numbers remains an open problem. However, it is currently known that an odd perfect number must have at least 75 prime factors and be larger than 10^{300} .

We conclude this section with a problem related to aliquot sums similar in nature to questions appearing in Chapter 10.

Investigation 5.55. The aliquot sum function is quite interesting in the fact that sometimes $s(n) \geq n$ and sometimes $s(n) < n$. Consequently, you can find interesting behavior when you apply the function several times in a row. For example,

$$s(220) = 284 \quad \text{and} \quad s(s(220)) = s(284) = 220.$$

- (a) What happens to most starting values n as you repeatedly apply $s(n)$?
- (b) What is the average number of distinct values in the sequence $(n, s(n), s(s(n)), \dots)$?
- (c) Can you find any other values that repeat, similar to the pair $(220, 284)$?

3.2. Sieving and the Number of Prime Divisors. We have neglected so far the number of prime divisors function. Its behavior is difficult to understand and is deeply related to the distribution of prime numbers. We content ourselves with a question and a brief digression into its relation to sieve theory.

Question 5.56. For a given n , can you determine if $\omega(n)$ is even or odd?

The basic idea of sieve theory is to count the number of integers with a given property in a given range; for example, fix a positive integer N and count the number of primes p with $N \leq p \leq 2N$. This is a rather hard problem due to the difficulty of determining any type of regular structure or pattern to the prime numbers.

In Chapter 1 we encountered the Sieve of Eratosthenes, whose result was the set of prime numbers in a given range. Legendre adapted this into a way to count the number of primes up to a given bound. For example, let's count the number of primes up to 100. We do this by removing all numbers that are *not* prime, the sieve step. Since $\sqrt{100} = 10$, we just need to sieve by primes 2, 3, 5, and 7.

We first remove all multiples of 2, 3, 5, and 7, as we did with the Sieve of Eratosthenes. For each given number, the largest integer (floor) function $\lfloor x \rfloor$ can be used to get the number of multiples up to that number. For example, the number

of multiples of 2 up to 100 is $\lfloor \frac{100}{2} \rfloor = 50$. Doing this for 2, 3, 5, and 7 separately is easy to compute, and it results in 117 numbers:

$$\left\lfloor \frac{100}{2} \right\rfloor + \left\lfloor \frac{100}{3} \right\rfloor + \left\lfloor \frac{100}{5} \right\rfloor + \left\lfloor \frac{100}{7} \right\rfloor = 50 + 33 + 20 + 14 = 117.$$

But wait—there are only 100 positive integers at most 100! The issue is that we have double counted. For example, 6 is both a multiple of 2 and a multiple of 3. So we must subtract the number of multiples of products of two primes $\{2 \cdot 3, 2 \cdot 5, 2 \cdot 7, 3 \cdot 5, 3 \cdot 7, 5 \cdot 7\}$. There are 45 such multiples. However, we still have an issue: we have now double “uncounted” numbers that are multiples of two of those products such as 60, which is a multiple of both 6 and 10. So, we must add back in the multiples that are products of three of the primes $\{2 \cdot 3 \cdot 5, 2 \cdot 3 \cdot 7, 2 \cdot 5 \cdot 7, 3 \cdot 5 \cdot 7\}$ of which there are 6. Again, we have double counted, so we “uncount” the multiples of four primes, of which there are none less than 100. So while the concept is simple, the details can become complicated. We summarize the steps in the following table.

operation	multiples of	value
add	2, 3, 5, 7	$\lfloor \frac{100}{2} \rfloor + \lfloor \frac{100}{3} \rfloor + \lfloor \frac{100}{5} \rfloor + \lfloor \frac{100}{7} \rfloor = 117$
subtract	6, 10, 14, 15, 21, 35	$\lfloor \frac{100}{6} \rfloor + \lfloor \frac{100}{10} \rfloor + \lfloor \frac{100}{14} \rfloor + \lfloor \frac{100}{15} \rfloor + \lfloor \frac{100}{21} \rfloor + \lfloor \frac{100}{35} \rfloor = 45$
add	30, 42, 70, 105	$\lfloor \frac{100}{30} \rfloor + \lfloor \frac{100}{42} \rfloor + \lfloor \frac{100}{70} \rfloor + \lfloor \frac{100}{105} \rfloor = 6$
subtract	210	$\lfloor \frac{100}{210} \rfloor = 0$

So the number of primes up to 100 is the total number of integers minus the composite numbers, but we have to be careful because in the addition and subtraction above, we removed 2, 3, 5, and 7 from our count and left in 1. So we have

$$100 - [117 - 45 + 6 - 0] + 4 - 1 = 100 - 78 + 3 = 25.$$

This inclusion-exclusion might remind you of the construction of the cyclotomic polynomials using the Möbius function. In fact, Legendre showed in 1808 that

$$\pi(x) = \left(\sum_{d|P(\sqrt{x})} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor \right) + \pi(\sqrt{x}) - 1,$$

where $\pi(x)$ is the number of primes at most x and $P(\sqrt{x})$ is the product of the primes at most \sqrt{x} .

Investigation 5.57.

- Use Legendre’s formula to compute $\pi(1000)$.
- Verify that it is correct by counting the number of primes up to 1000.

While the Legendre sieve does give an exact formula, in practice it is only useful for small numbers since it takes a very long time to compute. You could get an approximate answer by computing only a few of the additions and subtractions, but the error is then very large.

Modern sieve methods try to attack this type of problem: determining (good) upper and lower bounds for the number of certain kinds of integers in a given range. However, the *parity problem* says that if all the integers you are trying to count have all even or all odd number of prime divisors, then sieving methods are unable to find good upper or lower bounds. Consequently, the parity of $\omega(n)$ is an important quantity in sieve theory.

4. Partitions

We now consider our last arithmetic function for this chapter, the partition function.

Definition 5.58. A set of positive integers $\{a_1, \dots, a_k\}$ is a *partition* of a positive integer n if $n = \sum_{i=1}^k a_i$. Each a_i is called a *part* of the partition.

Two partitions are considered the same if they differ only in the order of their summands. For example, $1 + 3 = 4$ and $3 + 1 = 4$ so that $\{1, 3\}$ and $\{3, 1\}$ are the same partition.

Definition 5.59. We define the arithmetic function $p(n)$ to be the number of partitions of the positive integer n .

Example 5.60. The partitions of 4 are

$$4 = 1 + 1 + 1 + 1,$$

$$4 = 2 + 1 + 1,$$

$$4 = 3 + 1,$$

$$4 = 2 + 2,$$

$$4 = 4,$$

and

$$p(4) = 5.$$

Investigation 5.61. Find all the partitions of a few small values of n .

- (a) What are the possible values of $p(n)$?
- (b) Can you have $p(n) = p(m)$ for $n \neq m$?
- (c) How is $p(n)$ growing with n ?

Question 5.62. Given a positive integer n , can you determine $p(n)$ efficiently?

Integer partitions were first studied by Euler, who gave us a way to compute $p(n)$ with an infinite series (similar to the series used for $\sigma_k(n)$ in section 5.3). For convenience we define $p(0) = 1$.

Theorem 5.63 (Euler). We have the following equality of infinite series

$$\sum_{n=0}^{\infty} p(n)x^n = (1 + x + x^2 + x^3 + \cdots)(1 + x^2 + x^4 + \cdots)(1 + x^3 + x^6 + \cdots) \cdots$$

or equivalently,

$$\sum_{n=0}^{\infty} p(n)x^n = \prod_{k=1}^{\infty} \frac{1}{1-x^k}.$$

Proof. We can rewrite the product as

$$(1+x^{1\cdot 1}+x^{2\cdot 1}+x^{3\cdot 1}+\cdots)(1+x^{1\cdot 2}+x^{2\cdot 2}+x^{3\cdot 2}+\cdots)(1+x^{1\cdot 3}+x^{2\cdot 3}+\cdots)\cdots.$$

The monomial chosen from the i th parenthesis has an exponent of the form $a_i \cdot i$. The value a_i is the number of times the part i appears in the partition. This gives a one-to-one correspondence between partitions of n and instances of x^n in the product. Thus, the coefficient of x^n is the number of partitions of n . Example 5.64 below illustrates this argument.

The second statement uses the geometric series expansion

$$\frac{1}{1-x^k} = 1 + x^k + x^{2k} + \cdots. \quad \square$$

Example 5.64. To find all the partitions of n , we need all the terms with exponents at least n . For $n = 4$, we compute

$$\begin{aligned} (1+x+x^2+x^3+x^4+\cdots)(1+x^2+x^4+\cdots)(1+x^3+x^6+\cdots)\cdots \\ = 1 + x + 2x^2 + 3x^3 + 5x^4 + \cdots \end{aligned}$$

to see that the number of partitions of 4 is 5. Each partition comes from some product of terms from the polynomials on the left-hand side, one from each term. We show how each partition corresponds to a product of polynomial terms:

$$\begin{aligned} 1 + 1 + 1 + 1 &\longleftrightarrow (x^4)(1)(1)(1)\cdots \\ &\quad \text{four 1's correspond to } (x)^4 = x^4 \\ &\quad \text{no 2,3,4's correspond to } (1)(1)(1), \\ 2 + 1 + 1 &\longleftrightarrow (x^2)(x^2)(1)(1)\cdots \\ &\quad \text{two 1's correspond to } (x)^2 = x^2 \\ &\quad \text{one 2 corresponds to } x^2 \\ &\quad \text{no 3,4's correspond to } (1)(1), \\ 2 + 2 &\longleftrightarrow (1)(x^4)(1)(1)\cdots \\ &\quad \text{two 2's correspond to } (x^2)^2 = x^4 \\ &\quad \text{no 1,3,4's correspond to } (1)(1)(1), \\ 3 + 1 &\longleftrightarrow (x)(1)(x^3)(1)\cdots \\ &\quad \text{one 1 corresponds to } x \\ &\quad \text{one 3 corresponds to } x^3 \\ &\quad \text{no 2,3's correspond to } (1)(1), \\ 4 &\longleftrightarrow (1)(1)(1)(x^4)\cdots \\ &\quad \text{one 4 corresponds to } x^4 \\ &\quad \text{no 1,2,3's correspond to } (1)(1)(1). \end{aligned}$$

Investigation 5.65. While the total number of partitions is certainly interesting, what if we place some restrictions on what parts can appear in the partition?

- (a) Find all partitions of 30 using only the parts $\{1, 5, 10, 25\}$.
- (b) Find all partitions of 20 using only prime numbers.
- (c) What other restrictions might you place on partitions?

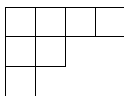
We will return to series representations of partitions shortly, but first we introduce another useful tool to the study of partitions: a Young³ diagram.

Definition 5.66. Given a partition $n = \sum a_i$, with $a_i \geq a_{i+1}$, we produce a diagram of equal sized boxes. In the first row, there are a_1 boxes. In the second row, there are a_2 boxes, etc. The resulting diagram is called a *Young diagram* for the partition.

Example 5.67. The partition

$$7 = 4 + 2 + 1$$

is represented with the diagram

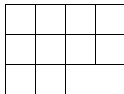


This may seem to be a very simple concept, and its power is in the visual representation of an algebraic concept. The next theorem illustrates its power.

Theorem 5.68. *The number of partitions of a positive integer n with largest part k is equal to the number of partitions of n with k parts.*

Proof. We create the Young diagram. Then we compare the partition coming from the standard orientation (rows) and the partition formed by the columns. The largest part k is the total number of columns, so the partition formed by the columns has k parts. Since the diagrams are in one-to-one correspondence with the partitions, the number of the two types of partitions is the same. \square

Example 5.69. We would represent the partition $10 = 4 + 4 + 2$ with largest part 4 as



The partition from the columns (with four parts) is

$$10 = 3 + 3 + 2 + 2.$$

The previous theorem is about partitions with restrictions either on the largest part or on the total number of parts.

³Alfred Young (1873–1940) was a British mathematician.

Question 5.70. Can you count the number of partitions with a given restriction (perhaps, partitions with largest part k)?

Definition 5.71. Let $p(n, k)$ be the number of partitions of n with largest part at most k (which is the same as the number of partitions with at most k parts).

It is actually quite simple to modify Euler's series for partitions $p(n, k)$.

Theorem 5.72. *We have the following equality of infinite series:*

$$\sum_{n=0}^{\infty} p(n, k) x^n = \frac{1}{1-x} \cdot \frac{1}{1-x^2} \cdots \frac{1}{1-x^k}.$$

Proof. We use geometric series to expand the right-hand side as

$$(1 + x + x^2 + x^3 + \cdots)(1 + x^2 + x^4 + \cdots) \cdots (1 + x^k + x^{2k} + \cdots).$$

Because we are taking one monomial from each set of parenthesis, the partitions have at most k parts. Note that we get fewer than k parts when the monomial 1 is used from a set of parentheses. \square

That is intriguing and leads to the following question.

Question 5.73. What other kinds of partitions can we count?

It seems that by modifying the exponents in

$$\prod_k \frac{1}{1-x^k},$$

we can change the number of parts or even which parts are allowed.

Example 5.74. The partitions that come from powers of 2 are counted with

$$\frac{1}{1-x} \cdot \frac{1}{1-x^2} \cdot \frac{1}{1-x^4} \cdot \frac{1}{1-x^8} \cdots.$$

An interesting application is currency counting.

Example 5.75. How many ways can you make \$1.00 given any number of quarters (\$0.25), dimes (\$0.10), nickels (\$0.05), and pennies (\$0.01)?

In other words, how many ways can you partition 100, using only 25, 10, 5, and 1. We are looking for the coefficient of x^{100} in the power series from multiples of $\{1, 5, 10, 25\}$.

$$\begin{aligned} \frac{1}{1-x} \cdot \frac{1}{1-x^5} \cdot \frac{1}{1-x^{10}} \cdot \frac{1}{1-x^{25}} &= 1 + x + x^2 + x^3 + x^4 + 2x^5 + \cdots \\ &\quad + 213x^{99} + 242x^{100} + 242x^{101} + \cdots. \end{aligned}$$

So there are 242 ways to make a dollar from coins.

With a little more effort we can also count partitions where the parts are all distinct.

Theorem 5.76. *The number of partitions with distinct parts $f(n)$ is given by*

$$\sum_{n=0}^{\infty} f(n)x^n = \prod_{i=1}^{\infty} (1+x^i).$$

Proof. Similar to the proof of Euler's original theorem, we take one monomial from each parentheses. However, the only monomials in each parentheses is either 1 or x^i . If x^i is used, then i can appear only once in the partition. \square

Example 5.77. The number of partitions of 6 with distinct parts is given by the coefficient of x^6 in

$$\begin{aligned} (1+x)(1+x^2)(1+x^3)(1+x^4)(1+x^5)(1+x^6) \\ = 1 + x + x^2 + 2x^3 + 2x^4 + 3x^5 + 4x^6 + \cdots, \end{aligned}$$

so the answer is 4. We can find them as

$$\begin{aligned} 6 &= 6, \\ 6 &= 5 + 1, \\ 6 &= 4 + 2, \\ 6 &= 3 + 2 + 1. \end{aligned}$$

We can combine these methods for even more restrictive questions.

Example 5.78. The number of partitions of 10 with parts restricted to powers of 2 is the coefficient of x^{10} in

$$\begin{aligned} (1+x^2+x^4+x^6+x^8+x^{10})(1+x^4+x^8)(1+x^8) \\ = 1 + x^2 + 2x^4 + 2x^6 + 4x^8 + 4x^{10} + \cdots, \end{aligned}$$


so the answer is 4. We can find them as

$$\begin{aligned} 10 &= 8 + 2, \\ 10 &= 4 + 4 + 2, \\ 10 &= 4 + 2 + 2 + 2, \\ 10 &= 2 + 2 + 2 + 2 + 2. \end{aligned}$$

We end this section with an open problem.

Question 5.79. Given a positive integer n , can you determine if $p(n)$ is even or odd without computing $p(n)$?

COMPUTATIONAL EXERCISES

 **5.1.** Compute the following values.


- a. $\varphi(210)$
- b. $\varphi(64)$
- c. $\varphi(54)$

- 5.2.** Determine all positive integers $n < 10^5$ such that $(\varphi(n) + 3)$ divides n .
- 5.3.** Find the positive integers n that solve the following equations.
- $\varphi(n) = 100$
 - $\varphi(n) = n/2$
 - $\varphi(n)$ divides $n + 1$
- 5.4.** The Goldbach⁴ conjecture states that every even integer greater than 2 can be written as the sum of two prime numbers.
- If the Goldbach conjecture is true, then for every positive integer m , there are primes p and q such that $\varphi(p) + \varphi(q) = 2m$. Find (p, q) for $1 \leq m \leq 10$.
 - Erdős⁵ asked if this holds for p and q not necessarily prime, but this relaxed form remains unproven even if we assume the Goldbach conjecture. Find (p, q) composite for $1 \leq m \leq 10$.

5.5. Consider the following conjecture of Michon.

Conjecture. *Every odd positive integer n such that $\gcd(n, \varphi(n)) = 1$ divides some Carmichael number.*

For each $3 \leq n < 30$ such that $\gcd(n, \varphi(n)) = 1$, find a Carmichael number which it divides.

 **5.6.** Compute the following values.

- $\mu(210)$
- $\mu(126)$
- $\mu(154)$

5.7.

- Given a positive integer B , write a program that outputs the squarefree positive integers at most B .
- Use this function to compute the average value of the Möbius function for $1 \leq n \leq 10^6$ and n squarefree.

5.8. Find the smallest n such that the cyclotomic polynomial $\Phi_n(x)$ has a coefficient other than 0, 1, or -1 .

5.9. Find the smallest triple of three consecutive positive integers each of which is squarefree with an even number of prime divisors.

5.10. Let p be a prime, and let S be the set of primitive roots for p . Verify for all primes $p < 1000$ that

$$\sum_{s \in S} s \equiv \mu(p-1) \pmod{p}.$$

5.11. Determine all positive integers n with $2 \leq n \leq 10,000$ that satisfy

$$\varphi(\sigma_1(n)) = n.$$

⁴Christian Goldbach (1690–1764) was a German mathematician.


⁵Paul Erdős (1916–1996) was a Hungarian mathematician.

5.12. A *nontotient* is a positive integer n for which $\varphi(x) = n$ has no solution. Determine the even nontotients $n \leq 100$.

5.13. Ford⁶ proved in 1999 that for every integer $k \geq 2$ there is a positive integer m for which the equation $\varphi(n) = m$ has exactly k solutions.

- a. For each $2 \leq k \leq 5$, determine an m and all n such that $\varphi(n) = m$ has k solutions.

However, no pair (m, n) is known for $k = 1$. Carmichael's totient function conjecture is that there are no positive integers m such that there is exactly one n with $\varphi(n) = m$.

 **5.14.** Compute the following values.

- a. $\sigma_2(15)$
 b. $\sigma_0(40)$
 c. $\omega(210)$

5.15. Write a function that takes as input positive integers k and n and returns $\sigma_k(n)$. Use it to compute the following.


- a. $\sigma_3(50)$
 b. $\sigma_{10}(8273)$
 c. $\sigma_0(182734823)$


5.16. Determine all positive integers k and n with $1 \leq n \leq 1000$ and $0 \leq k \leq 5$ such that $\sigma_k(n) = \sigma_k(n+1)$.

5.17. Determine the number of integers $1 \leq n \leq 10^5$ that have $\omega(n)$ even and the number that have $\omega(n)$ odd.

5.18. Determine the first 12 perfect numbers.

5.19. Use series to compute $p(n)$ for $1 \leq n \leq 20$.

 **5.20.** Determine all the partitions of 7 with at most three parts.

 **5.21.** Determine all the integers $n < 20$ that have a partition whose parts are either 5 or 7.

5.22. Determine the number of ways to represent \$1.32 given any number of quarters (\$0.25), dimes (\$0.10), nickels (\$0.05), and pennies (\$0.01).

5.23. Determine the number of ways you can partition 167 using only 5, 7, and 13 as parts.

5.24. Determine the number of partitions of 230 with at most five parts.

⁶Kevin Ford is an American mathematician.

THEORETICAL EXERCISES

5.25. Let k be a fixed positive integer. Prove that the function $f(n) = \gcd(n, k)$ is multiplicative.

5.26. Let m be a positive integer. Prove that

$$\varphi(2m) = \begin{cases} 2\varphi(m) & m \equiv 0 \pmod{2}, \\ \varphi(m) & m \equiv 1 \pmod{2}. \end{cases}$$

5.27. Let n be a positive integer. Prove that

$$\sum_{k=1}^n \sum_{d|k} \varphi(d) = \frac{n(n+1)}{2}.$$

5.28. Prove that for positive integers m and n that $m \mid n$ implies $\varphi(m) \mid \varphi(n)$.

5.29. Let m and n be positive integers with $d = \gcd(m, n)$. Prove that

$$\varphi(mn) = \varphi(m)\varphi(n)\frac{d}{\varphi(d)}.$$

5.30. Let n be a positive integer.

- a. Prove that if n has r distinct (odd) prime factors, then $2^r \mid \varphi(n)$.
- b. Prove that $\varphi(n)$ is even for $n \geq 3$.

5.31. Let n and k be positive integers with $k \geq 2$. Prove that

- a. $\varphi(n^k) = n\varphi(n^{k-1})$.
- b. $\varphi(\varphi(p^k)) = p^{k-2}\varphi((p-1)^2)$.

5.32. Fix a positive integer m . Prove that the set of solutions n to $\varphi(mn) = m\varphi(n)$ is $\{km : k \in \mathbb{N}\}$.

5.33. Let A , B , and C be nonempty finite sets. Prove the following inclusion-exclusion statements.

- a. $|A \cup B| = |A| + |B| - |A \cap B|$
- b. $|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|$

5.34. Prove that for a prime p and any positive integer e ,

$$\Phi_{p^e}(x) = \Phi_p(x^{p^{e-1}}).$$

5.35. Let n be a positive integer.

- a. Prove that $\deg(\Phi_n(x)) = \varphi(n)$.
- b. If n is odd, prove that $\Phi_{2n}(x) = \Phi_n(-x)$.

5.36. For a positive integer n , define

$$\sigma_{-1}(n) = \sum_{d|n} \frac{1}{d}.$$

- a. Prove that for any positive integer n ,

$$n\sigma_{-1}(n) = \sum_{d|n} \frac{n}{d} = \sum_{d|n} d.$$

- b. Use the previous part to prove that for any positive integer n ,

$$\frac{\sigma_1(n)}{n} = \sigma_{-1}(n).$$

- c. Conclude that if n is a perfect number,

$$\sigma_{-1}(n) = 2.$$

5.37. For a positive integer $n = pq$ with p and q prime numbers, prove that knowing n and either $\varphi(n)$ or $\sigma_1(n)$ is sufficient to determine p and q .

5.38. Prove that $\frac{\sigma_1(n)}{n}$ does not have a minimum value for positive integers n .

5.39. Let n be a positive integer. Prove that $\sigma_1(n) = 2^k$ for some positive integer k if and only if n is the product of distinct Mersenne primes.

5.40. Let m and n be positive integers. Prove that if $\gcd(m, n) = 1$, then $\omega(mn) = \omega(m) + \omega(n)$.

5.41. For a positive integer n , define $\Omega(n)$ as the number of prime divisors of n counted with multiplicity, e.g., $\Omega(4) = \#\{2, 2\} = 2$. Define $\lambda(n) = (-1)^{\Omega(n)}$. This is called the *Liouville function*.

- a. Prove that for any positive integers m and n that $\lambda(mn) = \lambda(m)\lambda(n)$, i.e., prove that $\lambda(n)$ is *completely* multiplicative.
 b. Let n be a positive integer. Prove that

$$\sum_{d|n} \lambda(d) = \begin{cases} 1 & n \text{ is a perfect square,} \\ 0 & \text{otherwise.} \end{cases}$$

5.42. Let n and k be positive integers. Prove that n^k has a partition with n parts.

5.43. Let n be a positive integer. Prove that the number of partitions of n which have all parts greater than 1 is given by

$$p(n) - p(n-1).$$

5.44. Let n be a positive integer. Prove that the number of partitions of n with distinct parts is equal to the number of partitions of n with odd parts.

5.45. Let n and k be positive integers with $k \leq n$. Prove that

$$p(n, k) = p(n-1, k-1) + p(n-k, k).$$

5.46. Prove that the number of ways to partition a positive integer n with consecutive parts is the same as the number of odd divisors (not including 1) of n . By consecutive parts we mean

$$n = r + (r+1) + \cdots + (r+k),$$

e.g.,

$$6 = 1 + 2 + 3.$$

5.47. Let m and n be positive integers, and let $f(m, n)$ be an arithmetic function of two variables. Assume that

$$\prod_{j=1}^{\infty} (1 + x^j + y^j) = \sum_{m,n} f(m, n) x^m y^n.$$

What is the function $f(m, n)$ counting?

EXPLORATION EXERCISES

5.48 (Coefficients of cyclotomic polynomials). R

- Find the first cyclotomic polynomial with a coefficient that is not one of $\{-1, 0, 1\}$.
- Find the first cyclotomic polynomial with absolute value at least k for $k \in \mathbb{N}$.
- Which numbers occur as a coefficient in a cyclotomic polynomial?
- Determine a growth rate for the coefficients of $\Phi_n(x)$ based on the number of prime factors of n .

5.49 (Repeated totient values).

- Can you find positive integers n such that $\varphi(n) = \varphi(n+1)$? How many such n are there?
- What about consecutive triples, $\varphi(n) = \varphi(n+1) = \varphi(n+2)$? How many such n are there?
- What is the longest consecutive sequence you can find with the same totient value?
- What about totient values from arithmetic progressions, i.e., sets of the form $\{an + b : n \in \mathbb{N}\}$ whose elements all have the same totient value?
- Can you find an arithmetic progression with a longer sequence of identical totient values than you found for consecutive sequences?

5.50 (Sum of divisors).

- What is the largest and smallest possible value of $\frac{\sigma_1(n)}{n}$?
- For which $a \in \mathbb{Q}$ does there exist an n such that $\sigma_1(n) = a$?
- Are there infinitely many n such that $\sigma_1(n) = a$ for each possible a ?

5.51 (Amicable and sociable numbers). Recall that given a positive integer n , we defined the sum of proper divisors of n as the *aliquot sum* denoted $s(n)$ (Definition 5.48).

Definition 5.80. We say that (m, n) are an *amicable pair*⁷ if $s(n) = m$ and $s(m) = n$. For example, $(220, 284)$ are an amicable pair.

More generally, an *aliquot sequence* is a sequence of positive integers (n_1, n_2, \dots) such that

$$s(n_i) = s(n_{i+1}) \quad \text{for all } i.$$

We say the sequence has *period* r if additionally

$$s(n_r) = s(n_1).$$

⁷Amicable numbers were known to the Pythagoreans, who credited them with many mystical properties. A general formula by which some of these numbers could be derived was invented circa 850 by Thābit ibn Qurra (826–901 C.E.).

An amicable pair is an aliquot sequence of period 2. Aliquot sequences of period larger than 2 are called *sociable numbers*.

- a. Can you find some more amicable pairs? What about sociable numbers of larger period?
- b. How large can the period be?
- c. Are there infinitely many sociable numbers of each possible period?
- d. Can you find a way to construct (or efficiently find) sociable numbers?

5.52 (Nonperfect numbers).

Definition 5.81. A number n is *almost perfect* if $s(n) = n \pm 1$, where $s(n) = \sigma_1(n) - n$ is the aliquot sum (Definition 5.48).

- a. What are all the almost perfect numbers?
- b. What about $s(n) = n \pm k$ for $k \in \mathbb{N}$?

Definition 5.82. We say that n is *abundant* if $s(n) > n$ and *deficient* if $s(n) < n$.

- c. Find some abundant/deficient numbers.
- d. Find some abundant/deficient numbers not divisible by 2 or 3.
- e. Find some examples of abundant/deficient numbers not divisible by the first k primes.
- f. What proportion of numbers are abundant/deficient?
- g. What is the largest value of $|s(n) - n|$?

5.53 (Number of prime divisors).

- a. Let k be a positive integer. There is a smallest integer n such that $\omega(n) \geq k$. Determine the growth rate of the smallest such n as $k \rightarrow \infty$.
- b. Determine the average value of $\omega(n)$ over a certain range.
- c. What is the longest sequence of consecutive n that have the same value for $\omega(n)$?
- d. How does the number of n with $\omega(n)$ having even parity compare to the number having odd parity over a given range?
- e. Can you find an infinite sequence of integers that have all even or all odd parity for $\omega(n)$?

Definition 5.83. For a positive integer n , define $\Omega(n)$ to be the number of prime divisors of n counted with multiplicity, e.g., $\Omega(4) = \#\{2, 2\} = 2$ and $\Omega(12) = \#\{2, 2, 3\} = 3$.

- f. Determine the growth rate for the smallest n such that $\Omega(n) \geq k$ for $k \in \mathbb{N}$.
- g. Determine the average value of $\Omega(n)$ over a certain range.

- h. What is the longest sequence of consecutive n that has the same value for $\Omega(n)$?
- i. Can you find an infinite sequence of numbers that all have the same value for $\Omega(n)$?

Definition 5.84. Define $\lambda(n) = (-1)^{\Omega(n)}$. This is called the *Liouville function*. The Liouville function returns the parity of $\Omega(n)$.

- j. Determine the average value of $\lambda(n)$ over a certain range.
- k. What is the longest sequence of consecutive n that has the same value for $\lambda(n)$?
- l. Can you find an infinite sequence of numbers that all have the same value for $\lambda(n)$?

5.54 (Prime partitions).

Definition 5.85. A partition of a number n is called a *prime partition* if every part is a prime number.

- a. Prime numbers can be written as prime partitions with one part. What numbers cannot be partitioned with fewer than k prime parts for $k \in \mathbb{N}$?
- b. Let $g(n)$ denote the number of prime partitions of n with two parts (Goldbach partitions). What can you say about this function?
- c. Determine a recursive relation for the number of prime partitions. In other words, can you determine the number of prime partitions of n if you know the number of prime partitions of all positive integers less than n ?

Algebraic Numbers

The set of integers represents a small but interesting subset of all numbers. The integers are closed under addition and multiplication. They have additive inverses but not multiplicative inverses. Extending the integers to include the multiplicative inverses, we get the set of fractions $\frac{a}{b}$ for $a, b \in \mathbb{Z}$ called the *rational numbers*.

Definition 6.1. The *rational numbers*, denoted \mathbb{Q} , is the set fractions of integers,

$$\mathbb{Q} = \left\{ \frac{a}{b} : a, b \in \mathbb{Z}, b \neq 0 \right\}.$$

In this chapter we extend our number system beyond integers and rational numbers. As motivation, consider the following question.

Question 6.2. Which primes $p \in \mathbb{Z}$ can be written as

$$p = a^2 + b^2$$

for integers a and b ?

The prime 5 is one such prime because $5 = 1^2 + 2^2$. You can also try all possible $a^2 + b^2$ for $a, b < 3$ to see that 7 cannot be written as the sum of two squares. What makes 5 different from 7? The key is to note that we can write

$$a^2 + b^2 = (a + b\sqrt{-1})(a - b\sqrt{-1}).$$

The value $\sqrt{-1}$ is an *algebraic number*, and the properties of such numbers (called *algebraic number theory*) is the focus of this chapter.

1. Algebraic or Transcendental

There is a fundamental difference between numbers such as π and $\sqrt{2}$. For example, $\sqrt{2}$ is a root of the polynomial $x^2 - 2$, but π is not the root of any polynomial with integer coefficients. This is an important distinction, and it is the difference between *algebraic* and *transcendental* numbers.

Definition 6.3. An *algebraic number* α is a root of a polynomial with integer coefficients. The irreducible polynomial that has α as a root is called the *characteristic polynomial* of α and the degree of the characteristic polynomial is the *degree* of the algebraic number α .

Transcendental numbers are the numbers that are not algebraic.

Example 6.4. The rational numbers $\mathbb{Q} = \{\frac{a}{b} : a, b \in \mathbb{Z}, b \neq 0\}$ are all algebraic numbers of degree 1. Each rational number $\frac{a}{b}$ is the root of the polynomial $bx - a$.

Example 6.5. The number $\sqrt{3}$ is a root of the polynomial $x^2 - 3$. The polynomial $x^2 - 3$ is irreducible so is the characteristic polynomial of $\sqrt{3}$. Thus, $\sqrt{3}$ is an algebraic number of degree 2.

The number $\sqrt[3]{3}$ is a root of the irreducible polynomial $x^3 - 3$ so is an algebraic number of degree 3.

Example 6.6. The cyclotomic polynomials of section 5.2 are the characteristic polynomials of the primitive n th roots of unity.

Question 6.7. Are there more algebraic numbers or transcendental numbers?

This may seem to be an odd question since there are infinitely many numbers in both sets. However, it is possible to show that there are “more” transcendental numbers. The notion of different infinite cardinalities was developed by Georg Cantor.¹

Question 6.8. How do you determine if a given number α is algebraic or transcendental?

To show a number is algebraic, one can simply exhibit its characteristic polynomial or, in fact, any polynomial with integer coefficients for which it is a root.

Example 6.9. Prove that $\alpha = \sqrt{1 + \sqrt{5}}$ is algebraic.

We see that

$$\alpha^2 = 1 + \sqrt{5},$$

so that

$$(\alpha^2 - 1)^2 = 5.$$

So we have

$$\alpha^4 - 2\alpha^2 - 4 = 0.$$

Thus, α is a root of

$$x^4 - 2x^2 - 4.$$

This polynomial is irreducible, so it is also the characteristic polynomial of α .

¹Georg Ferdinand Ludwig Philipp Cantor (1845–1918) was a German mathematician.

In general, it is very difficult to show that a number is transcendental. In 1844, Liouville² gave the first proof that a particular number is transcendental when he showed that

$$(19) \quad \sum_{k=1}^{\infty} 10^{-(k!)}$$

is transcendental. The number in (19) is called the *Liouville number*. Lindemann³ showed that π was transcendental in 1882; he also showed that e^α is transcendental for any algebraic (nonzero) α . Thus, $e = e^1$ is transcendental and, using the contrapositive, since $e^{i\pi} = -1$, π is transcendental.

For simplicity we consider only algebraic numbers of degree 2 for the rest of the chapter.

2. Quadratic Number Fields and Norms

Let α be a degree 2 algebraic number, a root of an irreducible polynomial $ax^2 + bx + c$ for some integers a , b , and c . In particular, α is not a rational number. We want to include α in our number system, so we must also include all combinations of α with rational numbers. For example, we want to include $1 + \alpha$ and $2 - 3\alpha$. The resulting object is similar to a \mathbb{Q} -vector space with basis $\{1, \alpha\}$ from linear algebra. However, since we also include the operations of multiplication and division, the resulting object is a field, not just a vector space. Notice that since $a\alpha^2 + b\alpha + c = 0$, we can write

$$\alpha^2 = -\frac{b}{a}\alpha - \frac{c}{a}.$$

In other words, α^2 can be written as a \mathbb{Q} -linear combination of the basis $\{1, \alpha\}$. The set of numbers resulting from extending the rational numbers by including α is called a *number field*.

Definition 6.10. For a squarefree integer d , we define

$$\mathbb{Q}(\sqrt{d}) = \{a + b\sqrt{d} : a, b \in \mathbb{Q}\},$$

that is, all the \mathbb{Q} -linear combinations of $\{1, \sqrt{d}\}$. If $d > 0$, we say $\mathbb{Q}(\sqrt{d})$ is a *real quadratic number field*. If $d < 0$, we say $\mathbb{Q}(\sqrt{d})$ is an *imaginary quadratic number field*.

Example 6.11. Consider the degree 2 algebraic number α defined as a root of $x^2 - 2$. In more familiar notation, $\alpha = \sqrt{2}$. The number α generates the real quadratic number field $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$. We can add two elements as

$$(1 + \sqrt{2}) + (3 + 2\sqrt{2}) = 4 + 3\sqrt{2}$$

and multiply two elements as

$$(1 + \sqrt{2})(3 + 2\sqrt{2}) = (3 + 2(\sqrt{2})^2) + (3 + 2)\sqrt{2} = 7 + 5\sqrt{2}.$$

²Joseph Liouville (1809–1882) was a French mathematician.

³Carl Louis Ferdinand von Lindemann (1852–1939) was a German mathematician.

Example 6.12. Consider the degree 2 algebraic number α defined as the root of $x^2 + x + 1$. By the quadratic equation, we can write α as the more familiar number

$$\alpha = \frac{1 + \sqrt{-3}}{2}.$$

In particular, α is a \mathbb{Q} -linear combination of $\{1, \sqrt{-3}\}$. Hence, α is an element of the imaginary quadratic field $\mathbb{Q}(\sqrt{-3})$.

What makes the notion of number fields interesting from a number theoretic perspective is that with the expansion of possible numbers, it may be that integers that were prime in \mathbb{Z} can now be factored.

Example 6.13. Consider $7 \in \mathbb{Q}(\sqrt{-3})$. Then we have

$$(2 - \sqrt{-3})(2 + \sqrt{-3}) = 7.$$

It is also possible for a number to have multiple distinct factorizations.

Example 6.14. Consider $21 \in \mathbb{Q}(\sqrt{-5})$. Then we have

$$21 = 3 \cdot 7,$$

$$21 = (1 + 2\sqrt{-5})(1 - 2\sqrt{-5}).$$

In addition to difficulty with unique factorization, the notion of less than or greater than does not make sense for algebraic numbers. For example, which is true,

$$(20) \quad -1 + \sqrt{-3} \leq 2 - \sqrt{-3} \quad \text{or} \quad -1 + \sqrt{-3} \geq 2 - \sqrt{-3}?$$

This ordering was essential when we were working with the integers, so we need something for algebraic numbers that will serve the same purpose.

Question 6.15. Can we salvage the notions of prime number, factorization, and ordering for number fields?

Addressing Question 6.15 is our main goal in this chapter.

Investigation 6.16. Examine factorization in number fields. However, be careful to avoid the equivalent of writing $5 = 1 \cdot 5$, which is not truly a new factorization. You may need to think about what makes 1 special in the integers.

- (a) For any given prime p , can you find a quadratic number field where you can factor p ?
- (b) For which quadratic number fields can you find a number with two different factorizations?

To solve the ordering problem, we define the *norm* of an algebraic number.

Definition 6.17. Define the *norm* of $a + b\sqrt{d}$ as $N(a + b\sqrt{d}) = a^2 - db^2$.

Since the norm of an algebraic number is rational, we can compare $N(\alpha)$ and $N(\beta)$ and produce an ordering.

Example 6.18. In equation (20) we tried unsuccessfully to compare the two algebraic numbers $-1 + \sqrt{-3}$ and $2 - \sqrt{-3}$. Using norms, we see

$$N(-1 + \sqrt{-3}) = 1 + 3 = 4,$$

$$N(2 - \sqrt{-3}) = 4 + 3 = 7,$$

so we can conclude that

$$2 - \sqrt{-3} > -1 + \sqrt{-3}.$$

However, this ordering is not the only purpose of the norm. We next prove two properties of norms that we will need in section 3 when we discuss prime numbers in quadratic fields. The first relates the norm to the characteristic polynomial, and the second shows that the norm is completely multiplicative.

Lemma 6.19. *Given $\alpha = u + v\sqrt{d}$ with characteristic polynomial $ax^2 + bx + c$ for integers a , b , and c ,*

$$N(\alpha) = \frac{c}{a}.$$

Proof. We first compute the norm of α as

$$N(\alpha) = u^2 - dv^2.$$

Now we need to show this expression is really $\frac{c}{a}$. To do this, we rewrite the polynomial $ax^2 + bx + c$ in terms of u , v , and d . The polynomial $ax^2 + bx + c$ has two roots, α and β , so we can write

$$ax^2 + bx + c = a(x - \alpha)(x - \beta) = a(x^2 - (\alpha + \beta)x + \alpha\beta).$$

Thus, $\alpha + \beta \in \mathbb{Q}$ and $\alpha\beta \in \mathbb{Q}$. Since $\alpha = u + v\sqrt{d}$, we must have that

$$\beta = u - v\sqrt{d}.$$

Consequently, the characteristic polynomial of α can be written as

$$a(x^2 - 2ux + (u^2 - dv^2)) = ax^2 - 2aux + a(u^2 - dv^2).$$

In particular, $c = a(u^2 - dv^2)$, and we have

$$\frac{c}{a} = u^2 - dv^2 = N(\alpha). \quad \square$$

Example 6.20. Let α be a root of $3x^2 + 4x + 6$. By Lemma 6.19, we have

$$N(\alpha) = \frac{6}{3} = 2.$$

Computing explicitly, we have

$$\alpha = \frac{-4 + \sqrt{16 - 72}}{6} = \frac{-2}{3} + \frac{1}{3}\sqrt{-14}.$$

We compute the norm as

$$\frac{4}{9} - (-14) \cdot \frac{1}{9} = \frac{18}{9} = 2.$$

Proposition 6.21. Given two numbers $\alpha, \beta \in \mathbb{Q}(\sqrt{d})$, we have

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

Proof. Let $\alpha = a + b\sqrt{d}$ and $\beta = u + v\sqrt{d}$. Then we compute

$$\begin{aligned} N(\alpha\beta) &= N((au + bvd) + (av + bu)\sqrt{d}) \\ &= (au + bvd)^2 - d(av + bu)^2 \\ &= a^2u^2 - db^2u^2 - dv^2a^2 + d^2v^2b^2. \end{aligned}$$

We also compute

$$N(\alpha)N(\beta) = (a^2 - b^2d)(u^2 - v^2d) = a^2u^2 - db^2u^2 - dv^2a^2 + d^2v^2b^2. \quad \square$$

Investigation 6.22. Proposition 6.21 proves that norms of algebraic numbers are multiplicative. Thus, if we are interested in which algebraic numbers are irreducible, it may be interesting to look at which algebraic numbers have norms that are prime numbers.

- (a) Fix a prime number p . Which algebraic numbers have norm equal to p ?
- (b) Do any of these numbers factor? What can you say about the norms of these factors?

3. Integers, Divisibility, Primes, and Irreducibles

To define the notions of prime number and factorization in number fields, we need notions of integers and divisibility. Recall from Chapter 1 that an integer d divides an integer n if there is an integer m such that $n = dm$. To make the same definition for algebraic numbers, we need a notion of “integer”.

Definition 6.23. A polynomial $a_nx^n + \cdots + a_1x + a_0$ is called *monic* if $a_n = 1$.

Example 6.24. The polynomial $x^2 + 1$ is monic, but $2x^2 + 1$ is not monic.

Definition 6.25. A number α is an *algebraic integer* if the characteristic polynomial of α is monic.

Example 6.26. The following are algebraic integers:

- $\sqrt{2}$ with characteristic polynomial $x^2 - 2$,
- $\sqrt{1 + \sqrt{5}}$ with characteristic polynomial $x^4 - 2x^2 - 5$.

The following are not algebraic integers:

- $\frac{\sqrt{2}}{3}$ with characteristic polynomial $3x^2 - 2$,
- $\frac{1}{1+\sqrt{5}}$ with characteristic polynomial $4x^2 - 2x + 1$.

Note that every $a \in \mathbb{Z}$ is an algebraic integer since its characteristic polynomial is $x - a$. Now we can generalize the definition of divisibility.

Definition 6.27. Given two algebraic integers α and β , we say that α *divides* β , denoted $\alpha \mid \beta$, if there is another algebraic integer γ such that

$$\beta = \alpha\gamma.$$

Example 6.28. For example, consider the following algebraic integers:

$$1 + \sqrt{3} \text{ with characteristic polynomial } x^2 - 2x - 2,$$

$$3 - 2\sqrt{3} \text{ with characteristic polynomial } x^2 - 6x - 3,$$

$$-3 + \sqrt{3} \text{ with characteristic polynomial } x^2 + 6x + 6.$$

They satisfy the relation

$$(1 + \sqrt{3})(3 - 2\sqrt{3}) = -3 + \sqrt{3},$$

so we can say

$$(1 + \sqrt{3}) \text{ divides } (-3 + \sqrt{3}).$$

We turn now to the notions of factorization and primes. Recall that since every integer is divisible by ± 1 , we considered

$$6 = 2 \cdot 3, \quad 6 = 1 \cdot 2 \cdot 3, \quad \text{and} \quad 6 = -1 \cdot -1 \cdot 2 \cdot 3$$

as the same factorization. Otherwise, we could add combinations of ± 1 to any factorization to get a “new” factorization. So we exclude ± 1 from factorizations. In $\mathbb{Q}(\sqrt{d})$ it is possible to get other “factors” of 1 that we must exclude. We call such numbers *units*, and they are characterized by having norm ± 1 .

Definition 6.29. A number $\alpha \in \mathbb{Q}(\sqrt{d})$ is a *unit* if $N(\alpha) = \pm 1$.

Example 6.30. The units in \mathbb{Q} are $\{\pm 1\}$.

The units in $\mathbb{Q}(\sqrt{-3})$ are

$$\left\{ \pm 1, \frac{\pm 1 \pm \sqrt{-3}}{2} \right\}.$$

We compute

$$N\left(\frac{\pm 1 \pm \sqrt{-3}}{2}\right) = \left(\frac{\pm 1}{2}\right)^2 - (-3)\left(\frac{\pm 1}{2}\right)^2 = \frac{1}{4} + \frac{3}{4} = 1.$$

Investigation 6.31. Every number field has ± 1 as units, but we saw that $\mathbb{Q}(\sqrt{-3})$ has additional units.

- (a) Find all the units for several different quadratic fields.
- (b) What field had the largest number of units? How many units did it have?
- (c) Can you find a quadratic field with infinitely many units?

In quadratic number fields, we make a distinction between irreducible numbers and prime numbers. We use the characterization in Lemma 1.38 for the notion of prime, which we restate in the following definition.

Definition 6.32. An algebraic integer p is *prime* if $p \mid ab$ implies $p \mid a$ or $p \mid b$.

An algebraic integer p is *irreducible* if its only divisors are itself and units.

For the integers, the notion of prime and irreducible coincide, so a factorization into irreducibles is a factorization into primes. However, given an algebraic integer, we can factor it into irreducibles, but the factors may or may not be prime.

Example 6.33. In $\mathbb{Q}(\sqrt{-5})$, 41 factors into two primes,

$$41 = (6 + \sqrt{-5})(6 - \sqrt{-5}).$$

However, 6 has two different factorizations into two irreducibles, none of which are prime:

$$6 = 2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5}).$$

In particular, the integer primes 2 and 3 are no longer prime in $\mathbb{Q}(\sqrt{-5})$. To see that 2 is not prime notice that

$$2 \mid (1 + \sqrt{-5})(1 - \sqrt{-5})$$

but

$$2 \nmid (1 + \sqrt{-5}) \quad \text{and} \quad 2 \nmid (1 - \sqrt{-5}).$$

The problem is that $\mathbb{Q}(\sqrt{d})$ may not have enough prime “numbers”, i.e., every irreducible should again split into primes, but those primes are not in $\mathbb{Q}(\sqrt{d})$. This issue is analogous to being unable to solve $x^2 = 2$ in \mathbb{Q} because the solution $\sqrt{2}$ is not rational. Kummer⁴ and Dedekind⁵ resolved this problem using a notion of prime ideals, and they recovered unique factorization for quadratic number fields, but this is well beyond our scope.

We consider instead the following questions.

Question 6.34. Given a quadratic number field $\mathbb{Q}(\sqrt{d})$, which primes $p \in \mathbb{Z}$ factor in $\mathbb{Q}(\sqrt{d})$?

Similarly we could ask the inverse question:

Question 6.35. Given a prime $p \in \mathbb{Z}$, for which squarefree integers d does p factor in $\mathbb{Q}(\sqrt{d})$?

Investigation 6.36.

- (a) Fix a squarefree integer d . Determine the primes $p \in \mathbb{Z}$ that factor in $\mathbb{Q}(\sqrt{d})$. Is this set of primes finite or infinite?
- (b) Fix a prime number $p \in \mathbb{Z}$. Determine all the quadratic number fields where p factors. Is this set of fields finite or infinite?

We use the multiplicativity of the norm to study irreducibles and Question 6.35.

Lemma 6.37. *If α is an algebraic integer, then $N(\alpha) \in \mathbb{Z}$.*

Proof. Let the characteristic polynomial of α be $ax^2 + bx + c$. However, since α is an algebraic integer, we have $a = 1$. Using Lemma 6.19, we compute the norm as

$$N(\alpha) = \frac{c}{a} = c$$

so that $N(\alpha) \in \mathbb{Z}$. □

⁴Ernst Eduard Kummer (1810–1893) was a German mathematician.

⁵Julius Wilhelm Richard Dedekind (1831–1916) was a German mathematician.

Notice that Lemma 6.37 is not an “if and only if” statement. It is possible to have an algebraic number with integer norm that is not an algebraic integer.

Example 6.38. The algebraic number with characteristic polynomial

$$2x^2 + x + 4$$

has norm 2 but is not an algebraic integer.

Theorem 6.39. *If an algebraic integer α has $N(\alpha)$ prime in \mathbb{Z} , then α is irreducible.*

Proof. We prove the contrapositive. From Lemma 6.37 we know that $N(\alpha)$ is an integer. From Proposition 6.21 we know that if α factors as

$$\alpha = ab$$

with a and b nonunits, then

$$N(\alpha) = N(a)N(b),$$

and since a and b are not units, the norms $N(a)$ and $N(b)$ are not ± 1 , so the integer $N(\alpha)$ is composite. \square

Example 6.40. In $\mathbb{Q}(\sqrt{-5})$,

$$N(3 + 2\sqrt{-5}) = 9 - (-5)4 = 29$$

so that $3 + 2\sqrt{-5}$ is irreducible.

Unfortunately, the reverse is not true.

Example 6.41. The number 5 is irreducible in $\mathbb{Q}(\sqrt{7})$, but

$$N(5) = 25$$

is composite.

We use quadratic residues (Definition 3.3) to answer Question 6.35.

Theorem 6.42. *Let p be an odd prime, and let d be a squarefree integer. If p factors in $\mathbb{Q}(\sqrt{d})$, then d is a quadratic residue modulo p .*

Proof. If p factors as $p = uv$ for algebraic integers $u, v \in \mathbb{Q}(\sqrt{d})$, by Theoretical Exercise 6.21, we have either

$$p = (a + b\sqrt{d})(a - b\sqrt{d}) \quad \text{or} \quad p = \left(\frac{a + b\sqrt{d}}{2}\right)\left(\frac{a - b\sqrt{d}}{2}\right),$$

with $a, b \in \mathbb{Z}$. Taking the norm of this equation, we have

$$p^2 = (a^2 - db^2)^2 \quad \text{or} \quad 16p^2 = (a^2 - db^2)^2,$$

which is the same as

$$\pm p = a^2 - db^2 \quad \text{or} \quad \pm 4p = a^2 - db^2.$$

Reducing modulo p , both equations become

$$a^2 \equiv db^2 \pmod{p}.$$

If $b \equiv 0 \pmod{p}$, then $a \equiv 0 \pmod{p}$ and p^2 divides the right-hand side, $a^2 - db^2$. Consequently, p^2 also divides the left-hand side, $\pm p$ or $\pm 4p$. For the first equation this is not possible; for the second it is only possible for $p = 2$, which is excluded by hypothesis.

If $b \not\equiv 0 \pmod{p}$, then b has an inverse mod p , and we can solve

$$d \equiv (a/b)^2 \pmod{p}.$$

In other words, d is a quadratic residue modulo p . □

Remark. The converse statement of Theorem 6.42 is also true, so that p factors in $\mathbb{Q}(\sqrt{d})$ if and only if d is a quadratic residue modulo p .

The case of $p = 2$ is examined in Exploration Exercise 6.23.

Example 6.43.

- 7 is a quadratic nonresidue modulo 5, so 5 remains irreducible in $\mathbb{Q}(\sqrt{7})$.
- 7 is a quadratic residue modulo 3, so 3 factors in $\mathbb{Q}(\sqrt{7})$ as

$$(2 - \sqrt{7})(-2 - \sqrt{7}) = -4 + 7 = 3.$$

We end with a question that has driven an entire field of study in modern number theory.

Question 6.44. For which number fields (not necessarily quadratic) $\mathbb{Q}(\alpha)$ does unique factorization fail?

4. Application: Sums of Two Squares

While we have barely scratched the surface in our investigations of primes and factorization for algebraic numbers, we have enough to answer Question 6.2, which motivated this chapter. We are trying to write a prime $p \in \mathbb{Z}$ as the sum of two integer squares $p = a^2 + b^2$.

Let's generate some data and see if we can find any patterns. The following are all the primes up to 100 that can be written as the sum of two squares.

p	$a^2 + b^2$	p	$a^2 + b^2$
2	$1^2 + 1^2$	41	$4^2 + 5^2$
5	$1^2 + 2^2$	53	$2^2 + 7^2$
13	$2^2 + 3^2$	61	$5^2 + 6^2$
17	$1^2 + 4^2$	73	$3^2 + 8^2$
29	$2^2 + 5^2$	89	$5^2 + 8^2$
37	$1^2 + 6^2$	97	$4^2 + 9^2$

There are many kinds of patterns we could look for, such as common residue classes modulo 4, but after you stare at the table for a little while, it may occur to you that the right-hand columns look a lot like norms ($a^2 - db^2$). In fact, given a number

$a + b\sqrt{-1} \in \mathbb{Q}(\sqrt{-1})$, we have $N(a + b\sqrt{-1}) = a^2 + b^2$. Furthermore, in $\mathbb{Q}(\sqrt{-1})$ we can write

$$p = a^2 + b^2 = (a + b\sqrt{-1})(a - b\sqrt{-1}),$$

so what we are really after is which primes factor in $\mathbb{Q}(\sqrt{-1})$.

Theorem 6.45. *Given a prime $p \in \mathbb{Z}$, we can write p as the sum of two squares if and only if p factors in $\mathbb{Q}(\sqrt{-1})$.*

Proof. One direction is the discussion we just finished: if we can write $p = a^2 + b^2$, then we can factor it as $(a + b\sqrt{-1})(a - b\sqrt{-1})$.

Now, assume that p factors in $\mathbb{Q}(\sqrt{-1})$ as $p = uv$ for algebraic integers u and v . Then, $N(p) = p^2 = N(u)N(v)$ so that $N(u) = N(v) = \pm p$. Then if $u = a + b\sqrt{-1}$, we have

$$\pm p = N(u) = a^2 + b^2.$$

However, since $a^2 + b^2$ is the sum of two square integers, it must be positive, and we have $p = a^2 + b^2$. \square

Using quadratic reciprocity (section 3.1), we can translate this into a congruence condition.

Corollary 6.46. *A prime $p \in \mathbb{Z}$ can be written as the sum of two squares if and only if $p = 2$ or $p \equiv 1 \pmod{4}$.*

Proof. For $p = 2$, we can write it as $2 = 1^2 + 1^2$.

For an odd prime p , we know from Theorem 6.42 that p remains prime if and only if

$$\left(\frac{-1}{p}\right) = 1,$$

where $\left(\frac{a}{p}\right)$ is the Legendre symbol. By Euler's criterion (Theorem 3.10), we compute

$$\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}}.$$

The exponent $\frac{p-1}{2}$ is even if and only if $p \equiv 1 \pmod{4}$. \square


Investigation 6.47. Fix an integer d . We want to know which primes p can be written as a sum

$$p = a^2 + db^2.$$

Corollary 6.46 answers the question for $d = 1$.

- (a) Which primes $p \in \mathbb{Z}$ can be written as $p = a^2 + 2b^2$ for integers a and b ?
- (b) Can you determine a pattern for the possible p for some other integers d ?


COMPUTATIONAL EXERCISES


 **6.1.** Find the characteristic polynomial of the following algebraic numbers. State whether or not they are algebraic integers.

a. $\frac{-1+\sqrt{-3}}{2}$

b. $\sqrt[3]{1+\sqrt{-7}}$


c. $\frac{\sqrt{5}}{\sqrt[3]{1+\sqrt{2}}}$

 **6.2.** Show that the algebraic integer $24+5\sqrt{-3}$ is divisible by $2+3\sqrt{-3}$ in $\mathbb{Q}(\sqrt{-3})$.

 **6.3.** Compute the norm of $6+\sqrt{-5}$ and $6-\sqrt{-5}$ in $\mathbb{Q}(\sqrt{-5})$ and conclude that both numbers are irreducible.

6.4. Find all the units in $\mathbb{Q}(\sqrt{-1})$.

6.5. Find the smallest positive integers x and y such that $x+y\sqrt{13}$ is a unit in $\mathbb{Q}(\sqrt{13})$.

 **6.6.** Find more than one factorization into irreducibles for each of the following numbers α in the given field.

a. $\mathbb{Q}(\sqrt{-3})$, $\alpha = 4$.

b. $\mathbb{Q}(\sqrt{-13})$, $\alpha = 14$.

6.7. Find all integers d with $|d| \leq 100$ such that 2 factors in $\mathbb{Q}(\sqrt{d})$.

6.8. For the field $\mathbb{Q}(\sqrt{-3})$, consider the numbers $a+b\sqrt{-3}$ for $-5 \leq a, b \leq 5$. Sort these numbers with respect to “ \leq ”.

6.9. For all integer primes $p \leq 100$, if possible write p as the sum of two integer squares.

6.10. Find all algebraic integers in $\mathbb{Q}(\sqrt{-13})$ with norm 133.

6.11. Given the polynomial $x^6 + 4x^4 + x^3 - 18x^2 - x - 2$, which is irreducible over \mathbb{Q} , find all primes $p < 1000$ for which f is reducible modulo p .

THEORETICAL EXERCISES

6.12. Prove that the characteristic polynomial of an algebraic number is unique (up to scaling).

6.13. If α is algebraic, prove that $\frac{1}{\alpha}$ is also algebraic.

6.14. Given two algebraic integers α and β , prove that if α divides β , then $N(\alpha)$ divides $N(\beta)$.

6.15. Prove that $\{2, 3, 1+\sqrt{-5}, 1-\sqrt{5}\}$ are irreducible in $\mathbb{Q}(\sqrt{-5})$.

6.16. For any squarefree integer d , prove that $\mathbb{Q}(\sqrt{d})$ is a field. For all $a, b, c \in \mathbb{Q}(\sqrt{d})$, prove the following properties:

a. Closed under addition and multiplication: $a+b \in \mathbb{Q}(\sqrt{d})$ and $a \cdot b \in \mathbb{Q}(\sqrt{d})$.

- b. Associativity of addition and multiplication: $(a + b) + c = a + (b + c)$ and $(ab)c = a(bc)$.
- c. Commutativity of addition and multiplication: $a + b = b + a$ and $ab = ba$.
- d. Existence of additive and multiplicative identities: $0, 1 \in \mathbb{Q}(\sqrt{d})$.
- e. Existence of additive and multiplicative inverses: $-a \in \mathbb{Q}(\sqrt{d})$ and $\frac{1}{a} \in \mathbb{Q}(\sqrt{d})$.
- f. Distributivity: $a(b + c) = ab + ac$.

6.17. Let d be a squarefree integer.

- a. Prove that if u is a unit in $\mathbb{Q}(\sqrt{d})$, then there is a unit v such that $uv = 1$.
- b. Prove that every algebraic integer $\alpha \in \mathbb{Q}(\sqrt{d})$ is divisible by every unit in $\mathbb{Q}(\sqrt{d})$.

6.18. Let d be a squarefree integer, and let c be any integer. Prove that the sets $\mathbb{Q}(\sqrt{d})$ and $\mathbb{Q}(\sqrt{c^2d})$ are the same.

6.19. Prove that $\{1, \sqrt[3]{2}, \sqrt[3]{4}\}$ is a basis for $\mathbb{Q}(\sqrt[3]{2})$. In other words, prove that every $\alpha \in \mathbb{Q}(\sqrt[3]{2})$ can be written as

$$\alpha = a + b\sqrt[3]{2} + c\sqrt[3]{4} \quad \text{with} \quad a, b, c \in \mathbb{Q}.$$

6.20. Prove that every algebraic integer can be factored into the product of a finite number of irreducibles (and a unit).

6.21. Prove that the following sets are all the algebraic integers in $\mathbb{Q}(\sqrt{d})$:

$$\begin{cases} \mathbb{Z}[\sqrt{d}] = \{a + b\sqrt{d} : a, b \in \mathbb{Z}\} & \text{if } d \equiv 2 \text{ or } 3 \pmod{4}, \\ \mathbb{Z}\left[\left(\frac{1+\sqrt{d}}{2}\right)\right] = \left\{\frac{a+b\sqrt{d}}{2} : a, b \in \mathbb{Z}\right\} & \text{if } d \equiv 1 \pmod{4}. \end{cases}$$

EXPLORATION EXERCISES

6.22 (Sums of squares). Consider a positive integer n . How many ways can n be written as the sum of k squares for some positive integer k ? For example for $k = 2$, we can write 50 in two different ways:

$$50 = 1^2 + 7^2 = 5^2 + 5^2.$$

- a. Find the smallest positive integer that cannot be written as a sum of two squares.
- b. Can all positive integers be written as the sum of three squares in at least one way?
- c. For which k can all positive integers be written as the sum of k squares in at least one way?
- d. For each $k \in \mathbb{N}$, which positive integers can be written as the sum of k squares in two ways? three ways? m ways?

6.23 (Factoring 2 in quadratic number fields). In Theorem 6.42 we determined that an odd prime factors in $\mathbb{Q}(\sqrt{d})$ if d is a quadratic residue modulo p . Determine the behavior of 2 in $\mathbb{Q}(\sqrt{d})$.

- For which d does 2 remain prime?
- For which d does 2 factor? Can you make a distinction between when 2 is a perfect square and when 2 factors into two distinct irreducibles?
- Conjecture the general conditions on d for each behavior.

6.24 (Prime splitting in number fields). Let $p \in \mathbb{Z}$ be a prime number. We are interested in how p factors in $\mathbb{Q}(\alpha)$ for an algebraic integer α .

- Consider α as a root of $x^2 - d$. Factor $f = x^2 - d$ modulo p . There are three possible cases: f remains irreducible, f splits into two distinct linear factors, or f is a square of one linear factor. How does each of these outcomes correspond to the factorization of p in $\mathbb{Q}(\sqrt{d})$? Compare your answer with the statement of Theorem 6.42.
- Consider now the number fields $\mathbb{Q}(\alpha)$ where α has minimal polynomial $f = x^3 - d$ for a positive integer d . Factor f modulo p and compare the result to the factorization of p in $\mathbb{Q}(\alpha)$.
- Consider an algebraic integer α of degree > 2 with minimal polynomial f . What does the factorization of f modulo p say about the factorization of p in $\mathbb{Q}(\alpha)$?

6.25 (Bases of number fields). Recall that every element of $\mathbb{Q}(\sqrt{d})$ can be written as $a + b\sqrt{d}$ for $a, b \in \mathbb{Q}$. We call $\{1, \sqrt{d}\}$ a \mathbb{Q} -basis for $\mathbb{Q}(\sqrt{d})$.

- Find a \mathbb{Q} -basis for $\mathbb{Q}(\sqrt[3]{d})$.

We can extend our definition of number fields as follows.

Definition 6.48. For numbers $\alpha_1, \dots, \alpha_n$ define

$$\mathbb{Q}(\alpha_1, \dots, \alpha_n) = \{a_0 + a_1\alpha_1 + \dots + a_n\alpha_n : a_0, \dots, a_n \in \mathbb{Q}\}.$$

For example,

$$\mathbb{Q}(\sqrt{2}, \sqrt{3}) = \{a + b\sqrt{2} + c\sqrt{3}\}.$$

- For $\mathbb{Q}(\sqrt{2}, \sqrt{3})$, can you find a single α such that the basis is $\{1, \alpha, \alpha^2, \alpha^3\}$?
- Can you find a basis generated by a single α when you change $\sqrt{2}$ and $\sqrt{3}$?
- What happens to the size of the basis when you add an additional square root?
- What about higher roots such as $\mathbb{Q}(\sqrt[n]{2}, \sqrt[k]{3})$?

Rational and Irrational Numbers

In the previous chapter we divided (complex) numbers into two categories, algebraic versus transcendental. We can also divide (real) numbers into two other categories: rational versus irrational. Our main focus will be approximating irrational numbers with rational numbers. For example, $\pi \approx \frac{355}{113}$.

1. Diophantine Approximation

Definition 7.1. Recall that the set of *rational numbers*, denoted \mathbb{Q} , is the set of fractions of integers:

$$\mathbb{Q} = \left\{ \frac{a}{b} : a, b \in \mathbb{Z}, b \neq 0 \right\}.$$

Notice that there are infinitely many ways to write a rational number $t \in \mathbb{Q}$ as a fraction of integers. For example,

$$\frac{1}{2} = \frac{2}{4} = \frac{3}{6} = \cdots.$$

A rational number $q = \frac{a}{b}$ is in *lowest terms* if $\gcd(a, b) = 1$.

Definition 7.2. If a real number r is not a rational number, it is called an *irrational number*. For example, $\sqrt{2}$ is irrational.

Theorem 7.3. $\sqrt{2}$ is irrational.

Proof. We proceed by contradiction and assume $\sqrt{2} \in \mathbb{Q}$. Then there are two relatively prime integers a and b such that

$$2 = \frac{a^2}{b^2}.$$

Multiplying both sides by b^2 , we have

$$2b^2 = a^2,$$

so a must be even. Write $a = 2a'$ for some integer a' . Then we have

$$2 = \frac{4(a')^2}{b^2},$$

and so

$$b^2 = 2(a')^2,$$

implying b must be even. However, $\gcd(a, b) = 1$ by assumption, so having both a and b even is a contradiction. Therefore, $\sqrt{2}$ must be irrational. \square

It is often difficult to prove that a given number is irrational. Johann Heinrich Lambert¹ proved in the 1760s that π is irrational. In fact, it is known that e and π are irrational; but, oddly enough, it is not known whether $\pi \cdot e$ or $\pi - e$ is irrational. It is known, however, that at least one of $\pi \cdot e$ and $\pi - e$ is irrational.

Question 7.4. Are there more rational numbers or irrational numbers?

As with algebraic versus transcendental, there are infinitely many numbers in both sets. However, because every rational number is algebraic (Example 6.4), every transcendental number is irrational, leading us to expect there to be more irrational numbers. It is possible to show that there are “more” irrational numbers using the notions of infinite cardinalities developed by Georg Cantor.

Instead of trying to prove which numbers are rational or irrational, we examine approximating irrational numbers with rational numbers. This area of number theory is called *Diophantine approximation* after Diophantus of Alexandria.² Diophantine approximation has a long history. For example, approximations of π have been known for thousands of years:

- (1900 B.C.E.) Ancient Babylonian’s $\pi \approx \frac{25}{8} = 3.125$.
- (1900 B.C.E.) Ancient Egyptian’s $\pi \approx \frac{256}{81} = 3.160$.
- (287–212 B.C.E.) Ancient Greece (Archimedes, using inscribed and circumscribed polygons)

$$3\frac{10}{71} < \pi < 3\frac{10}{70}.$$

The average value is 3.14185.

- (480 C.E.) A Chinese mathematician (with polygons) $\pi \approx \frac{355}{113} = 3.1415929$.

The modern computer algorithms for computing π to many decimals typically use some form of an infinite series that converges to π . Each finite partial sum represents π as a rational number. As we take more and more terms, the rational number gets closer and closer to π , but the numerator and denominator can get very large very quickly. For example, a naive rational approximation for π simply takes the first few digits of π divided by a power of 10,

$$\pi = 3.141592 \dots \approx \frac{3141592}{1000000} = \frac{392699}{125000}.$$

¹Johann Heinrich Lambert (1728–1777) was a Swiss mathematician.

²Diophantus of Alexandria was a Greek mathematician. The exact dates of his life are unknown, but a famous problem in the *Greek Anthology* puts his age of death at 84. Most historians place his life in the 3rd century C.E.

Intuitively, this is a “worse” approximation of π than $\frac{355}{113}$ because it gives the same number of correct digits but requires a fraction with a larger numerator and denominator. So our goal is to find the rational number with the smallest numerator and denominator that is closest in absolute value.

Definition 7.5. Let α be a real number. We say a rational number $\frac{p}{q}$ is a *best approximation of the first kind* if $1 \leq b \leq q$ and if $\frac{a}{b} \neq \frac{p}{q}$ implies that

$$\left| \alpha - \frac{a}{b} \right| > \left| \alpha - \frac{p}{q} \right|.$$

Besides finding a “best approximation”, we also want to bound the error between that approximation and the actual value.

Question 7.6. How well can an irrational number be approximated by a “small” rational number?

We will prove in Theorem 7.22 that if an irrational number can be approximated “very well” by rational numbers, then it must be transcendental!

To make Question 7.6 more precise, we need a definition for “small”.

Investigation 7.7. Pick an irrational number α and an error bound ϵ .

- (a) What is the rational number t with the smallest denominator such that $|\alpha - t| < \epsilon$?
- (b) Repeat this for successively smaller ϵ .
- (c) How does the choice of ϵ affect the resulting approximation t ?

2. Height of a Rational Number

While it was easy to determine the “size” of an integer (absolute value), determining the size of a rational number has several choices. One choice is to render $t = \frac{a}{b}$ as a decimal and consider $|t|$. Consider the following two rational numbers:

$$\left| \frac{1}{4} \right| = 0.25 \quad \text{and} \quad \left| \frac{100000001}{400000000} \right| = 0.2500000025.$$

Even though their absolute values are almost the same, the second number is much more complicated. Instead of using absolute value for size, we want to measure the complexity of the number. Accordingly, we introduce the notion of the height of a number.

Definition 7.8. We define the *height* of a rational number $t = \frac{a}{b}$ with $\gcd(a, b) = 1$ as

$$H\left(\frac{a}{b}\right) = \max(|a|, |b|).$$

We define the height of a list of rational numbers as

$$H(t_1, \dots, t_n) = \max(H(t_1), \dots, H(t_n)).$$

Example 7.9.

- (a) $H\left(\frac{1}{4}\right) = 4.$
- (b) $H\left(\frac{100000001}{400000000}\right) = 400000000.$
- (c) $H(-5) = 5.$
- (d) $H\left(\frac{2}{15}, \frac{7}{17}, -\frac{3}{2}\right) = 17.$

Remark. The height of a number is in direct correspondence to how many bits a computer would require to store the number.

A very useful property of heights is that there are only finitely many rational numbers with height less than a given bound, whereas there are infinitely many rational numbers whose absolute value is less than a given bound.

Example 7.10. We can list all rational numbers with height less than 4. They are the numbers whose numerator and denominator are at most 3 in absolute value. We have the 15 numbers

$$\left\{0, \pm 1, \pm 2, \pm \frac{1}{2}, \pm \frac{1}{3}, \pm \frac{2}{3}, \pm \frac{3}{2}, \pm 3\right\}.$$

Theorem 7.11. For any constant $B > 0$, there are finitely many $t \in \mathbb{Q}$ such that $H(t) < B$.

Proof. Let $t = \frac{a}{b}$ with $a, b \in \mathbb{Z}$ and $\gcd(a, b) = 1$. We know that $H\left(\frac{a}{b}\right) = \max(|a|, |b|)$, so for $H\left(\frac{a}{b}\right) < B$, we must have $|a| < B$ and $|b| < B$. Since there are only finitely many such choices for a and b , there are only finitely many $t \in \mathbb{Q}$ such that $H(t) < B$. \square

This finiteness allows us to define more general counting functions than we used in section 1.4.

Definition 7.12. Given a set S and a positive integer B , we define

$$N_S(B) = \#\{x \in S : H(x) < B\}.$$

Example 7.13. Counting integers, we have

$$N_{\mathbb{Z}}(B) = \#\{a \in \mathbb{Z} : |a| < B\} = 2(B-1) + 1 = 2B-1.$$

Example 7.14. We could also count solutions to equations. Define S to be rational solutions (x, y) to $x^2 + y^2 = 1$. We compute

$$N_S(B) = \#\{(x, y) : x, y \in \mathbb{Q} \text{ and } x^2 + y^2 = 1\}.$$

Note that a right triangle with integer sides (x, y, z) (a Pythagorean triple) corresponds to such a point since $x^2 + y^2 = z^2$ implies $\left(\frac{x}{z}\right)^2 + \left(\frac{y}{z}\right)^2 = 1$. The following

table shows the value of $N_S(B)$ for various B values.

B	$N_S(B)$	$\lfloor \frac{4B}{\pi} \rfloor$
10	12	12
50	60	63
100	132	127
500	644	636
1000	1268	1273
2000	2556	2546

The last column gives an approximate value of $N_S(B)$ in terms of B showing that $N_S(B)$ is linear in B ; in other words, the following limit converges to a nonzero constant:

$$\lim_{B \rightarrow \infty} \frac{N_S(B)}{B}.$$

Example 7.15. Counting the number of solutions to an equation by height can often result in striking geometric images. Figure 7.1 represents the projection to the xy -plane of rational points (x, y, z) such that $x^2 + y^2 + z^2 = 1$ up to height 500. The white discs are centered around points that cannot be well approximated by a rational number with height at most 500.

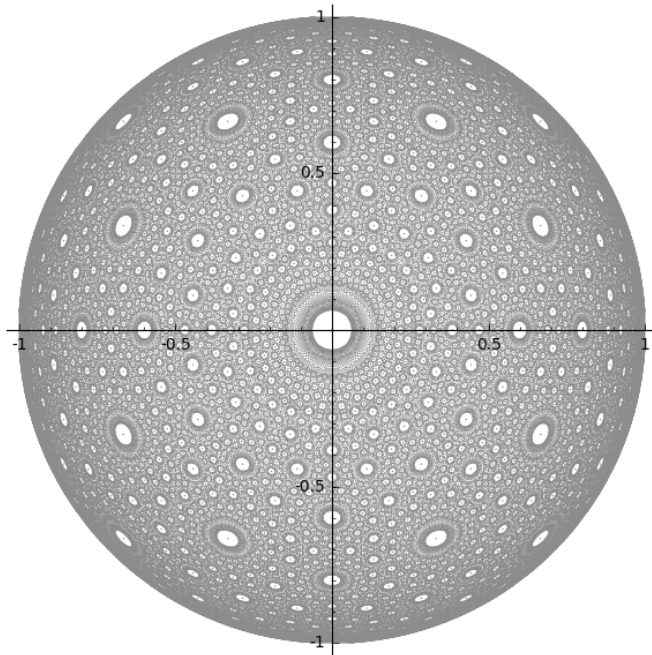


Figure 7.1. Rational points on the unit sphere

Investigation 7.16. A rough approximation of growth rates for $N_S(B)$ is the exponent t that makes the following limit converge to a nonzero constant

$$(21) \quad \lim_{B \rightarrow \infty} \frac{N_S(B)}{B^t}.$$

- (a) Determine the exponent t needed for the limit (21) to converge for $S = \mathbb{Q}$.
- (b) Determine the exponent t needed for the limit (21) to converge when S is the set of rational solutions to $y = x^2$.
- (c) Determine the exponent t needed for the limit (21) to converge when S is the set of rational solutions to $y^2 = x^3 + x + 1$.

3. Heights and Approximations

Now that we have defined the height of a rational number, we are ready to return to approximations. Let α be a real number, and let $\frac{p}{q}$ be an approximation. We can represent α as a decimal $b.b_0b_1b_2\dots$, where b is the integer part and b_i the decimal digits. Similarly, we can write $\frac{p}{q} = b + \frac{p'}{q}$ with $p' < q$. Consequently, when we talk about the size of a rational approximation, we use the absolute value of the denominator, which is the height of the approximation of the decimal part of α : $|q| = H\left(\frac{p'}{q}\right)$.

Example 7.17. Let's look at a few best approximations of the first kind (Definition 7.5) for π . The decimals on the right-hand side are carried out until the digits no longer agree with the actual value of π .

$$\begin{aligned} 3 &\approx 3\dots, \\ \frac{19}{6} &\approx 3.1\dots, \\ \frac{22}{7} &\approx 3.14\dots, \\ \frac{267}{85} &\approx 3.141\dots, \\ \frac{333}{106} &\approx 3.1415\dots, \\ \frac{355}{113} &\approx 3.141592\dots. \end{aligned}$$

What you should notice in Example 7.17 is that to get a better approximation, we need numbers with a larger height. In particular, if a rational number is very close in absolute value to an irrational number, then the height of that rational number must be large (recall the white disks in Figure 7.1). We now make Question 7.6 more precise.

Question 7.18. For an irrational number α , what is the minimum value of $\left|\alpha - \frac{a}{b}\right|$ for a rational number $\frac{a}{b}$ in terms of $|b|$?

Example 7.19. We examine the accuracy of the best approximations of the first kind for two irrational numbers.

α	$ b $	$\frac{a}{b}$	$ \alpha - \frac{a}{b} $
$\sqrt{2}$	$\approx 10^2$	$\frac{99}{70}$	$7.215 \cdot 10^{-5}$
	$\approx 10^3$	$\frac{577}{408}$	$2.124 \cdot 10^{-6}$
	$\approx 10^4$	$\frac{8119}{5741}$	$1.073 \cdot 10^{-8}$
	$\approx 10^5$	$\frac{47321}{33461}$	$3.158 \cdot 10^{-10}$
	$\approx 10^6$	$\frac{665857}{470832}$	$1.595 \cdot 10^{-12}$
	$\approx 10^7$	$\frac{9369319}{6625109}$	$8.055 \cdot 10^{-15}$
α	$ b $	$\frac{a}{b}$	$ \alpha - \frac{a}{b} $
$\sqrt{3}$	$\approx 10^2$	$\frac{97}{56}$	$9.205 \cdot 10^{-5}$
	$\approx 10^3$	$\frac{989}{571}$	$1.771 \cdot 10^{-6}$
	$\approx 10^4$	$\frac{5042}{2911}$	$3.407 \cdot 10^{-8}$
	$\approx 10^5$	$\frac{70226}{40545}$	$1.756 \cdot 10^{-10}$
	$\approx 10^6$	$\frac{978122}{564719}$	$9.052 \cdot 10^{-13}$
	$\approx 10^7$	$\frac{9973081}{5757961}$	$1.741 \cdot 10^{-14}$

It is important to note that if we simply took the first n digits of the number, i.e., denominator 10^n , we would end up with a much larger error than the best approximation.

α	$ b $	$\frac{a}{b}$	$ \alpha - \frac{a}{b} $
$\sqrt{2} \approx 1.41421356$	10	$\frac{14}{10}$	$1.421 \cdot 10^{-2}$
	10^2	$\frac{141}{100}$	$4.214 \cdot 10^{-3}$
	10^3	$\frac{1414}{1000}$	$2.136 \cdot 10^{-4}$
	10^4	$\frac{14142}{10000}$	$1.356 \cdot 10^{-5}$
	10^5	$\frac{141421}{100000}$	$3.562 \cdot 10^{-6}$
	10^6	$\frac{1414213}{1000000}$	$5.624 \cdot 10^{-7}$
	10^7	$\frac{14142135}{10000000}$	$6.237 \cdot 10^{-8}$

Investigation 7.20. Choose an irrational number α .

- Determine the best rational approximation of the first kind for α of height at most 10^n for $n = 1, 2, 3, \dots$
- Determine a relationship between the height bound and the error.

There is a long and interesting history of results for Question 7.18. We first prove Dirichlet's³ theorem, which shows that there are many rational approximations whose error is at most the square of the denominator.

Theorem 7.21 (Dirichlet, 1842). *Let α be irrational. Then there are infinitely many rational numbers $\frac{a}{b}$ with $\gcd(a, b) = 1$ and*

$$\left| \alpha - \frac{a}{b} \right| < \frac{1}{b^2}.$$

The proof of Dirichlet's theorem relies on the *pigeonhole principle*, which is used in proofs throughout number theory. In essence, the principle is that if you have more objects than categories, then two objects must be in the same category.

³Johann Peter Gustav Lejeune Dirichlet (1805–1859) was a German mathematician.

For example, if you have ten boxes and eleven marbles, then no matter how you distribute the marbles, at least one box will have at least two marbles. Or, if you have a paragraph with at least 27 words, at least two of them must start with the same letter. Or, in a town of at least $26^3 + 1 = 17577$ people, at least two of them will have the same first, middle, and last initials.

Recall that $\lfloor x \rfloor$ means the largest integer less than or equal to x so that $\lfloor 1.2 \rfloor = 1$. Denote $\{x\} = x - \lfloor x \rfloor$ so that $\{1.2\} = 0.2$. In particular, $0 \leq \{x\} < 1$ is the decimal part of (x) .

Proof of Theorem 7.21. For any positive integer N , divide the interval $[0, 1)$ into subintervals of length $\frac{1}{N}$. Now consider the $N + 1$ numbers

$$\{\{n\alpha\} : n = 0, 1, \dots, N\}.$$

By the pigeonhole principle, at least two of these numbers lie in the same subinterval. In particular, for some integers $0 \leq m, n \leq N$, we have

$$|\{m\alpha\} - \{n\alpha\}| < \frac{1}{N}.$$

Thus, there are integers p and q with $q = |m - n| \leq N$ such that

$$|\{m\alpha\} - \{n\alpha\}| = |q\alpha - p| < \frac{1}{N}.$$

Rearranging, we have

$$\left| \alpha - \frac{p}{q} \right| < \frac{1}{qN} \leq \frac{1}{q^2},$$

where the last inequality comes from the fact that $q \leq N$.

Since we can find integers p and q for any choice of N , there are infinitely many such approximations. \square

Dirichlet's theorem is true for any irrational number. However, there are some irrational numbers which have much better approximations. Liouville proved that if the number is algebraic, there cannot be too many good approximations.

Theorem 7.22 (Liouville, 1844). *Let α be an algebraic number of degree $d \geq 1$. Then, for any integer $n > d$, there are only finitely many rational numbers $\frac{a}{b}$ with $\gcd(a, b) = 1$ such that*

$$\left| \alpha - \frac{a}{b} \right| < \frac{1}{b^n}.$$

Proof. Fix a positive integer $n > d$. Assume that $\left| \alpha - \frac{a}{b} \right| < \frac{1}{b^n} \leq 1$.

Let $f(x) = a_dx^d + \dots + a_1x + a_0$ be the characteristic polynomial of α . In particular, $f(x)$ is an irreducible polynomial with integer coefficients and $f(\alpha) = 0$. Write $f(x) = (x - \alpha)g(x)$ for some polynomial $g(x)$. Additionally, since $f(x)$ is irreducible, we have $f\left(\frac{a}{b}\right) \neq 0$ for $\frac{a}{b} \neq \alpha$; otherwise, $bx - a$ would divide $f(x)$.

For $\frac{a}{b} \neq \alpha$, we compute

$$\left| \left(\frac{a}{b} - \alpha \right) g\left(\frac{a}{b} \right) \right| = \left| f\left(\frac{a}{b} \right) \right| = \frac{|a_da^d + \dots + a_1ab^{d-1} + a_0b^d|}{b^d} \geq \frac{1}{b^d}$$

so that

$$\left| \alpha - \frac{a}{b} \right| \geq \frac{1}{b^d |g(\frac{a}{b})|}.$$

By our initial assumption, $\frac{a}{b}$ is contained in the closed interval $[\alpha - 1, \alpha + 1]$. Since $g(x)$ is a polynomial, it is continuous and, hence, bounded on a closed interval. Let $C > 0$ be a constant such that $|g(x)| \leq C$ for all $x \in [\alpha - 1, \alpha + 1]$. Note that C depends only on α and not on $\frac{a}{b}$. We have

$$\left| \alpha - \frac{a}{b} \right| \geq \frac{1}{Cb^d}.$$

Combined with our initial assumption on $\frac{a}{b}$, we have

$$\frac{1}{Cb^d} \leq \left| \alpha - \frac{a}{b} \right| < \frac{1}{b^n}$$

so that

$$\frac{1}{C} < \frac{1}{b^{n-d}}.$$

In particular,

$$(22) \quad b^{n-d} < C.$$

Since C is a fixed constant and $n > d$, there are only finitely many integers b that satisfy equation (22). For each such b , there are only finitely many choices of a so that

$$\frac{a}{b} \in \left(\alpha - \frac{1}{b^n}, \alpha + \frac{1}{b^n} \right).$$

Hence, there are only finitely many rational numbers $\frac{a}{b}$ that approximate α to within $\frac{1}{b^n}$. \square

Liouville used his theorem to give the first proof that a number is transcendental. Consider the Liouville number

$$\beta = \sum_{k=1}^{\infty} 10^{-(k!)}.$$

As a decimal, the Liouville number has a 0 in every decimal place except

$$\{1!, 2!, 3!, 4!, 5!, \dots\},$$

where there is a 1,

$$\beta = 0.110001000000000000000000100 \dots$$

Corollary 7.23. *The Liouville number β is transcendental.*

Proof. We will show that for every positive integer n , there exists a rational number $\frac{a_n}{b_n}$ such that $\left| \beta - \frac{a_n}{b_n} \right| < \frac{1}{b_n^n}$.

Fix $n \in \mathbb{N}$. Define

$$b_n = 10^{n!} \quad \text{and} \quad a_n = b_n \sum_{k=1}^n \frac{1}{10^{k!}}.$$

Then we compute

$$0 < \left| \beta - \frac{a}{b} \right| = \sum_{k=n+1}^{\infty} \frac{1}{10^k!} \leq \frac{2}{10^{(n+1)!}} \leq \frac{2}{10 \cdot 10^{(n!)}} < \frac{1}{10^{n!}}.$$

To finish the proof, assume that β is algebraic of degree d . For any fixed $n > d$, the infinitely many rational numbers in the set

$$\left\{ \frac{a_m}{b_m} : m \geq n \right\}$$

all satisfy

$$\left| \beta - \frac{a_m}{b_m} \right| < \frac{1}{b_m^n} \leq \frac{1}{b_m^n}.$$

This contradicts Theorem 7.22, so β must be transcendental. \square

Over the following 100 years, mathematicians gradually lowered the exponent in Liouville's theorem until finally Klaus Roth⁴ proved that *any* increase to the exponent of 2 in Dirichlet's theorem leads to only finitely many approximations of an algebraic number. This theorem is one of the results that led to his receiving the Fields Medal⁵ in 1958. Its proof is well beyond our scope.

Theorem 7.24 (Roth's theorem, 1955). *Suppose $\epsilon > 0$ and α is irrational and algebraic. Then there are only finitely many rational numbers $\frac{a}{b}$ with $\gcd(a, b) = 1$, such that*

$$\left| \alpha - \frac{a}{b} \right| < \frac{1}{b^{2+\epsilon}}.$$

We now have a good idea of how well an irrational number can be approximated by a rational number, but we have not addressed the problem of finding the best possible approximation.

Question 7.25. Given an irrational number α and a bound B , how do you find the best rational approximation of α with height at most B ?

Of course, you can simply try every number with height at most B , but that is very inefficient.

4. Continued Fractions

The notion of continued fractions provides one way of constructing good approximations of a given irrational number.

⁴Klaus Friedrich Roth (1925–2015) was a British mathematician.

⁵The Fields Medal is one of the most prestigious awards in mathematics and is awarded only every four years. Several medals are awarded each time.

Definition 7.26. A (*simple*) *continued fraction* is an expression of the form

$$a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \cdots}}},$$

where a_i are integers and $a_i \geq 1$ for $i > 0$. A continued fraction is often written in the shorthand notation

$$[a_0; a_1, a_2, a_3, \dots].$$

Example 7.27. Consider the continued fraction

$$[1; 2, 3, 2] = 1 + \frac{1}{2 + \frac{1}{3 + \frac{1}{2}}}.$$

We can simplify the expression to get a simple fraction:

$$\begin{aligned} 1 + \frac{1}{2 + \frac{1}{3 + \frac{1}{2}}} &= 1 + \frac{1}{2 + \frac{1}{\frac{7}{2}}} = 1 + \frac{1}{2 + \frac{2}{7}} \\ &= 1 + \frac{1}{\frac{16}{7}} = 1 + \frac{7}{16} = \frac{23}{16}. \end{aligned}$$

As with sequences from calculus, a continued fraction $[a_0; a_1, a_2, \dots]$ may have infinitely many terms. It is not too hard to believe that every rational number has a continued fraction expansion, but what about irrational numbers?

Question 7.28. What real numbers α have a continued fraction expansion?

Let's start with the case of rational numbers. To find the continued fraction expansion of a rational number $\frac{a}{b}$, we repeat the process of splitting off the integer part and making the numerator 1. We will use the fact that

$$\frac{a}{b} = \frac{1}{\frac{b}{a}}.$$

Example 7.29. Consider the rational number $\frac{75}{33}$. We compute

$$\frac{75}{33} = 2 + \frac{9}{33} = 2 + \frac{1}{\frac{33}{9}}.$$

Now we continue with $\frac{33}{9}$ to have

$$\frac{33}{9} = 3 + \frac{6}{9} = 3 + \frac{2}{3} = 3 + \frac{1}{\frac{3}{2}}.$$

In particular, we have

$$\frac{75}{33} = 2 + \frac{1}{\frac{33}{9}} = 2 + \frac{1}{3 + \frac{1}{\frac{3}{2}}}.$$

Continuing now with $\frac{3}{2}$, we have

$$\frac{3}{2} = 1 + \frac{1}{2},$$

so that

$$\frac{75}{33} = 2 + \frac{1}{3 + \frac{1}{1 + \frac{1}{2}}}.$$

The right-hand side is in continued fraction form, so we stop here and would write

$$\frac{75}{33} = [2; 3, 1, 2].$$

Perhaps you noticed that at each step we are taking the quotient and remainder, and the remainder from the previous step becomes the dividend of the next. In particular, the continued fraction expansion of a rational number is exactly the same as the set of coefficients in the Euclidean algorithm (Theorem 1.29).

Example 7.30. Applying the Euclidean algorithm to $(75, 33)$, we get

$$\begin{aligned} 75 &= \boxed{2} \cdot 33 + 9, \\ 33 &= \boxed{3} \cdot 9 + 6, \\ 9 &= \boxed{1} \cdot 6 + 3, \\ 6 &= \boxed{2} \cdot 3 + 0. \end{aligned}$$

The terms in the boxes are the parts of the continued fraction

$$\frac{75}{33} = [2; 3, 1, 2].$$

What we need to consider is how the process in Example 7.29 would work if we started with an irrational number. At each step we split off the integer part and make the numerator 1 using a reciprocal. The process would still work for irrational as well as rational numbers since

$$\alpha = \frac{1}{1/\alpha}.$$

Recall that $\lfloor x \rfloor$, the floor function, is the largest integer less than or equal to x and that $\{x\}$ denotes the decimal part of x .

Example 7.31. Consider $\alpha = \sqrt{2} = 1.4142135 \dots$.

Then we have

$$\sqrt{2} = \lfloor \sqrt{2} \rfloor + \{\sqrt{2}\} = 1 + 0.4142135 \dots = 1 + \frac{1}{\frac{1}{0.4142135 \dots}} = 1 + \frac{1}{2.4142135 \dots}.$$

Continuing, we have

$$2.4142135\ldots = 2 + 0.4142135\ldots = 2 + \frac{1}{\frac{1}{0.4142135\ldots}} = 2 + \frac{1}{2.4142135\ldots}$$

so that

$$\sqrt{2} = 1 + \frac{1}{2 + \frac{1}{2.4142135\ldots}}.$$

Notice that the remaining denominator is the same as the previous denominator. Consequently, the continued fraction repeats and the continued fraction expansion of $\sqrt{2}$ is

$$\sqrt{2} = [1; 2, 2, 2, \ldots] = 1 + \frac{1}{2 + \frac{1}{2 + \frac{1}{2 + \cdots}}}.$$

In particular, the continued fraction expansion is infinite. As with decimals, we denote the repeating portion with a bar:

$$\sqrt{2} = [1; \overline{2}].$$

However, it may not be the case that we end up with a repeating continued fraction as in Example 7.31.

Example 7.32. Let $\alpha = \pi = 3.141592653589793238\ldots$. We compute

$$\pi = 3 + 0.141592653589793238 = 3 + \frac{1}{7.06251330593104579},$$

$$7.06251330593104579 = 7 + 0.06251330593104579 = 7 + \frac{1}{15.9965944066857150},$$

$$15.9965944066857150 = 15 + 0.9965944066857150 = 15 + \frac{1}{1.00341723101337755},$$

$$1.00341723101337755 = 1 + 0.00341723101337755 = \frac{1}{292.634591013971913}.$$

So we have that the continued fraction expansion of π starts as

$$\pi = 3 + \frac{1}{7 + \frac{1}{15 + \frac{1}{1 + \frac{1}{292.634591013971913}}}}.$$

Written in sequence form, this is $\pi = [3; 7, 15, 1, 292, \ldots]$.

Not only have we answered Question 7.28 by showing that every real number has a continued fraction, but we have also given an explicit algorithm to compute the continued fraction expansion. This algorithm is described in Algorithm 7.1.

Having seen examples of numbers with both finite and infinite continued fraction expansions, we consider the following question.

Question 7.33. Which numbers have finite continued fraction expansions?

Algorithm 7.1. Continued Fraction Expansion

Input: a real number b , a positive integer N **Output:** a simple continued fraction $[a_0; a_1, a_2, \dots, a_N]$ **Algorithm:**

- 1: Set $a_0 = \lfloor b \rfloor$.
 - 2: Compute $b = \lfloor b \rfloor + \frac{1}{b_1}$.
 - 3: Set $k = 1$.
 - 4: Repeat until b_k is an integer or $k \geq N$.
 - a: Set $a_k = \lfloor b_k \rfloor$.
 - b: Compute $k = k + 1$.
 - c: Set $b_k = \lfloor b_k \rfloor + \frac{1}{b_{k+1}}$.
 - 5: Set $a_k = b_k$.
 - 6: Return $[a_0; a_1, a_2, \dots, a_k]$.
-

Investigation 7.34.

- (a) Compute the continued fraction expansion for various rational and irrational numbers.
- (b) Do you notice any patterns developing for which expansions are finite versus infinite continued fractions?
- (c) For those that are infinite, do you notice any patterns developing for the repeating portion?

In one direction, Question 7.33 is actually quite easy. If you are given a finite continued fraction, then it must be a rational number. Recall in Example 7.27 we showed how a finite continued fraction can be represented as a simple fraction, that is, a rational number. The *contrapositive* statement then says that an irrational number must have an infinite continued fraction expansion. However, we have not ruled out the possibility that there is a rational number with an infinite continued fraction expansion. On the other hand, we were able to make a connection between continued fraction expansions of rational numbers and the Euclidean algorithm, so we can use what we know about the Euclidean algorithm.

Theorem 7.35. *Let α be a real number. Then α has a finite continued fraction expansion if and only if α is rational.*

Proof. A finite continued fraction can be written as a simple fraction, so it must be a rational number.

Conversely, assume that α is a rational number. We create a continued fraction expansion of α by applying the Euclidean algorithm to the numerator and denominator of the fraction. The Euclidean algorithm terminates in finitely many steps, so the continued fraction is finite. \square

Since finite continued fractions are rational, if we apply the continued fraction algorithm (Algorithm 7.1) to an irrational number but stop after completing only finitely many steps, then we have a rational approximation of the irrational number called a *convergent*. What we hope to show is that this “convergent” is a “good” approximation.

Definition 7.36. The k th convergent of $[a_0; a_1, a_2, \dots]$ is the rational number

$$C_k = [a_0; a_1, \dots, a_k].$$

Example 7.37. The fourth convergent of $\sqrt{2} = 1.41421356\dots$ is given by $[1; 2, 2, 2, 2]$, and we have

$$C_4 = [1; 2, 2, 2, 2] = \frac{41}{29} \approx 1.4137931,$$

which is accurate to two decimal places. If we take the sixth convergent, we have

$$C_6 = [1; 2, 2, 2, 2, 2, 2] = \frac{140}{99} \approx 1.414141,$$

which is accurate to three decimal places.

5. Approximating Irrational Numbers with Convergents

Question 7.38. Let α be an irrational number. How good an approximation is the k th convergent C_k to α ?

Investigation 7.39.

- Compute the first few convergents of π . Which convergent is needed for five decimals of accuracy? for ten decimals?
- Choose a different irrational and repeat (a).
- Are the same number of convergents needed for the same amount of accuracy for different irrationals?

The answer to Question 7.38 turns out to be “very good”, but we need a few more facts to get there. First we give another definition of “best approximation”.

Definition 7.40. A rational number $\frac{p}{q}$ is called a *best approximation of the second kind* to a real number α if $1 \leq b \leq q$ and $\frac{a}{b} \neq \frac{p}{q}$ implies

$$|b\alpha - a| > |q\alpha - p|.$$

In Theoretical Exercise 7.26 you will prove that a best approximation of the second kind is also a best approximation of the first kind (Definition 7.5).

Example 7.41. In Example 7.17 we gave the best approximations of the first kind for π . Here we show which are only of the first kind and which are also of the second kind.

Approx	Value	Kind
3	$\approx 3 \dots$	second
$\frac{19}{6}$	$\approx 3.1 \dots$	first
$\frac{22}{7}$	$\approx 3.14 \dots$	second
$\frac{267}{85}$	$\approx 3.141 \dots$	first
$\frac{333}{106}$	$\approx 3.1415 \dots$	second
$\frac{355}{113}$	$\approx 3.141592 \dots$	second

We will prove that best approximations of the second kind must be convergents of a continued fraction expansion of α and that all convergents are best approximations of the first kind. We start with a recursive formula for computing the numerator and denominator of the convergents.

Lemma 7.42. *Let $C_k = \frac{p_k}{q_k}$ be the rational expression of the k th convergent C_k in lowest terms. The following recursion with $p_1 = 1$, $p_0 = a_0$, $q_{-1} = 0$, and $q_0 = 1$ determines p_k and q_k in terms of the continued fraction expansion $[a_k]$:*

$$\begin{aligned} p_k &= a_k p_{k-1} + p_{k-2}, \\ q_k &= a_k q_{k-1} + q_{k-2}. \end{aligned}$$

Proof. We proceed by induction. For $k = 0$, we have

$$p_0 = a_0, \quad q_0 = 1.$$

For $k = 1$, the recursion gives

$$p_1 = a_0 a_1 + 1, \quad q_1 = a_1$$

so that

$$\frac{p_1}{q_1} = \frac{a_0 a_1 + 1}{a_1} = \frac{a_0 a_1}{a_1} + \frac{1}{a_1} = a_0 + \frac{1}{a_1} = C_1.$$

Thus, the base case of $k = 1$ is satisfied.

Now we assume the recursion is valid up to $k - 1$ and consider $[a_0; a_1, \dots, a_k]$. We can consider this as a continued fraction with $k - 1$ terms:

$$[a_0; a_1, \dots, a_{k-2}, a_{k-1}, a_k] = \left[a_0; a_1, \dots, a_{k-2}, a_{k-1} + \frac{1}{a_k} \right].$$

Then, by induction we have

$$\begin{aligned} C_k &= \frac{\left(a_{k-1} + \frac{1}{a_k}\right) p_{k-2} + p_{k-3}}{\left(a_{k-1} + \frac{1}{a_k}\right) q_{k-2} + q_{k-3}} \\ &= \frac{(a_{k-1} a_k + 1) p_{k-2} + a_k p_{k-3}}{(a_{k-1} a_k + 1) q_{k-2} + a_k q_{k-3}} \\ &= \frac{a_k (a_{k-1} p_{k-2} + p_{k-3}) + p_{k-2}}{a_k (a_{k-1} q_{k-2} + q_{k-3}) + q_{k-2}} \\ &= \frac{a_k p_{k-1} + p_{k-2}}{a_k q_{k-1} + q_{k-2}}, \end{aligned}$$

verifying the recursion. So, by induction, it is true for all k . \square

Now we show that the difference between successive convergents is getting smaller as k gets larger. In fact, we give an explicit value for the difference.

Theorem 7.43. *Let $C_k = \frac{p_k}{q_k}$ be the rational expression of the convergent C_k in lowest terms for the continued fraction expansion $[a_k]$.*

- (a) $C_k - C_{k-1} = \frac{(-1)^{k-1}}{q_k q_{k-1}}$
- (b) $C_k - C_{k-2} = \frac{(-1)^k a_k}{q_k q_{k-2}}$

Proof.

(a) The statement is

$$\frac{p_k}{q_k} - \frac{p_{k-1}}{q_{k-1}} = \frac{(-1)^{k-1}}{q_k q_{k-1}}.$$

Clearing denominators, this is equivalent to

$$(23) \quad p_k q_{k-1} - p_{k-1} q_k = (-1)^{k-1}.$$

We proceed by induction. For $k = 1$, we have

$$\begin{aligned} p_0 &= a_0 \quad \text{and} \quad q_0 = 1, \\ p_1 &= a_0 a_1 + 1 \quad \text{and} \quad q_1 = a_1. \end{aligned}$$

We compute

$$p_1 q_0 - p_0 q_1 = a_0 a_1 + 1 - a_0 a_1 = 1,$$

verifying equation (23), which is equivalent to the statement

$$C_1 - C_0 = \frac{1}{a_1} = \frac{1}{q_0 q_1}.$$

This proves the base case.

Now we assume the relation is true for $k - 1$ and (using Lemma 7.42) consider

$$\begin{aligned} p_k q_{k-1} - p_{k-1} q_k &= (a_k p_{k-1} + p_{k-2}) q_{k-1} - p_{k-1} (a_k q_{k-1} + q_{k-2}) \\ &= p_{k-2} q_{k-1} - p_{k-1} q_{k-2} \\ &= -(p_{k-1} q_{k-2} - p_{k-2} q_{k-1}) \\ &= -(-1)^{k-2} \\ &= (-1)^{k-1}, \end{aligned}$$

again verifying equation (23). By induction, the statement is true for all k .

(b) We verify this directly. We apply part (a) and Lemma 7.42 to see that

$$\begin{aligned} C_k - C_{k-2} &= (C_k - C_{k-1}) + (C_{k-1} - C_{k-2}) \\ &= \frac{(-1)^{k-1}}{q_k q_{k-1}} + \frac{(-1)^{k-2}}{q_{k-1} q_{k-2}} \\ &= (-1)^{k-1} \frac{q_{k-2} - q_k}{q_k q_{k-1} q_{k-2}} \\ &= (-1)^{k-1} \frac{-a_k q_{k-1}}{q_k q_{k-1} q_{k-2}} \\ &= (-1)^k \frac{a_k}{q_k q_{k-2}}. \end{aligned}$$

□

As a consequence, we have that the odd-numbered convergents are decreasing and the even-numbered convergents are increasing and that the convergents do, in fact, converge to the real number.

Corollary 7.44. *Let α be a real number, and let $\{C_k\}$ be its sequence of convergents. The sequence satisfies*

$$C_1 > C_3 > C_5 > \cdots,$$

$$C_0 < C_2 < C_4 < \cdots,$$

and

$$C_{2k} \leq \alpha \leq C_{2k+1} \quad \text{for all } k \geq 0.$$

Moreover,

$$\lim_{k \rightarrow \infty} C_k = \alpha.$$

Proof. We consider differences

$$(24) \quad C_k - C_{k-2}.$$

Recall that a_k is always positive. By Theorem 7.43, if k is odd, the difference in equation (24) is negative; if k is even, the difference in equation (24) is positive.

We now prove convergence. The sequence of odd convergents is an increasing sequence bounded above by α and so converges. The sequence of even convergents is a decreasing sequence bounded below by α and so converges. By Theorem 7.43, they must converge to the same value, since the difference $C_k - C_{k-1}$ goes to 0 as $k \rightarrow \infty$. Consequently, we must have that $\{C_k\}$ converges to α . \square

Remark. The sequence of convergents looks a lot like the partial sums of an alternating series. In particular, the convergents are oscillating around the limiting value with the even terms getting closer from below and the odd terms getting closer from above.

Example 7.45. Consider the convergents of $\sqrt{2} = 1.4142135 \dots$.

Even conv.	in \mathbb{Q}	as decimal	Odd conv.	in \mathbb{Q}	as decimal
C_0	2	2	C_1	$\frac{3}{2}$	1.5
C_2	$\frac{4}{3}$	$1.\overline{3}$	C_3	$\frac{7}{5}$	1.4
C_4	$\frac{17}{12}$	$1.41\overline{6}$	C_5	$\frac{24}{17}$	$1.411764 \dots$
C_6	$\frac{41}{29}$	$1.41379 \dots$	C_7	$\frac{99}{70}$	$1.414285 \dots$
C_8	$\frac{140}{99}$	$1.4141\overline{41}$	C_9	$\frac{239}{169}$	$1.414201 \dots$

Now we are ready to prove that the convergents of continued fractions are “good” rational approximations of irrational numbers.

Corollary 7.46. *Let α be a real number. Let $C_k = \frac{p_k}{q_k}$ be a convergent of the continued fraction expansion of α . Then,*

$$|\alpha - C_k| < \frac{1}{q_k^2}.$$

Proof. We have

$$|C_k - \alpha| \leq |C_k - C_{k+1}| = \frac{1}{q_k q_{k+1}} < \frac{1}{q_k^2}.$$

The last inequality is strictly “less than” since the sequence of denominators q_k is strictly increasing for $k \geq 1$ by the recursion in Lemma 7.42. \square

Dirichlet's theorem (Theorem 7.21) proves that there are infinitely many approximations with error at most the square of the denominator. Convergents give a specific (infinite) sequence of such numbers. While Roth's theorem (Theorem 7.24) says we cannot in general do better than exponent 2, we might hope to improve the bound with a constant such as

$$\left| \alpha - \frac{p}{q} \right| \leq \frac{1}{cq_k^2}$$

for some constant c .

We will need the following relationship between the arithmetic mean and the geometric mean of two positive real numbers.

Lemma 7.47 (Arithmetic Mean–Geometric Mean). *Let x and y be real numbers. Then*

$$\sqrt{xy} \leq \frac{x+y}{2}$$

with equality if and only if $x = y$.

Proof. We compute

$$\begin{aligned} \frac{x+y}{2} - \sqrt{xy} &= \frac{x - 2\sqrt{xy} + y}{2} \\ &= \frac{(\sqrt{x} - \sqrt{y})^2}{2} \geq 0. \end{aligned} \quad \square$$

Theorem 7.48. *Let α be a real number. Let $C_k = \frac{p_k}{q_k}$ be the rational expression of the convergent C_k in lowest terms for the continued fraction expansion of α . For every $k \geq 1$, at least one of $\frac{p}{q} = C_k$ or $\frac{p}{q} = C_{k+1}$ satisfies*

$$\left| \alpha - \frac{p}{q} \right| < \frac{1}{2q^2}.$$

Proof. From Theorem 7.43 we know that

$$|C_k - C_{k+1}| = \frac{1}{q_k q_{k+1}} \text{ for all } k.$$

Apply Lemma 7.47 to $\frac{1}{q_k^2}$ and $\frac{1}{q_{k+1}^2}$ to have

$$\frac{1}{q_k q_{k+1}} = \sqrt{\frac{1}{q_k^2 q_{k+1}^2}} < \frac{1}{2} \left(\frac{1}{q_k^2} + \frac{1}{q_{k+1}^2} \right) = \frac{1}{2q_k^2} + \frac{1}{2q_{k+1}^2},$$

where we use the strict form of Lemma 7.47 since $\frac{1}{q_k^2} \neq \frac{1}{q_{k+1}^2}$.

In particular, since α is between C_k and C_{k+1} , for every $k \geq 1$

$$(25) \quad |\alpha - C_k| + |\alpha - C_{k+1}| = |C_k - C_{k+1}| < \frac{1}{2q_k^2} + \frac{1}{2q_{k+1}^2}.$$

This implies

$$|\alpha - C_k| < \frac{1}{2q_k^2} \quad \text{or} \quad |\alpha - C_{k+1}| < \frac{1}{2q_{k+1}^2}$$

because if they were both greater than the corresponding term, the inequality (25) would not be true. \square

Example 7.49. We examine the first few convergents of π in comparison to Theorem 7.48.

$C_k = \frac{p_k}{q_k}$	$ \pi - C_k < \frac{1}{2q_k^2}$
3	True
$\frac{22}{7}$	True
$\frac{333}{106}$	False
$\frac{355}{113}$	True
$\frac{103993}{33102}$	False
$\frac{104348}{33215}$	True
$\frac{208341}{66317}$	False
$\frac{312689}{99532}$	True
$\frac{833719}{265381}$	False
$\frac{1146408}{364913}$	True
$\frac{4272943}{1360120}$	False

Investigation 7.50. Consider the inequality

$$\left| \alpha - \frac{p}{q} \right| < \frac{1}{cq^k}.$$

Dirichlet's theorem says for $c = 1$, we can find infinitely many solutions $\frac{p}{q}$. Theorem 7.48 says for $c = 2$, we can still find infinitely many solutions.

Find the largest possible c where there are still infinitely many solutions. (*Hint:* Consider the continued fraction expansion of $\frac{1+\sqrt{5}}{2}$.)

We are now ready to prove that the convergents are best approximations. To prove this statement, we need a standard result from analysis called the Triangle Inequality.

Lemma 7.51 (Triangle Inequality). For any real numbers a and b ,

$$|a + b| \leq |a| + |b|.$$

Proof. It suffices to show that $|a + b|^2 \leq (|a| + |b|)^2$. We compute

$$\begin{aligned}
 |a + b|^2 &= (a + b)(a + b) \\
 &= a^2 + b^2 + 2ab \\
 &= |a|^2 + |b|^2 + 2ab \\
 &\leq |a|^2 + |b|^2 + 2|a||b| \\
 &= (|a| + |b|)^2.
 \end{aligned}$$

□

Theorem 7.52. Let α be a real number. The convergents of a continued fraction expansion of α are best approximations of the first kind for α .

Proof. Let α be an irrational number, and let $C_k = \frac{p_k}{q_k}$ be the k th convergent for the continued fraction expansion of α .

Suppose that $\frac{p}{q}$ is a better approximation than $\frac{p_k}{q_k}$, i.e.,

$$\left| \alpha - \frac{p}{q} \right| < \left| \alpha - \frac{p_k}{q_k} \right| \quad \text{and} \quad q \leq q_k.$$

Then (using the Triangle Inequality)

$$\begin{aligned} \left| \frac{p}{q} - \frac{p_k}{q_k} \right| &= \left| \left(\frac{p}{q} - \alpha \right) + \left(\alpha - \frac{p_k}{q_k} \right) \right| \\ &\leq \left| \frac{p}{q} - \alpha \right| + \left| \alpha - \frac{p_k}{q_k} \right| \\ &< 2 \left| \alpha - \frac{p_k}{q_k} \right| \quad (\text{since } \frac{p}{q} \text{ is a better approximation}) \\ &\leq 2 \frac{1}{2q_k^2} = \frac{1}{q_k^2} \quad (\text{Theorem 7.48}). \end{aligned}$$

On the other hand, since $\frac{p}{q} \neq \frac{p_k}{q_k}$ so that $|pq_k - qp_k|$ is a positive integer, we have

$$\left| \frac{p}{q} - \frac{p_k}{q_k} \right| = \left| \frac{pq_k - qp_k}{qq_k} \right| \geq \frac{1}{qq_k} \geq \frac{1}{q_k^2}.$$

Since we cannot have both

$$\left| \frac{p}{q} - \frac{p_k}{q_k} \right| < \frac{1}{q_k^2} \quad \text{and} \quad \left| \frac{p}{q} - \frac{p_k}{q_k} \right| \geq \frac{1}{q_k^2},$$

this is a contradiction, so $\frac{p}{q}$ cannot exist. □

Example 7.53. Consider $\sqrt{2} \approx 1.4142 \dots$. We have already seen that

$$\sqrt{2} = [1; 2, 2, 2, \dots].$$

We compute

$$C_4 = [1; 2, 2, 2, 2] = \frac{41}{29} \approx 1.4138.$$

The next best approximation of $\sqrt{2}$ with denominator at most 29 is

$$\frac{24}{17} \approx 1.4118,$$

which is a worse approximation than C_4 .

While each convergent is a best approximation of the first kind, it is not true that every best approximation of the first kind is a convergent.

Example 7.54. We computed the first few best approximations of the first kind for π in Example 7.17. They are the following.

$$\begin{aligned} 3 &\approx 3 \cdots, \\ \frac{19}{6} &\approx 3.1 \cdots, \\ \frac{22}{7} &\approx 3.14 \cdots, \\ \frac{267}{85} &\approx 3.141 \cdots, \\ \frac{333}{106} &\approx 3.1415 \cdots, \\ \frac{355}{113} &\approx 3.141592 \cdots. \end{aligned}$$

Of these, $\frac{19}{6}$ and $\frac{267}{85}$ are not convergents!

However, it is true the every best approximation of the second kind is a convergent. In other words, if a rational number approximates an irrational number well enough, then it must be the convergent of a continued fraction.

Theorem 7.55. *If $\frac{p}{q}$ is a best approximation of the second kind to a real number α , then $\frac{p}{q}$ is a convergent of the continued fraction expansion of α .*

Proof. Assume that $\frac{p}{q}$ is not a convergent of α . Let $\frac{p_k}{q_k}$ be the convergents of $\alpha = [a_0; a_1, a_2, \dots]$. First we show that

$$\frac{p_0}{q_0} < \frac{p}{q} < \frac{p_1}{q_1}.$$

If $\frac{p}{q} < \frac{p_0}{q_0} = a_0$, then

$$|q\alpha - p| = q \left| \alpha - \frac{p}{q} \right| \geq \left| \alpha - \frac{p}{q} \right| > |\alpha - a_0|,$$

which contradicts that $\frac{p}{q}$ is a best approximation of the second kind.

Since $\alpha < \frac{p_1}{q_1}$, if $\frac{p}{q} > \frac{p_1}{q_1}$, then

$$|q\alpha - p| = q \left| \alpha - \frac{p}{q} \right| > q \left| \frac{p_1}{q_1} - \frac{p}{q} \right| \geq q \frac{1}{q_1 q} = \frac{1}{q_1} = \frac{1}{a_1}.$$

Since $|\alpha - a_0| < \frac{1}{a_1}$, this inequality is a contradiction to $\frac{p}{q}$ being a best approximation of the second kind.

So we can assume that $\frac{p}{q}$ lies between two convergents. In particular, $C_k < \frac{p}{q} < C_{k+2} < \alpha$ or $\alpha < C_{k+2} < \frac{p}{q} < C_k$. In either case,

$$\frac{1}{q_k q_{k+1}} = \left| \frac{p_{k+1}}{q_{k+1}} - \frac{p_k}{q_k} \right| > \left| \frac{p}{q} - \frac{p_k}{q_k} \right| \geq \frac{1}{q_k q}$$

so that $q > q_{k+1}$. Also

$$|q\alpha - p| = q \left| \alpha - \frac{p}{q} \right| > q \left| \frac{p_{k+2}}{q_{k+2}} - \frac{p}{q} \right| \geq q \frac{1}{q_{k+2} q} = \frac{1}{q_{k+2}}.$$

However,

$$|q_{k+1}\alpha - p_{k+1}| = q_{k+1} \left| \alpha - \frac{p_{k+1}}{q_{k+1}} \right| \leq q_{k+1} \frac{1}{q_{k+1}q_{k+2}} = \frac{1}{q_{k+2}}.$$

This contradicts that $\frac{p}{q}$ is a best approximation of the second kind. Hence, $\frac{p}{q}$ must be a convergent of α . \square

Corollary 7.56. *If*

$$\left| \frac{p}{q} - \alpha \right| < \frac{1}{2q^2},$$

the rational number $\frac{p}{q}$ is a convergent of the continued fraction expansion of α .

Proof. By Theorem 7.55, we just need to show that $\frac{p}{q}$ is a best approximation of the second kind.

Assume that

$$\left| \frac{p}{q} - \alpha \right| < \frac{1}{2q^2}$$

and, for some $\frac{a}{b} \neq \frac{p}{q}$, that

$$|b\alpha - a| \leq |q\alpha - p|.$$

We need to prove that $q < b$.

We have

$$\left| \alpha - \frac{a}{b} \right| = \frac{1}{b} |b\alpha - a| \leq \frac{1}{b} |q\alpha - p| < \frac{1}{2bq}.$$

Furthermore,

$$\left| \frac{aq - bp}{bq} \right| = \left| \frac{a}{b} - \frac{p}{q} \right| \leq \left| \frac{a}{b} - \alpha \right| + \left| \frac{p}{q} - \alpha \right| < \frac{1}{2bq} + \frac{1}{2q^2}.$$

Since $\frac{a}{b} \neq \frac{p}{q}$, $|aq - bp|$ is a positive integer so that

$$\frac{1}{bq} < \frac{1}{2bq} + \frac{1}{2q^2} = \frac{q + b}{2bq^2}.$$

Consequently,

$$1 < \frac{q + b}{2q}$$

and

$$q < b.$$

Hence, $\frac{p}{q}$ is a best approximation of the second kind and, thus, a convergent. \square

We have now answered the problem we set out to solve: We have a simple algorithm that computes best possible approximations to an irrational number (Algorithm 7.2).

While convergents offer the best possible rational approximations, there is still the question of how many convergents we must compute to get an approximation within a certain error bound (Theoretical Exercise 7.27). Unfortunately, an optimal answer depends on the irrational number we are trying to approximate. The following example demonstrates the worst-case scenario.

Algorithm 7.2. Best Rational Approximation**Input:** an irrational number α , a positive real number ϵ **Output:** a rational number that is the best approximation of the second kind for α with error at most ϵ .**Algorithm:****1:** Set $q_1 = 1$.**2:** Repeat until $2q_k^2 > \frac{1}{\epsilon}$.**a:** Compute p_k, q_k using Lemma 7.42.**3:** If $\left| \alpha - \frac{p_k}{q_k} \right| < \epsilon$ return p_k/q_k . Else return p_{k+1}/q_{k+1} .**Example 7.57.** Consider the irrational number

$$\phi = \frac{1 + \sqrt{5}}{2} \approx 1.618033988 \dots,$$

also known as the *golden ratio*. We compute its continued fraction expansion as

$$\phi = [1; \overline{1}] = [1; 1, 1, 1, \dots].$$

We can compute the associated convergents as shown below.

Convergent	as rational	as decimal
C_1	2	2
C_2	$\frac{3}{2}$	1.5
C_3	$\frac{5}{3}$	$1.6\overline{6}$
C_4	$\frac{8}{5}$	1.6
C_5	$\frac{13}{8}$	1.625
C_6	$\frac{21}{13}$	1.61538...
C_7	$\frac{34}{21}$	1.61904...
C_8	$\frac{55}{34}$	1.61764...
C_9	$\frac{89}{55}$	$1.618\overline{18}$
C_{10}	$\frac{144}{89}$	1.617977...
C_{11}	$\frac{233}{144}$	$1.61805\overline{5}$
C_{12}	$\frac{377}{233}$	1.618025...
C_{13}	$\frac{610}{377}$	1.618037...
C_{14}	$\frac{987}{610}$	1.6180327...
C_{15}	$\frac{1597}{987}$	1.6180344...
C_{16}	$\frac{2584}{1597}$	1.6180338...
C_{17}	$\frac{4181}{2584}$	1.6180340...
C_{18}	$\frac{6765}{4181}$	1.61803396...

Notice that C_{17} has only five decimals correct and C_{18} has only seven decimals correct. By comparison, the third convergent of π , $C_3 = \frac{355}{113}$, already has six decimals correct.

The reason it is so difficult to approximate ϕ with a rational number is that the continued fraction expansion contains only 1's. This causes the q_i to grow as slowly as possible (see the recursion in Lemma 7.42). Since the error is bounded as

$$|C_k - \alpha| < |C_k - C_{k+1}| \leq \frac{1}{q_k q_{k+1}},$$

when q_{k+1} is large compared to q_k (when a_{k+1} is large), the error associated to C_k is small compared to q_k . Since every a_i is 1 for the continued fraction expansion of ϕ , the errors are as “bad” as possible. Consequently, ϕ is sometimes called the *most irrational* number.

You may have noticed that the numerators and denominators of the convergents for ϕ look familiar. They are the Fibonacci numbers because the recursion from Lemma 7.42 with $a_k = 1$ is the generating recursion for the Fibonacci numbers.

COMPUTATIONAL EXERCISES


7.1. Determine the value of $N(\mathbb{Q}, 100)$. In other words, count the number of rational numbers with height at most 100.

7.2.


- Let S be the set of rational points on the unit circle $x^2 + y^2 = 1$. Determine $N(S, 100)$.
- Let S be the set of rational points on $x^3 + y^3 = 1$. Determine $N(S, 100)$.

7.3.

- Let S be the set of rational points on the unit sphere $x^2 + y^2 + z^2 = 1$. Determine $N(S, 100)$.
- Let S be the set of rational points on $x^3 + y^3 + z^3 = 1$. Determine $N(S, 100)$.

 **7.4.** Find a (simple) continued fraction expansion for the following numbers.

- $\frac{68}{23}$
- $\frac{23}{5}$

 **7.5.** Write the continued fraction as a rational number $\frac{a}{b}$ in simplest form.

- $1 + \frac{1}{4 + \frac{1}{3 + \frac{1}{5}}}$
- $[0; 1, 4, 2, 1]$

7.6. Find the best rational approximation of the first kind with height at most 1000 of the following irrational numbers.

- e
- $\sqrt{7}$

7.7. Use convergents to find a rational approximation $\frac{a}{b}$ of the following numbers α such that $|\alpha - \frac{a}{b}| < 10^{-6}$.

- a. $\alpha = \sqrt{3}$
- b. $\alpha = \ln 5$

7.8. Let α be an irrational number. Roth's theorem says that for any constant $\epsilon > 0$ there are only finitely many rational numbers $\frac{a}{b}$ with $\gcd(a, b) = 1$ such that

$$\left| \alpha - \frac{a}{b} \right| < \frac{1}{b^{2+\epsilon}}.$$

Find them all for $\alpha = \sqrt{2}$ and $\epsilon = \frac{1}{10}$.

7.9.

- a. Find the first 20 terms of the continued fraction expansion for Hilbert's number $2^{\sqrt{2}}$.
- b. Compute the first 20 convergents of the continued fraction expansion of $2^{\sqrt{2}}$.

7.10. Consider the infinite continued fraction

$$[4; \overline{2, 1, 3, 1, 2, 8}].$$

- a. Compute the first ten convergents.
- b. What irrational number is this continued fraction?

7.11. Find an approximate rational value with height at most 10,000 for the Liouville number by taking convergents of the partial sum up to $k = 5$:

$$\sum_{k=1}^{\infty} 10^{-(k!)}.$$

THEORETICAL EXERCISES

7.12.

- a. Prove that for any two rational numbers s and t with $s < t$, there is an irrational number α such that $s < \alpha < t$.
- b. Prove that for any two irrational numbers α and β with $\alpha < \beta$, there is a rational number t such that $\alpha < t < \beta$.

7.13. Prove or find a counterexample for each of the following statements.

- a. The sum of two rational numbers is rational.
- b. The sum of two irrational numbers is irrational.
- c. The sum of an irrational number and a rational number is irrational.

7.14. Prove or find a counterexample for each of the following statements.

- a. The product of two rational numbers is rational.
- b. The product of two nonzero irrational numbers is irrational.
- c. The product of a nonzero irrational number and a nonzero rational number is irrational.

7.15.

- a. Prove that \sqrt{p} is irrational for every prime p .
- b. What breaks down in the proof of Theorem 7.3 if we replace 2 by 4?

7.16.

- a. Prove that $\sqrt[3]{p}$ is irrational for every prime p .
- b. Prove that $\sqrt[3]{m}$ is irrational for any m not a perfect cube (i.e., $m \neq n^3$ for some $n \in \mathbb{Z}$).

7.17. Prove that $\sqrt[k]{m}$ is irrational for any m not a perfect k th power (i.e., $m \neq n^k$ for some $n \in \mathbb{Z}$).

7.18. Prove that for any positive integer n and rational numbers a_1, \dots, a_n ,

- a. $H(a_1 \cdots a_n) \leq H(a_1)H(a_2) \cdots H(a_n)$.
- b. $H(a_1 + \cdots + a_n) \leq nH(a_1)H(a_2) \cdots H(a_n)$.

7.19. Prove that a rational number $\frac{a}{b}$ has only finitely many rational approximations $\frac{p}{q}$ which satisfy

$$\left| \frac{a}{b} - \frac{p}{q} \right| < \frac{1}{q^2}.$$

7.20. Consider the real number

$$\alpha = \sum_{k=0}^{\infty} 2^{-k!}.$$

Prove that α is transcendental.

7.21. Let p and q be positive integers. Prove that if $p > q$ and $\frac{p}{q} = [a_0; a_1, \dots, a_k]$, then $\frac{q}{p} = [0; a_0, a_1, \dots, a_k]$.

7.22. Given a continued fraction $[a_0; a_1, \dots]$ with convergents $C_n = \frac{p_n}{q_n}$, prove that for all $n \geq 1$

$$\frac{p_n}{p_{n-1}} = [a_n; a_{n-1}, \dots, a_0] \quad \text{and} \quad \frac{q_n}{q_{n-1}} = [a_n; a_{n-1}, \dots, a_1].$$

7.23.

- a. Prove that $[a_0; a_1, \dots, a_k]$ and $[a_0; a_1, \dots, a_k - 1, 1]$ represent the same continued fraction.
- b. Prove that any rational number has exactly two continued fraction expansions.

7.24. Let (s_k) be a sequence of real numbers satisfying the following properties:

Property 1: For any positive integer k , $s_{2k} \leq s_{2k+2}$ and $s_{2k+1} \leq s_{2k-1}$; that is, the even subsequence is increasing and the odd subsequence is decreasing.

Property 2: For any positive integers k and j , $s_{2k} \leq s_{2j+1}$; that is, every even index term is less than every odd index term.

Property 3: $\lim_{k \rightarrow \infty} |s_k - s_{k-1}| = 0$.

Prove that the sequence (s_k) converges to a limit s that satisfies

$$s_{2k} \leq s \leq s_{2k+1} \quad \text{for all } k \geq 0.$$

7.25. Let α be an irrational number. Draw the line $y = \alpha x$ in the plane \mathbb{R}^2 . Prove that if you follow the line from $(0, 0)$, every time you encounter a point (p, q) with integer coordinates closer to the line than any previous point with integer coordinates, this is the next convergent in the continued fraction expansion of α .

7.26. Let α be a real number. Prove that if $\frac{p}{q} \in \mathbb{Q}$ is a best approximation of the second kind for α , then $\frac{p}{q}$ is a best approximation of the first kind.

7.27. Fix a positive real number $\epsilon > 0$. Give a lower bound on the positive integer k such that for any real number α with convergents C_k , $|\alpha - C_k| < \epsilon$.

EXPLORATION EXERCISES

7.28 (Counting points of bounded height). One application of height functions is the counting function $N_S(B)$, which counts the number of points in the set S up to height B . This function is used to give a quantitative measure of the number of points in an infinite set.

- Let B be a positive integer. Determine an upper bound in terms of B on $N_{\mathbb{Q}}(B)$, the number of rational numbers up to height B .
- What real number does the following limit approach?

$$\lim_{B \rightarrow \infty} \frac{N_{\mathbb{Q}}(B)}{B^2}$$

Hint: It involves π .

- Let $S = \{(x, y) : x, y \in \mathbb{Q}, y^2 = x^2 + 1\}$. For what positive integer n does the following limit exist?

$$\lim_{B \rightarrow \infty} \frac{N_S(B)}{B^n}$$

What value does it approach?

- Let $S = \{(x, y) : x, y \in \mathbb{Q}, y^2 = x^3 + 2\}$. For what rational number n does the following limit exist?

$$\lim_{B \rightarrow \infty} \frac{N_S(B)}{B^n}$$

What value does it approach?

- As you vary the constant c for

$$S_c = \{(x, y) : x, y \in \mathbb{Q}, y^2 = x^3 + c\},$$

how does the necessary exponent for convergence for the following limit to exist vary?

$$\lim_{B \rightarrow \infty} \frac{N_{S_c}(B)}{B^n}$$

- What about counting points on other curves?

7.29 (Integer lattice points). We can think of a rational number $\frac{a}{b}$ as a pair of integers (a, b) . Consequently, we are interested in points with integer coordinates in the plane \mathbb{R}^2 . We call the set of points $\{(x, y) : x, y \in \mathbb{Z}\}$ *lattice points*. We want to determine how lattice points interact with geometric objects.

- How many lattice points are contained in a square of side length r centered at the origin?
- How many lattice points are contained in a circle of radius r centered at the origin?
- How big a circle (in terms of its area) is needed to contain n lattice points?
- For which n can you find a circle with *exactly* n lattice points on its boundary? The circle need not be centered at the origin.
- Consider circles of radius \sqrt{n} centered at the origin, that is, $x^2 + y^2 = n$. How many lattice points are on the boundary of these circles?
- Consider ellipses of the form $x^2 + dy^2 = n$ for integer d and positive integer n . How many lattice points are on the boundary of these ellipses?

7.30 (Simultaneous Diophantine approximation). Given a list of real numbers $(\alpha_1, \dots, \alpha_n)$, we want to find the best approximations of all the α_i by rational numbers with the same denominator. In particular, find integers (p_1, \dots, p_n, q) such that the vector $(p_1 - \alpha_1 q, p_2 - \alpha_2 q, \dots, p_n - \alpha_n q)$ is “short”. We say (a_1, \dots, a_n) is *short* if the vector norm $\|\cdot\|$ is small:

$$\|(a_1, \dots, a_n)\| = \sqrt{a_1^2 + \dots + a_n^2}.$$

- Given the numbers $(\sqrt{2}, \sqrt{3}, \sqrt{5})$, find the best simultaneous approximation with denominator $q \leq 100$.
- Given the numbers $(\sqrt{2}, \sqrt[3]{2}, \sqrt[4]{2})$, find the best simultaneous approximation with denominator $q \leq 100$.
- Fix an irrational number α and consider the n numbers $(1, \alpha, \alpha^2, \alpha^3, \dots, \alpha^{n-1})$. For example, take $(1, \sqrt[n]{2}, \sqrt[n]{2^2}, \dots, \sqrt[n]{2^{n-1}})$. Are these numbers particularly easy to approximate or particularly hard; i.e., do you need a larger denominator for given error bounds compared to other sets of numbers?
- Try to find a bound on the error in terms of the denominator, such as in Dirichlet’s theorem (Theorem 7.21).

7.31 (Periodic continued fractions).

Definition 7.58. We say that a continued fraction expansion is *periodic* if there exist integers N and k such that $a_{n+k} = a_n$ for all $n \geq N$. It is denoted as

$$[a_0; a_1, \dots, a_{N-1}, \overline{a_N, a_{N+1}, \dots, a_{N+k}}],$$

and k is called the *period* of the continued fraction.

For example,

$$[0; 7, 1, 2, 3, 1, 2, 3, 1, 2, 3, \dots] = [0; 7, \overline{1, 2, 3}]$$

has period 3.

- Which numbers have periodic continued fraction expansions?

Consider now the continued fraction of \sqrt{d} for d an integer that is not a perfect square.

- b. Prove that d can be written uniquely in the form $d = m^2 + j$ for some $j \in \{1, 2, \dots, 2m\}$.
- c. Which values of j give $k = 1$? What is the continued fraction expansion in this case?
- d. Which values of j give $k = 2$? What is the continued fraction expansion in this case?
- e. What about other values of j ?

You can also consider square roots of rational numbers. Let $\frac{a}{b}$ be a rational number with $\gcd(a, b) = 1$ that is not a perfect square.

- f. How does the value of a_{N+k} compare to the value of $\sqrt{\frac{a}{b}}$?
- g. Can you say anything about the sequence a_N, \dots, a_{N+k-1} ?
- h. Give an upper bound on k in terms of a and b .

7.32 (Solving a Quadratic Equation). Consider a quadratic equation

$$ax^2 + bx + c = 0$$

with $a, b, c \in \mathbb{Z}$ and $a \neq 0$. We want to approximate the solutions using continued fractions.

- a. Solve for x in terms of a , b , and c . (It is OK to have x on both sides of the equation.)
- b. Recursively apply your equation from part (a) by substituting the value of x into the x on the right-hand side of the equation. The result is a continued fraction.
- c. Solve for x in the other possible way to get the second solution.
- d. Choose some specific quadratic equations and approximate their solutions using this method.
- e. Write a function using continued fractions that finds the solutions of a quadratic equation to within a user-specified error bound.

Diophantine Equations

1. Introduction and Examples

Diophantus of Alexandria wrote a series of books called *Arithmetica* in which he studied solutions to algebraic equations and systems of algebraic equations. The study of such problems is now called Diophantine equations since, as far as we know, Diophantus was the first to systematically study them. We focus on integer solutions to polynomial equations. The following are some examples that are probably familiar to you.

- (a) The linear equation problem from Chapter 1. We fix integers a and b and ask for which integers d can we find integers (x, y) so that

$$ax + by = d.$$

- (b) Pythagorean triples: What are the relatively prime integer triples (x, y, z) such that

$$x^2 + y^2 = z^2?$$

And, more generally, in the Fermat equation from Fermat's Last Theorem, are there any nonzero integer triples (x, y, z) such that

$$x^n + y^n = z^n?$$

- (c) Pell's equation: Fix an integer d . What are the integer pairs (x, y) that satisfy

$$x^2 - dy^2 = \pm 1?$$

The following two examples demonstrate how to find the solutions to two specific Diophantine equations.

Example 8.1. Using modular arithmetic from Chapter 2, we know that if an equation has a solution, then it must have a solution modulo every prime. Consequently, if we can find a prime p for which there is no solution to an equation modulo p , then the equation has no integer solutions. Consider

$$x^3 + y^3 = 11.$$

Working modulo 7, we have to solve

$$x^3 + y^3 \equiv 4 \pmod{7}.$$

We are taking the sums of cubes, so we compute all the cubes modulo 7. They are $\{0, 1, 6\}$. Taking all nine possible choices for the pair (x^3, y^3) modulo 7,

$$\{(0, 0), (0, 1), (0, 6), (1, 0), (1, 1), (1, 6), (6, 0), (6, 1), (6, 6)\},$$

we see that none sum to 4 modulo 7. Since there are no solutions modulo 7, there are no integer solutions.

The method of the previous example is useful since it reduces the problem to a finite number of steps: checking each residue class. However, we do not know for which primes there is no solution or even if there will be such a prime. Consequently, a computer search is more effective than paper and pencil since a computer can quickly search through many different primes.

The following example is more difficult because there is a solution modulo every prime, $(0, 0, 0)$. The goal is to show that this solution is the only solution, which is not possible with the methods of Example 8.1.

Example 8.2 ([12, §6.1.2]). Consider the equation

$$y^2 = -2x^4 - 12x^2z^2 + 6z^4.$$

This equation clearly has the solution $(0, 0, 0)$, and the goal is to show that $(0, 0, 0)$ is the *only* solution. We solve this problem by arriving at a contradiction.

Proof. Let (x, y, z) be a solution other than $(0, 0, 0)$. If $d = \gcd(x, z)$, then, since the terms on the right-hand side are all degree 4, we have that d^4 divides the right-hand side. Then, d^4 must also divide the left-hand side y^2 so that d^2 divides y . Consequently, the integers $(x/d, y/d^2, z/d)$ must also be a solution. Hence, we can assume that x and z are relatively prime.

If one of x or z is 0, then the other must be 1. However, the two possibilities, $(x, z) = (0, 1)$ and $(x, z) = (1, 0)$, do not have a corresponding y that gives a solution. So we may also assume that x and z are nonzero.

Reducing both sides of the equation modulo 4, we have

$$y^2 \equiv 2x^4 + 2z^4 \pmod{4}.$$

The only solutions modulo 4 with $\gcd(x, z) = 1$ are $(x, z) \in \{(1, 1), (1, 3), (3, 1), (3, 3)\}$. Hence, both x and z must be odd.

Reducing both sides of the equation modulo 2, we have $y^2 \equiv 0 \pmod{2}$, so y must be even. Then there is some integer y_1 such that $y = 2y_1$, and we write

$$(2y_1)^2 = -2x^4 - 12x^2z^2 + 6z^4$$

or

$$2y_1^2 = -x^4 - 6x^2z^2 + 3z^4.$$

Since the left-hand side is even, then so is the right-hand side. Recall that x and z are both odd and the square of an odd number is 1 modulo 8. So we have

$$2y_1^2 \equiv -1 - 6 + 3 \equiv -4 \pmod{8}$$

and

$$y_1^2 \equiv -2 \pmod{4}.$$

There are no such y_1 , so there are no nonzero solutions to the original equation. \square

In Example 8.1, the choice of prime depended on the equation, but no other part of the method depended on the equation. In Example 8.2, the whole process was specific to that equation. Consequently, being able to solve one particular Diophantine equation does not mean that you will be able to solve another using the same method. The search for general methods drove much of the development of modern number theory, and there is now a wide range of methods to address general types of Diophantine equations. The ultimate goal was a solution to David Hilbert's 10th problem, which asked for a method to determine if any given Diophantine equation has a solution.

Question 8.3. Given a Diophantine equation with any number of unknown quantities and with rational coefficients, is there a process according to which it can be determined in a finite number of operations whether the equation is solvable in rational integers?

Today, it is known that no such general algorithm exists. This result is the combined work of Martin Davis,¹ Yuri Matiyasevich,² Hilary Putnam,³ and Julia Robinson.⁴ This does not mean that all hope is lost, just that the methods for solving one Diophantine equation may not suffice for another.

In this chapter we expand on methods modulo primes and then examine a few specific Diophantine equations and methods for solving them.

2. Working Modulo Primes

We now return to the method of Example 8.1 where we showed that 11 cannot be written as the sum of two integer cubes, i.e.,

$$x^3 + y^3 = 11$$

has no integer solutions. We were able to check modulo 7 that there were no solutions. However, it is unclear whether we were just lucky to choose an equation for which there is a small prime with no solution. We might hope that there are no solutions modulo most (or even some positive proportion of) primes.

Question 8.4. If a Diophantine equation has no integer solutions, for what proportion of primes does it not have a solution modulo p ?

¹Martin David Davis (1928–) is an American mathematician.

²Yuri Vladimirovich Matiyasevich, (1947–) is a Russian mathematician.

³Hilary Whitehall Putnam (1926–2016) was an American philosopher and mathematician.

⁴Julia Hall Bowman Robinson (1919–1985) was an American mathematician.

Let's generate some data for our example equation

$$x^3 + y^3 = 11.$$

A short computation finds that 7 is the only prime smaller than 1,000,000 for which the equation does not have a solution. This is somewhat troubling for our method because, if we wish to use this method to show there are no solutions, we need to be able to find a prime that has no solutions. If there are very few such primes, a search is probably infeasible, even by computer. However, as a quick check using a computer for small primes, this method is very fast and can quickly rule out the existence of solutions.

Unfortunately, there are equations with no integer solutions for which there are solutions modulo every prime.

Example 8.5. The following two equations have no nonzero integer solutions, but they have nonzero solutions modulo every prime p . The first was studied by Carl-Erik Lind⁵ and Hans Reichardt;⁶ the second, by Ernst Selmer.⁷

- $x^4 - 17y^4 = z^2$,
- $3x^3 + 4y^3 + 5z^3 = 0$.

Investigation 8.6.

- (a) Consider the following families of Diophantine equations. Determine the constants k , $|k| < 20$, and primes $p < 1000$ for which there are no solutions to the equation.
 - (1) $x^k = 2$
 - (2) $y^3 + x^3 = k$
 - (3) $y^2 = x^5 + k$
 - (4) $y^2 = x^6 + k$
 - (5) $y^3 = x^4 + k$
- (b) Can you find your own Diophantine equation that has no solutions modulo some prime p ?
- (c) Can you find your own Diophantine equation that has no solutions modulo more than one prime?

The equations in Example 8.5 are special in the sense that they violate the *Hasse principle*. The Hasse principle asserts that if there are solutions modulo all primes, then we expect there are integer solutions. The Hasse principle is named after Helmut Hasse⁸ who investigated using local information (information modulo primes) to determine global information (integer or rational solutions). However, just as there are equations that violate the Hasse principle, there are situations where the local information completely determines the global information. The following theorem is due to Gauss.

⁵Carl-Erik Lind (1922–1993) was a Swedish mathematician.

⁶Hans Reichardt (1908–1991) was a German mathematician.

⁷Ernst Sejersted Selmer (1920–2006) was a Norwegian mathematician.

⁸Helmut Hasse (1898–1979) was a German mathematician.

Theorem 8.7. *Fix an integer a . Then the equation*

$$x^2 = a$$

is solvable if and only if it is solvable modulo all primes p .

Theorem 8.7 says that a number is a perfect square if and only if it is a perfect square modulo all primes.

Proof. If a is a perfect square, then it is a perfect square modulo all primes. We have only to prove the reverse implication.

Assume that $x^2 \equiv a \pmod{p}$ is solvable for all primes p and that a is not a square. Write $a = k^2 N$ for integers k and N with N squarefree. In particular, a is a quadratic residue for the same primes as N . We need to show that $N = 1$.

Factor N as $N = p_1 \cdots p_n$ for distinct primes p_i . If all the p_i were even, there would be only one prime and $N = 2$. Then reducing modulo 3 we see that 2 is a nonresidue modulo 3 so that N , and a , are nonresidues modulo 3; a contradiction. Hence, we know at least one of the p_i is odd, call it p_1 . Let b be a nonresidue of p_1 and c_i a quadratic residue of p_i for $2 \leq i \leq n$. The congruences

$$\begin{aligned} x &\equiv 1 \pmod{4}, \\ x &\equiv b \pmod{p_1}, \\ x &\equiv \begin{cases} c_i \pmod{p_i} & p_i \neq 2, \\ 1 \pmod{8} & p_i = 2, \end{cases} \end{aligned}$$

always have a solution z modulo $4N$ by the Chinese Remainder Theorem (Theorem 2.44) since the moduli are relatively prime. Then the sequence of general solutions, $4Nj + z$ for $j \in \mathbb{N}$, is an arithmetic progression. Any arithmetic progression has infinitely many prime values by Dirichlet's theorem (Theorem 1.62). Let q be a prime value in the sequence of general solutions that does not divide k . Then we have

$$\begin{aligned} \left(\frac{b}{p_1}\right) &= \left(\frac{q}{p_1}\right) = \left(\frac{p_1}{q}\right) = -1 \quad \text{and} \\ \left(\frac{c_i}{p_i}\right) &= \left(\frac{q}{p_i}\right) = \left(\frac{p_i}{q}\right) = 1 \quad \text{for } 2 \leq i \leq n \end{aligned}$$

so that

$$\left(\frac{a}{q}\right) = \left(\frac{N}{q}\right) = \left(\frac{b}{q}\right) \prod_{i=2}^n \left(\frac{c_i}{q}\right) = -1,$$

contradicting that a is a square modulo all primes. Hence, we must have $N = 1$ and a is a perfect square. \square

2.1. Finding Integer Solutions by Working Modulo Primes. So far we have mainly examined ways to show that an equation has no solutions. In this section, we will use information modulo primes to find integer solutions. The first method relies on the Chinese Remainder Theorem (Theorem 2.44) and is a sieving method. The idea is that we want to separate the integers that solve a given equation from the set of all integers. Reducing modulo a prime gives a congruence condition on these solutions. Reducing modulo several primes gives multiple congruence conditions

on these solutions. The Chinese Remainder Theorem allows us to recombine the multiple congruence conditions into a single condition modulo a composite number. We still must search this congruence class for the solution, but this is a much more manageable task than searching over all possible integers!

Example 8.8. We want to solve the equation

$$x^3 - 15x^2 + 11x + 42 = 0.$$

Working modulo primes 2, 5, and 7, we see that the only solutions are

$$x \equiv 0 \pmod{2},$$

$$x \equiv 4 \pmod{5},$$

$$x \equiv 0 \pmod{7}.$$

So we know x must be even, must have remainder 4 when divided by 5, and must be divisible by 7. We apply the Chinese Remainder Theorem to the system of congruences to see that

$$x \equiv 14 \pmod{70}.$$

In fact, 14 is a solution to the original equation.

You may wonder why we did not just search for conditions modulo 70 from the start. It is more efficient to search modulo 2, 5, and 7 and combine, than to search modulo 70. In the example, we needed to check $2 + 5 + 7 = 14$ values to get the three conditions and then do one Chinese Remainder Theorem calculation. Working modulo 70 from the start would require checking 70 values.

There are two key points in Example 8.8 that made it work particularly well:

- (1) The product of the moduli, $2 \cdot 5 \cdot 7 = 70$ is larger than the integer solution 14.
- (2) There is a unique solution for x modulo each of the primes 2, 5, and 7.

While neither point is truly essential, they are very helpful for the practicality of the method.

- If the product of the primes is not large enough, we would need to search through the residue class for the solution.
- If there is more than one solution modulo each prime, then we have multiple combinations to check. We can solve the Chinese Remainder Theorem problem for each possible set of congruences and get a list of possible solutions, but there are potentially very many possibilities.

This next example shows how to address both points.

Example 8.9. Consider the equation

$$x^4 - 85x^3 + 1427x^2 + 6893x - 34220 = 0.$$

A computer search modulo 2, 3, and 5 gives the solutions

$$x \equiv 0 \text{ or } 1 \pmod{2},$$

$$x \equiv 2 \pmod{3},$$

$$x \equiv 0, 2, \text{ or } 4 \pmod{5}.$$

Since any particular solution can satisfy only one condition for each prime, this set of conditions is actually six separate possibilities. For example, we could have

$$\begin{aligned}x &\equiv 0 \pmod{2}, \\x &\equiv 2 \pmod{3}, \\x &\equiv 0 \pmod{5}.\end{aligned}$$

We apply the Chinese Remainder Theorem to the six possible systems of three linear congruences to have the possible conditions

$$x \equiv 2, 5, 14, 17, 20, \text{ or } 29 \pmod{30}.$$

Searching the residue classes of these solutions modulo 30,

$$\begin{aligned}\overline{2} &= \{\dots, -58, -28, 2, 32, 62, \dots\}, \\ \overline{5} &= \{\dots, -55, -25, 5, 35, 65, \dots\}, \\ \overline{14} &= \{\dots, -46, -16, 14, 44, 74, \dots\}, \\ \overline{17} &= \{\dots, -43, -13, 17, 47, 77, \dots\}, \\ \overline{20} &= \{\dots, -40, -10, 20, 50, 80, \dots\}, \\ \overline{29} &= \{\dots, -31, -1, 29, 59, 89, \dots\},\end{aligned}$$

we find that 29 and 62 are solutions. So we have

$$x^4 - 85x^3 + 1427x^2 + 6893x - 34220 = (x - 29)(x - 62)(x^2 + 3x - 20),$$

where the quadratic $x^2 + 3x - 20$ is irreducible. So we have found the only two integer solutions.

While this sieving method is often effective, the combinatorics for multiple solutions can become cumbersome. Somewhat surprisingly, we can often be more efficient just using a single prime. The key to this method is called *Hensel lifting* (or Hensel's Lemma). The idea is to approximate an integer by working modulo p^n for a fixed prime p as the exponent n increases.

Example 8.10. For the equation in Example 8.8,

$$x^3 - 15x^2 + 11x + 42 = 0,$$

we found the integer solution $x = 14$ by applying the Chinese Remainder Theorem to solutions modulo several different primes. If we instead consider one prime, $p = 2$, and reduce modulo powers of 2, we can find the conditions

$$\begin{aligned}x &\equiv 0 \pmod{2}, \\x &\equiv 2 \pmod{4}, \\x &\equiv 6 \pmod{8}, \\x &\equiv 14 \pmod{16}, \\x &\equiv 14 \pmod{32}, \\&\vdots\end{aligned}$$

If there is an integer solution, as soon as p^n gets bigger than that solution, the sequence of values stabilizes.

While it is very fast to search for solution modulo 2, searching for a solution modulo 2^n for large n is not efficient. The key is to think of these solutions written as base 2 numbers. In other words, we can write

$$14 = 0 \cdot 2^0 + 2 \cdot 2^1 + 1 \cdot 2^2 + 1 \cdot 2^3.$$

However, we may need infinitely many terms, so we instead think of these as power series in the base 2. We compute the values of $-1 \pmod{2^n}$ for each n to have

$$-1 = 1 + 1 \cdot 2 + 1 \cdot 2^2 + 1 \cdot 2^3 + 1 \cdot 2^4 + \cdots$$

as an infinite expansion base 2. Such expansions are called *p-adic numbers* and are written as

$$x = a_0 + a_1p + a_2p^2 + a_3p^3 + \cdots.$$

Exploration Exercise 8.40 explores a little of the general theory of *p*-adic numbers.

Finding a_0 for a reasonably small prime is a quick computer search. It turns out that, in many cases, once you know a_0 , then each a_i for $i > 0$ is uniquely determined and can be explicitly computed. We emphasize that the a_i for $i > 0$ can be **computed**, so there is no brute force searching. This fact is the content of Hensel's Lemma.

Theorem 8.11 (Hensel's Lemma). *Let $f(x)$ be a polynomial with integer coefficients. Fix a prime p . Suppose that a is a solution to the congruence*

$$f(x) \equiv 0 \pmod{p^n}$$

for some integer n such that

$$f'(a) \not\equiv 0 \pmod{p}.$$

Then, there is a unique integer $b \in \{0, \dots, p-1\}$ such that $a + bp^n$ is a solution to

$$f(x) \equiv 0 \pmod{p^{n+1}}.$$

Moreover, we have

$$b = -\frac{f(a)}{p^n f'(a)} \pmod{p}.$$

We will need a lemma to prove this.

Lemma 8.12. *Let $f(x)$ be a polynomial with integer coefficients. Then for indeterminants x and y , we have*

$$f(x+y) = f(x) + f'(x)y + g(x,y)y^2$$

for some polynomial $g(x,y)$ with integer coefficients.

Proof. Let $f(x) = \sum_{i=0}^d c_i x^i$. Then, by expanding the first two terms of the binomial formula, we have

$$f(x+y) = \sum_{i=0}^d c_i (x+y)^i = c_0 + \sum_{i=1}^d c_i (x^i + ix^{i-1}y + h(x,y)y^2)$$

for some polynomial $h(x, y)$ with integer coefficients. We can now split the sum and see what results.

$$\begin{aligned} f(x+y) &= c_0 + \sum_{i=1}^d c_i(x^i + ix^{i-1}y + h(x, y)y^2) \\ &= c_0 + \sum_{i=1}^d c_i x^i + \sum_{i=1}^d ix^{i-1}y + \sum_{i=1}^d c_i h(x, y)y^2 \\ &= \sum_{i=0}^d c_i x^i + \left(\sum_{i=1}^d ix^{i-1} \right) y + h(x, y)y^2 \sum_{i=1}^d c_i. \end{aligned}$$

Letting $g(x, y) = h(x, y) \sum_{i=1}^d c_i$ and noting that

$$\begin{aligned} f(x) &= \sum_{i=0}^d c_i x^i, \\ f'(x) &= \sum_{i=1}^d ic_i x^{i-1}, \end{aligned}$$

we have

$$f(x+y) = f(x) + f'(x)y + g(x, y)y^2. \quad \square$$

Proof of Theorem 8.11. We apply Lemma 8.12 to $f(a + bp^n)$ to see that

$$f(a + bp^n) = f(a) + f'(a)bp^n + g(a, bp^n)b^2p^{2n}$$

for some polynomial $g(x, y)$ with integer coefficients. Reducing modulo p^{n+1} , we have

$$f(a + bp^n) \equiv f(a) + f'(a)bp^n \pmod{p^{n+1}}.$$

Recall that $f(a) \equiv 0 \pmod{p^n}$ so that p^n divides $f(a)$. We are trying to find b so that

$$f(a + bp^n) \equiv f(a) + f'(a)bp^n \equiv 0 \pmod{p^{n+1}}.$$

Since $f'(a)bp^n$ contains a term p^n , we need only compute b modulo p . We solve the previous equation for b to have

$$b \equiv -\frac{f(a)}{p^n f'(a)} \pmod{p}.$$

Since p^n divides $f(a)$, the part $\frac{f(a)}{p^n}$ is an integer. By assumption, $f'(a) \not\equiv 0 \pmod{p}$ so it is invertible modulo p . Hence, b is well-defined modulo p . \square

To emphasize, the condition $f'(a) \not\equiv 0 \pmod{p}$ is independent of n and needs to be checked only once!

Example 8.13. Let's revisit Example 8.10. We are considering the equation

$$f(x) = x^3 - 15x^2 + 11x + 42 = 0.$$

We compute

$$f'(x) = 3x^2 - 30x + 11.$$

A quick computation yields that

$$f(0) \equiv 0 \pmod{2} \quad \text{and} \quad f'(0) \equiv 1 \pmod{2}.$$

So we satisfy the hypotheses of Hensel's Lemma. Recall we are notating the solution as

$$x = a_0 + a_1p + a_2p^2 + \cdots,$$

so we have $a_0 = 0$. We compute

$$-\frac{f(0)}{2} \frac{1}{f'(0)} = -\frac{42}{2} \frac{1}{11} = \frac{-21}{11}$$

so that

$$a_1 = \frac{-21}{11} \pmod{2} = 1.$$

At this stage we have

$$\begin{aligned} x &= a_0 + a_1 \cdot 2 = 0 + 1 \cdot 2 = 2, \\ f(2) &\equiv 0 \pmod{4}. \end{aligned}$$

We apply Hensel's Lemma with $a = 2$ to have

$$-\frac{f(2)}{2^2} \frac{1}{f'(2)} = -\frac{12}{4} \frac{1}{-37} = \frac{3}{37}$$

so that

$$a_2 = \frac{3}{37} \pmod{2} = 1$$

and

$$\begin{aligned} x &= a_0 + a_1 \cdot 2 + a_2 \cdot 4 = 0 + 1 \cdot 2 + 1 \cdot 4 = 6 \\ f(6) &\equiv 0 \pmod{8}. \end{aligned}$$

Continuing, we have

$$\begin{aligned} a_3 &= -\frac{f(6)}{8f'(6)} \pmod{2} = \frac{27}{61} \pmod{2} = 1, \\ a_4 &= -\frac{f(14)}{16f'(14)} \pmod{2} = \frac{0}{179} \pmod{2} = 0, \\ a_5 &= -\frac{f(14)}{32f'(14)} \pmod{2} = \frac{0}{179} \pmod{2} = 0, \\ &\vdots \end{aligned}$$

So we have

$$x = 0 + 1 \cdot 2 + 1 \cdot 4 + 1 \cdot 8 = 14.$$

We say that the solution 0 modulo 2 *lifts* to the solution 14.

Example 8.14. Consider the equation

$$f(x) = x^4 - 134x^3 - 2x^2 + 270x - 268 = 0.$$

We see that it has derivative

$$f'(x) = 4x^3 - 402x^2 - 4x + 270$$

so that

$$f'(x) \equiv 0 \pmod{2}.$$

We cannot apply Hensel's Lemma modulo 2, so we work modulo 3. We find that

$$\begin{aligned}f(0) &\equiv 1 \pmod{3}, \\f(1) &\equiv 2 \pmod{3}, \\f(2) &\equiv 0 \pmod{3}.\end{aligned}$$

However,

$$f'(2) \equiv 0 \pmod{3}$$

so that we cannot use Hensel's Lemma modulo 3. Working modulo 5, we find that the only solution is

$$f(4) \equiv 0 \pmod{5},$$

which satisfies

$$f'(4) \equiv 3 \pmod{5}.$$

Applying Hensel's Lemma with $a = 4$, we have

$$a_1 = -\frac{f(4)}{5} \frac{1}{f'(4)} \pmod{5} = -\frac{1508}{5922} \pmod{5} = 1$$

so that

$$a = 4 + 1 \cdot 5 = 9.$$

Applying Hensel's Lemma with $a = 9$, we have

$$a_2 = -\frac{f(9)}{5^2} \frac{1}{f'(9)} \pmod{5} = 0.$$

So we again use $a = 9$ to have

$$a_3 = -\frac{f(9)}{5^3} \frac{1}{f'(9)} \pmod{5} = 1$$

so that

$$a = 4 + 1 \cdot 5 + 1 \cdot 5^3 = 134.$$

We see that

$$f(134) = 0,$$

and we are done.

Note that similar to the sieving methods, if you choose a prime p which has multiple solutions modulo p , then you must lift each solution to find all integer solutions to the equation.

Investigation 8.15. In this investigation we examine some other behaviors of Hensel's Lemma.

- Choose a polynomial that you know has two distinct integer roots a and b . Find a prime where you can use Hensel's Lemma to find both roots.
- Pick a prime p such that $a \equiv b \pmod{p}$. What part of Hensel's Lemma fails?
- Choose an irreducible polynomial of degree at least 2. Find a prime at which you can apply Hensel's Lemma. Apply Hensel's Lemma. What can you say about the resulting sequence of a_i ?

This ends our discussion of methods using modular arithmetic. We have only scratched the surface of what is known and what can be done. The topics of p -adic numbers and p -adic analysis go much deeper into the subject.

3. Pythagorean Triples

We begin our discussion of several specific Diophantine equations with the problem of Pythagorean triples.

Question 8.16. What are the integer solutions to

$$x^2 + y^2 = z^2?$$

Integer solutions to the Diophantine equation in Question 8.16 are called *Pythagorean triples* since such solutions (x, y, z) represent the side lengths of a right triangle by the Pythagorean theorem. Instead of trying to work modulo primes, we find all solutions through geometry.

We are considering the equation

$$x^2 + y^2 = z^2.$$

If we divide both sides by z^2 and rename $X = \frac{x}{z}$ and $Y = \frac{y}{z}$, this is equivalent to finding rational solutions to

$$X^2 + Y^2 = 1.$$

This object is a circle of radius 1 centered at $(0, 0)$, and the points on the axes represent four obvious solutions:

$$\{(0, \pm 1), (\pm 1, 0)\}.$$

From one of these points we can now find all other solutions. Take, for example, $(-1, 0)$ and draw a straight line through this point with slope m . The equation of this line is given by

$$Y = m(X + 1).$$

Since two points make a distinct line, every point on $X^2 + Y^2 = 1$ is on exactly one such line. So, every point on the circle is a solution to the system of equations

$$\begin{aligned} X^2 + Y^2 &= 1, \\ Y &= m(X + 1). \end{aligned}$$

Substituting for Y into the first equation, we see that

$$X^2 + m^2(X^2 + 2X + 1) = 1,$$

which simplifies to

$$(m^2 + 1)X^2 + 2m^2X + (m^2 - 1) = 0.$$

This is a quadratic polynomial, and we want to know its two roots. We could apply the quadratic equation, but recall that we already know that $X = -1$ is a root, which means that $X + 1$ divides the polynomial. Proceeding with the division, we have

$$(m^2 + 1)X^2 + 2m^2X + (m^2 - 1) = (X + 1)((m^2 + 1)X + (m^2 - 1)).$$

So the two X values of points on both the line and the circle are

$$X \in \left\{ -1, \frac{1-m^2}{1+m^2} \right\}.$$

We obtain the Y values from the line equation

$$Y \in \left\{ 0, \frac{2m}{1+m^2} \right\}.$$

Since $m = \frac{\Delta Y}{\Delta X}$, the second point on the line is rational if and only if m is rational. Thus, we obtain *every* rational point by allowing m to range over all rational values. To go from rational points $(\frac{x}{z}, \frac{y}{z})$ on the circle to integer Pythagorean triples, we clear the denominators to get

$$(x, y, z) = (1 - m^2, 2m, 1 + m^2).$$

Since m is rational, we write $m = \frac{a}{b}$ for integers a, b with $b \neq 0$ and multiply through by b^2 to get

$$(1 - m^2, 2m, 1 + m^2) \mapsto (b^2 - a^2, 2ab, b^2 + a^2).$$

We have now proven the following theorem.

Theorem 8.17. *The set of integer Pythagorean triples is*

$$\{(b^2 - a^2, 2ab, b^2 + a^2) : a, b \in \mathbb{Z}, b \neq 0\}.$$

Example 8.18. We write down a few Pythagorean triples:

- $(a, b) = (0, 1) \mapsto (x, y, z) = (1, 0, 1)$, which satisfies

$$1^2 + 0^2 = 1^2.$$

- $(a, b) = (1, 1) \mapsto (x, y, z) = (0, 2, 2)$, which satisfies

$$0^2 + 2^2 = 2^2.$$

- $(a, b) = (1, 2) \mapsto (x, y, z) = (3, 4, 5)$, which satisfies

$$3^2 + 4^2 = 5^2.$$

Investigation 8.19. Fix a height bound B and consider all integers a, b with $-B \leq a, b \leq B$. There are $(2B + 1)^2$ such pairs of integers.

- (a) How many distinct Pythagorean triples are made from these $(2B + 1)^2$ pairs of integers? What is the counting function of Pythagorean triples in terms of B ?
- (b) Even if we are just counting the (a, b) that produce distinct triples, we are still not counting right triangles with integer sides since some of those triples contain values ≤ 0 . Can you count just the triples corresponding to triangles?

4. Fermat's Last Theorem

Having solved the case of Pythagorean triples, we generalize the problem.

Question 8.20. For which n does the equation

$$x^n + y^n = z^n$$

have a nontrivial integer solution? By nontrivial we mean that all three of x , y , and z are nonzero.

The equation

$$x^n + y^n = z^n$$

is called the *Fermat equation* and is the subject of Fermat's Last Theorem, proven by Wiles and Taylor⁹ in the 1990s.

Theorem 8.21. *For $n \geq 3$, the Fermat equation has no nontrivial solutions in integers.*

The proof of this theorem is one of the biggest achievements in number theory in the 20th century, and it is well beyond the scope of this book. We will look only at the case $n = 4$ as the proof illustrates the powerful method of *descent* developed by Pierre de Fermat. The idea goes as follows:

- Assume there is one “positive” integer solution (x, y, z) .
- From that solution, we derive a new “smaller” solution (X, Y, Z) .
- We can repeat this process indefinitely. Thus, there are infinitely many smaller positive solutions.
- This contradicts the well ordering property (page 5) of the positive integers, so the first solution cannot exist.

4.1. $n = 4$ and the method of descent. We use Fermat's method of descent to show there are no nontrivial solutions to

$$(26) \quad x^4 + y^4 = z^4.$$

⁹Richard Taylor (1962–) is a British mathematician.

We instead consider the equation

$$(27) \quad x^4 + y^4 = z^2.$$

If there are no solutions to equation (27), then there are no solutions to equation (26) since a solution (x, y, z) to equation (26) corresponds to a solution (x, y, z^2) to equation (27).

Theorem 8.22. *There are no nontrivial integer solutions to*

$$x^4 + y^4 = z^2.$$

Proof. Assume that x , y , and z are all not 0. Then we know that $z > 1$ since $x^4 + y^4 \geq 2$. Let (x, y, z) be the positive integer solution with the smallest z . From our study of Pythagorean triples we know there are integers s, t with $s \neq 0$ such that

$$\begin{aligned} x^2 &= 2st, \\ y^2 &= s^2 - t^2, \\ z &= s^2 + t^2, \end{aligned}$$

since

$$(x^2)^2 + (y^2)^2 = z^2.$$

The y^2 equation gives another Pythagorean triple, $y^2 + t^2 = s^2$. So there exist integers a and b with $a \neq 0$ such that

$$\begin{aligned} t &= 2ab, \\ y &= a^2 - b^2, \\ s &= a^2 + b^2, \end{aligned}$$

with $\gcd(a, b) = 1$. Writing x^2 in terms of a and b , we see that

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

Any prime q that divides ab must divide a or b but not both since $\gcd(a, b) = 1$. Then, q cannot divide $a^2 + b^2$. Thus, $\gcd(ab, a^2 + b^2) = 1$ and they must both be squares since their product is a square. So there is an integer Z such that

$$Z^2 = a^2 + b^2 = s.$$

Since $\gcd(a, b) = 1$ and ab is a square, then a and b both must also be squares (say $X^2 = a$ and $Y^2 = b$), and we have a new solution,

$$X^4 + Y^4 = Z^2.$$

Now we show that the new solution (X, Y, Z) is smaller.

$$Z^2 = X^4 + Y^4 = a^2 + b^2 = s \leq s^2 + t^2 = z < z^2,$$

where the last inequality follows from $z > 1$. In particular,

$$Z < z,$$

contradicting the assumption that (x, y, z) is the smallest solution. □

Corollary 8.23. *There are no nontrivial solutions to the Fermat equation for $n = 4$.*

Proof. If there were a nontrivial integer solution (x, y, z) to $x^4 + y^4 = z^4$, then (x, y, z^2) would be a nontrivial integer solution to $x^4 + y^4 = z^2$. Since there are no nontrivial integer solutions to $x^4 + y^4 = z^2$, this is a contradiction. \square

Corollary 8.24. *There are no nontrivial integers solutions to*

$$x^{4n} + y^{4n} = z^{4n}$$

for every positive integer n .

Proof. If there were a nontrivial integer solution (x, y, z) to $x^{4n} + y^{4n} = z^{4n}$, then (x^n, y^n, z^n) would be a nontrivial integer solution to $x^4 + y^4 = z^4$. Since there are no nontrivial integer solutions to $x^4 + y^4 = z^4$, this is a contradiction. \square

5. Pell's Equation and Fundamental Units

In this section, we study the quadratic equation called *Pell's equation*

$$x^2 - dy^2 = 1,$$

where d is a fixed integer. The solutions to Pell's equation are directly related to the quadratic number fields studied in Chapter 7, and they were first studied in ancient Greece for $d = 2$ to give good rational approximations to $\sqrt{2}$. In the middle ages (628 C.E., including Brahmagupta) Indian mathematicians determined the first ad hoc and later general methods to solve Pell's equation. Several centuries later, there was renewed interest in Pell's equation when Pierre de Fermat (re-)discovered how to solve Pell's equation in 1657 and challenged the mathematical community to find a solution. The name is due to Euler's mistakenly attributing to John Pell¹⁰ the work of William Brouncker¹¹, who (re-)discovered the first general solution in 1657 in response to Fermat's challenge.

Question 8.25. For which d does Pell's equation have an integer solution?

If $d < 0$, then

$$x^2 - dy^2 = 1$$

does not have any nontrivial integer solutions since 1 is not the sum of any two positive integers. So we restrict to d positive. If $d = c^2$ is a square, then by substituting $z = cy$ the equation becomes $x^2 - z^2 = 1$, which is a Pythagorean triple, and we know all such solutions. More generally, if $d = c^2 D$ where D is the squarefree portion of d , we again substitute $z = cy$ to have the equation

$$x^2 - Dz^2 = 1.$$

Consequently, we assume that d is positive and squarefree.

¹⁰John Pell (1611–1685) was an English mathematician.

¹¹William Brouncker (1620–1684) was an English mathematician.

To see the connection to quadratic number fields, consider the field $\mathbb{Q}(\sqrt{d}) = \{a + b\sqrt{d} : a, b, \in \mathbb{Q}\}$. In this field, we can factor the equation as

$$(x + y\sqrt{d})(x - y\sqrt{d}) = 1.$$

Recall that the norm of $x + y\sqrt{d}$ in $\mathbb{Q}(\sqrt{d})$ is given by

$$N(x + y\sqrt{d}) = x^2 - dy^2.$$

So solutions to Pell's equations are elements of norm 1 in $\mathbb{Q}(\sqrt{d})$. We summarize this with the following lemma.

Lemma 8.26. *Integer solutions to $x^2 - dy^2 = 1$ are exactly the units of $\mathbb{Z}[\sqrt{d}] = \{a + b\sqrt{d} : a, b \in \mathbb{Z}\}$ of norm 1.*

Proof. If $x + y\sqrt{d}$ is a unit with norm 1, then

$$N(x + y\sqrt{d}) = x^2 - dy^2 = 1.$$

If $x^2 - dy^2 = 1$, then $x + y\sqrt{d}$ has norm 1. □

Since norms are multiplicative (Proposition 6.21), if we multiply two elements $\mathbb{Z}[\sqrt{d}]$ of norm 1, we get another element of $\mathbb{Z}[\sqrt{d}]$ of norm 1. So, given one unit with norm 1, every power of it is another unit with norm 1 and a new solution to Pell's equation. It turns out that not only are units powers of other units, but there is a single unit whose powers form every unit (for $d > 0$). In other words, there is some unit ϵ such that every other unit u can be written as

$$u = \pm \epsilon^m$$

for some positive integer m . This generating unit is called the *fundamental unit* of $\mathbb{Q}(\sqrt{d})$.

Definition 8.27. For $d > 0$, the *fundamental unit* of $\mathbb{Q}(\sqrt{d})$ is the smallest unit with $N(x + y\sqrt{d}) = 1$ and $x + y\sqrt{d} > 1$.

Example 8.28. The fundamental unit of $\mathbb{Q}(\sqrt{2})$ is $1 + \sqrt{2}$. Of the four units with norm 1,

$$\pm 1 \pm \sqrt{2},$$

their numerical values are

$$1 + \sqrt{2} \approx 2.414,$$

$$1 - \sqrt{2} \approx -0.414,$$

$$-1 - \sqrt{2} \approx -2.414,$$

$$-1 + \sqrt{2} \approx 0.414.$$

Hence, $1 + \sqrt{2}$ is the fundamental unit since it is the smallest unit with numerical value at least 1.

Theorem 8.29. *For $d > 0$, let ϵ be the fundamental unit of $\mathbb{Q}(\sqrt{d})$. Then every other unit can be written as*

$$u = \pm \epsilon^m$$

for some positive integer m .

Proof. Let ϵ be the fundamental unit of $\mathbb{Q}(\sqrt{d})$. Assume there is some unit $u \neq \pm \epsilon^m$ for every integer m . If $u = x + y\sqrt{d}$, by taking $\pm x \pm y\sqrt{d}$, we may assume that $|u| > 1$. Since ϵ is minimal and $\epsilon > 1$, we have

$$\lim_{m \rightarrow \infty} \epsilon^m = \infty,$$

and there is some m such that

$$\epsilon^m < |u| < \epsilon^{m+1}.$$

Thus,

$$1 < \epsilon^{-m} |u| < \epsilon$$

and

$$N(\epsilon^{-m} u) = 1.$$

This contradicts the minimality of ϵ so such a u cannot exist. \square

While this is a nice theorem about the fundamental unit, we have yet to establish the fact that such a unit always exists. It could be that there are no units of norm 1 and, thus, no fundamental unit. Or perhaps, infinitely many units of norm 1, for which none has minimal absolute value. Courtesy of Lemma 8.26, this is the same as asking whether Pell's equation always has a solution, bringing us back to Question 8.25.

Theorem 8.30. *For d positive and squarefree, Pell's equation always has a solution.*

Proof. We use Dirichlet's Diophantine approximation theorem and the pigeon-hole principle to obtain a solvable system of congruences.

Dirichlet's Diophantine approximation theorem (Theorem 7.21) tells us that

$$\left| x - y\sqrt{d} \right| < \frac{1}{y}$$

has infinitely many solutions in integers (x, y) . Let (x, y) be such an integer solution. Then since $\left| x - y\sqrt{d} \right| < \frac{1}{y}$, we have

$$x < y\sqrt{d} + \frac{1}{y}.$$

Note that for all integers y and d we have

$$\frac{1}{y} < y\sqrt{d},$$

which implies

$$x < 2y\sqrt{d},$$

so that

$$x + y\sqrt{d} < 3y\sqrt{d}.$$

Then we see that

$$\begin{aligned} N(x + \sqrt{d}y) &= |x^2 - dy^2| = (x - y\sqrt{d})(x + y\sqrt{d}) \\ &< |x - y\sqrt{d}| 3y\sqrt{d} < \frac{1}{y} 3y\sqrt{d} = 3\sqrt{d}. \end{aligned}$$

This last bound is independent of (x, y) , so there is some integer a , $|a| < 3\sqrt{d}$ such that $x^2 - dy^2 = a$ for infinitely many choices of (x, y) . Since there are infinitely many pairs (x, y) and only finitely many residue classes modulo a , using the pigeonhole principle, we find two pairs (x_1, y_1) and (x_2, y_2) such that

$$(28) \quad x_1 \equiv x_2 \pmod{a} \quad y_1 \equiv y_2 \pmod{a}.$$

Assuming that $x_1 + y_1\sqrt{d} > x_2 + y_2\sqrt{d}$, define

$$x + y\sqrt{d} = \frac{x_1 + y_1\sqrt{d}}{x_2 + y_2\sqrt{d}}.$$

Then we have

$$\begin{aligned} x + y\sqrt{d} &= \frac{x_1 + y_1\sqrt{d}}{x_2 + y_2\sqrt{d}} = \frac{(x_1 - \sqrt{d}y_1)(x_2 + y_2\sqrt{d})}{x_2^2 - dy_2^2} \\ &= \frac{(x_1 - \sqrt{d}y_1)(x_2 + y_2\sqrt{d})}{a} = \frac{(x_1x_2 - dy_1y_2) - (x_1y_2 - x_2y_1)\sqrt{d}}{a}. \end{aligned}$$

From the congruence assumptions (28), we have

$$\begin{aligned} x_1x_2 - dy_1y_2 &\equiv x_1^2 - dy_1^2 \equiv 0 \pmod{a}, \\ x_1y_2 - x_2y_1 &\equiv 0 \pmod{a}. \end{aligned}$$

Therefore, a divides both terms in the numerator and $x + y\sqrt{d} \in \mathbb{Z}[\sqrt{d}]$, giving an integer solution (x, y) to $x^2 - dy^2 = 1$. \square

Now that we know there is at least one solution, by the well ordering property, there must be a smallest solution. In terms of Pell's equation, this solution is called the *fundamental solution*, and it corresponds to the fundamental unit of $\mathbb{Q}(\sqrt{d})$.

Corollary 8.31. *For d positive and squarefree, $\mathbb{Q}(\sqrt{d})$ always has a fundamental unit.*

Proof. The fundamental solution to Pell's equation corresponds to the fundamental unit of $\mathbb{Q}(\sqrt{d})$. \square

While Theorem 8.30 proves the existence of a solution and, hence, a smallest solution, it says nothing about how large the smallest solution may be.

Example 8.32. The equation

$$x^2 - 61y^2 = 1$$

has smallest solution

$$(x, y) = (1766319049, 226153980).$$

Question 8.33. How can you find the fundamental unit of $\mathbb{Q}(\sqrt{d})$ or, equivalently, the fundamental solution to Pell's equation?

Investigation 8.34. Apply continued fractions to Pell's equation $x^2 - dy^2 = 1$. In other words, consider the convergents of \sqrt{d} as $\frac{x}{y}$.

- (a) Is every convergent of \sqrt{d} a solution to Pell's equation?
- (b) Is every solution to Pell's equation a convergent of \sqrt{d} ?
- (c) Which convergent of \sqrt{d} is the fundamental solution of Pell's equation? (*Hint:* The answer depends on the period of the continued fraction expansion of \sqrt{d} .)

Example 8.35. Consider

$$x^2 - 7y^2 = 1.$$

The continued fraction expansion of $\sqrt{7}$ is

$$\sqrt{7} = [2, \overline{1, 1, 4}].$$

For $C_k = \frac{p_k}{q_k}$, we compute

cont. frac.	p_k	q_k	$p_k^2 - 7q_k^2$
[2]	2	1	-3
[2; 1]	3	1	2
[2; 1, 1]	5	2	-3
[2; 1, 1, 1]	8	3	1
[2; 1, 1, 1, 4]	37	14	-3
[2; 1, 1, 1, 4, 1]	45	17	2
[2; 1, 1, 1, 4, 1, 1]	82	31	-3
[2; 1, 1, 1, 4, 1, 1, 1]	127	48	1

By looking at the last column, we see that C_3 and C_7 are solutions to $x^2 - 7y^2 = 1$, but the others are not.

Proposition 8.36. Every solution (x, y) to Pell's equation $x^2 - dy^2 = 1$ corresponds to a convergent $\frac{x}{y}$ of \sqrt{d} .

Proof. Rearranging

$$x^2 - dy^2 = 1,$$

we get

$$\frac{x^2}{y^2} - d = \frac{1}{y^2}.$$

In particular, $\frac{x}{y} > \sqrt{d}$. Using the fact that $x^2 - dy^2 = 1$, we have

$$\begin{aligned} \frac{x}{y} - \sqrt{d} &= \frac{x - y\sqrt{d}}{y} = \frac{x - y\sqrt{d}}{y(x^2 - dy^2)} = \frac{1}{y(x + y\sqrt{d})} \\ &= \frac{1}{y^2(\frac{x}{y} + \sqrt{d})} < \frac{1}{y^2 2\sqrt{d}} < \frac{1}{2y^2}. \end{aligned}$$

By Theorem 7.55, $\frac{x}{y}$ must be a convergent of \sqrt{d} . □

Knowing that the smallest solution must come from a convergent of the continued fraction expansion of \sqrt{d} certainly narrows down the possibilities, but we still do not know which convergent to use. From Example 8.35 we might guess that C_{k-1} is the fundamental solution, where k is the period of the continued fraction expansion of \sqrt{d} . To test this guess, let's look at some more examples.

d	cont. frac.	solutions
3	$[1; \overline{1, 2}]$	C_1, C_3, C_5, \dots
5	$[2; \overline{4}]$	C_1, C_3, C_5, \dots
6	$[2; \overline{2, 4}]$	C_1, C_3, C_5, \dots
7	$[2; \overline{1, 1, 1, 4}]$	C_3, C_7, C_{11}, \dots
10	$[3; \overline{6}]$	C_1, C_3, C_5, \dots
11	$[3; \overline{3, 6}]$	C_1, C_3, C_5, \dots
13	$[3; \overline{1, 1, 1, 1, 6}]$	$C_9, C_{19}, C_{29}, \dots$
14	$[3; \overline{1, 2, 1, 6}]$	C_3, C_7, C_{11}, \dots
15	$[3; \overline{1, 6}]$	C_1, C_3, C_5, \dots
17	$[4; \overline{8}]$	C_1, C_3, C_5, \dots
19	$[4; \overline{2, 1, 3, 1, 2, 8}]$	$C_5, C_{11}, C_{17}, \dots$
43	$[6; \overline{2, 2, 12}]$	$C_5, C_{11}, C_{17}, \dots$

We chose 43 for its odd period since it seems that odd periods behave differently than even periods. Our original guess of C_{k-1} was not quite right, and we adjust it with the following conjecture.

Conjecture 8.37. *The fundamental solution of Pell's equation $x^2 - dy^2 = 1$ is given by C_{k-1} for $k = r$ or $2r$, where r is the period of the continued fraction expansion of \sqrt{d} . Moreover, if C_t gives the fundamental solution, then all solutions are given by C_{t+mr} , where $m \in \mathbb{N}$ if r is even and $m \in 2\mathbb{N}$ if r is odd.*

This conjecture is, in fact, true, but we will not prove it. Assuming it is true, we have Algorithm 8.1 to solve Pell's equation.

Algorithm 8.1. Fundamental Solution to Pell's Equation

Input: d a squarefree positive integer

Output: fundamental solution to $x^2 - dy^2 = 1$

Algorithm:

- 1: Compute the continued fraction expansion of \sqrt{d} , $[a_0; \overline{a_1, \dots, a_r}]$.
 - 2: Compute the $(r-1)$ -th convergent, p_{r-1}/q_{r-1} .
 - 3: If (p_{r-1}, q_{r-1}) is a solution, then return (p_{r-1}, q_{r-1}) .
 - 4: Otherwise, compute the $(2r-1)$ -th convergent p_{2r-1}/q_{2r-1} and return (p_{2r-1}, q_{2r-1}) .
-

Example 8.38. Consider the equation from Example 8.32:

$$x^2 - 61y^2 = 1.$$

To find the smallest solution, we compute the continued fraction expansion of $\sqrt{61}$ as

$$\sqrt{61} = [7; \overline{1, 4, 3, 1, 2, 2, 1, 3, 4, 1, 14}].$$

This has period 11, so we check the tenth convergent

$$C_{10} = \frac{29718}{3805},$$

but

$$29718 + 3805\sqrt{61}$$

is not a solution. So we need the twenty-first convergent

$$C_{21} = \frac{1766319049}{226153980},$$

which gives the correct fundamental solution

$$(x, y) = (1766319049, 226153980).$$

Investigation 8.39. For $d > 0$ and squarefree, compute the fundamental unit of $\mathbb{Q}(\sqrt{d})$.

- (a) Which small d has the largest fundamental solution?
- (b) Conversely, can you create a d using continued fractions that is guaranteed to have either a small or large fundamental solution to Pell's equation?

6. Waring Problem

Proposed by Edward Waring¹² in *Meditationes algebraicae* in 1770, the Waring problem is to find the positive integer $g(k)$ such that any positive integer can be represented as the sum of $g(k)$ nonnegative k th powers.

Question 8.40. Does there exist an integer $g(k)$ such that every positive integer n can be written as the sum of $g(k)$ k th powers? If $g(k)$ does exist, what is its value?

In particular, we want to know when we can solve

$$x_1^k + x_2^k + \cdots + x_{g(k)}^k = n$$

for all positive integers n . At first glance it is not even clear that $g(k)$ should exist.

Example 8.41. We can write

$$2 = 1^2 + 1^2$$

so that $g(2) \geq 2$. However, 6 cannot be written as the sum of fewer than three squares

$$6 = 2^2 + 1^2 + 1^2$$

¹²Edward Waring (1736–1798) was an English mathematician.

so that $g(2) \geq 3$. But 7 cannot be written as the sum of fewer than four squares

$$7 = 2^2 + 1^2 + 1^2 + 1^2$$

so that $g(4) \geq 4$. Maybe this sequence keeps growing and $g(k)$ does not exist.

Investigation 8.42.

- (a) What other positive integers cannot be expressed as the sum of fewer than three integer squares? Can you find a pattern?
- (b) What other positive integers cannot be expressed as the sum of fewer than four integer squares? Can you find a pattern?

The problem for $k = 2$ was proposed by Diophantus in *Arithmetica*, and Lagrange proved in 1770 that $g(2) = 4$. In other words,

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = n$$

is solvable for every positive integer n . David Hilbert answered the first part of Question 8.40 in 1909 by proving that $g(k)$ exists for every k .

It is, in fact, known that $g(2) = 4$, $g(3) = 9$, and $g(4) = 19$ (and many more values). We will prove that $g(2) = 4$. In particular, we will show first that every positive integer can be written as the sum of four squares, and second that there is at least one integer that cannot be written as the sum of fewer than four squares. The second statement is much easier than the first, and we have already discussed (without proof) that 7 cannot be written as fewer than four squares.

6.1. Four Squares ($k = 2$). In 1770, Lagrange solved the Waring problem for $k = 2$ by showing that every positive integer n can be written as

$$n = a^2 + b^2 + c^2 + d^2,$$

and any $n \equiv 7 \pmod{8}$ cannot be represented as a sum of three squares, so $g(2) = 4$.

Example 8.43. We represent the integers $1, \dots, 10$ as the sum of four squares:

$$\begin{aligned} 1 &= 1^2 + 0^2 + 0^2 + 0^2, \\ 2 &= 1^2 + 1^2 + 0^2 + 0^2, \\ 3 &= 1^2 + 1^2 + 1^2 + 0^2, \\ 4 &= 2^2 + 0^2 + 0^2 + 0^2, \\ 5 &= 2^2 + 1^2 + 0^2 + 0^2, \\ 6 &= 2^2 + 1^2 + 1^2 + 0^2, \\ 7 &= 2^2 + 1^2 + 1^2 + 1^2, \\ 8 &= 2^2 + 2^2 + 0^2 + 0^2, \\ 9 &= 3^2 + 0^2 + 0^2 + 0^2, \\ 10 &= 3^2 + 1^2 + 0^2 + 0^2. \end{aligned}$$

Remark. Legendre improved the theorem in 1798 by stating that a positive integer can be expressed as the sum of three squares if and only if it is not of the form $4^k(8m+7)$. His proof was incomplete but was later finished by Gauss.

We will need several lemmas simplifying the problem before we can prove the main theorem.

Lemma 8.44 (Euler four-square identity). *The product of two sums of four squares is a sum of four squares:*

$$\begin{aligned} (a^2 + b^2 + c^2 + d^2)(w^2 + x^2 + y^2 + z^2) \\ = (aw + bx + cy + dz)^2 + (ax - bw - cz + dy)^2 \\ + (ay + bz - cw - dx)^2 + (az - by + cx - dw)^2. \end{aligned}$$

Proof. Direct computation. □

Lemma 8.44 allows us to reduce to the case of writing primes as the sum of four squares. Since any integer can be factored into the product of primes, if all the primes can be written as the sum of four squares, then every (positive) integer is the sum of four squares. We can easily see that $2 = 1^2 + 1^2 + 0^2 + 0^2$, so we need only consider odd primes.

Lemma 8.45. *Let p be an odd prime number. There are integers a, b, c, d , and m with $0 < m < \frac{p}{2}$ such that*

$$a^2 + b^2 + c^2 + d^2 = mp.$$

Proof. Let $p = 2n + 1$ be an odd prime, and consider the two sets of residues

$$X = \{x^2 \pmod{p} : 0 \leq x \leq n\} \quad \text{and} \quad Y = \{-y^2 - 1 \pmod{p} : 0 \leq y \leq n\}.$$

The set X contains $n + 1$ different elements since, given $a^2, b^2 \in X$,

$$\text{if } p \mid a^2 - b^2, \text{ then } p \mid (a + b) \text{ or } p \mid (a - b).$$

But since

$$a - b \leq 2n < p \quad \text{and}$$

$$a + b \leq 2n < p,$$

we must have $a = b$. Similarly for $(-a^2 - 1), (-b^2 - 1) \in Y$, if $p \mid (-a^2 - 1) - (-b^2 - 1) = b^2 - a^2$, then $a = b$. So Y must also have $n + 1$ distinct elements.

Then, since there is a total of $2(n + 1)$ elements in the two sets but only $2n + 1 = p$ residue classes, by the pigeonhole principle, the two sets have at least one element in common. In other words, there are a and b such that

$$a^2 + b^2 + 1 \equiv 0 \pmod{p}.$$

In addition, we have

$$0 < a^2 + b^2 + 1 = mp \leq 2n^2 + 1.$$

But since $p^2 = (2n + 1)^2 = 4n^2 + 4n + 1$, we have

$$0 < a^2 + b^2 + 1 = mp \leq 2n^2 + 1 \leq np < \frac{p}{2} \cdot p,$$

so that m is at most $\frac{p}{2}$. □

Lemma 8.46. *Let m be a positive integer. If $2m$ is the sum of four integer squares, then so is m .*

Proof. Let $2m = x^2 + y^2 + z^2 + w^2$. Since $2m$ is even, then 0, 2, or 4 of $\{x, y, z, w\}$ are even. So we can (after possibly renaming) assume that x, y and z, w have the same parity. In that case, $x \pm y$ and $z \pm w$ are even and

$$m = \left(\frac{x+y}{2}\right)^2 + \left(\frac{x-y}{2}\right)^2 + \left(\frac{z+w}{2}\right)^2 + \left(\frac{z-w}{2}\right)^2$$

is the sum of four (integer) squares. \square

Proposition 8.47. *Every odd prime p can be written as the sum of four integer squares.*

Proof. Let p be an odd prime. Then, from Lemma 8.45, there are integers a, b, c, d , and m such that

$$a^2 + b^2 + c^2 + d^2 = mp$$

for some $0 < m < \frac{p}{2}$.

If $m = 1$, we are done, so assume $m \geq 2$. We will use the method of descent to show that $m = 1$.

If m is even, then by Lemma 8.46, we may write $\frac{m}{2}p$ as the sum of four squares. Thus, we may assume that m is odd.

Further, we assume that m is minimal with the property that mp can be written as the sum of four squares:

$$a^2 + b^2 + c^2 + d^2 = mp.$$

Choose integers $-\frac{m}{2} < x, y, z, w < \frac{m}{2}$ such that

$$\begin{aligned} x &\equiv b \pmod{m}, \\ y &\equiv c \pmod{m}, \\ z &\equiv d \pmod{m}, \\ w &\equiv a \pmod{m}. \end{aligned}$$

Then we have

$$x^2 + y^2 + z^2 + w^2 < 4\left(\frac{m}{2}\right)^2 = m^2.$$

Notice that since

$$a^2 + b^2 + c^2 + d^2 \equiv 0 \pmod{m},$$

by the choice of (x, y, z, w)

$$x^2 + y^2 + z^2 + w^2 \equiv 0 \pmod{m}.$$

So we have

$$x^2 + y^2 + z^2 + w^2 = km$$

for some integer $k < m$. We take the product

$$(29) \quad (a^2 + b^2 + c^2 + d^2)(x^2 + y^2 + z^2 + w^2) = km^2p.$$

From Lemma 8.44, we can write km^2p as the sum of four squares. Examining each of them, we have

$$\begin{aligned}aw + bx + cy + dz &\equiv w^2 + x^2 + y^2 + z^2 \equiv 0 \pmod{m}, \\ax - bw - cz + dy &= (ax - bw) + (dy - cz) \equiv 0 \pmod{m}, \\ay + bz - cw - dx &= (ay - cw) + (bz - dx) \equiv 0 \pmod{m}, \\az - by + cx - dw &= (az - dw) + (cx - by) \equiv 0 \pmod{m}.\end{aligned}$$

Thus, we can divide equation (29) by m^2 to have kp is the sum of four squares. Since $k < m$, this contradicts the minimality of m , so we must have $m = 1$. \square

Theorem 8.48 (Lagrange). $g(2) = 4$.

Proof. Let n be a positive integer. We can factor n into the product of primes. Since $2 = 1^2 + 1^2 + 0^2 + 0^2$ and, by Proposition 8.47, any odd prime may be written as the sum of four squares, we can write every factor of n as a product of four squares. Then by Lemma 8.44, n can be written as the sum of four squares.

To finish the proof, we show that 7 cannot be expressed as the sum of fewer than four squares. If 7 can be written as the sum of three or fewer squares, one must be 2 since $1^2 + 1^2 + 1^2 < 7$. But $2^2 + 2^2 > 7$, so exactly one number must be 2. Then we have $7 = 2^2 + 1^2 + 1^2 + 1^2$ is the only representation. \square

We close with a further question. You may have noticed in our examples at the beginning of the section that there may be more than one way to write a number as the sum of four squares:

$$\begin{aligned}9 &= 3^2 + 0^2 + 0^2 + 0^2 \\&= 2^2 + 2^2 + 1^2 + 0^2\end{aligned}$$

and

$$\begin{aligned}10 &= 3^2 + 1^2 + 0^2 + 0^2 \\&= 2^2 + 2^2 + 1^2 + 1^2.\end{aligned}$$

Question 8.49. How many ways can a number n be represented as the sum of k squares?

Investigation 8.50.


- (a) Which positive integers can be written as the sum of four squares in only one way?
- (b) Which positive integers n can be written as the sum of four squares in more than one way?
- (c) What about more than two ways? or more than three ways? ...
- (d) Is there an upper bound on the number of ways a positive integer can be written as the sum of four squares?

COMPUTATIONAL EXERCISES

8.1. Show that the following equations have no integer solutions.

a. $4x^3 - 7y^3 = 2003$

b. $x^3 + y^4 = 7$

 **8.2.** Use the Chinese Remainder Theorem at primes 3 and 5 to find an integer solution to the polynomial equation

$$x^3 - 11x^2 + 37x - 400 = 7.$$

8.3. Use the Chinese Remainder Theorem to find an integer solution to the following equations.

a. $x^3 - 33x^2 + 35x - 96$

b. $x^4 - 3562x^3 - 2x^2 + 7133x - 32058$


8.4. Represent the following integers as a sequence

$$a_0 + a_1p + a_2p^2 + \cdots$$

with $0 \leq a_i < p$ for the given prime p .

a. $n = 1365$ for $p = 3$

b. $n = 231365$ for $p = 5$

 **8.5.** Use Hensel's Lemma modulo 3 to find an integer solution to

$$x^3 - 119x - 22 = 0.$$

8.6. Find an integer solution to the following equations by applying Hensel's Lemma at an appropriate prime.

a. $x^3 - 67x^2 - 283x - 71 = 0$

b. $2x^4 - 1462x^3 + 41x^2 - 29968x - 2193 = 0$

8.7. Find all the Pythagorean triples (x, y, z) with $z \leq 40$ and $\gcd(x, y, z) = 1$.

8.8. Find solutions to

$$x^3 + y^3 = z^3 - n$$

for $-1 \leq n \leq 2$. These are called *Fermat near misses* since they are almost solutions to the Fermat equation.

8.9. Find the fundamental unit of $\mathbb{Q}(\sqrt{17})$.

8.10. Find the fundamental solution of $x^2 - 11y^2 = 1$.

8.11. Some integers can be written as the sum of two squares (see section 6.4). For example, $2 = 1^2 + 1^2$ and $5 = 1^2 + 2^2$.

a. Find the smallest integer that can be written as the sum of two positive integer squares in two different ways.

b. Find the smallest positive integer that can be written as the sum of two positive integer cubes in two different ways.

8.12. Fix a positive integer a and consider the sequence of numbers

$$\{a^n + 1 : n \in \mathbb{N}\}.$$

- a. For $a = 2$, which members of the sequence with $1 \leq n \leq 20$ can be written as the sum of two squares?
- b. Find a value of a where all members of the sequence with $1 \leq n \leq 20$ can be written as the sum of two squares.

8.13. Compute the number of ways you can write n as the sum of four squares for $1 \leq n \leq 100$. Count only distinct ways; i.e., $5 = 2^2 + 1^2$ and $5 = 1^2 + 2^2$ only count as 1.

Which number can be written in the most different ways? How many ways can it be written?

8.14. Find the first four pairs (m, n) such that the n th triangle number equals the m th square:

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

8.15. Sums of cubes.

- a. Write each integer $1 \leq n \leq 40$ as the sum of cubes.
- b. Find an integer that cannot be written as the sum of 8 (or fewer) cubes.

8.16. Sums of fourth powers.

- a. Write each integer $1 \leq n \leq 40$ as the sum of fourth powers.
- b. Find an integer that cannot be written as the sum of 18 (or fewer) fourth powers.

8.17. Find all integer solutions to $x^3 + y^3 = z^2$ of height at most 10.

8.18. Find the positive integer solutions to the Catalan equation

$$x^n - y^m = 1$$

for $n, m \geq 2$.

8.19. Find all solutions up to height 200 for

$$x^2 + 7 = y^m$$

with $2 \leq m \leq 15$.

THEORETICAL EXERCISES

8.20. Prove that $w^3 + y^3 + z^3 = x^2 + y^2 + z^2$ has infinitely many integer solutions.

8.21. Prove that the only integer solutions of $y^2 + 2 = x^3$ are $(x, y) = (3, \pm 5)$.

8.22. Let p be prime. Prove that

$$x^3 + py^3 + p^2z^3 = 0$$

has no (integer) solutions other than $(0, 0, 0)$.

8.23. Let $f(x)$ be a polynomial with integer coefficients that has an integer root. Prove there are only finitely many primes that do not satisfy the hypotheses of Hensel's Lemma.

8.24. Reformulate and prove Hensel's Lemma so that if $f(x_i) \equiv 0 \pmod{p^k}$, then $f(x_{i+1}) \equiv 0 \pmod{p^{2k}}$.

8.25. Reformulate and prove Hensel's Lemma for a system of two variable equations,

$$f(x, y) = g(x, y) = 0.$$

8.26. Determine the conditions on (a, b) in the formula for Pythagorean triples in Theorem 8.17 so that you get only the primitive triples (those whose greatest common divisor is 1).

8.27. Prove that the only integer solutions of $x^4 + 9 = y^2$ are $(x, y) = (\pm 2, \pm 5)$ and $(0, \pm 3)$.

8.28. Prove that there is no Pythagorean triple in which the smaller two terms are squares.

8.29. Use descent to prove that $x^3 + 2y^3 = 4z^3$ has no integer solutions other than $(0, 0, 0)$.

8.30. Use Pythagorean triples (and descent) to prove that the only integer solutions to $y^2 = x^4 + 1$ are $(0, \pm 1)$.

8.31. Prove that Fermat's Last Theorem is equivalent to

$$x^p + y^p = z^p \quad \text{has no nonzero integer solutions for all primes } p > 2.$$

8.32. If $d \neq 1$ is a square integer, prove that $(\pm 1, 0)$ are the only integer solutions to $x^2 - dy^2 = 1$.

8.33. If $d \leq 0$, prove that

$$x^2 - dy^2 = 1$$

has only finitely many integer solutions.

8.34. For d a positive and squarefree integer, prove that in $\mathbb{Q}(\sqrt{d})$, the fundamental unit $x + y\sqrt{d}$ has $x, y > 0$.

8.35. Let m and n be positive integers. Recall that a triangular number is of the form $1 + 2 + 3 + \cdots + n$. The following equation represents when the n th triangle number equals the m th square:

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

Make a change of variables to convert this equality to a Pell's equation, and find the fundamental solution.

8.36. For an integer $d > 0$, prove that there are infinitely many units in $\mathbb{Q}(\sqrt{d})$.

8.37. Let n be a positive integer. Prove that if $n \equiv 7 \pmod{8}$, then n cannot be expressed as the sum of three squares.

EXPLORATION EXERCISES

8.38 (Generalized four-squares). Consider the generalized four-square problem: For which nonnegative integers a, b, c , and d can we find integers (x_1, x_2, x_3, x_4) for every positive integer n such that

$$(30) \quad n = ax_1^2 + bx_2^2 + cx_3^2 + dx_4^2?$$

Note that when $a = b = c = d = 1$, this is Lagrange's four-square theorem and there are always solutions.

- The case $(a, b, c, d) = (1, 2, 5, 5)$ has a single integer n for which there is no (x_1, x_2, x_3, x_4) . Find it.
- For tuples (a, b, c, d) with $1 \leq a, b, c, d \leq 5$ for how many can you solve equation (30) for all n ?
- For the tuples that fail, can you separate the ones that fail for many n from the ones that fail for few n ?
- Find some other tuples (a, b, c, d) that work for all n or almost all n .

8.39 (Number of solutions modulo p). Given an equation $f(x_1, \dots, x_n) = 0$, there are at most p^n solutions modulo p since there are p possible choices for each x_i . However, we expect there to be significantly fewer since “most” values are not solutions. Define N_p as the number of solutions modulo p . Can you find polynomials with many solutions and polynomials with few solutions for certain primes?

Example 8.51. $y^2 = x^5 - x$ satisfies $|N(p) - p| < 4\sqrt{p}$. In other words, there are neither too many nor too few solutions for any prime.

In two variables, $f(x, y) = 0$ represents a curve in the plane. The following are some interesting families of curves you may consider.

- Lines: $y = mx + b$ for some integers m, b .
- Circles: $y^2 + x^2 = c^2$ for some integer c .
- Parabolas: $y^2 = mx^2$ for some integer m .
- Elliptic curves: $y^2 = x^3 + c$ for some integer c .
- Hyperelliptic curves: $y^2 = x^d + c$ for $d > 3$ and some integer c .
- Fermat curves: $y^n + x^n = z^n$ for positive integers n .

8.40 (p -adic numbers). The notion of writing an integer as a sum of powers of a prime number

$$a_0 + a_1 \cdot p + a_2 \cdot p^2 + \cdots + a_k \cdot p^k = \sum_{i=0}^k a_i p^i$$

is only a short step away from the notion of a p -adic number.

Definition 8.52. Let p be a prime number. A p -adic integer is a formal series

$$\sum_{i=0}^{\infty} a_i p^i = a_0 + a_1 p + a_2 p^2 + \cdots \quad \text{with} \quad 0 \leq a_i < p \text{ for all } i.$$

We denote the set of p -adic integers as \mathbb{Z}_p .

- a. Pick a few integers n and primes p and find their representation as a p -adic integer.

On the other hand, negative numbers pose a problem since $0 \leq a_i < p$ for all i . However, we already have a way around this issue with Hensel's Lemma, which allows us to "lift" a root modulo p to a root modulo p^n for any n (under certain conditions).

- b. Use Hensel's Lemma to find the p -adic representation of -1 for a prime p . Does this work for all primes?
- c. Choose a few other integers (positive and negative), and find their p -adic representation using Hensel's Lemma.

We can now write all integers as p -adic integers, but what about rational numbers? Let $x = \frac{a}{b}$ be a rational number. Since we are working modulo p , we only have difficulty when b is not relatively prime to p . However, in that case, $b = p^k m$ for some finite integer k with $\gcd(m, p) = 1$. Since we have only finitely many powers of p in the denominator for any given b , we expand our notion of p -adic integer to include finitely many negative powers of p .

Definition 8.53. Let p be a prime number. A p -adic number is a formal series

$$\sum_{i=-k}^{\infty} a_i p^i = \frac{a_{-k}}{p^k} + \frac{a_{1-k}}{p^{k-1}} + \cdots + a_0 + a_1 p + a_2 p^2 + \cdots \quad \text{with} \quad 0 \leq a_i < p \text{ for all } i$$

for some positive integer k .

We denote the set of p -adic numbers as \mathbb{Q}_p .

For $x = \frac{a}{b}$ with $b = p^k m$, we have $p^k x = \frac{a}{m}$ with $\gcd(m, p) = 1$. We can use Hensel's Lemma to find a p -adic representation of $\frac{a}{m}$ and then multiply by p^{-k} to get a representation of $\frac{a}{b}$.

- d. Find the p -adic representation of $\frac{1}{2}$.
- e. Choose a few other rational numbers, and find their p -adic representation using Hensel's Lemma.

All rational numbers can now be represented as p -adic numbers for all primes p . What about irrational numbers?

- f. Can you determine how to represent $\sqrt{2}$ as a p -adic number? Can you do it for all primes?
- g. What about other irrational numbers?
- h. Are there " p -adic" solutions to equations?

8.41 (Fermat near misses). We are interested in triples (x, y, z) that “almost” satisfy the equation for Fermat’s Last Theorem

$$x^n + y^n = z^n.$$

However, the notion of “almost” is open to interpretation. One definition is to consider the equation

$$(31) \quad x^n + y^n = z^n + k$$

for small integers k .

a. For which pairs (n, k) can you solve equation (31)?

Another way to think about this problem is that $x^n + y^n$ is “almost” an n th power. So instead of fixing k , you could ask for pairs (x, y) such that the difference between $x^n + y^n$ and z^n is particularly small in comparison to x and y . For example, the equation

$$1782^{12} + 1841^{12} = 1922^{12}$$

appears in the 1995 “Halloween” episode of *The Simpsons*. While the equation is not actually correct, it is “almost” correct. In particular,

$$\sqrt[12]{1782^{12} + 1841^{12}} \approx 1921.999999956.$$

b. For various integers $n > 2$, find pairs (x, y) so that $x^n + y^n$ is almost an n th power. You’ll need to determine what “almost” means, but since

$$1^n + x^n \approx x^n$$

for any x , you do not want to consider $(1, x, x)$ as a Fermat near miss.

8.42 (Pell-like equations). Pell’s equation has a connection to fundamental units in quadratic number fields, but as a Diophantine equation, we could consider generalizations. For example, consider the following two problems.

- a.** Determine when $x^2 - dy^2 = -1$ is solvable in integers.
- b.** What about $x^2 - dy^2 = n$ for other integers n ?

8.43 (Congruent number problem).

Definition 8.54. A rational number n is called a *congruent number* if there is a right triangle with rational side lengths and area n .

In other words, n is a congruent number if there exists $a, b, c \in \mathbb{Q}$ with the following properties.

Property 1: $a, b, c > 0$

Property 2: $a^2 + b^2 = c^2$

Property 3: $\frac{ab}{2} = n$

Since scaling the sides (a, b, c) of a rational right triangle by a factor d ,

$$(a, b, c) \mapsto (da, db, dc),$$

changes the area by d^2 , we may clear denominators of any congruent number and look for integers n that are congruent numbers.

- a. Given a Pythagorean triple (a, b, c) , compute the associated congruent number. What happens to n when you consider the triple (da, db, dc) ?
- b. Construct some congruent numbers from Pythagorean triples.
- c. Which congruent numbers can you construct from Pythagorean triples?
- d. Are there congruent numbers that result from more than one triple?
- e. Fix a positive integer B . Which integers $1 \leq n \leq B$ are congruent numbers?
- f. How many positive integers are not congruent numbers?

8.44 (Waring problem). We know that $g(2) = 4$, $g(3) = 9$ and $g(4) = 19$.

- a. Conjecture some other values for $g(k)$.
- b. Try to prove upper or lower bounds for $g(k)$ in terms of k .

You can also consider the Waring problem modulo primes. In particular, find the smallest positive integer n so that every residue class modulo p can be written as a sum of n k th powers.

- c. Find n for several specific k and small p .
- d. For fixed k , is there a value n that works for all p ?
- e. What about composite moduli?

Elliptic Curves

1. Introduction

In this chapter we study a particular family of Diophantine equations called *elliptic curves*. The subject of elliptic curves is vast, and we content ourselves with examining just a single aspect of the theory: rational points of finite order (Question 9.16).

Definition 9.1. An *elliptic curve* is the set of points (x, y) such that

$$y^2 = x^3 + ax + b$$

for constants a and b with $4a^3 + 27b^2 \neq 0$.

The condition $4a^3 + 27b^2$ is exactly the condition that the curve is nonsingular, i.e., when $x^3 + ax + b$ has distinct roots (Theoretical Exercise 9.15). For example, $y^2 = x^3$ would not be an elliptic curve because x^3 has the root 0 repeated three times. Geometrically, the curve $y^2 = x^3$ has what is called a *cusp* at $(0, 0)$; see Figure 9.1.

To determine when $x^3 + ax + b$ has multiple roots, we use the notion of a discriminant. You may recall that associated to a quadratic equation $ax^2 + bx + c$ is the *discriminant*: $b^2 - 4ac$. It occurs under the square root in the quadratic formula

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Its significance is that the discriminant is 0 if and only if the two roots of $ax^2 + bx + c$ are the same, i.e., if

$$ax^2 + bx + c = \left(x + \frac{b}{2a}\right)^2.$$

For a cubic polynomial, we can also find an expression, the *discriminant*, which is not zero if and only if the three roots are all distinct. The discriminant of the cubic $ax^3 + bx^2 + cx + d$ is given by

$$b^2c^2 - 4ac^3 - 4b^3d - 27a^2d^2 + 18abcd.$$

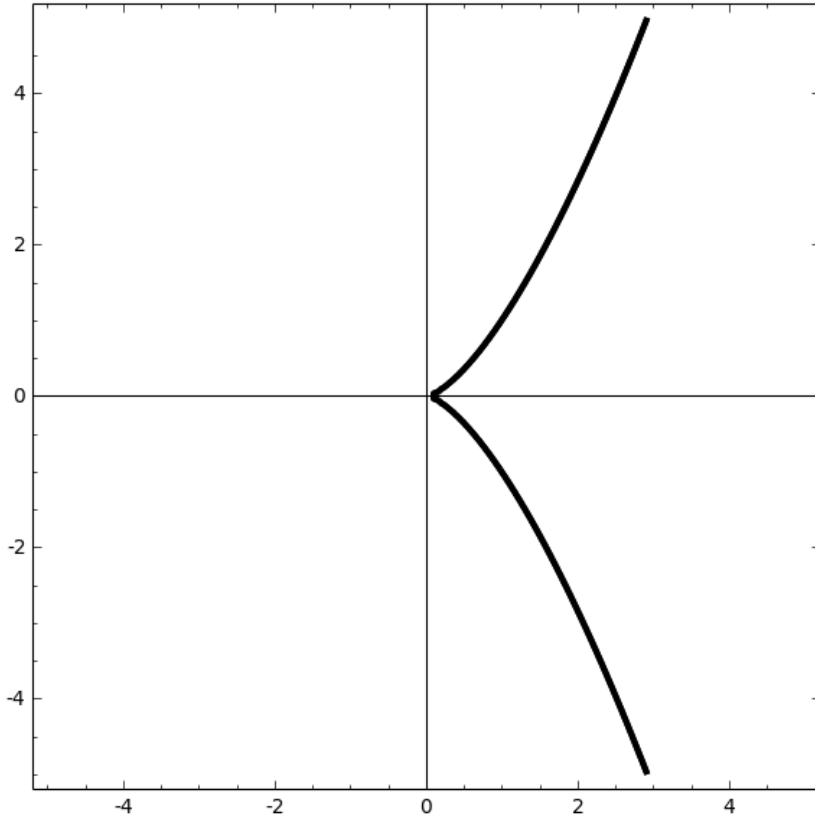


Figure 9.1. Cusp

The cubic polynomial associated to an elliptic curve is just $x^3 + ax + b$, which has a much simpler discriminant.

Definition 9.2. Given an elliptic curve $E : y^2 = x^3 + ax + b$, we define the *discriminant* of E to be

$$D = 4a^3 + 27b^2.$$

Example 9.3. The discriminant of

$$y^2 = x^3 + 2x + 3$$

is

$$D = 4 \cdot 2^3 + 27 \cdot 3^2 = 275.$$

Example 9.4. The discriminant of the cusp equation

$$y^2 = x^3$$

is

$$D = 0,$$

so this is not an elliptic curve.

Elliptic curves appear in Diophantus's *Arithmetica*. For example, we might set $a = 0$ to have the equation

$$(32) \quad y^2 = x^3 + b,$$

which describes the pairs of squares and cubes that differ by the constant b . For example, if $b = 1$, there is the solution

$$3^2 = 2^3 + 1.$$

In 1650 Fermat challenged fellow mathematicians to prove that the only integer solutions to $y^2 = x^3 - 2$ are $(3, \pm 5)$. He also claimed to have a proof of this fact, which he never published. In 1730 Euler gave a proof that assumed algebraic integers in $\mathbb{Q}(\sqrt{-2})$ have a unique factorization, which was not yet known at that time. Finally, Thue¹ proved in 1909 that for any b , there are only finitely many integer solutions to equation (32).

Investigation 9.5. Consider the elliptic curve $y^2 = x^3 + b$ for various b . These are called *Mordell curves* after Louis Mordell.²

- (a) Is there an integer solution (x, y) for every b ?
- (b) Can you find a b with many integer solutions?
- (c) How many more points do you find if you consider rational solutions instead of just integer solutions?

If we consider rational solutions, it turns out that there are infinitely many solutions for some values of b . For example, $b = -2$ has infinitely many rational solutions such as

$$\left(\frac{129}{100}, -\frac{383}{1000}\right) \quad \text{and} \quad \left(\frac{164323}{29241}, -\frac{66234835}{5000211}\right).$$

Even more surprisingly, just as with Pythagorean triples in section 8.3 and with Pell's equation in section 8.5, given one solution, we are able to generate additional solutions! In the case of elliptic curves, this is done with a geometric construction: *addition of points*.

We pose the following main question.

Question 9.6. Given an elliptic curve E , how many integer and rational points does it have?

Our first step in answering this question is to see how to create a new solution from a known solution.

¹Axel Thue (1863–1922) was a Norwegian mathematician.

²Louis Joel Mordell (1888–1972) was a British mathematician.

2. Addition of Points

Given an elliptic curve E , let $P = (x_0, y_0)$ and $Q = (x_1, y_1)$ be two different rational points on the curve, i.e., $x_0, x_1, y_0, y_1 \in \mathbb{Q}$ and

$$y_i^2 = x_i^3 + ax_i + b \quad \text{for } i \in \{0, 1\}.$$

These two points are connected by a unique line

$$(y - y_0) = \frac{y_1 - y_0}{x_1 - x_0}(x - x_0),$$

whose slope is rational if $x_0 \neq x_1$. We will ignore the case $x_0 = x_1$ for the moment. The line intersects E at the two points P and Q and, since E is given by a degree 3 equation in x , the line has a third intersection point with the curve. Take this third point as $R = (x_2, y_2)$. Since E is defined by an equation of the form $y^2 = x^3 + ax + b$, then $R' = (x_2, -y_2)$ is also on the curve.

Definition 9.7. Using the notation from the previous paragraph, we define

$$P + Q = R'.$$

Figure 9.2 shows this geometrically.

Example 9.8. Consider the elliptic curve

$$y^2 = x^3 - x + 4$$

and the two points $P = (-7/4, -5/8)$ and $Q = (0, 2)$. We take the line through those two points

$$y = \frac{3}{2}x + 2$$

and see that the line intersects the curve in the third point $R = (4, 8)$. Then the addition is

$$P + Q = (4, -8).$$

For the case $x_0 = x_1$ but $P \neq Q$, we have that P and Q are connected by a vertical line, and we imagine another point \mathcal{O} called *the point at infinity*. Think of \mathcal{O} as the point where all vertical lines meet at the horizon.

Definition 9.9. If $P = (x_0, y_0)$ and $Q = (x_0, -y_0)$, we define $P + Q = \mathcal{O}$.

We have only one last case to consider: $P = Q$. In this case, we consider the tangent line to P , which intersects the elliptic curve in one other point, and proceed as above. What is interesting about this process is that points now have an addition law with \mathcal{O} taking the place of 0 (we, in fact, get an abelian group). In other words,

$$P + \mathcal{O} = P$$

and

$$P + Q = Q + P.$$

We can also get negative values. In particular, if P and Q are joined by a vertical line, we have

$$P + Q = \mathcal{O}.$$

Then for $P = (x_0, y_0)$, we define $-P = (x_0, -y_0)$ so that

$$P + (-P) = \mathcal{O}.$$

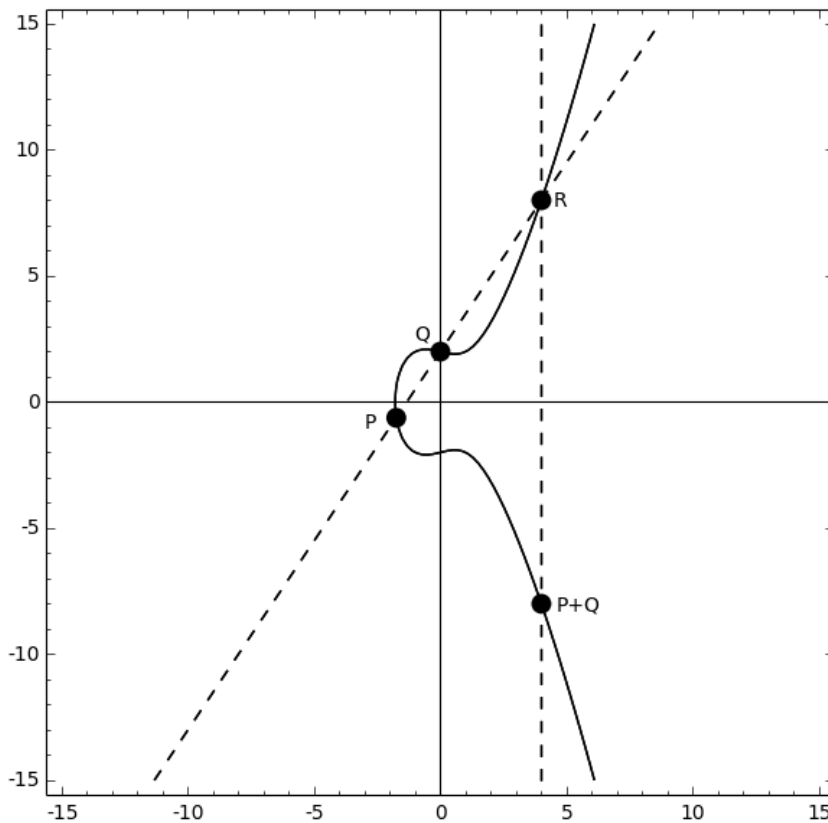


Figure 9.2. Addition of points

Question 9.10. If P and Q are integer or rational points on an elliptic curve, is $P + Q$ an integer or a rational point?

To answer this question, we need to look more closely at the process. We have an elliptic curve $E : y^2 = x^3 + ax + b$ with $a, b \in \mathbb{Q}$ and two rational points $P = (x_0, y_0)$ and $Q = (x_1, y_1)$. We have two cases to consider: P and Q are distinct or $P = Q$.

For P and Q distinct, they are connected by a unique line whose slope is rational. Write the line as

$$y = mx + t,$$

where m and t are rational numbers. To find the third point, we substitute the line equation into the equation for E

$$(mx + t)^2 = x^3 + ax + b$$

to have

$$(33) \quad x^3 - m^2x^2 + (a - 2tm)x + b - t^2 = 0.$$

This is a cubic equation so it has three roots: the three points on the line. Two of those points are P and Q , which have x -coordinates x_0 and x_1 . Label the third root as x_2 . The three roots of a cubic polynomial determine the coefficients in the following way:

$$(34) \quad (x-x_0)(x-x_1)(x-x_2) = x^3 - (x_0+x_1+x_2)x^2 + (x_0x_1+x_0x_2+x_1x_2)x - x_0x_1x_2.$$

By equating the coefficient of x^2 in equation (33) and the right-hand side of equation (34), we know that

$$x_0 + x_1 + x_2 = m^2$$

so that

$$x_2 = m^2 - x_0 - x_1.$$

Consequently, x_2 is rational. Knowing the x -coordinate, we can find the y -coordinate from the line equation as

$$y_2 = mx_2 + t,$$

which must also be rational.

Example 9.11. For

$$y^2 = x^3 + 8,$$

the two points on the curve $P = (46, 312)$ and $Q = (1, 3)$ are joined by a line with slope

$$m = \frac{312 - 3}{46 - 1} = \frac{309}{45} = \frac{103}{15},$$

so that

$$x_2 = \left(\frac{103}{15}\right)^2 - 46 - 1 = \frac{34}{225}.$$

We compute

$$t = y_1 - mx_1 = 3 - \frac{103}{15} = -\frac{58}{15},$$

so that

$$y_2 = \frac{103}{15} \cdot \frac{34}{225} - \frac{58}{15} = -\frac{9548}{3375}.$$

Finally, we take the negative of the y component to get

$$P + Q = \left(\frac{34}{225}, \frac{9548}{3375}\right).$$

You are invited to verify that the point $P + Q$ is, in fact, on the curve by checking that

$$\left(\frac{9548}{3375}\right)^2 = \left(\frac{34}{225}\right)^3 + 8.$$

If $P = Q$, we have to use the tangent line whose slope (by implicit differentiation) is

$$m = \frac{3x^2 + a}{2y}.$$

Since x , y , and a are rational, so is m . We can again work through the point addition process to have

$$x_2 = m^2 - 2x_0$$

is rational, so y_2 is also rational.

Example 9.12. Consider the elliptic curve

$$y^2 = x^3 - 2$$

which has the point $P = (3, 5)$. We find the tangent line at $(3, 5)$ to have slope

$$m = \frac{f'(x)}{2y} = \frac{27}{10}$$

with equation

$$y - 5 = \frac{27}{10}(x - 3).$$

We compute

$$x_2 = \left(\frac{27}{10}\right)^2 - 2 \cdot 3 = \frac{129}{100}$$

so that the other intersection point is

$$\left(\frac{129}{100}, \frac{383}{1000}\right).$$

So we have

$$P + P = \left(\frac{129}{100}, -\frac{383}{1000}\right).$$

Since all the values and equations are rational, it is clear that the addition of two rational points is either \mathcal{O} or rational. We summarize what we know in the following theorem.

Theorem 9.13. *Let $E : y^2 = x^3 + ax + b$ be an elliptic curve, and let $P = (x_0, y_0)$ and $Q = (x_1, y_1)$ be points on E with rational coordinates. Then*

$$\begin{cases} P + Q = \mathcal{O} & \text{if } Q = -P, \\ P + Q = (x_2, -(mx_2 + t)) & \text{otherwise,} \end{cases}$$

where

$$y = mx + t = \begin{cases} \text{the line through } P, Q & P \neq Q, \\ \text{the tangent line to } E \text{ at } P & P = Q, \end{cases}$$

and

$$x_2 = m^2 - x_0 - x_1.$$

Investigation 9.14. Consider the family of Mordell curves

$$y^2 = x^3 + b, \quad b \neq 0.$$

Use a combination of brute force searching and point addition to find the rational points for each curve. The counts include \mathcal{O} .

- For $b = 0$, there are six points.
- For $b = 2$ there are four points with $H(x) < 10,000$.
- For $b = 24$ there are 27 points with $H(x) < 10,000$.
- Can you find another value of b with “many” small points?

You may have noticed in Investigation 9.14 that while each addition of points results in a rational point on the elliptic curve, it may not be a “new” rational point. In particular, there are elliptic curves with only finitely many rational points.

Definition 9.15. We define

$$nP = P + \cdots + P \quad n\text{-times.}$$

Question 9.16. Given an elliptic curve E , are there any rational points for which $nP = P$ for some positive integer n ?

Answering Question 9.16 will be our main focus for the rest of this chapter.

Example 9.17. Consider the elliptic curve

$$E : y^2 = x^3 + x + 2$$

and the point $P = (1, 2)$. Computing the addition of points, we find

$$\begin{aligned} P &= (1, 2), \\ 2P &= (-1, 0), \\ 3P &= (1, -2), \\ 4P &= \mathcal{O}, \\ 5P &= (1, 2) = P, \end{aligned}$$

so that the set of points

$$\{nP : n \in \mathbb{Z}\}$$

is a finite set of points. It turns out that these four points are the only four rational points on E .

Investigation 9.18. While Question 9.16 is looking for P with $nP = P$, we can ask a related question. Given points P and Q , can you find n such that

$$nP = Q,$$

assuming that such an n exists? This is the elliptic curve equivalent of the discrete log problem from section 4.3.

- (a) For $E : y^2 = x^3 + 9x + 17$ working modulo 23, solve the discrete log problem for $P = (16, 5)$ and $Q = (4, 5)$.
- (b) For $E : y^2 = x^3 + 2x + 7$ working modulo 353, solve the discrete log problem for $P = (2, 103)$ and $Q = (343, 80)$.
- (c) Create your own curves and points and explore the discrete log problem for elliptic curves.

If we want to use elliptic curves for public key cryptography, we need the discrete log problem to be difficult.

- (d) Do you think the discrete log problem for elliptic curves is easier or harder than the discrete log problem for modular arithmetic?

3. Points of Finite Order

Definition 9.19. We say that a point P on an elliptic curve E is a *torsion point* of order n if

$$nP = \mathcal{O} \quad \text{or equivalently} \quad (n+1)P = P.$$

We say a point P that is not a torsion point has *infinite order*.

Question 9.16 is asking about the existence of rational torsion points. To get started, let us try to find points of particular orders.

Order One: The points of order one are the points with

$$P = \mathcal{O}.$$

Clearly, the only such point is \mathcal{O} itself.

Order Two: The points of order two are the points with

$$2P = \mathcal{O}$$

or

$$P = -P.$$

Writing $P = (x_0, y_0)$, this is the same as

$$(x_0, y_0) = (x_0, -y_0).$$

Hence, we must have $y_0 = 0$ and the order two points are the roots of

$$x^3 + ax + b.$$

Example 9.20. Consider the curve

$$y^2 = x^3 - x + 2.$$

Then for $y = 0$, we have

$$0 = x^3 - x + 2 = (x - 1)(x^2 + x + 2).$$

So the point $(1, 0)$ is order two, i.e., the tangent line at $(1, 0)$ is vertical.

Order > 2 : For higher order points, we can find explicit formulas for the addition of points (Computational Exercise 9.1), so we could solve increasingly complicated equations to find the order n points. For example, we could find order three points by solving

$$2P = -P$$

or

$$\frac{x^4 - 2ax^2 - 8bx + a^2}{4x^3 + 4ax + 4b} = x,$$

which is

$$x^4 - 2ax^2 - 8bx + a^2 = 4x^4 + 4ax^2 + 4bx,$$

which is

$$3x^4 + 6ax^2 + 12bx - a^2 = 0.$$

Example 9.21. Consider the elliptic curve

$$E : y^2 = x^3 + 4.$$

To find the points of order three we need a root of

$$3x^4 + 6ax^2 + 12bx - a^2 = 3x^4 + 48x = 3x(x^3 + 16),$$

which is $x = 0$. Then for $P = (0, 2)$, we check

$$P = (0, 2),$$

$$2P = (0, -2) = -P,$$

$$3P = \mathcal{O},$$

so that P has order three.

We could continue these explicit calculations and find the points of each possible order. However, there is a serious issue with this method: the algebra is getting more and more complicated and will quickly become infeasible as the order increases.

There is a second potential issue: If the goal is to find *all* rational torsion points, how do we know when to stop? Without a bound on the possible order of a torsion point, we would have to continue indefinitely checking all orders $n \in \mathbb{N}$. However, there is a very deep theorem of Barry Mazur³ from 1977 that bounds the order of a rational torsion point at 12, i.e., all rational torsion points with $nP = \mathcal{O}$ satisfy $n \leq 12$.

With the order bounded at 12, we could conceivably do the algebra (with the aid of a computer algebra system), but we would like a more efficient method. Somewhat surprisingly, the answer turns out to be related to integer points.

Investigation 9.22. Consider the family of Mordell curves

$$y^2 = x^3 + b, \quad b \neq 0.$$

- (a) Find all the torsion points for the curves with $b \in \{0, 2, 24\}$, i.e., which of the points from Investigation 9.14 were torsion?
- (b) What value of b gives you the largest number of torsion points?

4. Integer Points and the Nagel–Lutz Theorem

Question 9.23. How many integer points can be on an elliptic curve?

We start by writing the discriminant, $D = 4a^3 + 27b^2$, as a linear combination of $f(x) = x^3 + ax + b$ and its derivative $f'(x) = 3x^2 + a$.

Lemma 9.24. Let $E : y^2 = f(x) = x^3 + ax + b$, then we can write

$$D = (27b - 18ax)f(x) + (6ax^2 - 9bx + 4a^2)f'(x).$$

Proof. Direct computation. □

³Barry Charles Mazur (1937–) is an American mathematician.

Note that Lemma 9.24 is valid with x as an indeterminate, so it is also valid for any value of x .

Lemma 9.25. *Given an elliptic curve $E: y^2 = x^3 + ax + b$ and a point $P = (x_0, y_0)$ on E , if P and $2P$ have integer coordinates, then $y_0 = 0$ or $y_0 \mid D$.*

Proof. Assume that P and $2P$ have integer coordinates and assume that $y_0 \neq 0$. We need to show that $y_0 \mid D$. If $y_0 \neq 0$, then $2P \neq \mathcal{O}$, and we write $2P = (x_2, y_2)$. By the formula for $2P$ (Theorem 9.13), we have

$$2x_0 + x_2 = m^2,$$

where $m = \frac{f'(x_0)}{2y_0}$ is the slope of the tangent line at P . Since x_0 and x_2 are both integers, we must have m is an integer. So we must have $2y_0 \mid f'(x_0)$ and, hence, $y_0 \mid f'(x_0)$. Also, $y_0 = f(x_0)$, so $y_0 \mid f(x_0)$. Since D is a linear combination of f and f' by Lemma 9.24, we must have $y_0 \mid D$. \square

We will show that a point of finite order is an integer point by showing that no prime divides its denominators.

Lemma 9.26. *Let E be an elliptic curve, and let P be a rational point on E . Then there are integers x , y , and z such that*

$$P = \left(\frac{x}{z^2}, \frac{y}{z^3} \right).$$

Proof. Let $P = (x, y)$, and let p be a prime. Write $x = \frac{u}{vp^t}$ for some integers u , v , and t with $\gcd(uv, p) = 1$ and $y = \frac{c}{dp^s}$ for some integers c , d , and s with $\gcd(cd, p) = 1$, where we may have $s, t = 0$. Then, from the equation of E , we have

$$\frac{c^2}{d^2p^{2s}} = \frac{u^3}{v^3p^{3t}} + a\frac{u}{vp^t} + b,$$

so we have

$$\frac{c^2}{d^2p^{2s}} = \frac{u^3 + av^2p^{2t} + bv^3p^{3t}}{v^3p^{3t}}.$$

We see that p does not divide the numerator of the right-hand side so that $2s = 3t$. Thus, a prime which divides the denominator of x or y must also divide the denominator of the other. In particular, we must have p^2 divides the denominator of x and p^3 divides the denominator of y . \square

Example 9.27. Recall the rational points we have seen in examples so far in this chapter.

- $y^2 = x^3 - x + 4$:

$$\left(-\frac{7}{4}, -\frac{5}{8} \right) = \left(-\frac{7}{2^2}, -\frac{5}{2^3} \right).$$

- $y^2 = x^3 - 2$:

$$\left(\frac{129}{100}, -\frac{383}{1000} \right) = \left(\frac{129}{10^2}, -\frac{383}{10^3} \right).$$

To show that the powers are not always exactly 2 and 3 consider

$$y^2 = x^3 - 2$$

and the point

$$\left(\frac{164323}{29241}, -\frac{66234835}{5000211} \right) = \left(\frac{164323}{3^4 \cdot 19^2}, -\frac{66234835}{3^6 \cdot 19^3} \right).$$

Definition 9.28. For $t > 0$ an integer, define sets

$$E(p^t) = \{(x, y) \mid p^{2t} \text{ divides the denominator of } x \text{ or} \\ p^{3t} \text{ divides the denominator of } y\}.$$

Our goal is to show a point of finite order cannot be in $E(p)$ for any prime p . Note that

$$E(\mathbb{Q}) \supset E(p) \supset E(p^2) \supset \cdots,$$

where $E(\mathbb{Q})$ is the set of all rational points on E .

Proposition 9.29. *Let E be an elliptic curve, let p be a prime number, and let $R_p \subset \mathbb{Q}$ be the set of rational numbers with denominators relatively prime to p .*

- (a) *If $P, Q \in E(p^t)$, then $P + Q \in E(p^t)$.*
- (b) *The map*

$$\phi : E(p^t) \pmod{E(p^{3t})} \rightarrow p^t R_p \pmod{p^{3t}} \\ (x, y) \mapsto \frac{x}{y}$$

is one-to-one and respects the addition of points, i.e.,

$$\phi(nP) \equiv n\phi(P).$$

The second part of the proposition is really a statement about quotient rings, and it looks rather intimidating in this notation, so let's examine it a little more closely before proving the proposition. The domain of the map ϕ are those points in $E(p^t)$ that are not also in $E(p^{3t})$, in other words, the points (x, y) where p^{2t} divides x (or p^{3t} divides y), but p^{6t} does not (or p^{9t} does not). The image of the map is in the integers modulo p^{3t} . This is possible since the denominator of any element in R_p is relatively prime to p , so it will have an inverse modulo p^{3t} ; hence, any rational number $p^t R_p$ can be expressed modulo p^{3t} . The content of the statement is that this map is one-to-one and respects addition of points.

Proof. Define new coordinates

$$u = \frac{x}{y} \quad v = \frac{1}{y},$$

so that if p^{2t} divides the denominator of x and p^{3t} divides the denominator of y , then p^t divides u and p^{3t} divides v .

In these new coordinates, the elliptic curve becomes

$$\frac{1}{v^2} = \frac{u^3}{v^3} + a \frac{u}{v} + b,$$

which simplifies to

$$v = u^3 + auv^2 + bv^3.$$

Notice also that \mathcal{O} becomes $(0, 0)$ in uv -coordinates (and the points of order two on E become \mathcal{O} in uv -coordinates). So we need to be a little careful with the addition law. First, notice that lines go to lines: if $ax + by = c$, then $au + b = cv$ so that lines in the xy -coordinates correspond to lines in uv -coordinates. So a line connecting P and Q in xy -coordinates goes to a line connecting their image in uv -coordinates. The last step where we take $-R$ instead of R is a vertical line connecting R with \mathcal{O} , so in uv -coordinates we find the third point on the line through $(0, 0)$ instead of \mathcal{O} . To prove what we need to prove, we need to work through the details of addition of points in the uv -coordinates.

Let $P' = (u_0, v_0)$ and $Q' = (u_1, v_1)$. We need the line through these two points and (after much calculation) find that

$$m = \frac{v_1 - v_0}{u_1 - u_0} = \frac{u_1^2 + u_0u_2 + u_0^2 + av_1^2}{1 - au_0(v_1 + v_0) - b(v_1^2 + v_1v_0 + v_0^2)},$$

or if $P = Q$, we take the tangent line to get

$$m = \frac{3u_0^2 + av_0^2}{1 - 2au_0v_0 - 3bv_0^2}.$$

We write the line as

$$v = mu + w.$$

The point of this change of variables is to get this 1 in the denominator of the slope so that if p divides u and v , then p does not divide the denominator of the slope. Similar to the original addition method, we substitute back into the uv -equation for the curve and look at the coefficient of u^2 to end up with

$$u_0 + u_1 + u_2 = -\frac{2amw + 3bm^2w}{1 + am^2 + bm^3}.$$

The final step, where we find the third point in the intersection of the line through (u_2, v_2) and $(0, 0)$ with the curve yields, the point $(-u_2, -v_2)$.

Now we consider the power of p dividing the new point $(-u_2, -v_2)$. If P', Q' are the image of points in $E(p^t)$, then p^t divides u and p^{3t} divides v so that p^t divides

$$u_0 + u_1 + u_2.$$

Since, p^t divides u_0 and u_1 , we must have p^t divides u_2 (and, hence, $-u_2$).

By the same assumptions, we also have

$$p^{2t} \mid m,$$

and writing

$$w = mu_0 - v_0,$$

we see that

$$p^{3t} \mid w.$$

Consequently, if p^t divides u_2 , then since

$$v_2 = mu_2 - w,$$

we have p^{3t} divides v_2 . What we have shown is that writing $u(P)$ for the u -coordinate of P' (the image of P in uv -coordinates)

$$u(P) + u(Q) - u(P + Q) \equiv 0 \pmod{p^{3t}}.$$

In particular, for the map $\phi : E(p^t) \rightarrow p^t R_p$, the points that map to 0 are the points that are in $E(p^{3t})$. The addition of points is respected modulo p^{3t} , so we have a one-to-one map

$$E(p^t) \pmod{E(p^{3t})} \rightarrow p^t R_p \pmod{p^{3t} R_p}. \quad \square$$

Theorem 9.30 (Nagel and Lutz). *Let E be an elliptic curve. The rational points (x, y) of finite order on E satisfy*

- (a) x, y are integers,
- (b) $y = 0$ or $y \mid D$, and
- (c) there are finitely many.

Proof.

- (a) Let $P \in E$ be a point of order n . Let p be a prime. Suppose $P \in E(p)$. Then we have that $P \in E(p^t)$ for finitely many t , since the denominators can only be divisible by a finite power of p . Let t be such that

$$P \in E(p^t) \quad \text{and} \quad P \notin E(p^{t+1}).$$

If $p \nmid n$, then by Proposition 9.29

$$\phi(nP) \equiv n\phi(P) \pmod{p^{3t}}.$$

Since $nP = \mathcal{O}$, we have $\phi(nP) = \phi(\mathcal{O}) = 0$. Then since $p \nmid n$, we must have

$$\phi(P) \equiv 0 \pmod{p^{3t}},$$

which contradicts the choice of t .

Now assume that $p \mid n$. Write $n = mp^k$ for $\gcd(m, p) = 1$ and consider $Q = p^k P$. Then $Q \in E(p)$ and Q has order m that is relatively prime to p , and we are in the previous case.

Hence, there is no prime p for which $P \in E(p)$, so P must be an integer point.

- (b) If P is finite order, then so is $2P$, both having integer coordinates by the first part, and so we must have $y = 0$ or $y \mid D$ by Lemma 9.25.
- (c) The integer D has only finitely many divisors, so there are only finitely many points that can have finite order. \square

It is important to recognize that the Nagel–Lutz theorem is not “if and only if”: not every divisor of D gives a point on the curve, and not every integer point is a torsion point.

Example 9.31. Consider the curve

$$y^2 = x^3 + 1,$$

which has

$$D = 27 = 3^3.$$

The positive divisors of D are

$$\{1, 3, 9, 27\}.$$

We examine each divisor.

- We let $y = 1$ and solve

$$1^2 = x^3 + 1$$

so that $x = 0$ to have the points

$$(0, \pm 1).$$

- We let $y = 3$ and solve

$$3^2 = x^3 + 1$$

so that $x = 2$, and we have the points

$$(2, \pm 3).$$

- For $y = 9$ and $y = 27$ there are no rational solutions.

Finally, we also consider $y = 0$ to have the point $(-1, 0)$. This is a collection of five points,

$$\{(0, \pm 1), (2, \pm 3), (-1, 0)\},$$

which together with \mathcal{O} give the six points of finite order.

Point	Order
\mathcal{O}	1
$(-1, 0)$	2
$(0, \pm 1)$	3
$(2, \pm 3)$	6

Definition 9.32. The set of points of finite order is called the *torsion subgroup* of E .

Mazur’s 1977 theorem not only bounds the order of a rational point as twelve, but also says the torsion subgroup must be one of only fifteen possibilities.

Theorem 9.33 (Mazur). *Let E be an elliptic curve. Then the set of points of finite order form one of the following fifteen groups:*

$$\begin{cases} \mathbb{Z}/n\mathbb{Z} & \text{for } 1 \leq n \leq 10 \text{ and } n = 12, \\ \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2n\mathbb{Z} & \text{for } 1 \leq n \leq 4. \end{cases}$$

We must be careful in turning the Nagel–Lutz theorem into an algorithm since it is not “if and only if”. In other words, there may be integer points with y dividing D that are not torsion. How do we check? With an upper bound on the number of torsion points from Mazur’s theorem, we could simply compute nP for each solution P for $1 \leq n \leq 12$. If nP is \mathcal{O} for some n , then P is torsion. Otherwise, P is not torsion. We describe finding the torsion subgroup in Algorithm 9.1.

We now have a good understanding of points of finite order, but since Nagel–Lutz is not “if and only if”, we still only partially understand integer points. A deep theorem of Siegel⁴ from 1928 says there are only finitely many integer points, but its proof is beyond our scope.

⁴Carl Ludwig Siegel (1896–1981) was a German mathematician.

Algorithm 9.1. Torsion Subgroup of an Elliptic Curve

Input: an elliptic curve E **Output:** a list of rational torsion points**Algorithm:**1: Compute D .2: For each divisor d of D (including 0), find integer solutions to

$$d^2 = x^3 + ax + b.$$

3: For each solution P , for $1 \leq n \leq 12$, check the following.a: If $nP = \mathcal{O}$, then P is order n .b: If $nP = (x_n, y_n)$ satisfies x_n or y_n are not integers or $y_n \nmid D$, then P is not torsion.

Investigation 9.34.

- (a) Consider the Mordell curves with $b \in \{0, 2, 24\}$ from Investigation 9.14. Use the Nagel–Lutz theorem to find all the torsion points.
 - (b) Consider the curve $y^2 = x^3 - 1386747x + 368636886$. Is the Nagel–Lutz theorem better or worse than a brute force search for torsion points?
-

5. Mordell–Weil Group and Points of Infinite Order

We have thoroughly investigated the points of finite order and now very briefly mention points of infinite order, i.e., points where the set $\{nP : n \in \mathbb{Z}\}$ is infinite. We would like to ask how many points of infinite order there are, but that question is too imprecise since any single rational point of infinite order generates infinitely many rational points. To count more carefully, given two points of infinite order P and Q , we want to count them as distinct only if

$$\{\pm P, \pm 2P, \pm 3P, \dots\} \cap \{\pm Q, \pm 2Q, \pm 3Q, \dots\} = \emptyset.$$

If those two sets are not disjoint, then we can find integers n, m such that

$$nP + mQ = \mathcal{O},$$

and we say that P and Q are *linearly dependent*. Otherwise, we say P and Q are *linearly independent*.

Example 9.35. Consider the curve $y^2 = x^2 + x + 9$. The two points

$$P = \left(\frac{1}{36}, -\frac{649}{216}\right) \quad \text{and} \quad Q = (46728, 10101033)$$

are linearly dependent. This is because

$$3P = 2Q = \left(\frac{4767698377793431873}{408123470668356}, \frac{10410297971218651651276095767}{8244937290591192023496}\right),$$

whereas the two points

$$P = (0, 3) \quad \text{and} \quad Q = (8, 23)$$

are linearly independent.

Table 9.1. Rank records

Rank	Year	Discoverer
3	1945	Billings
4	1945	Wiman
6	1974	Penney–Pomerance
7	1975	Penney–Pomerance
8	1977	Grunewald–Zimmert
9	1977	Brumer–Kramer
12	1982	Mestre
14	1986	Mestre
15	1992	Mestre
17	1992	Nagao
19	1992	Fermigier
20	1993	Nagao
21	1994	Nagao–Kouya
22	1997	Fermigier
23	1998	Martin–McMillen
24	2000	Martin–McMillen
28	2006	Elkies

Question 9.36. How many linearly independent rational points of infinite order are there on an elliptic curve?

Definition 9.37. The number of linearly independent rational points is called the *rank* of E .

Example 9.38. The curve $y^2 = x^3 + x + 9$ from Example 9.35 has rank 2. In particular, any point of infinite order is a linear combination of

$$P = (0, 3) \quad \text{and} \quad Q = (8, 23).$$

It is not currently known how large the rank can be. Table 9.1 gives the largest known ranks and the years they were discovered.

We content ourselves to merely state the main theorem in this area proven by Louis Mordell in 1922 and later proven for number fields by André Weil.⁵

Theorem 9.39 (Mordell–Weil theorem). *Given an elliptic curve E , there are only finitely many linearly independent rational points.*

We call the set of rational points on E , $E(\mathbb{Q})$, the *Mordell–Weil group*. In the language of group theory, we would say $E(\mathbb{Q})$ is finitely generated.

6. Application: Congruent Numbers

Because the set of rational points of an elliptic curve is so well understood, whenever we can convert a Diophantine problem to a problem about rational points on an elliptic curve, it is extremely helpful. Consider the problem of congruent numbers.

⁵André Weil (1906–1998) was a French mathematician.

Definition 9.40. We say that an integer n is a *congruent number* if n is the area of a right triangle with rational side lengths. In other words, there exist rational numbers a , b , and c such that

$$\begin{cases} c^2 = a^2 + b^2, \\ n = \frac{1}{2}ab. \end{cases}$$

Example 9.41. Six is a congruent number since it is the area of the right triangle with side lengths $(3, 4, 5)$.

This problem was studied by the ancient Greeks and by Arab scholars in the tenth century. More recently, Fibonacci proved that 5 is a congruent number and Euler, that 7 is a congruent number. Fermat proved that 1, 2, and 3 are not congruent numbers.

By examining other Pythagorean triples, we can generate more congruent numbers.

Triple	Congruent Number
(3, 4, 5)	6
(6, 8, 10)	24
(5, 12, 13)	30
(8, 15, 17)	60
(7, 24, 25)	84
(12, 16, 20)	96
(10, 24, 26)	120
(9, 40, 41)	180
(20, 21, 29)	210
(16, 30, 34)	240

Using these integer triples, we can also get additional rational triples that form congruent numbers. We see that for the triple $(9, 40, 41)$, the triangle has area

$$\frac{1}{2} \cdot 9 \cdot 40 = 180 = 5 \cdot 6^2.$$

So if we divide the triple by 6, we get 5 as a congruent number:

$$\left(\frac{3}{2}, \frac{20}{3}, \frac{41}{6}\right) \rightarrow \frac{1}{2} \cdot \frac{3}{2} \cdot \frac{20}{3} = 5.$$

Investigation 9.42. Which congruent numbers can you construct from Pythagorean triples?

While we can probably generate infinitely many congruent numbers in this way, the construction gives us no way to determine if a given number is congruent.

Question 9.43. Given an integer n , is it a congruent number?

This can be a hard problem since for seemingly simple numbers, the triangle with smallest height sides can be complicated.

Example 9.44. For $n = 157$, the triangle with side lengths

$$\left(\frac{6803298487826435051217540}{411340519227716149383203}, \frac{411340519227716149383203}{21666555693714761309610}, \frac{224403517704336969924557513090674863160948472041}{8912332268928859588025535178967163570016480830} \right)$$

is the right triangle with smallest height side lengths with area 157.

Such an example makes our task seem much more daunting and rules out a brute force search. Fortunately, elliptic curves come to the rescue.

Theorem 9.45. *There is a one-to-one correspondence between the sets*

$$\{(a, b, c) : a^2 + b^2 = c^2, \frac{1}{2}ab = n\} \quad \text{and} \quad \{(x, y) : y^2 = x^3 - n^2x, y \neq 0\}.$$

The correspondence is given by

$$(a, b, c) \mapsto \left(\frac{nb}{c-a}, \frac{2n^2}{c-a} \right) \quad \text{and} \quad (x, y) \mapsto \left(\frac{x^2 - n^2}{y}, \frac{2nx}{y}, \frac{x^2 + n^2}{y} \right).$$

Proof. We verify by direct calculation.

$$y^2 = \left(\frac{2n^2}{c-a} \right)^2 = \frac{4n^4}{(c-a)^2} = \frac{2n^3ab}{(c-a)^2}$$

and

$$\begin{aligned} x^3 - n^2x &= \left(\frac{nb}{c-a} \right)^3 - n^2 \frac{nb}{c-a} \\ &= \frac{n^3b^3 - n^3b(c-a)^2}{(c-a)^3} \\ &= \frac{n^3b(c^2 - a^2) - n^3b(c-a)^2}{(c-a)^3} \\ &= \frac{n^3b(2ac - 2a^2)}{(c-a)^3} \\ &= \frac{2n^3ab(c-a)}{(c-a)^3} \\ &= \frac{2n^3ab}{(c-a)^2} \\ &= y^2. \end{aligned}$$

In the other direction we have

$$\begin{aligned}\frac{1}{2}ab &= \frac{1}{2} \frac{x^2 - n^2}{y} \frac{2nx}{y} \\ &= \frac{(x^2 - n^2)(nx)}{y^2} = \frac{n(x^3 - xn^2)}{y^2} \\ &= \frac{n(y^2)}{y^2} = n.\end{aligned}$$

Also

$$a^2 + b^2 = \frac{(x^2 - n^2)^2}{y^2} + \frac{4n^2x^2}{y^2} = \frac{x^4 - 2n^2x^2 + n^4}{y^2} = \frac{(x^2 + n^2)^2}{y^2} = c^2. \quad \square$$

Thus, an efficient algorithm for determining if the elliptic curve

$$y^2 = x^3 - n^2x$$

has a rational point answers the question of whether n is a congruent number. The problem of finding rational points on elliptic curves has been intensely studied and there are many advanced algorithms devoted specifically to solving this problem. However, there is no currently known algorithm that can find all the rational points in all cases.

Example 9.46. We show that 1 is not a congruent number.

The elliptic curve $y^2 = x^3 - x$ has four rational points

$$\{\mathcal{O}, (-1, 0), (0, 0), (1, 0)\}.$$

By Theorem 9.45 none of these points correspond to a triple (a, b, c) that forms a right triangle with area 1. Any such triangle must be a point on this curve, so no such triangle exists.

Example 9.47. We find an explicit right triangle that has area 6.


Consider the elliptic curve $y^2 = x^3 - 36x$. This curve has the rational point $(12, 36)$. Applying the map from Theorem 9.45 we find


$$(12, 36) \mapsto \left(\frac{144 - 36}{36}, \frac{144}{36}, \frac{144 + 36}{36} \right) = (3, 4, 5).$$

Computing the area we have

$$\frac{3 \cdot 4}{2} = 6.$$

COMPUTATIONAL EXERCISES

 **9.1.** Given an elliptic curve $y^2 = x^3 + ax + b$ and a point $P = (x_0, y_0)$, find an explicit expression for the x coordinate of $2P$ in terms of a , b , and x_0 .

 **9.2.** Let $E : y^2 = x^3 + x + 1$ be an elliptic curve with rational points $P = (0, 1)$ and $Q = (72, 611)$. Compute the following point additions.

a. $2P$

b. $P + Q$

9.3. Determine whether $P = (0, 3)$ and $Q = (3, 6)$ are finite or infinite order on the elliptic curve

$$y^2 = x^3 + 9.$$

9.4. Find the 16 points on the elliptic curve $y^2 = x^3 + 17$ which have integer coordinates.

9.5. For the elliptic curve

$$y^2 = x^3 + 15$$


and the points

$$P = (1, 4), \\ Q = (109, 1138),$$

write

$$R_1 = \left(\frac{76229458789}{314969010841}, -\frac{684939947894780378}{176767223233196861} \right), \\ R_2 = \left(\frac{495936485521}{117302140036}, -\frac{382345211934663449}{40175279149489784} \right)$$

as linear combinations of P and Q .

 **9.6.** Use the Nagel–Lutz theorem to find all torsion points on the elliptic curve $y^2 = x^3 - x$.

9.7. Find all the rational torsion points for each of the following elliptic curves.

- a. $y^2 = x^3 + 4x$
- b. $y^2 = x^3 + 9$
- c. $y^2 = x^3 - 58347x + 3954150$

9.8. Find a rational point of infinite order on each of the following elliptic curves.

- a. $y^2 = x^3 + 2$
- b. $y^2 = x^3 - x + 3$
- c. $y^2 = x^3 - 2x + 4$

9.9. The point $P = (-1, 1)$ has infinite order on the elliptic curve

$$y^2 = x^3 + 2.$$

If we work modulo a prime p , every point has finite order. Determine the prime $p < 1000$ for which P has largest order. What is that order?

9.10. What is the average rank of the Mordell curves

$$y^2 = x^3 + b$$

for $-200 \leq b \leq 200$, $b \neq 0$?

9.11. For the elliptic curves

$$E : y^2 = x^3 + n^2,$$

find the n with $1 \leq n \leq 1000$ for which E has the most points modulo 19.

9.12. Consider the elliptic curve

$$E : y^2 = x^3 - 2x.$$

For which primes $1 < p < 100$ is the set of points modulo p cyclic (i.e., generated by a single point)?

9.13. Find triangles that realize each $n \in \{5, 6, 7, 13, 14\}$ as a congruent number.

9.14. Which of the integers $1 \leq n \leq 100$ are congruent numbers?

THEORETICAL EXERCISES

9.15. Prove directly that $4a^3 + 27b^2$ is not zero if and only if the three roots of $x^3 + ax + b$ are distinct.

9.16. Given a cubic polynomial $f(x) = ax^3 + bx^2 + cx + d$, we defined the discriminant as

$$D(f) = b^2c^2 - 4ac^3 - 4b^3d - 27a^2d^2 + 18abcd.$$

If the three roots of $f(x)$ are $\{x_0, x_1, x_2\}$, we can also compute the discriminant as

$$D(f) = a^4(x_0 - x_1)^2(x_0 - x_2)^2(x_1 - x_2)^2.$$

You will prove the two formulas are the same. We will first assume that $a = 1$ so that $f(x) = x^3 + bx^2 + cx + d$ is monic.

- a. Write $(x_0 - x_1)(x_0 - x_2)(x_1 - x_2)$ as a difference $A - B$ for polynomials A, B in terms of x_0, x_1, x_2 . Then we have

$$D(f) = (A - B)^2 = (A + B)^2 - 4AB.$$

- b. Write $A + B$ in terms of b, c, d by recalling that

$$-b = x_0 + x_1 + x_2,$$

$$c = x_0x_1 + x_0x_2 + x_1x_2,$$

$$-d = x_0x_1x_2.$$

- c. For any y_1, y_2, y_3 prove that

$$y_1^3 + y_2^3 + y_3^3 = (y_1 + y_2 + y_3)^3 - 3(y_1 + y_2 + y_3)(y_1y_2 + y_2y_3 + y_3y_1) + 3y_1y_2y_3.$$

- d. Write AB in terms of b, c, d by applying the previous part twice for appropriate choices of y_1, y_2, y_3 .

- e. Compute $D(f) = (A + B)^2 - 4AB$.

- f. Prove the general formula by noting that

$$a \left(x^3 + \frac{b}{a}x^2 + \frac{c}{a}x + \frac{d}{a} \right) = ax^3 + bx^2 + cx + d.$$

9.17. Prove that a polynomial $f(x)$ has a multiple root if and only if $f(x)$ and $f'(x)$ share a common root.

9.18. Prove that an equation of the form $y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$ (general Weierstrass form) can be transformed into an elliptic curve $y^2 = x^3 + ax + b$ with a change of variables as long as you are not working modulo 2 or 3.

9.19. Prove that $(0, 0)$ is the only rational point on $y^2 = x^3 + x$.

9.20. Let $c \neq 0$ be a fixed constant. Prove that

$$u^3 + v^3 = c$$

is an elliptic curve using the change of variables

$$u + v = \frac{12c}{x}, \quad u - v = \frac{y}{3x}.$$

9.21. Let $p \in \mathbb{Z}$ be a prime number. Prove that the point $(0, p)$ is order 3 on the elliptic curve

$$y^2 = x^3 + p^2.$$

9.22. Let E be an elliptic curve, and let P be a torsion point of order n . Prove that $n \leq 6D$.

9.23. Let $p \in \mathbb{Z}$ be an odd prime with $p \equiv 2 \pmod{3}$. Prove that

$$y^2 = x^3 + b$$

has exactly $p + 1$ points (including \mathcal{O}) modulo p for $0 \leq b \leq p - 1$. (*Hint:* First prove that $x \mapsto x^3 + b$ permutes the residues classes modulo p .)

9.24. Let $p \in \mathbb{Z}$ be a prime with $p \equiv 1 \pmod{4}$. Prove that the elliptic curve $y^2 = x^3 + x$ has $p + 1$ points (including \mathcal{O}) modulo p . (*Hint:* Consider $x \mapsto -x$ and use the fact that -1 is not a quadratic residue modulo p .)

9.25. Prove that there are infinitely many congruent numbers.

9.26. The n th pyramidal number is the sum of the first n squares

$$1^2 + 2^2 + \cdots + n^2.$$

Prove that the pyramidal numbers that are also perfect squares, correspond to the integer points of an elliptic curve.

9.27. Prove that the rational numbers a , b , and c that satisfy

$$a + b + c = n = abc$$

for some positive integer n correspond to rational points on an elliptic curve.

EXPLORATION EXERCISES

9.28 (Mordell curves). Mordell curves are elliptic curves of the form $y^2 = x^3 + b$ for $b \in \mathbb{Z}$ with $b \neq 0$.

- For which b are there no integer points?
- Is there an upper bound on the number of integer points independent of b ?
- What is the average number of torsion points as b varies?
- Is there an upper bound on the rank as b varies?
- What is the average rank as b varies?

9.29 (Integer family). Consider the family of elliptic curves $y^2 = x^3 - x + m^2$ for $m \in \mathbb{Z}$.

- There are 15 integer points on these curves for every m . Find formulas for them in terms of m .
- For specific m , there can be additional integer points. For which m can you find the most integer points?
- Are there any points of infinite order that occur for all m ?

9.30 (Prime family). Consider the family of elliptic curves $y^2 = x^3 + px$ with $p \in \mathbb{Z}$ prime.

- What are the possible ranks of these curves?
- Determine congruence conditions on p for those ranks.

9.31 (Elliptic curves over finite fields). Consider an elliptic curve

$$E : y^2 = x^3 + ax + b$$

and a prime $p \in \mathbb{Z}$ with

$$4a^3 + 27b^2 \not\equiv 0 \pmod{p}.$$

Since there are finitely many pairs (x, y) modulo p , there will be finitely many points on E modulo p .

- What is the average number of points on E in terms of p ?
- Can you find curves with significantly more or significantly fewer points than the average?
- Determine upper and lower bounds in terms of p for the number of points on E .

In addition to looking at the number of points, you can look at the group structure of the points modulo p . Recall that these groups will always be finite abelian groups.

- What group structures can you get modulo p ? In particular, can you get a cyclic group?

9.32 (Elliptic curve Diffie–Hellman key exchange). Consider an elliptic curve $E : y^2 = x^3 + ax + b$ modulo a prime $p \in \mathbb{Z}$. We will use this curve for another version of the Diffie–Hellman key exchange algorithm (section 4.3).

For this exploration, you will need a partner. Perform the following five steps.

Step 1: Agree on an elliptic curve E modulo a prime p and a point P with prime order n , i.e., $nP = \mathcal{O}$ and n is prime. The information (E, P, p) is public.

Step 2: Have each partner choose a private key d as a random integer $1 \leq d \leq n - 1$. Compute $Q = dP$. Now each partner has a pair (d_1, Q_1) and (d_2, Q_2) .

Step 3: Exchange the public keys Q_1 and Q_2 .

Step 4: Each partner computes the mutual point

$$R = d_1Q_2 = d_2Q_1.$$

Step 5: The mutual key can now be some derived from the point R , say the x -coordinate.

The following are some questions to consider.

- a. Having exchanged keys Q_1 and Q_2 , how difficult is it to determine d_1 and d_2 from the public information (E, P, p) ?
- b. For the resulting mutual key to be strong enough for secure encryption, how large must p and n be? Is this still a feasible algorithm?
- c. Is this more or less secure than the standard Diffie–Hellman algorithm?

9.33 (Lenstra⁶ elliptic curve factorization). Elliptic curves provide another factorization algorithm. The method is as follows.

Step 1: Input: $n \in \mathbb{Z}$.

Step 2: Choose a random elliptic curve

$$E : y^2 = x^3 + ax + b$$

with

$$\gcd(D, n) = 1$$

and a point $P \in E$ with $P \neq \mathcal{O}$.

Step 3: Consider some number B which is a product of powers of small primes (say the factorial of some small integer).

Step 4: Compute BP (working modulo n).

- If we were able to finish the computation, then start again with another curve and point.
- If we encountered $mP = \mathcal{O}$ for some intermediate m , then start again with another curve and point.
- If we encounter a failure of point addition (a noninvertible element modulo n), that means that the slope of the line between P and Q has $1 < \gcd(m, n) \leq n$. If it is $< n$, then we get a nontrivial factor of n .

The following are some questions to consider.

- a. Implement Lenstra factorization and factor a few large integers.
- b. How does this compare to other factorization methods?

9.34 (Occurrence of torsion). Mazur's theorem (Theorem 9.33) states there are exactly 15 possibilities for the \mathbb{Q} -rational torsion subgroups of an elliptic curve.

- a. Determine the frequency with which each of these 15 groups occur.
- b. Can you find an infinite family of curves all with the same torsion subgroup?

⁶Hendrik Lenstra (1949–) is a Dutch mathematician.

9.35 (Generalizations of congruent numbers).

- a. An integer n is called a $\frac{2\pi}{3}$ congruent number if $n\sqrt{3}$ occurs as the area of a triangle with rational side lengths whose largest angle is $\frac{2\pi}{3}$.
1. Find an elliptic curve that describes this problem.
 2. Which positive integers are $\frac{2\pi}{3}$ congruent numbers?
- b. An integer n is called a $\frac{\pi}{3}$ congruent number if $n\sqrt{3}$ occurs as the area of a triangle with rational side lengths whose largest angle is $\frac{\pi}{3}$.
1. Find an elliptic curve that describes this problem.
 2. Which positive integers are $\frac{\pi}{3}$ congruent numbers?
- c. More generally, a *Heron triangle* is a triangle with rational side lengths and rational area; named after Heron of Alexandria.⁷ For example, a triangle with sides (13, 14, 15) has area 84. More specifically, since we must have

$$a^2 = b^2 + c^2 - 2bc \cos \theta \quad \text{and} \quad 2n = bc \sin \theta,$$

we need to have $(\cos \theta, \sin \theta)$ are rational. In particular, there is a rational number t so that

$$(\cos \theta, \sin \theta) = \left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2} \right).$$

Therefore, we call an integer n a *t-congruent number* if there are positive rational numbers (a, b, c) such that

$$a^2 = b^2 + c^2 - 2bc \frac{1-t^2}{1+t^2} \quad \text{and} \quad 2n = bc \frac{2t}{1+t^2}.$$

For $t = 1$, this is the congruent number problem.

1. Reformulate *t*-congruent numbers in terms of rational points on elliptic curves.
2. Find some examples of (n, t) which solve the generalized problem.

⁷Heron of Alexandria (circa 10–70 C.E.) was a Greek mathematician.

Dynamical Systems

1. Discrete Dynamical Systems

The points of finite order on an elliptic curve in Chapter 9 is a specific instance of the more general notion of periodic points of a dynamical system.

Definition 10.1. A (*discrete*) *dynamical system* is a set S with a self-map $f : S \rightarrow S$.

Since f maps S to S , we can compose f with itself any number of times.

Definition 10.2. For a positive integer n , we define the *n th iterate of f* denoted $f^{\circ n}$ as the composition of f with itself n times:

$$f^{\circ n} = \underbrace{f \circ \cdots \circ f}_{n \text{ times}}.$$

By convention, $f^{\circ 0}$ is the identity map.

Example 10.3. Given the polynomial $f(x) = x^2$, we can consider it as a self-map on the rational numbers $f : \mathbb{Q} \rightarrow \mathbb{Q}$. We have

$$f^{\circ 3}(x) = f(f(f(x))) = x^8.$$

Example 10.4. Given an elliptic curve $E : y^2 = x^3 + ax + b$, we can define the doubling function on rational points

$$\begin{aligned} [2] : E(\mathbb{Q}) &\rightarrow E(\mathbb{Q}) \\ P &\mapsto 2P. \end{aligned}$$

Then we have

$$[2]^{\circ n} P = 2^n P.$$

While there is a rich theory on the properties of functions under iteration, our focus will be on the properties of points under iteration.

Definition 10.5. Let $f : S \rightarrow S$ and $z \in S$. We define the (*forward*) *orbit* of z by f as

$$\mathcal{O}_f(z) = \{z, f(z), f^{\circ 2}(z), f^{\circ 3}(z), \dots\}.$$

We think of the orbit as repeatedly applying f to the point z to get a sequence of points $\mathcal{O}_f(z)$. We often denote this sequence with a directed graph,

$$z \xrightarrow{f} f(z) \xrightarrow{f} f(f(z)) \xrightarrow{f} \dots$$

Example 10.6. Consider the polynomial map

$$\begin{aligned} f : \mathbb{Q} &\rightarrow \mathbb{Q} \\ x &\mapsto x^2 - 1. \end{aligned}$$

We compute the orbit of $x = 2$ as

$$\mathcal{O}_f(2) = \{2, f(2), f^{\circ 2}(2), f^{\circ 3}(2), \dots\} = \{2, 3, 8, 63, \dots\}.$$

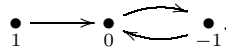
We can represent this orbit with the (infinite) directed graph

$$2 \xrightarrow{f} 3 \xrightarrow{f} 8 \xrightarrow{f} 63 \xrightarrow{\quad} \dots$$

If we instead consider the orbit of $x = 1$, we compute

$$\mathcal{O}_f(1) = \{1, 0, -1, 0, -1, 0, -1, \dots\}.$$

We can represent this orbit with the (finite) directed graph



Example 10.6 shows the two main possibilities for orbits: they can have finitely many distinct values or infinitely many distinct values. We say the orbit is *finite* or *infinite*, respectively.

Definition 10.7. Points z with finite orbit are called *preperiodic*, and there exist integers $n > k \geq 1$ such that

$$f^{\circ n}(z) = f^{\circ k}(z).$$

In other words, the orbit eventually loops back on itself. If, in fact,

$$f^{\circ n}(z) = z,$$

then we say that z is *periodic* with period n .

If a point is not preperiodic, we say that it is *wandering*.

In Example 10.6, for the function $f(x) = x^2 - 1$, we found that 2 is wandering and that $\{0, 1, -1\}$ are preperiodic with $\{0, -1\}$ periodic. We can say that $\{0, -1\}$ are both periodic with period 2, but since

$$0 = f^{\circ 2}(0) = f^{\circ 4}(0) = f^{\circ 6}(0) = \cdots,$$

0 has period $2n$ for any positive integer n . To deal with this nonuniqueness of periods, we introduce more precision in our terminology.

Definition 10.8. Let $f : S \rightarrow S$, and let $z \in S$ be a preperiodic point for f . If there is a pair of integers (m, n) , $n > 0$, $m \geq 0$ such that

$$f^{\circ(m+n)}(z) = f^{\circ m}(z),$$

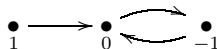
we say that z has *cycle structure* (m, n) and call m the *tail* of z and n the *period* of z . If (m, n) is the smallest such pair of integers, we call (m, n) the *minimal cycle structure*, m the *minimal tail*, and n the *minimal period*.

Notice that periodic points have minimal cycle structure $(0, n)$ for some positive integer n .

Example 10.9. For

$$\begin{aligned} f : \mathbb{Q} &\rightarrow \mathbb{Q} \\ x &\mapsto x^2 - 1, \end{aligned}$$

we have the directed graph of rational preperiodic points



and, hence, the following preperiodic points and their minimal cycle structures.

x	(m, n)
1	(1, 2)
0	(0, 2)
-1	(0, 2)

In this chapter, we consider dynamical systems that are defined by polynomials or rational functions.

Definition 10.10. A *rational function* $f(x) \in \mathbb{Q}(x)$ is a fraction of polynomials with rational coefficients

$$f(x) = \frac{g(x)}{h(x)} = \frac{a_n x^n + \cdots + a_1 x + a_0}{b_m x^m + \cdots + b_1 x + b_0}.$$

We may clear denominators of the coefficients and assume that $g(x), h(x) \in \mathbb{Z}[x]$. We define the *degree* of f as

$$\deg(f(x)) = \max(\deg(g(x)), \deg(h(x))).$$

Investigation 10.11. Consider the family of functions

$$\begin{aligned} f_c : \mathbb{C} &\rightarrow \mathbb{C} \\ f_c(z) &= z^2 + c, \end{aligned}$$

where $c \in \mathbb{C}$ can be any constant.

- (a) Determine the values of $c \in \mathbb{C}$ for which elements in the orbit of 0 are bounded in absolute value, i.e.,

$$\mathcal{M} = \{c \in \mathbb{C} : |f^{on}(0)| \leq B, \text{ for some constant } B, \text{ for all } n \in \mathbb{N}\}.$$

In practice, you can choose $B = 4$. Why is 4 sufficient?

- (b) Plot the values $c \in \mathcal{M}$ on the plane as (x, y) for $c = x + iy$.
 (c) Which points in the plot correspond to the case that 0 is periodic?
 (d) What else can you say about the structure of this set?

The set \mathcal{M} is called the *Mandelbrot set*.

The point doubling map on elliptic curves in Chapter 9 gives our first example of a dynamical system with a rational function.

Example 10.12. Given an elliptic curve $E : y^2 = x^3 + ax + b$, we can define the doubling function on the x -coordinate of points

$$P \mapsto 2P$$

by

$$\begin{aligned} f : \mathbb{Q} &\rightarrow \mathbb{Q} \\ f(x) &= \frac{x^4 - 2ax^2 - 8bx + a^2}{4x^3 + 4ax + 4b}. \end{aligned}$$

The iterates will give us the collection of x -coordinates for

$$\{P, 2P, 4P, 8P, \dots\}.$$

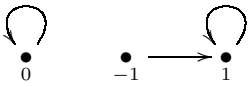
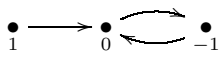
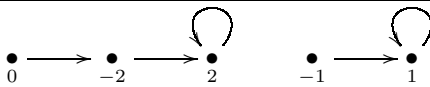

The examples we have examined so far in this chapter have all had rational periodic and preperiodic points. However, it is not clear what we should expect in general. Figure 10.1 computes the rational periodic and preperiodic points for a few functions.

Though Figure 10.1 represents a very small data set, there seem to be relatively few, if any, rational preperiodic points for a given function.

Question 10.13. Given a rational function $f(x) \in \mathbb{Q}(x)$, how many preperiodic points are defined over \mathbb{Q} ?

Our goal in this chapter is to answer Question 10.13 computationally. In particular, given a rational function f , we want to compute all the rational preperiodic points.

Table 10.1. Rational preperiodic structures

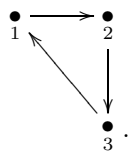
$f(x)$	preperiodic structure
x^2	
$x^2 - 1$	
$x^2 + 1$	—
$x^2 - 1/2$	—
$x^2 + 1/2$	—
$x^2 - 2$	
$x^2 + 2$	—
$x^2 - 1/3$	—
$x^2 - 2/3$	—
$x^2 - 3$	
$x^2 + 1/3$	—
$x^2 + 2/3$	—
$x^2 + 3$	—
$x^2 + 3/2$	—

1.1. Constructing Polynomials with Specified Dynamics. However, before we determine the rational preperiodic points for a given function, we consider the inverse problem.

Question 10.14. Given a directed graph, can you find a polynomial that acts as the arrows of that directed graph?

Not every directed graph is possible since we need it to come from a function. The directed graphs that come from functions are the ones where every node has a single arrow leaving it.

Example 10.15. Consider the following directed graph



A polynomial f with that preperiodic structure satisfies

$$f(1) = 2, \quad f(2) = 3, \quad \text{and} \quad f(3) = 1.$$

Since a polynomial is uniquely determined by its coefficients, we need to find the coefficients of

$$f(x) = a_n x^n + \cdots + a_1 x + a_0.$$

We have three equations

$$\begin{aligned} f(1) &= a_n + a_{n-1} + \cdots + a_1 + a_0 = 2, \\ f(2) &= a_n 2^n + a_{n-1} 2^{n-1} + \cdots + a_1 2 + a_0 = 3, \\ f(3) &= a_n 3^n + a_{n-1} 3^{n-1} + \cdots + a_1 3 + a_0 = 1. \end{aligned}$$

These are (independent) linear equations so, as long as there are at least three coefficients for f , then there is a solution. When there are exactly three coefficients for f , there is a unique solution. There are three coefficients for f when $\deg(f) = 2$. We solve

$$\begin{aligned} f(1) &= a_2 + a_1 + a_0 = 2, \\ f(2) &= 4a_2 + 2a_1 + a_0 = 3, \\ f(3) &= 9a_2 + 3a_1 + a_0 = 1 \end{aligned}$$

to get the polynomial

$$f(x) = -\frac{3}{2}x^2 + \frac{11}{2}x - 2.$$

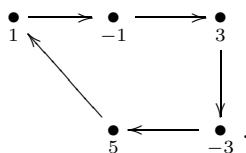
The method used in Example 10.15 is called *Lagrange interpolation*, and the resulting unique polynomial is called a *Lagrange polynomial*. Lagrange interpolation is the generalization of finding the unique line that goes through two points in the plane.

Theorem 10.16. *For every positive integer n and nonnegative integer m , there exists a polynomial $f \in \mathbb{Q}[x]$ of degree $m + n - 1$ which has a rational point with minimal cycle structure (m, n) .*

Proof. Apply Lagrange interpolation. □

Since we can find a polynomial of some degree with any dynamical behavior, we can find a polynomial with a rational periodic point with large minimal period. However, the polynomial will have a correspondingly large degree. In particular, Lagrange interpolation tells us that there is a degree d polynomial with a rational periodic point with minimal period $d + 1$. Example 10.15 constructs an explicit degree 2 polynomial with a point with period 3. Similarly, you can construct a degree 3 polynomial with a point with minimal period 4. However, the following example shows that there are degree 3 polynomials with a rational periodic point with minimal period larger than can be explained by Lagrange interpolation.

Example 10.17. The polynomial $f(x) = \frac{x^3 - 25x + 12}{12}$ has the periodic cycle



We could ask the following general question.

Question 10.18. Let d be a positive integer. Does there exist a constant $C(d)$ depending on d so that *any* polynomial $f \in \mathbb{Q}[x]$ of degree d has no periodic points defined over \mathbb{Q} with minimal period larger than $C(d)$?

Example 10.19. Bjorn Poonen¹ has conjectured that when $d = 2$, $C(2)$ can be taken to be 3.

More generally, one can also ask for a bound on the total number of rational preperiodic points.

Question 10.20. Let d be a positive integer. Does there exist a constant $M(d)$ depending on d so that *any* polynomial $f \in \mathbb{Q}[x]$ of degree d has at most $M(d)$ preperiodic points defined over \mathbb{Q} ?

The existence of such a constant is the dynamical systems version of Mazur's theorem (Theorem 9.33) on the torsion points of an elliptic curve. The constant $M(d)$ was conjectured to exist by Patrick Morton² and Joseph Silverman³ in 1994 but remains an unsolved problem.

Investigation 10.21. Consider polynomial functions $f \in \mathbb{Q}[x]$.

- (a) Find examples of polynomials f of degree 2 with a periodic point defined over \mathbb{Q} with minimal period 1, 2, and 3.
- (b) Conjecture a bound for the number of preperiodic points defined over \mathbb{Q} for a degree 2 polynomial.
- (c) Find examples of polynomials f of degree 3 with a periodic point defined over \mathbb{Q} with minimal period 1, 2, 3, and 4.
- (d) Conjecture a bound for the minimal period of a rational periodic point for a degree 3 polynomial.

¹Bjorn Poonen (1968–) is an American mathematician.

²Richard Patrick Morton is an American mathematician.

³Joseph H. Silverman (1955–) is an American mathematician.

2. Dynatomic Polynomials

One way to find periodic and preperiodic points for a given function is by solving the equation that defines them. Since we are working with polynomial or rational functions, this comes down to factoring polynomials. For example, if z is periodic with period n , then it must satisfy the equation

$$f^{\circ n}(z) = z.$$

Example 10.22. Let $f(x) = x^2 - 1$. Then we find the fixed points (points with period 1) by solving

$$x^2 - 1 = x,$$

which is

$$(35) \quad x^2 - x - 1 = 0.$$

The polynomial $x^2 - x - 1$ is irreducible over \mathbb{Q} , so equation (35) has no solutions in \mathbb{Q} , i.e., f has no fixed points defined over \mathbb{Q} .

We find points with period 2 by solving

$$\begin{aligned} f^{\circ 2}(x) &= x, \\ x^4 - 2x^2 &= x, \end{aligned}$$

which factors as

$$(36) \quad x(x+1)(x^2 - x - 1) = 0.$$

So we have two rational periodic points with period 2: $\{0, -1\}$.

We solve for points with cycle structure $(1, 2)$ by solving

$$f^{\circ 3}(x) = f(x).$$

This is the equation

$$x^8 - 4x^6 + 4x^4 - 1 = x^2 - 1.$$

We get

$$(37) \quad x^2(x-1)(x+1)(x^2 - x - 1)(x^2 + x - 1) = 0.$$

So we have three rational points $\{0, 1, -1\}$. The point 1 is the only point that has *minimal* cycle structure $(1, 2)$.

Remark. Notice that the irreducible degree 2 polynomial $x^2 - x - 1$ appeared in all three equations (35), (36), and (37). This is because $x^2 - x - 1$ represents fixed points, which also have cycle structure (m, n) for any m and n .

We can get similar equations for rational functions by clearing denominators in

$$f^{m+n}(x) = f^m(x).$$

Example 10.23. For the function

$$f(x) = \frac{-5x^2 + 4}{4x},$$

we solve for the fixed points as

$$\frac{-5x^2 + 4}{4x} = x,$$

which is

$$-5x^2 + 4 = 4x^2,$$

to have

$$9x^2 - 4 = 0$$

and the fixed points

$$x = \left\{ \pm \frac{2}{3} \right\}.$$

For points with period 2, we have

$$f^{\circ 2}(x) = \frac{-125x^4 + 264x^2 - 80}{-80x^3 + 64x},$$

and we find

$$-125x^4 + 264x^2 - 80 = x(-80x^3 + 64x),$$

which factors as

$$5(x-2)(x+2)(3x-2)(3x+2) = 0.$$

The four points with period 2 are

$$\left\{ \pm 2, \pm \frac{2}{3} \right\},$$

which includes the two fixed points.

While we are successfully finding the rational preperiodic points with a given cycle structure, it would be better if we could find the points where (m, n) is the minimal cycle structure. To accomplish this, we apply the same inclusion-exclusion technique as for cyclotomic polynomials (section 5.2) for determining primitive roots of unity.

Definition 10.24. Let $f(x) \in \mathbb{Q}(x)$. For each integer m , define polynomials g_m and h_m as

$$f^{\circ m}(x) = \frac{g_m(x)}{h_m(x)}.$$

We define the n th *dynatomic polynomial* $\Phi_n(f)$ as

$$\Phi_n(f) = \prod_{d|n} (g_d(x) - xh_d(x))^{\mu(n/d)},$$

where μ is the Möbius function from section 5.2.

For the case that f is a polynomial, this becomes

$$\Phi_n(f) = \prod_{d|n} (f^{\circ d}(x) - x)^{\mu(n/d)}.$$

To see what is going on, let's consider a few positive integers n and a polynomial function f . For $n = 2$, the only divisor is 1, and we compute

$$\mu(2) = -1 \quad \text{and} \quad \mu(1) = 1.$$

Therefore,

$$\Phi_2(f) = \frac{f^{\circ 2}(x) - x}{f(x) - x}.$$

The interpretation is that we are taking all the points of period 2 and “removing” those that are period 1 by dividing by $f(x) - x$.

For $n = 4$, we have divisors 1, 2, and 4 with Möbius values

$$\mu(4) = 0, \quad \mu(2) = -1, \quad \text{and} \quad \mu(1) = 1.$$

Therefore,

$$\Phi_4(f) = \frac{f^{\circ 4}(x) - x}{f^{\circ 2}(x) - x}.$$

The interpretation is that we are taking all the points of period 4 and “removing” those that are period 2, which includes the fixed points, by dividing by $f^{\circ 2}(x) - x$.

For $n = 6$, we have divisors 1, 2, 3, and 6 with Möbius values

$$\mu(6) = 1, \quad \mu(2) = \mu(3) = -1, \quad \text{and} \quad \mu(1) = 1.$$

Therefore,

$$\Phi_6(f) = \frac{(f^{\circ 6}(x) - x)(f(x) - x)}{(f^{\circ 2}(x) - x)(f^{\circ 3}(x) - x)}.$$

The interpretation is that we are taking all the points of period 6 and “removing” those that are period 2 and 3. But, the period 2 and 3 points both include the fixed points, so we have “removed” the fixed points twice. Since there is only one copy of the fixed points in $f^{\circ 6}(x) - x$, we need to add $f(x) - x$ to the numerator to end up with $\Phi_6(f)$ as a polynomial.

Example 10.25. Let $f(x) = x^2 - 7/4$. Then we have

$$\Phi_2(f) = \frac{f^{\circ 2}(x) - x}{f(x) - x} = x^2 + x - \frac{3}{4} = (2x - 1)(2x + 3).$$

Example 10.26. Let $f(x) = 5/24x^3 - 53/24x + 1$. Then we have

$$\Phi_4(f) = \frac{f^{\circ 4}(x) - x}{f^{\circ 2}(x) - x} = x(x - 3)(x - 1)(x + 1)g(x),$$

where $g(x)$ is an irreducible polynomial of degree 68.

All our examples have resulted in polynomials, but we have only considered polynomial functions.

Example 10.27. Consider the function

$$f(x) = \frac{x^2 - 2}{x}.$$

We have

$$g_1(x) = x^2 - 2 \quad \text{and} \quad h_1(x) = x$$

so that

$$\Phi_1(f) = (x^2 - 2) - x(x) = -2,$$

and there are no fixed points. To find the points with period 2, we first compute

$$f^{\circ 2}(x) = \frac{x^4 - 6x^2 + 4}{x^3 - 2x}$$

so that

$$g_2(x) = x^4 - 6x^2 + 4 \quad \text{and} \quad h_2(x) = x^3 - 2x.$$

We compute

$$\begin{aligned}\Phi_2(f) &= \frac{g_2(x) - xh_2(x)}{g(x) - xh(x)} = \frac{x^4 - 6x^2 + 4 - x(x^3 - 2x)}{x^2 - 2 - x(x)} \\ &= \frac{-4x^2 + 4}{-2} = 2x^2 - 2.\end{aligned}$$

We solve $2x^2 - 2 = 0$ to get the points $\{\pm 1\}$ as rational periodic points of minimal period 2.

In all our examples, the end result was a polynomial defining exactly the points that we were interested in. It is not clear we will always get a polynomial, much less exactly the correct polynomial. The potential problem is the existence of multiple roots, i.e., repeated factors. We did not have this concern for cyclotomic polynomials since there were no repeated roots—every root of unity is distinct. The following example illustrates the problem.

Example 10.28. Let $f(x) = x^2 - 3/4$. Then

$$\begin{aligned}f(x) - x &= (2x - 3)(2x + 1), \\ f^{\circ 2}(x) - x &= (2x - 3)(2x + 1)^3.\end{aligned}$$

We still have a polynomial

$$\Phi_2(f) = (2x + 1)^2,$$

but its roots are not points with minimal period 2; they are fixed points. In fact, there are no points with minimal period 2 for f !

Definition 10.29. We call z a point of *formal period* n for f if $\Phi_n(f)(z) = 0$.

While it is true that all minimal periodic points of period n are formal periodic points of period n , the converse is not true.

Theorem 10.30 (Morton and Silverman 1994). Let $f(x) \in \mathbb{Q}(x)$.

- (a) $\Phi_n(f)$ is a polynomial for all $n \in \mathbb{N}$.
- (b) If z is a simple root (multiplicity 1) of $\Phi_n(f)$, then z is periodic with minimal period n .

Although the proof does not involve anything beyond what we know, it is rather lengthy, involving the analysis of several different cases.⁴

Investigation 10.31. Let $f(x)$ be a polynomial. Consider the construction

$$\Psi_{n,m}(f) = \frac{\Phi_n(f^{\circ m})}{\Phi_n(f^{\circ(m-1)})}.$$

- (a) Is Ψ always a polynomial in x ?
- (b) What can you say about its roots?
- (c) Can you find a similar construction if $f(x)$ is a rational function?

⁴We leave the interested reader to consult [48, Theorem 4.5].

Dynatomic polynomials are useful for determining periodic points of small period, but there are two issues. Since the degrees of $f^{\circ n}(x)$ grow exponentially, $\Phi_n(f)$ quickly becomes unmanageable. Also, we are interested in determining all the rational periodic points, so we must somehow know when to stop looking. We focus on the second issue, which is a simpler version of Question 10.18 since it asks for a bound for a specific function instead of a uniform bound.

Question 10.32. Given a function $f(x) \in \mathbb{Q}(x)$, can we find an upper bound on the minimal period of a periodic point defined over \mathbb{Q} ?

It is not yet clear to us that such a bound should even exist. Consequently, we look to reduction modulo primes to gain some additional information.

3. Resultant and Reduction Modulo Primes

Our goal is to obtain information about rational periodic points for a function f by looking at the periodic points of f modulo a prime p .

Definition 10.33. Given a rational function $f(x) = \frac{a_n x^n + \cdots + a_1 x + a_0}{b_m x^m + \cdots + b_1 x + b_0}$, we define the reduction of f modulo p to be the reduction of each coefficient:

$$\overline{f}(x) = \frac{\overline{a}_n x^n + \cdots + \overline{a}_1 x + \overline{a}_0}{\overline{b}_m x^m + \cdots + \overline{b}_1 x + \overline{b}_0}.$$

Example 10.34. Given $f(x) = \frac{x^2+3x}{6x+1}$ and $p = 3$, we have

$$\overline{f}(x) = x^2.$$

Example 10.35. Given $f(x) = 2x^2 + 4x + 3$ and $p = 2$, we have

$$\overline{f}(x) = 1.$$

Example 10.35 is especially important to us. It shows that reducing modulo a prime may not preserve the dynamical structure; \overline{f} is a constant map, but f is not constant. Since our goal is to obtain dynamical information about f by looking at the dynamics of \overline{f} , Example 10.35 shows we need to be careful about which primes we use.

We are interested in dynamical behavior, so we need reduction to commute with iteration, i.e.,

$$\overline{f^{\circ n}(x)} \equiv \overline{f}^{\circ n}(\overline{x}) \pmod{p}.$$

To determine which primes work and which do not, we need to look at resultants.

Definition 10.36. We define the *resultant* of two polynomials

$$\begin{aligned} g(x) &= a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0, \\ h(x) &= b_m x^m + b_{m-1} x^{m-1} + \cdots + b_1 x + b_0, \end{aligned}$$

denoted $\text{Res}(g, h)$, as the determinant of the $(m + n) \times (m + n)$ matrix

$$(38) \quad \begin{pmatrix} a_n & a_{n-1} & \cdots & a_1 & a_0 & & & \\ & a_n & a_{n-1} & \cdots & a_1 & a_0 & & \\ & & a_n & a_{n-1} & \cdots & a_1 & a_0 & \\ & & & a_n & a_{n-1} & \cdots & a_1 & a_0 \\ b_m & b_{m-1} & \cdots & b_1 & b_0 & & & \\ & b_m & b_{m-1} & \cdots & b_1 & b_0 & & \\ & & b_m & b_{m-1} & \cdots & b_1 & b_0 & \\ & & & b_m & b_{m-1} & \cdots & b_1 & b_0 \end{pmatrix},$$

where the blank entries are all 0. Notice that if $g, h \in \mathbb{Z}[x]$, then $\text{Res}(g, h) \in \mathbb{Z}$.

Example 10.37. Consider the two polynomials $g(x) = 3x^2 + 1$, and $h(x) = 6x^2 + 2x$. Then we compute the resultant as

$$\text{Res}(g, h) = \det \begin{pmatrix} 3 & 0 & 1 & 0 \\ 0 & 3 & 0 & 1 \\ 6 & 2 & 0 & 0 \\ 0 & 6 & 2 & 0 \end{pmatrix} = 9.$$

Proposition 10.38. *The resultant of two polynomials g and h is 0 if and only if g and h have a common zero.*

Proof. Assume that $\deg(g) = n$ and $\deg(h) = m$.

Two polynomials have a common zero if and only if they have a common factor. They have a common factor if and only if there are two polynomials $c(x)$ and $d(x)$ of degrees at most $n - 1$ and $m - 1$, respectively, such that

$$(39) \quad d(x)g(x) = c(x)h(x).$$

Treating the coefficients of $c(x)$ and $d(x)$ as unknowns, equation (39) induces a system of linear equations by equating the coefficients on each side

$$\begin{aligned} a_0 d_0 &= b_0 c_0 \\ a_1 c_0 + c_1 a_0 &= b_1 d_0 + d_1 b_0 \\ &\vdots \end{aligned}$$

Writing these equations as a matrix, we get the transpose of (38). A linear system of equations has a solution if and only if the determinant of the associated matrix is nonzero. \square

Definition 10.39. Let $f(x) = \frac{g(x)}{h(x)}$ with $g, h \in \mathbb{Z}[x]$. Define

$$\text{Res}(f) = \text{Res}(g(x), h(x)),$$

where we consider g and h as both having the same degree.

Example 10.40. Let $f(x) = \frac{x^3+1}{x^2+2x}$. Then we have $\deg(f) = 3$ and $g(x) = x^3 + 1$ and $h(x) = x^2 + 2x$. We think of $h(x)$ as a degree 3 polynomial as $h(x) = 0x^3 + x^2 + 2x + 0$, so the resultant is

$$\text{Res}(f) = \text{Res}(g, h) = \det \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 & 0 \end{pmatrix} = -7.$$

Proposition 10.41. Let $f(x) = \frac{g(x)}{h(x)}$ and $\bar{f}(x) = \frac{\bar{g}(x)}{\bar{h}(x)}$. The following are equivalent:

- (a) $\deg f = \deg \bar{f}$,
- (b) $g(x)$ and $h(x)$ have no common zeros modulo p ,
- (c) $\text{Res}(f) \not\equiv 0 \pmod{p}$.

Proof. The first two are equivalent since the value of $\deg(\bar{f})$ is the degree of f minus any cancellation that occurs in $\frac{\bar{g}(x)}{\bar{h}(x)}$. Each common factor corresponds to a common zero. In other words,

$$\deg(\bar{f}) = \deg(f) - \#\{z \in \mathbb{Z}/p\mathbb{Z} : \bar{g}(z) \equiv \bar{h}(z) \equiv 0 \pmod{p}\}.$$

The last two are equivalent by Proposition 10.38. □

Definition 10.42. Let $f(x)$ be a rational function. We say that a prime p is a prime of *good reduction* if either condition of Proposition 10.41 is satisfied. Otherwise, we say p is a prime of *bad reduction*.

Example 10.43. Consider the rational function

$$f(x) = \frac{x^2 + x}{x + 3}.$$

We compute the resultant as

$$\text{Res}(f) = 6 = 2 \cdot 3.$$

Then f has bad reduction at 2 and 3.

At $p = 2$, we have

$$\bar{f}(x) = \frac{x^2 + x}{x + 1} = \frac{x(x + 1)}{x + 1} = x$$

so that $\deg(f) > \deg(\bar{f})$. By Proposition 10.41, this makes 2 a prime of bad reduction for f . We see that $f(1) = \frac{1}{2}$, which cannot be reduced modulo 2, so the dynamical structure of f and the dynamical structure of \bar{f} are different.

At $p = 3$, we have

$$\bar{f}(x) = \frac{x^2 + x}{x + 3} = \frac{x(x + 1)}{x} = x + 1$$

so that $\deg(f) > \deg(\bar{f})$. By Proposition 10.41 this makes 3 a prime of bad reduction for f . We see that $f(0) = 0$ but that $\bar{f}(0) \equiv 1 \pmod{3}$ so that the dynamical structure of f and the dynamical structure of \bar{f} are different.

All other primes are primes of good reduction.

We proved in Chapter 2 that reduction modulo primes respects addition and multiplication, so any solution to an equation in integers will still be a solution modulo a prime (we used this extensively in section 8.2). It may seem that Example 10.43 contradicts the properties of modular arithmetic, but it does not. What happens is that there is a common factor in the numerator and denominator modulo p . Write $f(x) = \frac{g(x)}{h(x)}$ so that the fixed point equation is $g(x) - xh(x) = 0$. Modulo p we write

$$\begin{aligned}\bar{g}(x) &\equiv u(x)v(x) \pmod{p}, \\ \bar{h}(x) &\equiv w(x)v(x) \pmod{p}\end{aligned}$$

to have the fixed point equation reducing to

$$(40) \quad g(x) - xh(x) \equiv v(x)(u(x) - w(x)) \pmod{p}.$$

However, first computing

$$\bar{f}(x) \equiv \frac{u(x)}{w(x)} \pmod{p},$$

we have the fixed point equation

$$(41) \quad u(x) - w(x) \equiv 0 \pmod{p}.$$

Equations (40) and (41) differ by a factor of $v(x)$ modulo p representing the lost information. So we have not broken the laws of modular arithmetic, nor does the dynamical structure of $f(x)$ reflect the dynamical structure of $\bar{f}(x)$.

Proposition 10.44. *Let $f(x) \in \mathbb{Q}(x)$ be a rational function. Then, f has only finitely many primes of bad reduction.*

Proof. The rational function $f(x)$ has bad reduction at p if and only if p divides the resultant $\text{Res}(f)$. We can clear denominators of the coefficients of f (considering the denominator of f as a degree 0 polynomial when f is a polynomial) so that all coefficients are integers and, consequently, $\text{Res}(f) \in \mathbb{Z}$.

Since any integer has only finitely many prime factors, there are only finitely many primes of bad reduction. \square

We conclude this section by proving that reduction commutes with iteration for primes of good reduction.

Theorem 10.45. *Let $f(x) = \frac{g(x)}{h(x)}$ be a rational function, and let $\bar{f} = \frac{\bar{g}(x)}{\bar{h}(x)} \pmod{p} = \frac{\bar{g}(x)}{\bar{h}(x)}$. If p is a prime of good reduction, then*

$$\overline{f^{\circ n}(x)} \equiv \bar{f}^{\circ n}(\bar{x}) \pmod{p}.$$

Proof. We know that

$$\begin{aligned}\overline{g(x)} &\equiv \overline{g(\overline{x})} \pmod{p}, \\ \overline{h(x)} &\equiv \overline{h(\overline{x})} \pmod{p}.\end{aligned}$$

Thus, the only way for

$$\frac{g(x)}{h(x)} \not\equiv \frac{\overline{g(x)}}{\overline{h(x)}} \pmod{p}$$

is for $\overline{g(x)}$ and $\overline{h(x)}$ to have a common factor. \square

Investigation 10.46. The prime factors of the resultant of f are the primes of bad reduction for f . What about primes of bad reduction for $f^{\circ n}$?

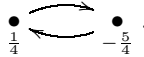
- Determine a relationship between $\text{Res}(f)$ and $\text{Res}(f^{\circ n})$ when $\deg(f) = 2$.
- Determine a relationship between $\text{Res}(f)$ and $\text{Res}(f^{\circ n})$ when $\deg(f) = 3$.
- Determine the general relationship between $\text{Res}(f)$ and $\text{Res}(f^{\circ n})$.
- What does this say about primes of bad reduction for the iterates $f^{\circ n}$?

4. Periods Modulo Primes

An important consequence of Theorem 10.45 is that periodic points reduce to periodic points for primes of good reduction.

Question 10.47. Is there a relationship between the minimal period of a point and the minimal period of its reduction modulo a prime of good reduction?

Example 10.48. Consider $f(x) = x^2 - \frac{21}{16}$. Then we see that $\frac{1}{4}$ has minimal period 2,



The only prime of bad reduction is 2. Let's start at $p = 3$ and see what happens to the period of $\frac{1}{4}$ modulo primes.

p	(minimal) period of $1/4$ modulo p
3	1
5	2
7	2
11	2
13	2
17	2
19	2

Except for $p = 3$ where $1/4$ is a fixed point, we have that the minimal period remains unchanged modulo p . What happens is that $\frac{1}{4} \equiv -\frac{5}{4} \pmod{3}$, so that the 2-cycle collapses to a fixed point.

Corollary 10.49. *Let $f(x)$ be a rational function with good reduction at p . If z is periodic with minimal period n , then \bar{z} is periodic with minimal period m and $m \mid n$.*

Proof. We know that \bar{z} is periodic since

$$\overline{f^{\circ n}(z)} \equiv \overline{f}^{\circ n}(\bar{z}) \pmod{p}.$$

Thus, the period of \bar{z} must be a divisor of n . \square

Recall that our objective is to determine the periodic points of $f(x)$ by examining the dynamics of $\overline{f}(x)$. It is an excellent start that the minimal period of \bar{z} must divide the minimal period of z . We can compute the finitely many periodic points for \overline{f} simply by computing the orbit of each of the finitely many residues classes. However, each period m for a reduced point \bar{z} gives an infinite set of possible minimal periods n for a point z :

$$n \in \{m, 2m, 3m, 4m, \dots\}.$$

Thus we need a more precise answer to Question 10.47. To obtain it, we define another property associated with a periodic point.

Definition 10.50. Let $f(x)$ be a rational function, and let z be a periodic point with minimal period n . Then the *multiplier* of z is

$$\lambda_z = (f^{\circ n})'(z),$$

where $(f^{\circ n})'$ is the derivative of the n th iterate of f .

Example 10.51. For $f(x) = x^2 - 2$, we have $x = 2$ is a fixed point, i.e., $n = 1$. We compute $f'(x) = 2x$ and so

$$\lambda_2 = f'(2) = 4.$$

Example 10.52. The function $f(x) = x^2 - \frac{21}{16}$ has the 2-cycle

$$\begin{array}{ccc} \bullet & \xleftrightarrow{\quad} & \bullet \\ \frac{1}{4} & & -\frac{5}{4} \end{array}.$$

We compute the multiplier of $x = \frac{1}{4}$ as

$$\lambda_{1/4} = (f^{\circ 2})'(1/4).$$

We have

$$f^{\circ 2}(x) = x^4 - \frac{21}{8}x^2 + \frac{105}{256},$$

whose derivative is

$$(f^{\circ 2})'(x) = 4x^3 - \frac{21}{4}x.$$

Finally, we have the multiplier as

$$\lambda_{1/4} = (f^{\circ 2})'(1/4) = \frac{1}{16} - \frac{21}{16} = -\frac{5}{4}.$$

We could also compute

$$\lambda_{-5/4} = (f^{\circ 2})'(-5/4) = -\frac{125}{16} + \frac{105}{16} = -\frac{5}{4}.$$

The fact that

$$\lambda_{1/4} = \lambda_{-5/4}$$

is not a coincidence. The multipliers of points in the same cycle are always the same (Theoretical Exercise 10.25).

We can now give a precise answer to Question 10.47 on the relationship between m and n .

Theorem 10.53. *Let $f(x)$ be a rational function of degree at least 2. Let z be a periodic point for f . Let p be a prime of good reduction, and define*

- n the minimal period of z ,
- m the minimal period of \bar{z} ,
- r the multiplicative order of $\overline{\lambda_z}$ modulo p .

Then

$$n = m \quad \text{or} \quad n = mrp^e$$

for some integer $e \geq 0$.

Proof. We know from Corollary 10.49 that $m \mid n$, so replace f by $f^{\circ m}$ and n by n/m . If z is fixed by f , then we are done, so assume not.

After a change of variables, we may assume that $z = 0$. Write

$$f(x) = \frac{a_d x^d + \cdots + a_1 x + a_0}{b_d x^d + \cdots + b_1 x + b_0}$$

with at least one of a_d and $b_d \neq 0$. Since 0 is fixed modulo p , we know that

$$\frac{a_0}{b_0} \equiv 0 \pmod{p},$$

so we may multiply the numerator and denominator of f by b_0^{-1} and rewrite f as

$$f(x) = \frac{a_d x^d + \cdots + a_1 x + a_0}{b_d x^d + \cdots + b_1 x + 1}.$$

Now we take the Taylor series expansion of f at $x = 0$ to get

$$f(x) = \mu + \lambda x + \frac{g(x)}{1 + h(x)} x^2$$

for some polynomials $g(x), h(x) \in \mathbb{Z}[x]$ and constants μ and λ . Notice that $\mu \equiv 0 \pmod{p}$ and $\lambda = f'(0)$.

By induction, we can show that

$$f^{\circ i}(0) \equiv \mu(1 + \lambda + \cdots + \lambda^{i-1}) \pmod{\mu^2}.$$

Since $f^{\circ n}(0) = 0$ and $\mu \equiv 0 \pmod{p}$, we have that

$$1 + \lambda + \cdots + \lambda^{n-1} \equiv 0 \pmod{p}.$$

This implies that $\lambda^n \equiv 1 \pmod{p}$ and so $r \mid n$. If $r = n$, then we are done and $e = 0$. Otherwise, replace f by $f^{\circ r}$ and n by n/r . Once again we can write $f(x)$ as a Taylor series

$$f(x) = \mu + \lambda x + \frac{g(x)}{1 + h(x)} x^2$$

and

$$1 + \lambda + \cdots + \lambda^{n-1} \equiv 0 \pmod{p}.$$

But now we have $\lambda = 1$, so we have

$$n \equiv 0 \pmod{p}.$$

We may now replace f with $f^{\circ p}$ and n by n/p and repeat the argument until $n = 1$. \square

Remark. If $\lambda_z = 0$, then the only possibility is $n = m$.

Example 10.54. Let $f(x) = x^2 - 21/16$. We saw in Example 10.48 that $\frac{1}{4}$ is periodic with minimal period 2 and fixed modulo 3 (a prime of good reduction). We compute the multiplier modulo 3

$$\lambda_{1/4} = f'(1/4) = 1/2 \equiv -1 \pmod{3}$$

so that $r = 2$. Thus, $e = 0$ and

$$n = mr.$$

Example 10.55. For $f(x) = x^3 + 2x^2 - x - 2$, we have that 2 is periodic with minimal period 2. The prime $p = 2$ is a prime of good reduction and 2 is fixed modulo 2. So we have $n = 2$ and $m = 1$. We compute λ_2 modulo 2 as

$$\lambda_2 = f'(2) = 3 \cdot 4 + 4 \cdot 2 - 1 \equiv 1 \pmod{2},$$

so that $r = 1$. Thus, $e = 1$, and we have

$$n = mp.$$

Remark. The integer e can be bounded, but the proof is much more difficult. For $f(x) \in \mathbb{Q}(x)$, we have

$$e \leq \begin{cases} 1 & p = 2, 3, \\ 0 & p > 3. \end{cases}$$

With the precise description of the reduced period, Theorem 10.53 allows us to answer Question 10.32 by bounding the minimal period of a periodic point for a given rational function $f(x)$.

Corollary 10.56. *Let p and q be distinct primes of good reduction for a rational function f . Then the minimal period of a periodic point defined over \mathbb{Q} is bounded as*

$$n \leq (p^2 - p)(q^2 - q).$$

Proof. For the prime p , Theorem 10.53 states that there are integers m_p, r_p , and e_p such that

$$n = m_p \quad \text{or} \quad n = p^{e_p}.$$

Let $q \neq p$ be a second prime of good reduction. Theorem 10.53 states that there are integers m_q, r_q , and e_q such that

$$n = m_q \quad \text{or} \quad n = q^{e_q}.$$

Since $\gcd(p, q) = 1$, we must have

$$n \leq m_p \quad \text{or} \quad n \leq m_p r_p m_q r_q \quad \text{or} \quad n \leq m_q,$$

depending on whether λ_z is zero modulo p or q .

Now we bound m_p and r_p in terms of p . Since there are only p residue classes modulo p , we can bound $m_p \leq p$. If the multiplier λ_z is zero modulo p , we have

$$n = m_p \quad \text{which implies} \quad n \leq p.$$

If the multiplier λ_z is not zero modulo p , then Fermat's Little Theorem (Theorem 2.21) implies that

$$\lambda_z^{p-1} \equiv 1 \pmod{p}$$

so that

$$r_p \leq p - 1.$$

We have the similar statements for q .

Consequently,

$$n \leq p(p-1)q(q-1). \quad \square$$

Remark. If we assume the stronger statement that the exponent e is bounded, the proof of Corollary 10.56 could be modified to use only one prime of good reduction.

While Corollary 10.56 answers Question 10.32, it does not answer the more general Question 10.18 since the primes of good and bad reduction depend on the function f . Question 10.18 remains an unsolved problem.

Investigation 10.57. While Theorem 10.53 describes what happens to the period of a periodic point modulo primes of good reduction, it says nothing about the tail of a preperiodic point. For the following questions, let $f(x) \in \mathbb{Q}[x]$ be a polynomial, and let z be a point with minimal cycle structure (m, n) with $m \neq 0$. Let (m', n') be the minimal cycle structure of \bar{z} .

- (a) Find a polynomial f , a point z , and a prime p of good reduction such that $m' = m$.
- (b) Find a polynomial f , a point z , and a prime p of good reduction such that $m' = 0$.
- (c) Find a polynomial f , a point z , and a prime p of good reduction such that $0 < m' < m$.
- (d) What does this say about a version of Theorem 10.53 for tails?

5. Algorithms for Rational Periodic and Preperiodic Points

While we have computed an explicit bound on the period of a rational periodic point, we have not yet determined how to find all the rational periodic and preperiodic points for a given function.

Algorithm 10.1 uses Theorem 10.53 and dynatomic polynomials to determine all rational periodic points for a given map.

Algorithm 10.1. Determining Rational Periodic Points by Good Reduction**Input:** a rational function $f \in \mathbb{Q}(x)$ **Output:** the set of periodic points defined over \mathbb{Q} **Algorithm:**

- 1: Compute $\text{Res}(f)$.
- 2: The primes of bad reduction S are the prime divisors of $\text{Res}(f)$.
- 3: Choose at least two (small) primes not in S .
- 4: For each prime use Theorem 10.53 to compute the list of possible periods.
- 5: Intersect the lists of possible periods to get a list N .
- 6: For each $n \in N$.
 - a: Compute the n th dynatomic polynomial, $\Phi_n(f)$.
 - b: Factor $\Phi_n(f)$ to determine any rational roots.
 - c: Add any rational roots to the list of periodic points.
- 7: Return the list of periodic points.

Example 10.58. Consider $f(x) = x^2 + 3x - 1$. There are no primes of bad reduction.We consider first $p = 2$. We compute

$$\begin{aligned} f(0) &\equiv 1 \pmod{2}, \\ f(1) &\equiv 1 \pmod{2} \end{aligned}$$

so that $x = 1$ is a fixed point modulo 2. We compute the multiplier as

$$f'(1) = 5 \equiv 1 \pmod{2},$$

whose multiplicative order is 1, so $r = 1$. The possible periods so far are

$$\{1, 2, 2^2, 2^3, \dots\}.$$

Working modulo 3, we have

$$\begin{aligned} f(0) &\equiv 2 \pmod{3}, \\ f(1) &\equiv 0 \pmod{3}, \\ f(2) &\equiv 0 \pmod{3}, \end{aligned}$$

so $x = 0$ is a point of period 2 modulo 3. We compute the multiplier as

$$(f^{\circ 2})'(0) = 0$$

so that the only possible period is

$$\{2\}.$$

Taking the intersection of the two lists of possible periods, we find that the only possible period is 2.

Now we compute $\Phi_2(f)$ and factor it as

$$\Phi_2(f) = (x + 1)(x + 3)$$

to have

$$\begin{array}{ccc} \bullet & \xrightarrow{\quad} & \bullet \\ -1 & \xleftarrow{\quad} & -3 \end{array}.$$

Example 10.59. Consider the polynomial

$$f(x) = \frac{x^3}{12} - \frac{25x}{12} + 1.$$

The resultant is

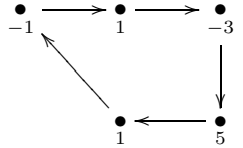
$$\text{Res}(f) = 1728 = 2^6 \cdot 3^3.$$

The primes of bad reduction are 2 and 3, so we start with $p = 5$. At $p = 5$ we get that 0 is periodic with period 5 with multiplier 0, so the list of possible periods is $\{5\}$. Since the list is already a single period, we move on to factoring the dynatomic polynomial.

Factoring the fifth dynatomic polynomial, we see that

$$\Phi_5(x) = (x-5)(x-3)(x-1)(x+1)(x+3)g(x),$$

where $g(x)$ is an irreducible polynomial of degree 235. Consequently, there is a rational 5 cycle and it is



Our original question (Question 10.13) asked for not just the rational periodic points, but the rational preperiodic points as well. We modify Algorithm 10.1 to determine the rational preperiodic points by finding the rational preimages of each rational periodic point z , i.e., the rational points such that $f(x) = z$. Assuming there are finitely many rational preperiodic points, this process will end after finitely many steps. Algorithm 10.2 describes the process.

Algorithm 10.2. Determining Rational Preperiodic Points by Good Reduction

Input: a rational function $f \in \mathbb{Q}(x)$

Output: the set of preperiodic points defined over \mathbb{Q}

Algorithm:

- 1: Let L be an empty set.
 - 2: Let S be the set of rational periodic points from Algorithm 10.1.
 - 3: Repeat the following until S is empty.
 - a: For each z in S .
 - i: Add z to L .
 - ii: Find the rational preimages of z by clearing denominators and factoring

$$f(x) = z.$$
 - iii: For each rational preimage not already in L , add the preimage to S .
 - iv: Remove z from S .
 - 4: Return the list L of rational preperiodic points.
-

Example 10.60. Consider $f(x) = x^2 + 3x - 1$ from Example 10.58. We have already seen that the rational periodic points are $\{-1, -3\}$. Now we compute the rational

preperiodic points. We start with $L = \emptyset$ and $S = \{-1, -3\}$. Starting with $z = -1$, we solve

$$\begin{aligned} f(x) &= -1, \\ x^2 + 3x - 1 &= -1, \\ x^2 + 3x &= x(x + 3) = 0, \end{aligned}$$

to have two rational preimages $\{-3, 0\}$. We add -1 to L and 0 to S to have

$$L = \{-1\} \quad S = \{0, -3\}.$$

Next we compute the rational preimages of $z = 0$ to have

$$\begin{aligned} f(x) &= 0, \\ x^2 + 3x - 1 &= 0. \end{aligned}$$

This polynomial is irreducible, so there are no rational preimages, and we add 0 to L to have

$$L = \{0, -1\}, \quad S = \{-3\}.$$

We compute the rational preimages of -3 as

$$\begin{aligned} f(x) &= -3, \\ x^2 + 3x - 1 + 3 &= (x + 1)(x + 2) = 0. \end{aligned}$$

We add -3 to L and the preimage -2 to S to have

$$L = \{0, -1, -3\}, \quad S = \{-2\}.$$

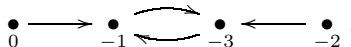
We compute the rational preimages of -2 as

$$\begin{aligned} f(x) &= -2, \\ x^2 + 3x + 1 &= 0. \end{aligned}$$


This polynomial is irreducible so there are no rational preimages, and we add -2 to L to have

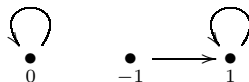
$$L = \{0, -1, -2, -3\} \quad S = \emptyset.$$

The algorithm terminates with $S = \emptyset$, and we have the directed graph

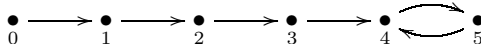


COMPUTATIONAL EXERCISES

 **10.1.** Find a polynomial with the following dynamical structure.



10.2. Find a polynomial with the following preperiodic cycle.




 **10.3.** Find all the fixed points defined over \mathbb{Q} for the function

$$f(x) = \frac{x^2 + 2x}{3x + 1}.$$

10.4. Use the dynatomic polynomial to find all the points defined over \mathbb{Q} with minimal period 2 for

$$f(x) = \frac{4}{3}x^3 - 2x^2 - \frac{7}{3}x + 2.$$

 **10.5.** What is the degree of $\Phi_{12}(f)$ for $f(x) = x^2 - 1$?

10.6. Find all the periodic points for $f(x) = x^2 - 29/16$ modulo 17.

10.7. Compute the discriminant $D = \text{Res}(f(x), f'(x))$ of the following polynomials.

a. $f(x) = ax^2 + bx + c$


b. $f(x) = x^3 + ax + b$

c. $f(x) = ax^3 + bx^2 + cx + d$

10.8. Consider the family of polynomials $f_c(x) = x^2 + c$. For $1 \leq n \leq 5$, find all $c \in \mathbb{C}$, where $f_c^{\circ n}(x) - x$ has a multiple root.

10.9. Find the primes of bad reduction for

$$f(x) = \frac{3x^3 + 2x + 5}{9x^3 + x + 9}.$$

 **10.10.** Compute the multiplier of $x = \frac{1}{2}$ for $f(x) = x^2 - \frac{7}{4}$.

10.11. Consider the rational maps


$$f(x) = \frac{9x + 81}{27x + 1} \quad \text{and} \quad g(x) = \frac{x^2 + 27x}{27x^2 + 3}.$$

Compute the primes of bad reduction for f , g , and $g \circ f$. Conclude that even if both f and g have bad reduction at a certain prime p , the composition may have good reduction at that prime.

10.12. Consider the collection of polynomials $f_c(x) = x^2 + c$ with $c \in \mathbb{Q}$ with height at most 50, i.e., $H(c) \leq 50$.

a. For which $f_c(x)$ are both 0 and 1 preperiodic?

b. For which $f_c(x)$ are both $\frac{5}{2}$ and $-\frac{3}{2}$ preperiodic?

 **10.13.** Use Theorem 10.53 with the primes 5 and 7 to find the possible periods of a periodic point defined over \mathbb{Q} for the polynomial

$$f(x) = x^2 - \frac{13}{9}.$$

10.14. Find the quadratic number field K with the following property. There exists a $c \in K$ such that 0 is preperiodic with minimal cycle structure (2, 2) for the polynomial $f(x) = x^2 + c$.

10.15. Using Algorithm 10.1, find all rational periodic points for

$$f(x) = x^2 - \frac{21}{16}.$$

10.16. Using Algorithm 10.2, find all rational preperiodic points for

$$f(x) = x^2 - \frac{13}{9}.$$

THEORETICAL EXERCISES

10.17. Let S be a finite set and $f : S \rightarrow S$. Prove that f is bijective if and only if every point of S is periodic for f .

10.18. Prove that the orbit of 3 by $f(x) = x^2 - 2x + 2$ is the set of Fermat numbers.

10.19. Prove that a rational function $f \in \mathbb{Q}(x)$ of degree $d > 1$ has at most $d^n + 1$ points with (not necessarily minimal) period n .

10.20. Prove that $f(x) = x^d + c$ for $c \in \mathbb{C}$ has no (real) points with minimal period greater than 1 if $d > 1$ is odd.

10.21. Prove that if $f(x) = x^2 + c$ for $c \in \mathbb{Q}$ has a periodic point in \mathbb{Q} , then the denominator of c must be a perfect square.

10.22. Let p be a prime and n a positive integer. Prove that

$$\Phi_n(f^{\circ p}) = \Phi_n(f)\Phi_{np}(f).$$

10.23. Let $f(x)$ be a rational function of degree $d > 1$. Prove that f can have at most $T(n)$ points with minimal period n , where

$$T(n) = \begin{cases} d + 1 & n = 1, \\ \sum_{k|n} \mu(n/k) d^k & n > 1. \end{cases}$$

10.24. Let $f(x), g(x) \in \mathbb{Q}(x)$ be rational functions. Prove that if $f(x)$ has good reduction at p and $g(x)$ has good reduction at p , then the composition $f \circ g$ also has good reduction at p .

10.25. Let $f(x)$ be a rational function and z a periodic point of period n .

a. Prove that

$$\lambda_z = \prod_{i=0}^{n-1} f'(f^{\circ i}(z)).$$

b. Conclude that multipliers for points in the same periodic orbit are the same.

10.26. For a polynomial $f(x)$ with no multiple roots, we define the Newton map as

$$F(x) = x - \frac{f(x)}{f'(x)}.$$

Prove that the fixed points of the Newton map have the following properties.

a. The fixed points of $F(x)$ are the roots of $f(x)$.

b. The multipliers of the fixed points of $F(x)$ are 0.

10.27. Let $f(x) \in \mathbb{Q}(x)$ be a rational function. Prove that if f has a periodic point defined over \mathbb{Q} with minimal period 3, then f has bad reduction modulo 2.

10.28. Prove that a monic polynomial with integer coefficients can have periodic points with minimal period only 1, 2, or 4.

10.29. Prove that any rational function $f(x) \in \mathbb{Q}(x)$ has only finitely many periodic points defined over \mathbb{Q} .

EXPLORATION EXERCISES

10.30 (Existence of periodic points). The map $f(x) = x^2 - 3/4$ has

$$f(x) - x = \frac{1}{4}(2x - 3)(2x + 1),$$

$$f^{\circ 2}(x) - x = \frac{1}{16}(2x - 3)(2x + 1)^3.$$

In other words, there are no periodic points with minimal period 2 (even over the complex numbers). This can only happen when there are periodic points with multiplicity > 1 .

- a. Consider the family $f(x) = x^2 + c$. Find c values where $f^{\circ n}(z) - z$ has multiple roots.
- b. Consider the family $f(x) = x^3 + cx$. Find c values where $f^{\circ n}(z) - z$ has multiple roots.
- c. Find a relationship between multipliers of periodic points and multiplicity of roots of $f^n(x) - x$.
- d. Look at the multiplicities of roots in $\Phi_n(f)$ for your c values from parts a and b. Do the formal periodic points which are multiple roots have minimal period n ?
- e. Find some other functions $f(x)$ for which $f^{\circ n}(z) - z$ has multiple roots.

10.31 (Multipliers of Lattès maps). In Chapter 9 we defined point addition on elliptic curves. Consider the map

$$[2] : E \rightarrow E$$

$$P \mapsto 2P.$$

Restricting to the x -coordinate of elliptic curves, this is a rational map of degree 4 called a *Lattès map*. In particular, this is a dynamical system.

- a. Pick a few different elliptic curves and find the fixed points (over \mathbb{C}) and multipliers of the associated Lattès map. Did you notice anything peculiar?
- b. Do you see the same phenomenon with periodic points of larger period?
- c. For the map

$$[m] : E \rightarrow E$$

$$P \mapsto mP$$

find the associated Lattès map by determining the rational function that takes the x -coordinate of P to the x -coordinate of mP .

- d. Determine what the multipliers should be for the multiplication by m Lattès map.

10.32 (Post-critically finite). Given a rational map $f(x)$, the *critical points* are the points where

$$f'(x) = 0.$$

The dynamical behavior of the critical points has a significant impact on the dynamical behavior of every point. The simplest possible dynamical behavior of a point is to have a finite forward orbit (i.e., preperiodic).

Definition 10.61. We say $f(x)$ is *post-critically finite* if all its critical points are preperiodic.

- a. Verify that $x^2 - 2$ is post-critically finite.
- b. Consider the family of polynomials $f_c(x) = x^2 + c$. Find polynomials in c whose roots are the c values where $f_c(x)$ is post-critically finite.
- c. What about the families $f_c(x) = x^d + c$ for $d > 2$?

Consider the family of cubic polynomials given by

$$g(x) = x^3 - 3a^2x + (2a^3 + v)$$

for constants a and v .

- d. What are the critical points of $g(x)$?
- e. Find values for a and v so that all critical points are fixed.
- f. What other values of a and v make $g(x)$ post-critically finite?

More generally,

- g. Verify that Lattès maps are post-critically finite.
- h. Can you find any other post-critically finite maps?

10.33 (Morton–Silverman uniform boundedness). Questions 10.18 and 10.20 ask for bounds on the minimal period of a periodic point defined over \mathbb{Q} ($C(d)$) and the number of preperiodic points defined over \mathbb{Q} ($M(d)$) for a rational map f that depends only on the degree of f . Morton and Silverman conjecture that these constants do exist. Conjecture values of the constants $C(d)$ and $M(d)$ for the following families.

- a. $f(x) = x^2 + c$.
- b. $f(x) = kx + \frac{b}{x}$.
- c. $f(x) = x^3 + ax + b$.
- d. What about other families?

10.34 (Conjugation).

Definition 10.62. A *linear fractional transformation* is a map of the form

$$\alpha(x) = \frac{ax + b}{cx + d}.$$

Such transformations are invertible if and only if $ad - bc \neq 0$.

Definition 10.63. For a rational function $f(x)$, we define

$$f^\alpha = \alpha \circ f \circ \alpha^{-1}$$

as the *conjugate* of f .

Example 10.64. Consider

$$f(x) = x^2 - 5x + 1 \quad \text{and} \quad \alpha = x - 5/2.$$

Then we have

$$f^\alpha(x) = x^2 - \frac{31}{4}.$$

- a. Is every degree 2 polynomial $f(x) = ax^2 + bx + c$ conjugate to

$$g(x) = x^2 + t$$

for some t ? In other words, does there exist an α with $f^\alpha = g$?

- b. How does the set of fixed points of f compare to the set of fixed points of f^α ?
 c. How does the set of multipliers of fixed points of f compare to the set of multipliers of fixed points of f^α ?
 d. What about periodic points and multipliers of periodic points with larger period?

10.35 (Canonical height and points of small height).

Definition 10.65. The *logarithmic height* of a rational number $\frac{a}{b}$ with $\gcd(a, b) = 1$ is given by

$$h(a/b) = \log(H(a/b)).$$

Definition 10.66. We define the *canonical height* of z with respect to f as

$$\hat{h}_f(z) = \lim_{n \rightarrow \infty} \frac{h(f^{\circ n}(z))}{(\deg(f))^n}.$$

Fix a function $f(x)$, say $f(x) = x^2 - 1$.

- a. Compare $\hat{h}_f(z)$ and $\hat{h}_f(f(z))$. How does the difference depend on f ?
 b. Can you bound the difference $|h(z) - \hat{h}_f(z)|$ independently of z ?
 c. Can you find any points with $\hat{h}_f(z) = 0$?
 d. What is the smallest value of $\hat{h}_f(z)$ that is not 0?
 e. Use parts b and c to give another algorithm for determining all rational preperiodic points. Apply it to $f(x)$.
 f. What is different when you change f ?

Polynomials

In this chapter we consider polynomials from the perspective of number theory.

Question 11.1. What number theoretic properties of integers have analogies for polynomials, e.g., primes, unique factorization, or modular arithmetic?

We first examine some of the basic properties of polynomials. Then we establish an analogy between number theory for integers and number theory for polynomials with monic irreducible polynomials playing the role of prime numbers. Finally, we examine some Diophantine equations with polynomial solutions.

1. Introduction to Polynomials

Definition 11.2. A *polynomial* is an expression of the form

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0,$$

where x is a variable and a_0, \dots, a_n are constants called the *coefficients* of $f(x)$.

We say f is defined over a ring R , notated $f \in R[x]$, when f is a polynomial with coefficients in R ; e.g., $f \in \mathbb{Z}[x]$ is a polynomial with integer coefficients.

Definition 11.3. If $a_n \neq 0$, we say $f(x)$ has *degree* n . A polynomial is *monic* if $a_n = 1$.

Example 11.4.

- (a) $x^2 + 1$ is a monic degree 2 polynomial in $\mathbb{Z}[x]$.
- (b) $3x^3 + \frac{7x}{2} - 5$ is a degree 3 polynomial in $\mathbb{Q}[x]$.
- (c) 5 is a degree 0 polynomial.

Definition 11.5. We say α is a *root* of the polynomial $f(x)$ if $f(\alpha) = 0$.

Definition 11.6. For $f = a_n x^n + \cdots + a_1 x + a_0$, with $f \in \mathbb{Z}[x]$, we define f modulo a prime p by reducing each coefficient of f modulo p :

$$\bar{f} = \bar{a}_n x^n + \cdots + \bar{a}_1 x + \bar{a}_0.$$

We say that $\bar{f} \in \mathbb{F}_p[x]$.

Definition 11.7. Let R be a ring. A nonzero polynomial $f \in R[x]$ that can be factored as $f = gh$ for nonunits $g, h \in R[x]$ is called *reducible over R* . Otherwise, f is *irreducible over R* .

Example 11.8.

- (a) $x^2 - 1$ is reducible over \mathbb{Z} since $x^2 - 1 = (x + 1)(x - 1)$.
- (b) $2x^2 + 2 = 2(x^2 + 1)$ is reducible over \mathbb{Z} since 2 is not a unit in \mathbb{Z} .
- (c) $x^2 + 1$ is irreducible over \mathbb{Z} .

Similar to factoring algebraic numbers in Chapter 7, irreducibility of polynomials depends on what kind of numbers we are working with.

Example 11.9.

- (a) $x^2 + 1 = (x + 1)^2$ is reducible in \mathbb{F}_2 (i.e., modulo 2).
- (b) The polynomial $f(x) = x^2 - 2$ is irreducible over \mathbb{Q} , but it can be factored as

$$(x - \sqrt{2})(x + \sqrt{2})$$

over $\mathbb{Q}(\sqrt{2})$.

- (c) The Fundamental Theorem of Algebra states that every polynomial of degree at least 2 is reducible over \mathbb{C} .

Investigation 11.10. Consider the monic quadratic polynomials $f_c(x) = x^2 + c$ with $c \in \mathbb{Z}$.

- (a) For which c is $f_c(x)$ irreducible over \mathbb{Z} ?
- (b) Are any of the reducible polynomials from part (a) irreducible modulo a prime $p \in \mathbb{Z}$?
- (c) Are any of the irreducible polynomials from part (a) reducible modulo a prime $p \in \mathbb{Z}$?

It is often a difficult task to prove that a polynomial is irreducible. We start by proving that irreducibility over \mathbb{Z} is equivalent to irreducibility over \mathbb{Q} , and then we give a criterion.

Definition 11.11. The greatest common divisor of the coefficients of a polynomial $f \in \mathbb{Z}[x]$ is called the *content* of f .

A polynomial $f \in \mathbb{Z}[x]$ is *primitive* if its content is 1.

Example 11.12.

- (a) $2x^2 + 6x + 8$ has content 2, so it is not primitive.
- (b) $3x^2 + 4x + 6$ has content 1, so it is primitive.

Lemma 11.13. *Let f and g be polynomials with integer coefficients. If f and g are both primitive, then their product fg is primitive.*

Proof. Suppose that $p \in \mathbb{Z}$ is a prime which divides all the coefficients of fg . Then $\overline{fg} \equiv 0 \pmod{p}$. Since reduction modulo a prime respects addition and multiplication, we have $\overline{fg} = \overline{f} \overline{g}$. Then we have $\overline{f} \overline{g} \equiv 0 \pmod{p}$, so we must have either $\overline{f} \equiv 0 \pmod{p}$ or $\overline{g} \equiv 0 \pmod{p}$. In other words, p divides the content of f or the content of g . Since f and g are both assumed to be primitive, no such prime can exist, so fg is also primitive. \square

Lemma 11.14. *Let $f \in \mathbb{Q}[x]$. There exists a constant $c \in \mathbb{Q}$ such that $f = cf_0$ for $f_0 \in \mathbb{Z}[x]$ primitive.*

Proof. By taking the least common multiple of the denominators of the coefficients of f , we have $mf \in \mathbb{Z}[x]$ for some integer m . Let n be the content of mf . Then we have

$$f_0 = \frac{m}{n}f$$

so

$$f = \frac{n}{m}f_0,$$

where $f_0 \in \mathbb{Z}[x]$ is primitive. \square

Theorem 11.15. *Let f be a primitive polynomial with integer coefficients. Then, $f(x)$ is irreducible over \mathbb{Q} if and only if $f(x)$ is irreducible over \mathbb{Z} .*

Proof. If f is irreducible over \mathbb{Q} , since $\mathbb{Z} \subset \mathbb{Q}$, the only way f can be reducible over \mathbb{Z} is to factor out a constant which is not a unit. But since f is primitive, it has content 1, so this is not possible.

Assume now that f is irreducible over \mathbb{Z} but reducible over \mathbb{Q} . Say $f = gh$ for $g, h \in \mathbb{Q}[x]$ each with degree at least 1. Using Lemma 11.14, write $g = cg_0$ and $h = dh_0$ for g_0, h_0 primitive and $c, d \in \mathbb{Q}$. Lemma 11.13 tells us that g_0h_0 must also be primitive. Then we have

$$f = cdg_0h_0.$$

However, since f , g_0 , and h_0 all have integer coefficients, we must have $cd \in \mathbb{Z}$, so f factors as $f = (cd)g_0h_0$ over \mathbb{Z} . This contradicts that f is irreducible over \mathbb{Z} , so f must be irreducible over \mathbb{Q} . \square

Since irreducibility is equivalent over \mathbb{Q} and \mathbb{Z} , we generally assume that we are working with (primitive) integer polynomials.

We give one criteria for the irreducibility of polynomials that plays a role in dynamical systems (Theoretical Exercise 11.16).

Theorem 11.16 (Eisenstein's criteria). *Let $f(x) = a_nx^n + \cdots + a_1x + a_0$ be a polynomial with integer coefficients. If there exists a prime $p \in \mathbb{Z}$ such that $p \nmid a_n$ and $p \mid a_0, a_1, \dots, a_{n-1}$ and $p^2 \nmid a_0$, then f is irreducible over \mathbb{Z} .*

Proof. Assume that f can be written as the product of polynomials with integer coefficients

$$f = (b_rx^r + \cdots b_1x + b_0)(c_sx^s + \cdots c_1x + c_0).$$

By assumption, $p \mid b_0 c_0 = a_0$ and $p^2 \nmid b_0 c_0$. Thus, $p \mid b_0$ and $p \nmid c_0$ (or $p \mid c_0$ and $p \nmid b_0$). If all the b_i are divisible by p , then so is a_n , which is a contradiction. Let b_j , $j < n$, be the first coefficient not divisible by p . Consider

$$a_j = b_j c_0 + b_{j-1} c_1 + \cdots + b_0 c_j.$$

Every term on the right-hand side except $b_j c_0$ is divisible by p , so $p \nmid a_j$. This is a contradiction and, therefore, f is irreducible. \square

Example 11.17. Consider the polynomial $f(x) = x^5 + 3x^3 - 18x + 6$. For $p = 3$, we have

$$p \nmid 1 \quad \text{and} \quad p \mid 3, -18, 6$$

and that

$$p^2 \nmid 6.$$

Thus, f is irreducible over \mathbb{Z} by Eisenstein's criteria.

This criteria is by no means all encompassing.

Example 11.18. The polynomial $f(x) = x^5 + 3x^2 + 9$ does not satisfy the hypotheses of Eisenstein's criteria, yet it is irreducible over \mathbb{Z} .

Investigation 11.19. Let $p \in \mathbb{Z}$ be a prime. There are only finitely many polynomials in $\mathbb{F}_p[x]$ of a fixed degree, so there can only be finite many irreducible polynomials in $\mathbb{F}_p[x]$ of a fixed degree.

- Count the number of irreducible polynomials of degree 2 modulo several different primes. What proportion of all degree 2 polynomials are irreducible?
- How many of these irreducible polynomials are monic?
- Repeat for degrees 3, 4, and 5.
- Can you make any general statements?

2. Factorization and the Euclidean Algorithm

As with integers, we define notions of divisibility, greatest common divisor, and the Euclidean algorithm. You will find both the statements and the proofs in this section strikingly similar to those in Chapter 1.

Definition 11.20. Let R be \mathbb{Z} , \mathbb{Q} , or \mathbb{F}_p and $f, g \in R[x]$ be polynomials. We say that f divides g if there is a polynomial $h \in R[x]$ such that $g = fh$. We denote divisibility as $f \mid g$.

Example 11.21. We see that $x - 1$ divides $x^2 - 1$ since

$$x^2 - 1 = (x - 1)(x + 1).$$

Example 11.22. A constant (a degree 0 polynomial) divides any polynomial $f \in \mathbb{Q}[x]$. If $f = a_n x^n + \cdots + a_1 x + a_0$, then

$$3 \left(\frac{a_n x^n}{3} + \cdots + \frac{a_0}{3} \right) = a_n x^n + \cdots + a_0 \Rightarrow 3 \mid f.$$

Theorem 11.23 (Polynomial division algorithm). *Let R be \mathbb{Q} or \mathbb{F}_p . If $f, g \in R[x]$ and $f \neq 0$, then there exists unique $q, r \in R[x]$ such that $g = qf + r$ with $\deg(r) < \deg(f)$.*

Proof. First we show existence. If $\deg(f) > \deg(g)$, we set $q = 0$ and $r = g$. Otherwise, let $d = \deg(g) - \deg(f)$. For notation, let

$$\begin{aligned} f &= a_m x^m + \cdots + a_1 x + a_0, \\ g &= b_n x^n + \cdots + b_1 x + b_0. \end{aligned}$$

We proceed by induction on d . If $d = 0$, then $m = n$ and we set

$$\begin{aligned} q &= \frac{b_n}{a_m}, \\ r &= g - fq. \end{aligned}$$

So we have

$$g = qf + r \quad \text{with} \quad \deg(r) < \deg(g).$$

Let k be a positive integer, and assume the theorem is true for $d < k$. Let $d = k$, and define

$$g_1 = g - \frac{b_n}{a_m} x^d f$$

so that $\deg(g_1) < \deg(g)$. By induction, there exists q_1 and r such that

$$g_1 = q_1 f + r \quad \text{with} \quad \deg(r) < \deg(f).$$

Therefore,

$$q_1 f + r = g_1 = g - \frac{b_n}{a_m} x^d f$$

so that

$$g = \left(q_1 + \frac{b_n}{a_m} x^d \right) f + r \quad \text{with} \quad \deg(r) < \deg(f).$$

Now for uniqueness, suppose that

$$g = q_1 f + r_1 = q_2 f + r_2 \quad \text{with} \quad \deg(r_1), \deg(r_2) < \deg(f).$$

Rearranging, we have

$$r_1 - r_2 = (q_2 - q_1)f,$$

so $f \mid (r_1 - r_2)$. But since $\deg(r_1), \deg(r_2) < \deg(f)$, this can happen only if $r_1 - r_2 = 0$. In this case

$$(q_2 - q_1)f = 0,$$

and we have $q_1 = q_2$. □

Example 11.24. Consider $f = x^2 + 1$ and $g = x^5 + 2x^4 - 1$. Then

$$x^5 + 2x^4 - 1 = (x^3 + 2x^2 - x - 2)(x^2 + 1) + (x + 1).$$

So we have

$$\begin{aligned} q &= x^3 + 2x^2 - x - 2, \\ r &= x + 1. \end{aligned}$$

We saw in Example 11.22 that a constant divides any polynomial. Consequently, to define a unique greatest common divisor of two polynomials, we need to restrict to monic divisors.

Definition 11.25. Let R be \mathbb{Q} or \mathbb{F}_p . The *greatest common divisor* of two polynomials $f, g \in R[x]$, denoted $\gcd(f, g)$, is the monic polynomial with coefficients in R of highest degree that divides both f and g .

Example 11.26. We have

$$\gcd(x^5 + 1, x^3 + 1) = x + 1.$$

Theorem 11.27. Let R be \mathbb{Q} or \mathbb{F}_p and $f, g \in R[x]$. The greatest common divisor of f and g is the monic polynomial of the smallest degree that is a linear (polynomial) combination of f and g .

Proof. Let $d \in R[x]$ be the smallest degree monic polynomial that is a linear combination of f and g . Write $d = af + bg$ for polynomials $a, b \in R[x]$. If $d = 1$, we are done. Otherwise, we need to prove that $d \mid f$ and $d \mid g$.

Using the division algorithm (Theorem 11.23), we write $f = qd + r$ for $0 \leq \deg(r) < \deg(d)$. Then we have

$$r = f - qd = f - q(af + bg) = (1 - qa)f - qbg.$$

Thus, r is also a linear combination of f and g . Since d is the smallest degree monic polynomial that is linear combination and $\deg(r) < \deg(d)$, we must have $r = 0$. Since r is the remainder of f after division by d , $r = 0$ implies that $d \mid f$. Similarly, we can show $d \mid g$.

Therefore, $d \mid \gcd(f, g)$. □

Example 11.28. Let $f = x^5 + 1$ and $g = x^3 + 1$. We write the greatest common divisor $x + 1$ as

$$xf + (1 - x^3)g = x + 1.$$

Corollary 11.29. Let R be \mathbb{Q} or \mathbb{F}_p and $d, f, g \in R[x]$. If $d \mid f$ and $d \mid g$, then $d \mid \gcd(f, g)$

Proof. We can write $\gcd(f, g) = af + bg$ for some $a, b \in R[x]$. Since d divides the right-hand side, it must also divide the left-hand side. □

Theorem 11.30 (Polynomial Euclidean algorithm). Let R be \mathbb{Q} or \mathbb{F}_p and $f, g \in R[x]$. Let $r_0 = f$ and $r_1 = d$, where $f = dg + r$. We construct a sequence of polynomials using the division algorithm

$$r_i = q_{i+1}r_{i+1} + r_{i+2}, \quad i \geq 0.$$

The last nonzero r_i is the greatest common divisor of f and g .

Proof. We proceed in two steps. First we show that the algorithm terminates, then we show that the terminating value is the greatest common divisor.

We proceed by the division algorithm. At each step

$$0 \leq \deg(r_{i+1}) < \deg(r_i)$$

so that after finitely many steps $\deg(r_i) = 0$. When $\deg(r_i) = 0$, r_i is a constant. Every constant exactly divides every polynomial when working over \mathbb{Q} so that $r_{i+1} = 0$ and the algorithm terminates.

We need to show that $\gcd(r_i, r_{i+1}) = \gcd(r_{i+1}, r_{i+2})$. Let $a = \gcd(r_i, r_{i+1})$ and $b = \gcd(r_{i+1}, r_{i+2})$. We will show $a \mid b$ and $b \mid a$.

Since $a \mid r_i$, $a \mid r_{i+1}$, and $r_i = q_{i+1}r_{i+1} + r_{i+2}$, we must also have $a \mid r_{i+2}$. By Corollary 11.29, we have $a \mid \gcd(r_{i+1}, r_{i+2}) = b$.

Similarly, $b \mid r_{i+1}$, $b \mid r_{i+2}$, and $r_i = q_{i+1}r_{i+1} + r_{i+2}$, so $b \mid r_i$. By Corollary 11.29, we have $b \mid \gcd(r_i, r_{i+1}) = a$.

We have $a \mid b$ and $b \mid a$ with both a and b monic, so $a = b$. \square

Example 11.31. We compute

$$\begin{aligned} \gcd(x^5 + 1, x^3 + 1) &= \gcd(-x^2 + 1, x^3 + 1) \\ &= \gcd(-x^2 + 1, x + 1) \\ &= \gcd(0, x + 1) \\ &= x + 1. \end{aligned}$$

Investigation 11.32. We saw in Theorem 1.35 that the Fibonacci numbers are the worst-case scenario in the number of steps needed to find the greatest common divisor for the Euclidean algorithm for integers. What is the worst-case scenario for the polynomial Euclidean algorithm?

The next lemma shows that irreducible polynomials have the same role in divisibility as prime numbers.

Lemma 11.33. *Let R be \mathbb{Q} or \mathbb{F}_p . Let $f, g, h \in R[x]$. If f is irreducible and $f \mid gh$, then $f \mid g$ or $f \mid h$.*

Proof. Assume that $f \nmid g$. Then $\gcd(f, g) = 1$. We can find polynomials $a, b \in R[x]$ such that

$$af + bg = 1.$$

Multiplying both sides by h , we have

$$afh + bgh = h.$$

Since f divides both terms on the left-hand side, it must also divide the right-hand side, i.e., $f \mid h$. \square

Theorem 11.34 (Polynomial unique factorization). *Let R be \mathbb{Q} or \mathbb{F}_p . Every polynomial $f \in R[x]$ can be written uniquely (up to reordering) in the form*

$$f = cp_1 \cdots p_r,$$

where $c \in R$ and $p_i \in R[x]$ are monic and irreducible.

Proof. We proceed by induction on the degree of f . If $\deg(f) = 1$, then f is linear and is irreducible. We can factor out the leading coefficient to make f monic.

Let k be a positive integer. Assume all polynomials with degree less than k can be factored into monic irreducible polynomials. Let $\deg(f) = k$. If f is irreducible, we are done, so assume that f is reducible. Then there exist polynomials g and h with $f = gh$ and $\deg(g), \deg(h) < \deg(f)$. By induction, g and h can be factored into irreducibles. Thus, f can be factored into irreducibles. To make them monic, we factor out the leading coefficients.

Now we show that this factorization is unique. Assume we have two factorizations,

$$\begin{aligned} f &= cp_1 \cdots p_r, \\ f &= dq_1 \cdots q_s. \end{aligned}$$

We know $p_1 \mid q_1 \cdots q_s$, and by Lemma 11.33 we must have $p_1 \mid q_i$ for some i . After relabeling, we may assume $i = 1$. So we have

$$cp_2 \cdots p_r = dq_2 \cdots q_s.$$

We continue in this process to see that $p_i = q_i$ for $1 \leq i \leq r$ and $r = s$. Finally, we have $c = d$. \square

Example 11.35.

- $x^2 - 1 = (x + 1)(x - 1)$.
- $x^3 + 2x^2 - 2x + 3 = (x^2 - x + 1)(x + 3)$.

Investigation 11.36. Consider the values of polynomials with integer coefficients. In particular, if $f(x) = x^2 + 1$, we say $f(2) = 5$ is a *value* of f . Explore how the irreducibility of f affects the occurrence of prime values for a polynomial.

- (a) Consider the reducible polynomial $f(x) = x^2 - 1$. Which integer values of x give prime values for f ?
- (b) Consider the irreducible polynomial $f(x) = x^2 + 1$. Which integer values of x give prime values for f ?
- (c) Can a reducible polynomial have infinitely many prime values?

3. Modular Arithmetic for Polynomials

We have already seen a few justifications that irreducible polynomials are similar to prime numbers.

- (a) *Divisibility*: For f irreducible, if $f \mid gh$, then $f \mid g$ or $f \mid h$.
- (b) gcd: For f monic and irreducible, $\gcd(f, g) = 1$ or $\gcd(f, g) = f$.
- (c) unique factorization into monic irreducibles.

Some other key property of primes occur in modular arithmetic, i.e., studying the integers modulo a prime. Recall that we are denoting \mathbb{F}_p as the set of residue classes of the integers modulo p . Then

- (d) \mathbb{F}_p is finite.
- (e) Every nonzero element in \mathbb{F}_p has a multiplicative inverse.
- (f) For $a, b \in \mathbb{F}_p$, if $ab = 0$, then $a = 0$ or $b = 0$.

We leave properties (e) and (f) to the exercises (Theoretical Exercises 11.22 and 11.23) and focus on property (d) and its consequences.

For polynomials f and g , f modulo g is the notion of remainder after division. This notion is well-defined since the division algorithm for polynomials (Theorem 11.23) proves that there is a unique remainder. We give the same definition for modular arithmetic of polynomials as we did for modular arithmetic of integers.

Definition 11.37. Let m , f , and g be polynomials. We say that

$$f \equiv g \pmod{m}$$

if m divides $f - g$.

Example 11.38.

$$x^7 + 3x^6 - 2x^5 + x^3 - 2x^2 + x - 3 \equiv 2x + 1 \pmod{x^3 + 1}.$$

As with integers, modular arithmetic for polynomials respects addition and multiplication.

Theorem 11.39. Let a , b , c , d , and m be polynomials. Assume that $a \equiv b \pmod{m}$ and $c \equiv d \pmod{m}$.

- (a) $a + c \equiv b + d \pmod{m}$.
- (b) $ac \equiv bd \pmod{m}$.

If in addition we denote $g = \gcd(c, m)$, then

- (c) if $ac \equiv bc \pmod{m}$, then $a \equiv b \pmod{m/g}$.

Proof. Identical to Theorem 2.10. □

Definition 11.40 generalizes the fact that for a positive integer n , there are $|n|$ residue classes of \mathbb{Z} modulo n .

Definition 11.40. Let $f \in \mathbb{F}_p[x]$. Define $|f| = p^{\deg(f)}$.

Example 11.41.

- (a) $|x^2 + 2| = 3^2 = 9$ in $\mathbb{F}_3[x]$.
- (b) $|x^3 - x + 1| = 2^3 = 8$ in $\mathbb{F}_2[x]$.
- (c) $|7| = p^0 = 1$ in $\mathbb{F}_p[x]$.

Proposition 11.42. Let $f \in \mathbb{F}_p[x]$. There are $|f|$ polynomials in $\mathbb{F}_p[x]$ modulo f .

Proof. Let $d = \deg(f)$. Every polynomial in $\mathbb{F}_p[x]$ is congruent to a polynomial of degree less than d by the division algorithm (Theorem 11.23). So a complete set of residue classes of $\mathbb{F}_p[x]$ modulo f are of the form $a_{d-1}x^{d-1} + \cdots + a_1x + a_0$ for $a_0, \dots, a_{d-1} \in \mathbb{F}_p$. There are p choices for each a_i , and there are d coefficients a_i . So there are $p^d = |f|$ possible residue classes. □

It is essential in Proposition 11.42 that the coefficients are in \mathbb{F}_p . If we were to reduce a polynomial in $\mathbb{Q}[x]$ by a monic irreducible polynomial f , we would still get that each residue class has a representative of degree strictly less than the degree of f . However, there are now infinitely many choices for each coefficient!

Example 11.43. The polynomials in $\mathbb{F}_2[x]$ modulo x^4+x+1 are all the polynomials of degree at most 3 with coefficients in $\{0, 1\}$. There are 16 of them:

$$\begin{aligned} \text{degree 0: } & \{0, 1\}, \\ \text{degree 1: } & \{x, x+1\}, \\ \text{degree 2: } & \{x^2, x^2+1, x^2+x, x^2+x+1\}, \\ \text{degree 3: } & \{x^3, x^3+1, x^3+x, x^3+x+1, x^3+x^2, x^3+x^2+1, \\ & x^3+x^2+x, x^3+x^2+x+1\}. \end{aligned}$$

Since there are finitely many polynomials of bounded degree, we can also define an Euler totient function.

Definition 11.44. Let $f \in \mathbb{F}_p[x]$. Define $\varphi(f)$ to be the number of nonzero polynomials of degree strictly less than $\deg(f)$ which are relatively prime to f .

Example 11.45. Proposition 11.42 proves that for $f \in \mathbb{F}_p[x]$ irreducible,

$$\varphi(f) = |f| - 1.$$

Investigation 11.46. Fix a monic irreducible polynomial $m \in \mathbb{F}_p[x]$. For each $f \in \mathbb{F}_p[x]$ nonzero modulo m , there is an exponent e such that

$$f^e \equiv 1 \pmod{m},$$

called the *multiplicative order* of f modulo m .

- (a) For $m = x^4 + x + 1 \in \mathbb{F}_2[x]$, compute the multiplicative order of each residue class. Compare your answers with $\varphi(m)$.
- (b) Choose a different p and m and repeat part (a).
- (c) Can you make a general statement?

Theorem 11.47.

- (a) Let f be an irreducible polynomial, and let $e \geq 1$ be an integer. Then

$$\varphi(f^e) = |f|^{e-1} (|f| - 1) = |f|^e \left(1 - \frac{1}{|f|}\right).$$

- (b) Let f and g be polynomials such that $\gcd(f, g) = 1$. Then

$$\varphi(fg) = \varphi(f)\varphi(g).$$

- (c) Let f be a polynomial and $p_1 \cdots p_r$ be the distinct irreducible factors of f . Then

$$\varphi(f) = |f| \left(1 - \frac{1}{|p_1|}\right) \cdots \left(1 - \frac{1}{|p_r|}\right).$$

Proof.

- (a) From the set of polynomials of degree at most $\deg(f^e) = \deg(f)^e$, the ones not relatively prime to f^e are exactly those divisible by f , in other words, the set $\{gf : \deg(g) \leq \deg(f)^{e-1}\}$. There are $|f|^{e-1}$ of these, so the total number relatively prime is

$$|f|^e - |f|^{e-1} = |f|^{e-1}(|f| - 1) = |f|^e \left(1 - \frac{1}{|f|}\right).$$

- (b) If $\gcd(f, g) = 1$, then each irreducible divisor of fg divides only one of f or g . Or, to be relatively prime to fg , a polynomial must be relatively prime to both f and g .
- (c) Using the first two portions of the theorem and given $f = p_1^{e_1} \cdots p_r^{e_r}$, we compute

$$\begin{aligned} \varphi(f) &= \varphi(p_1^{e_1} \cdots p_r^{e_r}) \\ &= \varphi(p_1^{e_1}) \cdots \varphi(p_r^{e_r}) \\ &= |p_1|^{e_1} \left(1 - \frac{1}{|p_1|}\right) \cdots |p_r|^{e_r} \left(1 - \frac{1}{|p_r|}\right) \\ &= |f| \left(1 - \frac{1}{|p_1|}\right) \cdots \left(1 - \frac{1}{|p_r|}\right). \end{aligned} \quad \square$$

Example 11.48. Let $f = x^5 - x^4 - x - 1 \in \mathbb{F}_5[x]$. Then we have

$$f = (x^2 + 1)(x - 1)^2(x + 1)$$

so that

$$\varphi(f) = 5^2 \cdot (5^2 - 5) \cdot 5 = 5^4 \cdot 4 = 2500.$$

Just like Fermat's Little Theorem (Theorem 2.21) and the more general Euler's formula (Theorem 2.31) for integers, every polynomial relatively prime to the modulus has multiplicative order dividing $\varphi(f)$.

Theorem 11.49 (Polynomial Euler's formula). *If $f \in \mathbb{F}_p[x]$ is nonzero and $g \in \mathbb{F}_p[x]$ is relatively prime to f , then*

$$g^{\varphi(f)} \equiv 1 \pmod{f}.$$

Proof. Let $A = \{a_1, \dots, a_{\varphi(f)}\}$ be the set of polynomials relatively prime to f in $\mathbb{F}_p[x]$ modulo f . Since g is relatively prime to f , we also have that

$$B = \{ga_1, \dots, ga_{\varphi(f)}\}$$

are all relatively prime to f . If

$$ga_i \equiv ga_j \pmod{f},$$

then

$$f \mid (a_i - a_j).$$

However, since $\deg(a_i) < \deg(f)$ and $\deg(a_j) < \deg(f)$, we must have $(a_i - a_j) = 0$ so that $i = j$. Thus, the sets A and B are the same, so the product of their elements is the same:

$$\prod_{a \in A} ga \equiv \prod_{a \in A} a \pmod{f}.$$

We collect the g 's in the product and apply the inverse of each $a \in A$ to have

$$\begin{aligned} g^{\varphi(f)} \prod_{a \in A} a &\equiv \prod_{a \in A} a \pmod{f}, \\ g^{\varphi(f)} &\equiv 1 \pmod{f}. \end{aligned} \quad \square$$

Corollary 11.50 (Polynomial Fermat's Little Theorem). *Let f be irreducible. Then*

$$g^{|f|-1} \equiv 1 \pmod{f}.$$

Investigation 11.51. Fix a polynomial $m \in \mathbb{F}_p[x]$. For each f in $\mathbb{F}_p[x]$ relatively prime to m , there is an exponent e such that

$$f^e \equiv 1 \pmod{m},$$

called the *multiplicative order* of f modulo m .

- (a) Define a notion of primitive root modulo m .
- (b) When m is irreducible, is there a primitive root?
- (c) When m is a product of distinct irreducibles, is there a primitive root?
- (d) What about for general m ?

3.1. Counting Irreducible Polynomials. An important question in number theory is the density and number of prime numbers. We consider the following question for polynomials.

Question 11.52. Which monic polynomials in $\mathbb{F}_p[x]$ are irreducible? How many are irreducible?

Let's examine all polynomials of degree 3 or less in $\mathbb{F}_2[x]$, i.e., $p = 2$.

Degree	Irreducible	Reducible
1	x $x + 1$	
2	$x^2 + x + 1$	$x^2 + x$ $x^2 + 1$ x^2
3	$x^3 + x + 1$ $x^3 + x^2 + 1$	x^3 $x^3 + 1$ $x^3 + x$ $x^3 + x^2$ $x^3 + x^2 + x$ $x^3 + x^2 + x + 1$

While this list does give some irreducible polynomials, the pattern of which ones are irreducible is not readily apparent.

Before we can count the number of monic irreducible polynomials, we need a couple of helpful lemmas.

Lemma 11.53. *Let p be a prime number, and let n and d be integers. Then, $x^{p^d} - x$ divides $x^{p^n} - x$ if and only if d divides n .*

Proof. The roots of $x^n - 1$ are the n th roots of unity, and the roots of $x^d - 1$ are the d th roots of unity. The d th roots of unity are n th roots of unity if and only if $d \mid n$. We now need to show that

$$p^d - 1 \mid p^n - 1 \quad \text{if and only if} \quad d \mid n.$$

If $p^d - 1 \mid p^n - 1$, then we perform long division:

$$\begin{array}{r} p^d - 1 \overline{) \begin{array}{l} p^{n-d} + p^{n-2d} + \dots \\ p^n \phantom{- p^{n-d}} - 1 \\ \hline p^n - p^{n-d} \\ p^{n-d} \phantom{- p^{n-2d}} - 1 \\ \hline p^{n-d} - p^{n-2d} \\ \hline p^{n-2d} \phantom{- p^{n-4d}} - 1 \\ \hline \vdots \end{array}} \end{array}$$

Assuming divisibility means the remainder is 0, we must have $p^n = kd$ for some k . Thus, $d \mid n$.

If $d \mid n$, then

$$p^n - 1 = p^{dk} - 1 = (p^d - 1)(1 + p^d + \dots + p^{d(k-1)}). \quad \square$$

Lemma 11.54. *Let $f \in \mathbb{F}_p[x]$ be an irreducible polynomial of degree d . If α is a root of f , then α is a root of $x^{p^d} - x$ modulo p .*

Proof. We know there are $|f| = p^d$ elements in $\mathbb{F}_p[x]$ modulo f . From Corollary 11.50, we know that for any $g \in \mathbb{F}_p[x]$,

$$g^{p^d} - g \equiv 0 \pmod{f}.$$

In particular,

$$\alpha^{p^d} - \alpha \equiv 0 \pmod{f}.$$

Since α is a constant, for any $n \in \mathbb{N}$,

$$\alpha^n \pmod{f} = \alpha^n \pmod{p}.$$

Thus, α is a root of $x^{p^d} - x$ in $\mathbb{F}_p[x]$. \square

The following theorem does not give us information about specific irreducible polynomials; rather, it gives us information about the entire set of irreducible polynomials of some fixed degree.

Theorem 11.55. *Let $I_d(x)$ be the product of all monic irreducible polynomials of degree d in $\mathbb{F}_p[x]$. Then for any positive integer n ,*

$$(42) \quad x^{p^n} - x = \prod_{d \mid n} I_d(x).$$

Proof. We first show that $x^{p^n} - x$ is squarefree (i.e., there are no repeated factors). Assume there is an f such that $x^{p^n} - x = f^2 g$. Differentiate both sides to get

$$-1 = 2ff'g + f^2g',$$

so f divides -1 , which is a contradiction.

Now we show that the irreducible factors of both sides of equation (42) are the same.

Assume that $f \mid x^{p^n} - x$. If α is a root of f , then by Lemma 11.54 α is a root of $x^{p^d} - x$. Since $\alpha \neq 0$, we must have $x^{p^d-1} - 1 \mid x^{p^n-1} - 1$, so $d \mid n$ by Lemma 11.53.

Let f be an irreducible polynomial of degree d with $d \mid n$, so $f \mid I_d(x)$. Let α be a root of f . By Lemma 11.54 α is a root of $x^{p^d} - x$, and then, by Lemma 11.53, α is a root of $x^{p^n} - x$. Since f is irreducible, it is the characteristic polynomial of α , so we must have $f \mid x^{p^n} - x$. \square

Knowing information about the product of all monic irreducible polynomials of fixed degree allows us to count their number.

Corollary 11.56. *Let $N_p(d)$ be the number of monic irreducible polynomials of degree d in $\mathbb{F}_p[x]$. Then for any positive integer n ,*

$$p^n = \sum_{d \mid n} d N_p(d).$$

Proof. Equate the degrees on both sides of equation (42). \square

Corollary 11.57. *Let n be a positive integer. There are*

$$N_p(n) = \frac{1}{n} \sum_{d \mid n} \mu(d) p^{n/d}$$

monic irreducible polynomials of degree n in $\mathbb{F}_p[x]$.

Proof. Apply the Möbius inversion formula (Theorem 5.20) to the previous corollary. \square

Example 11.58. There are 78,120 monic irreducible polynomials of degree 5 in $\mathbb{F}_5[x]$.

4. Diophantine Equations for Polynomials

When working with Diophantine equations in Chapter 8, we could just as easily have allowed for equations with polynomial coefficients and polynomial solutions. Surprisingly, some of the problems that were very hard for integers turn out to be simple for polynomials. For example, in section 4.2 we will prove Fermat's Last Theorem for polynomials!

Note that in this section we use x as a variable in the Diophantine equations and not as the polynomial variable.

4.1. Pythagorean triples. We wish to solve

$$x^2 + y^2 = z^2,$$

where x , y , and z are polynomials with integer coefficients. In section 8.3, we saw that all Pythagorean triples are of the form

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

for any integers a and b with $b \neq 0$. If we think of this as a two-variable polynomial with variables a and b and integer coefficients, then we say that

$$x^2 + y^2 = z^2$$

has the unique solution

$$(43) \quad (a^2 - b^2, 2ab, a^2 + b^2)$$

in $\mathbb{Z}[a, b]$. Choosing any integer pair (a, b) gives an integer solution by substituting the pair into the polynomial solution. Notice that if a Diophantine equation with integer coefficients has a polynomial solution, then there are infinitely many integer solutions.

Thinking of equation (43) as a single solution represents a significant shift in our thought process. We are no longer looking at $(a^2 - b^2, 2ab, a^2 + b^2)$ as giving an infinite set of integer values that are Pythagorean triples. Instead we are viewing $(a^2 - b^2, 2ab, a^2 + b^2)$ as a single two-variable polynomial Pythagorean triple.

4.2. Fermat's Last Theorem.

Theorem 11.59. *If n is a positive integer with $n > 2$, the equation $x^n + y^n = z^n$ has no solutions in polynomials of degree at least 1 with integer coefficients.*

Proof. Assume that $f^n + g^n = h^n$ with $\gcd(f, g) = 1$. Take the derivative of both sides to get

$$(44) \quad f'f^{n-1} + g'g^{n-1} = h'h^{n-1}.$$

Now subtract f times equation (44) from f' times the original Fermat equation to get

$$f'g^n - fg'g^{n-1} = f'h^n - fh'h^{n-1}.$$

In particular,

$$g^{n-1}(f'g - fg') = h^{n-1}(fh' - f'h).$$

Suppose $f'g - fg' = 0$ so that $f'g = fg'$. Since $\gcd(f, g) = 1$, this implies that $g \mid g'$ and $f \mid f'$. But that is not possible if $\deg(f), \deg(g) > 0$. So we can (similarly) assume that $f'g - fg'$, $f'h - fh'$, and $g'h - gh'$ are all not 0.

Since $\gcd(g, h) = 1$, we must have $g^{n-1} \mid f'h - fh'$ and $h^{n-1} \mid f'g - fg'$. In particular, this implies

$$\deg(g^{n-1}) \leq \deg(f) + \deg(h) - 1,$$

$$\deg(h^{n-1}) \leq \deg(f) + \deg(g) - 1.$$

Doing a slightly different subtraction of the two original equations, we could also get $f^{n-1} \mid g'h - gh'$ so that

$$\deg(f^{n-1}) \leq \deg(g) + \deg(h) - 1.$$

So we have

$$\begin{aligned} n \deg(f) &\leq \deg(f) + \deg(g) + \deg(h) - 1, \\ n \deg(g) &\leq \deg(f) + \deg(g) + \deg(h) - 1, \\ n \deg(h) &\leq \deg(f) + \deg(g) + \deg(h) - 1. \end{aligned}$$

Adding all three together, we get

$$n(\deg(f) + \deg(g) + \deg(h)) \leq 3(\deg(f) + \deg(g) + \deg(h)) - 3,$$

which is

$$(n - 3)(\deg(f) + \deg(g) + \deg(h)) \leq -3.$$

This is a contradiction. \square

4.3. Pell's Equation. We are interested in nontrivial (degree at least 1) polynomial solutions to

$$x^2 - dy^2 = 1,$$

where x , y , and d are polynomials with integer coefficients. We will use the variable t for the polynomial variable.

Question 11.60. For which $d(t)$ does Pell's equation have a nontrivial solution?

Unfortunately, the problem is much harder in polynomials, and the set of polynomials $d(t)$ for which there is a solution remains an open problem.

We solve one specific Pell equation.

Theorem 11.61 (Nathanson). *The equation*

$$x^2 - (t^2 + d)y^2 = 1$$

has solutions in $\mathbb{Z}[t]$ only for $d = \pm 1, \pm 2$.

We consider the case $d = -1$, so we are looking for solutions to

$$x^2 - (t^2 - 1)y^2 = 1.$$

Definition 11.62. Define polynomials T_n, U_n as

$$T_n + U_n \sqrt{t^2 - 1} = (t + \sqrt{t^2 - 1})^n.$$

The T_n are called *Chebyshev polynomials of the first kind*, and the U_n are called *Chebyshev polynomials of the second kind*.

Theorem 11.63. *The pairs of polynomials (T_n, U_n) are solutions to the Pell equation*

$$x^2 - (t^2 - 1)y^2 = 1.$$

Proof. We can factor the equation as

$$(x - \sqrt{t^2 - 1}y)(x + \sqrt{t^2 - 1}y) = 1.$$

For $n = 1$ we get $(t, 1)$, which is a solution. So assume that (x_n, y_n) is a solution. Then we compute

$$(x_n + \sqrt{t^2 - 1}y_n)(t + \sqrt{t^2 - 1}) = (tx_n + (t^2 - 1)y_n) + (x_n + ty_n)\sqrt{t^2 - 1}.$$

We have

$$(x_{n+1}, y_{n+1}) = (tx_n + (t^2 - 1)y_n, x_n + ty_n).$$

Substituting into the Pell equation, we have

$$x_{n+1}^2 - (t^2 - 1)y_{n+1}^2 = x_n^2 - (t^2 - 1)y_n^2 = 1. \quad \square$$

Question 11.64. What are the polynomials T_n and U_n ?

Proposition 11.65. *We have the following recursive formulas.*

(a) *For Chebyshev polynomials of the first kind,*

$$\begin{aligned} T_0 &= 1, \\ T_1 &= t, \\ T_n &= 2tT_{n-1} - T_{n-2}. \end{aligned}$$

(b) *For Chebyshev polynomials of the second kind,*

$$\begin{aligned} U_0 &= 0, \\ U_1 &= 1, \\ U_n &= 2tU_{n-1} - U_{n-2}. \end{aligned}$$

Proof. Recall that we have defined the polynomials T_n, U_n as

$$T_n + U_n\sqrt{t^2 - 1} = (t + \sqrt{t^2 - 1})^n.$$

We show the statement by induction. Consider the base case $n = 2$:

$$(t + \sqrt{t^2 - 1})^2 = 2t^2 - 1 + 2t\sqrt{t^2 - 1},$$

so we have

$$\begin{aligned} T_2 &= 2t^2 - 1 = 2tT_1 - T_0, \\ U_2 &= 2t = 2tU_1 - U_0. \end{aligned}$$

Now assume the formulas hold up to (T_n, U_n) . Consider

$$\begin{aligned} (t + \sqrt{t^2 - 1})^{n+1} &= (t + \sqrt{t^2 - 1})^{n-1}(t + \sqrt{t^2 - 1})^2 \\ &= (t + \sqrt{t^2 - 1})^{n-1}(2t^2 - 1 + 2t\sqrt{t^2 - 1}) \\ &= (t + \sqrt{t^2 - 1})^{n-1}(2t(t + \sqrt{t^2 - 1}) - 1) \\ &= 2t(t + \sqrt{t^2 - 1})^n - (t + \sqrt{t^2 - 1})^{n-1} \\ &= 2t(T_n + U_n\sqrt{t^2 - 1}) - (T_{n-1} + U_{n-1}\sqrt{t^2 - 1}) \\ &= 2tT_n - T_{n-1} + (2tU_n - U_{n-1})\sqrt{t^2 - 1}, \end{aligned}$$

which are the desired formulas. \square

4.4. Waring Problem.

Question 11.66. Given a positive integer n , let $G_p(n) = k$ be the smallest integer such that any polynomial $g \in \mathbb{F}_p[x]$ can be written as

$$g = f_1^n + \cdots + f_k^n$$

for $f_i \in \mathbb{F}_p[x]$.

It is not at all clear that $G_p(n)$ exists for all p and n . In fact, we start by showing some specific choices of p and n for which $G_p(n)$ does not exist.

Lemma 11.67. For any prime p ,

$$(a_n x^n + \cdots + a_1 x + a_0)^p = a_n^p x^{pn} + \cdots + a_1^p x^p + a_0^p$$

in $\mathbb{F}_p[x]$.

Proof. We proceed by induction on the number of terms m . For $m = 1$, this is trivial:

$$(a_0)^p = a_0^p.$$

Assume it holds up to $m = n$, and consider $m = n + 1$. Grouping terms, we have

$$(a_n x^n + \cdots + a_1 x + a_0)^p = (a_n x^n + (a_{n-1} x^{n-1} + \cdots + a_1 x + a_0))^p.$$

Consider the right-hand side as $(x + y)^p = x^p + y^p$, and we have

$$(a_n x^n + \cdots + a_1 x + a_0)^p = a_n^p x^{np} + (a_{n-1} x^{n-1} + \cdots + a_1 x + a_0)^p.$$

Using the induction hypothesis on the right-hand side gives the desired result:

$$(a_n x^n + \cdots + a_1 x + a_0)^p = a_n^p x^{pn} + \cdots + a_1^p x^p + a_0^p. \quad \square$$

Lemma 11.68. We have $G_p(n) \leq k$ if and only if there are polynomials $f_1, \dots, f_k \in \mathbb{F}_p[x]$ such that $x = f_1^n + \cdots + f_k^n$.

Proof. Assume first that $G_p(n) \leq k$. Then we can write

$$x = f_1^n + \cdots + f_k^n.$$

Now assume that we have

$$x = f_1^n + \cdots + f_k^n.$$

Let $g \in \mathbb{F}_p[x]$ be any polynomial. Then we have

$$g = (f_1^n \circ g) + \cdots + (f_k^n \circ g). \quad \square$$

We now prove existence of $G_p(n)$ in few specific cases.

Theorem 11.69. For all primes $p \in \mathbb{Z}$, if $\gcd(p, n) \neq 1$, then $G_p(n)$ does not exist.

Proof. Assume we can write $n = kp$. We compute with Lemma 11.67

$$(a_n x^n + \cdots + a_1 x + a_0)^n = ((a_n x^n + \cdots + a_1 x + a_0)^p)^k,$$

so that the resulting polynomial has monomials with exponents only powers of p . In particular x , cannot be written as any number of n th powers. \square

Theorem 11.70. Let $p \in \mathbb{Z}$ be a prime number. If $p \neq 2$ and $\left(\frac{-1}{p}\right) = 1$, then $G_p(2) = 2$.

Proof. We have

$$\left(x^2 + \frac{1}{4}\right)^2 - \left(x^2 - \frac{1}{2}\right)^2 = x. \quad \square$$

Corollary 11.71. For all odd primes $p \in \mathbb{Z}$, $G_p(2) \leq p$.

Proof. Over \mathbb{Q} we have

$$x = \left(x^2 + \frac{1}{4}\right)^2 - \left(x^2 - \frac{1}{2}\right)^2.$$

Then, for any $p \neq 2$ we have

$$x = \left(x^2 + \frac{1}{4}\right)^2 + (p-1) \left(x^2 - \frac{1}{2}\right)^2,$$

which is a sum of p squares. \square

Theorem 11.72. We have $G_3(2) = 3$.

Proof. By Lemma 11.68, we consider x . Since x is not a square, we have $G_3(2) \geq 2$. So we try to write x as the sum of two squares. First observe that in the polynomial

$$(a_n x^n + \cdots + a_1 x + a_0)^2 + (b_n x^n + \cdots + b_1 x + b_0)^2$$

the coefficients of $\{1, x\}$ are affected only by $\{a_0, a_1, b_0, b_1\}$. So we may consider just the sum of squares of linear polynomials. There are finitely many such polynomials (in fact, nine of them), so we can simply check all 81 possible combinations to see that none equals x . Then, $G_3(2) \geq 3$. Now, we see that

$$x = (x+1)^2 + (x+2)^2 + (x+2)^2,$$

so that $G_3(2) = 3$. \square

Investigation 11.73. We can also ask how many ways a given polynomial can be written as the sum of k n th powers in $\mathbb{F}_p[x]$. Consider the case $G_5(2) = 2$.


- (a) Find a polynomial in $\mathbb{F}_5[x]$ that can be written as a sum of two squares in more than one way. How many ways can it be written?
- (b) Do you get the same number of ways for a different polynomial?
- (c) What polynomials can you find that can be written in the most and the least number of ways?

COMPUTATIONAL EXERCISES


11.1. Consider the polynomial

$$f(x) = x^5 - 2x^4 - 4x^3 + x^2 + 4x + 6.$$


- a. Factor $f(x)$ into irreducibles over \mathbb{Q} .
- b. Factor $f(x)$ into irreducibles over $\mathbb{Q}(\sqrt{2})$.
- c. Factor $f(x)$ into irreducibles over \mathbb{C} .

 **11.2.** Find an irreducible polynomial that does not satisfy the hypotheses of Eisenstein's criteria.

11.3. Compute the greatest common divisors of $x^n - 1$ and $x^m - 1$ for $1 \leq n < m \leq 10$. Compare them to $\gcd(m, n)$.

 **11.4.** Use the Euclidean algorithm to find polynomials $a, b \in \mathbb{Z}[x]$ such that

$$x = a(x^2 + x + 1) + b(x^2 + 1).$$

 **11.5.** As polynomials with integer coefficients, compute

$$(x^5 + 2x^4 - 2x - 1)^2 + (x^2 - 1) \pmod{x^3 - 1}.$$

11.6. Let $f(x) = x^2 - x - 1$. Find a polynomial $g \in \mathbb{F}_3[x]$ that has multiplicative order $\varphi(f)$ modulo f .

11.7. Find all monic irreducible polynomials of degree 3 over $\mathbb{F}_5[x]$.

11.8. Determine the number of monic irreducible polynomials of degree 6 in $\mathbb{F}_{11}[x]$.

11.9. For a positive integer n , write a function that returns the n th Chebyshev polynomial of either the first or second kind.

11.10. Find a (nontrivial) solution to $x^2 - (t^2 + 1)y^2 = 1$.

11.11. Write x as the sum of three squares in $\mathbb{F}_7[x]$. How many different ways can it be done with polynomials of degree at most 2?

THEORETICAL EXERCISES

11.12. Let R be \mathbb{Z} , \mathbb{Q} , or \mathbb{F}_p . Let f and g be polynomials in $R[x]$.

- a. Prove that if $f \mid g$ and $g \mid f$, then $\deg(f) = \deg(g)$.
- b. Prove that if f and g are also monic, then $f = g$.

11.13. Let R be \mathbb{Q} or \mathbb{F}_p . Let f , g , and h be polynomials in $R[x]$. If $\gcd(f, g) = 1$, prove that $\gcd(h, fg) = \gcd(h, f)\gcd(h, g)$.

11.14. For $f \in \mathbb{Z}[x]$, prove that $f(x)$ is irreducible if and only if $f(x+1)$ is irreducible.

11.15. Prove that the cyclotomic polynomials $\Phi_p(x)$ are irreducible over \mathbb{Z} when p is prime.

11.16. Let $f \in \mathbb{Z}[x]$. Prove that if f is irreducible by Eisenstein's criteria, then so is $f \circ f$.

11.17. Prove that in $\mathbb{Q}[x]$ there are infinitely many irreducible polynomials of every degree.

11.18. Prove the following generalization of Eisenstein's criteria.

Let $f(x) = a_n x^n + \cdots + a_1 x + a_0$ be a polynomial with integer coefficients. Assume that $p \mid a_i$ for $0 \leq i < n$ and $p \nmid a_n$. Let k be the smallest positive integer such that $p^2 \nmid a_k$ for some $0 \leq k < n$. Then if $f = gh$, we have $\min(\deg(g), \deg(h)) \leq k$.

11.19. Let R be \mathbb{Q} or \mathbb{F}_p . Mimic Euclid's proof (Theorem 1.46) that there are infinitely many prime numbers to prove that there are infinitely many irreducibles in $R[x]$.

11.20. Let $f \in \mathbb{Z}[x]$ with $\deg(f) \geq 1$, and let p be a prime. Prove that if \bar{f} is irreducible in $\mathbb{F}_p[x]$ and $\deg(\bar{f}) = \deg(f)$, then f is irreducible over \mathbb{Q} .

11.21. Let R be \mathbb{Q} or \mathbb{F}_p . Prove that for any polynomial $f \in R[x]$, if $\gcd(f, f') = 1$ where f' is the derivative of f , then f is squarefree.

11.22. Let R be \mathbb{Q} or \mathbb{F}_p . Let $f \in R[x]$ be an irreducible polynomial. Prove that every polynomial $g \in R[x]$ with f not dividing g has an inverse modulo f .

11.23. Let R be \mathbb{Q} or \mathbb{F}_p . Let $f \in R[x]$ be an irreducible polynomial and $g, h \in R[x]$. Prove that if $gh \equiv 0 \pmod{f}$, then at least one of $g \equiv 0 \pmod{f}$ or $h \equiv 0 \pmod{f}$.

11.24. Prove that for any $f \in \mathbb{F}_p[x]$

$$f^p - f = \prod_{a \in \mathbb{F}_p} f - a.$$

11.25. Let $f \in \mathbb{F}_p[x]$ be an irreducible polynomial of degree d . Let t be an indeterminate.

a. Prove that

$$t^{|f|-1} - 1 \equiv \prod_{0 \leq \deg(g) < d} t - g \pmod{f}.$$

Use (a) to prove the following two facts.

b. (Polynomial Wilson's theorem) Prove that

$$\prod_{0 \leq \deg(g) < d} g \equiv -1 \pmod{f}.$$

c. If $d \mid |f| - 1$, prove that $x^d \equiv 1 \pmod{f}$ has exactly d solutions.

11.26. If p, q are odd primes, prove that the number of monic irreducible polynomials of degree q in $\mathbb{F}_p[x]$ is

$$\frac{p^q - p}{q}.$$

11.27. Let p be a prime. Prove that for each $n \in \mathbb{N}$, there is a monic irreducible polynomial of degree n in $\mathbb{F}_p[x]$.

11.28. Let $d \in \mathbb{Z}[t]$. Prove for d a square or $\deg(d)$ odd that

$$x^2 - dy^2 = 1$$

has no solutions (other than constants).

EXPLORATION EXERCISES

11.29 (Generalized Fermat's Last Theorem). Let a , b , and c be positive integers at least 2. The generalized Fermat problem asks for solutions to

$$x^a + y^b = z^c$$

with relatively prime polynomials x , y , and z . For $x, y, z \in \mathbb{C}[t]$, the possible triples (a, b, c) are

$$\begin{aligned} (2, 2, c) & \text{ for } c \geq 2, \\ (2, 3, c) & \text{ for } c \in \{3, 4, 5\}. \end{aligned}$$

Instead consider the generalized Fermat problem for $x, y, z \in \mathbb{F}_p[t]$.

- Find some solutions to the case $(2, 2, 2)$ for various p .
- Consider the case $(2, 2, c)$ for $c \geq 2$. For which p can you find solutions?
- Consider the case $(2, 3, c)$ for $c \in \{3, 4, 5\}$. For which p can you find solutions?
- Fix a different triple (a, b, c) . For which p can you find solutions?
- Fix a prime p . For which triples (a, b, c) can you find solutions?

11.30 (Mahler measure).

Definition 11.74. Given a polynomial $f = c(x - a_1) \cdots (x - a_n)$, where we may have $a_i, c \in \mathbb{C}$, we define the *Mahler measure of f* as

$$M(f) = |c| \prod_{i=1}^n \max(1, |a_i|).$$

- Is the Mahler measure completely multiplicative, i.e., does

$$M(fg) = M(f)M(g)?$$

- Find a polynomial with $M(f) = 1$. Can you find all such polynomials in $\mathbb{Z}[x]$?
- Lehmer's conjecture states that there is a constant $\mu > 1$ such that if $f \in \mathbb{Z}[x]$ is irreducible, then

$$M(f) = 1 \quad \text{or} \quad M(f) > \mu.$$

Conjecture a value for μ .

11.31 (Quadratic Reciprocity for Polynomials). Let p be a prime number. Let f be an irreducible polynomial in $\mathbb{F}_p[x]$. Consider the residue classes modulo f .

- What can you say about the number of quadratic residues? Does this depend on f or p ?
- Does an equivalent of the Legendre symbol exist in $\mathbb{F}_p[x]$? What properties does it have?

- c. Is there an equivalent of Euler's criterion (Theorem 3.10)?
- d. Formulate a Law of Quadratic Reciprocity (Theorem 3.14).
- e. What about higher order residues, i.e., solutions to

$$g^d \equiv a \pmod{f}?$$

11.32 (Continued fractions and Pell's equation). In section 8.5 we were able to use the continued fraction expansion of \sqrt{d} to find solutions to Pell's equation. In particular certain convergents of the periodic continued fraction expansion of \sqrt{d} corresponded to solutions (Conjecture 8.37). We can also define continued fraction expansions for rational functions.

Definition 11.75. Consider the following Laurent series with variable t :

$$F = c_m t^m + c_{m-1} t^{m-1} + \cdots + c_1 t + c_0 + \frac{c_{-1}}{t} + \frac{c_{-2}}{t^2} + \cdots,$$

where m is a positive integer. Define the floor function to take the polynomial portion of the Laurent series, i.e.,

$$\lfloor F \rfloor = c_m t^m + c_{m-1} t^{m-1} + \cdots + c_1 t + c_0.$$

We can then determine a continued fraction expansion of F as

$$a_0 = \lfloor F \rfloor,$$

$$F_1 = \frac{1}{F - a_0}$$

and recursively

$$a_i = \lfloor F_i \rfloor,$$

$$F_{i+1} = \frac{1}{F_i - a_i}.$$

The continued fraction expansion of F is then

$$[a_0; a_1, a_2, \dots].$$

Example 11.76. Consider $F = \sqrt{t^2 + 1}$. We can write this as

$$F = t \sqrt{1 + \frac{1}{t^2}}.$$

Taking a truncated power series expansion of $\sqrt{1+t}$ as

$$\sqrt{1+t} \approx 1 + \frac{t}{2} - \frac{t^2}{8} + \frac{t^3}{16} - \frac{5t^4}{128} + \frac{7t^5}{256} - \frac{21t^6}{1024},$$

we have

$$F = \sqrt{t^2 + 1}$$

$$\approx \frac{t^{12} + t^{10}/2 - t^8/8 + t^6/16 - 5t^4/128 + 7t^2/256 - 21/1024}{t^{11}}.$$

We write this as quotient and remainder to have

$$F = t + \frac{t^{10}/2 - t^8/8 + t^6/16 - 5t^4/128 + 7t^2/256 - 21/1024}{t^{11}}$$

so that $a_0 = t$. Then we compute

$$F_1 = \frac{1}{F - a_0} \approx \frac{t^{11}}{t^{10}/2 - t^8/8 + t^6/16 - 5t^4/128 + 7t^2/256 - 21/1024}.$$

Writing this as quotient and remainder, we have

$$F_1 = 2t + \frac{t^9/4 - t^7/8 + 5t^5/64 - 7t^3/128 + 21t/512}{t^{10}/2 - t^8/8 + t^6/16 - 5t^4/128 + 7t^2/256 - 21/1024}.$$

We can continue this process to get

$$\sqrt{t^2 + 1} \approx [t; 2t, 2t, 2t, \dots].$$

In fact, we have exactly the periodic continued fraction expansion

$$\sqrt{t^2 + 1} = [t; \overline{2t}].$$

- Compute the continued fraction expansions for a few $\sqrt{d(t)}$. Are they always periodic? You should assume that d is monic and of even degree.
- What periods are possible for continued fraction expansions of $\sqrt{d(t)}$?
- Do any of the convergents of $\sqrt{d(t)}$ correspond to solutions to $x^2 - dy^2 = 1$?
- Is there a difference between the existence of solutions in $\mathbb{Z}[t]$ and $\mathbb{Q}[t]$?

11.33 (Waring problem).

- For primes p where -1 is a quadratic residue modulo p , we know that $G_p(2) = 2$ (Theorem 11.70). For primes where -1 is not a quadratic residue modulo p , find an upper bound on $G_p(2)$ by writing x as a sum of squares (Lemma 11.68).
- Determine upper bounds for $G_p(3)$ for some primes p .
- What about upper bounds for other pairs (p, n) ?
- Can you conjecture an exact value of $G_p(n)$ for any particular pair (p, n) ?

11.34 (Elliptic curves). We can define an elliptic curve with polynomial coefficients:

$$E_t : y^2 = x^3 + a(t)x + b(t).$$

We can think of this either as an elliptic curve over polynomials or as a family of elliptic curves parameterized by values of t .

- Choose an elliptic curve E_t , say, $y^2 = x^3 + t$. Does every t value result in an elliptic curve (i.e., it has a nonzero discriminant)?
- Consider the elliptic curve $y^2 = x^3 - x + t^2$. Can you find any (polynomial) points on this curve?
- Consider the elliptic curve $y^2 = x^3 + x + t^2$. Notice that $(0, t)$ is a point of finite order. For specific values of t , the curve can have additional torsion points. How many?
- Consider the elliptic curve $y^2 = x^3 - t^2x + 1$. Find two independent points on this curve. Are there values of t where there are more independent rational points?

Bibliography

- [1] Titu Andreescu and Bogdan Enescu, *Mathematical olympiad treasures*, Birkhäuser Boston, Inc., Boston, MA, 2004. MR2025063
- [2] David Bailey, Peter Borwein, and Simon Plouffe, *On the rapid computation of various polylogarithmic constants*, Math. Comp. **66** (1997), no. 218, 903–913, DOI 10.1090/S0025-5718-97-00856-9. MR1415794
- [3] Matthew Baker and Laura DeMarco, *Preperiodic points and unlikely intersections*, Duke Math. J. **159** (2011), no. 1, 1–29, DOI 10.1215/00127094-1384773. MR2817647
- [4] L’umbomíra Balková and Arnaka Hrušková. *Continued fractions of quadratic numbers*. [arxiv.org/1302.0521](https://arxiv.org/abs/1302.0521), 2014.
- [5] Robert L. Benedetto, Benjamin Dickman, Sasha Joseph, Benjamin Krause, Daniel Rubin, and Xinwen Zhou, *Computing points of small height for cubic polynomials*, Involve **2** (2009), no. 1, 37–64, DOI 10.2140/involve.2009.2.37. MR2501344
- [6] R. P. Brent, G. L. Cohen, and H. J. J. te Riele, *Improved techniques for lower bounds for odd perfect numbers*, Math. Comp. **57** (1991), no. 196, 857–868, DOI 10.2307/2938723. MR1094940
- [7] Ezra Brown and Bruce T. Myers, *Elliptic curves from Mordell to Diophantus and back*, Amer. Math. Monthly **109** (2002), no. 7, 639–649, DOI 10.2307/3072428. MR1917222
- [8] Georg Cantor, *Contributions to the founding of the theory of transfinite numbers*, Dover Publications, Inc., New York, N. Y., 1952. Translated, and provided with an introduction and notes, by Philip E. B. Jourdain. MR0045635
- [9] R. D. Carmichael, *Note on Euler’s φ -function*, Bull. Amer. Math. Soc. **28** (1922), no. 3, 109–110, DOI 10.1090/S0002-9904-1922-03504-5. MR1560520
- [10] D. Chaum, E. Van Heijst, and B. Pfitzmann. *Cryptographically strong undeniable signatures, unconditionally secure for the signer*, vol. 576 of Lecture Notes in Computer Science. Springer, 1992.
- [11] G. L. Cohen and P. Hagis Jr., *On the number of prime factors of n if $\varphi(n)(n-1)$* , Nieuw Arch. Wisk. (3) **28** (1980), no. 2, 177–185. MR582925
- [12] Henri Cohen, *Number theory. Vol. I. Tools and Diophantine equations*, Graduate Texts in Mathematics, vol. 239, Springer, New York, 2007. MR2312337

- [13] Jordan Ellenberg, mathoverflow. <http://mathoverflow.net/questions/70371/greatest-common-divisor-of-a2n-1-and-b2n-1>, July 2011.
- [14] Andreas-Stephan Elsenhans and Jörg Jahnel, *New sums of three cubes*, Math. Comp. **78** (2009), no. 266, 1227–1230, DOI 10.1090/S0025-5718-08-02168-6. MR2476583
- [15] Claire Ferguson, *Helaman Ferguson: Mathematics in stone and bronze*, Meridian Creative Group, Erie, PA, 1994. With a foreword by Richard Waller. MR1263655
- [16] Kevin Ford, *The number of solutions of $\phi(x) = m$* , Ann. of Math. (2) **150** (1999), no. 1, 283–311, DOI 10.2307/121103. MR1715326
- [17] Richard K. Guy, *Unsolved problems in number theory*, 3rd ed., Problem Books in Mathematics, Springer-Verlag, New York, 2004. MR2076335
- [18] Kevin G. Hare, *More on the total number of prime factors of an odd perfect number*, Math. Comp. **74** (2005), no. 250, 1003–1008, DOI 10.1090/S0025-5718-04-01683-7. MR2114661
- [19] Robert Harron and Andrew Snowden, *Counting elliptic curves with prescribed torsion*, J. Reine Angew. Math. **729** (2017), 151–170, DOI 10.1515/crelle-2014-0107. MR3680373
- [20] David Hilbert. *Mathematische probleme*. Gött. Nach., 1900:253–197, 1900.
- [21] David Hilbert, *Mathematical problems*, Bull. Amer. Math. Soc. **8** (1902), no. 10, 437–479, DOI 10.1090/S0002-9904-1902-00923-3. MR1557926
- [22] Dale Husemoller, *Elliptic curves*, 2nd ed., Graduate Texts in Mathematics, vol. 111, Springer-Verlag, New York, 2004. With appendices by Otto Forster, Ruth Lawrence and Stefan Theisen. MR2024529
- [23] Benjamin Hutz and Adam Towsley, *Misiurewicz points for polynomial maps and transversality*, New York J. Math. **21** (2015), 297–319. MR3358544
- [24] Benjamin Hutz, *Determination of all rational preperiodic points for morphisms of PN* , Math. Comp. **84** (2015), no. 291, 289–308, DOI 10.1090/S0025-5718-2014-02850-0. MR3266961
- [25] Kenneth Ireland and Michael Rosen, *A classical introduction to modern number theory*, 2nd ed., Graduate Texts in Mathematics, vol. 84, Springer-Verlag, New York, 1990. MR1070716
- [26] joro. mathoverflow. <http://mathoverflow.net/questions/70548/factors-of-gcd22n-1-32n-1>, July 2011.
- [27] H. W. Lenstra Jr., *Factoring integers with elliptic curves*, Ann. of Math. (2) **126** (1987), no. 3, 649–673, DOI 10.2307/1971363. MR916721
- [28] D. H. Lehmer, *Factorization of certain cyclotomic functions*, Ann. of Math. (2) **34** (1933), no. 3, 461–479, DOI 10.2307/1968172. MR1503118
- [29] Neil Lyall and Alex Rice, *Polynomial differences in the primes*, Combinatorial and additive number theory—CANT 2011 and 2012, Springer Proc. Math. Stat., vol. 101, Springer, New York, 2014, pp. 129–146, DOI 10.1007/978-1-4939-1601-6_10. MR3297076
- [30] Michelle Manes, *\mathbb{Q} -rational cycles for degree-2 rational maps having an automorphism*, Proc. Lond. Math. Soc. (3) **96** (2008), no. 3, 669–696, DOI 10.1112/plms/pdm044. MR2407816
- [31] Yu. V. Matiyasevich, *Desyataya problema Gil'berta* (Russian, with Russian summary), Matematicheskaya Logika i Osnovaniya Matematiki [Monographs in Mathematical Logic and Foundations of Mathematics], vol. 26, VO “Nauka”, Moscow, 1993. MR1247235

- [32] Maurice Mignotte, *How to share a secret*, Cryptography (Burg Feuerstein, 1982), Lecture Notes in Comput. Sci., vol. 149, Springer, Berlin, 1983, pp. 371–375, DOI 10.1007/3-540-39466-4_27. MR707286
- [33] Patrick Morton and Joseph H. Silverman, *Rational periodic points of rational functions*, Internat. Math. Res. Notices **2** (1994), 97–110, DOI 10.1155/S1073792894000127. MR1264933
- [34] Patrick Morton and Joseph H. Silverman, *Periodic points, multiplicities, and dynamical units*, J. Reine Angew. Math. **461** (1995), 81–122, DOI 10.1515/crll.1995.461.81. MR1324210
- [35] Nikolay G. Moshchevitin. *On some open problems in diophantine approximation*, [arXiv:1202.4539](#), 2012.
- [36] W. Narkiewicz, *On a class of monic binomials*, Proc. Steklov Inst. Math. **280** (2013), no. suppl. 2, S65–S70, DOI 10.1134/S0081543813030073. MR3447581
- [37] Melvyn B. Nathanson, *Polynomial Pell’s equations*, Proc. Amer. Math. Soc. **56** (1976), 89–92, DOI 10.2307/2041581. MR0401641
- [38] P. T. Bateman, J. L. Selfridge, and S. S. Wagstaff Jr., *The new Mersenne conjecture*, Amer. Math. Monthly **96** (1989), no. 2, 125–128, DOI 10.2307/2323195. MR992073
- [39] Mathematical Association of America. *American Invitational Mathematics Exam*, 1987.
- [40] Carl D. Olds. *Continued fractions*, volume 9 of New Mathematical Library. Mathematical Association of America, 1992.
- [41] Bjorn Poonen, *The classification of rational preperiodic points of quadratic polynomials over \mathbf{Q} : a refined conjecture*, Math. Z. **228** (1998), no. 1, 11–29, DOI 10.1007/PL00004405. MR1617987
- [42] Victor V. Prasolov, *Polynomials, Algorithms and Computation in Mathematics*, vol. 11, Springer-Verlag, Berlin, 2004. Translated from the 2001 Russian second edition by Dimitry Leites. MR2082772
- [43] Michael Rosen, *Number theory in function fields*, Graduate Texts in Mathematics, vol. 210, Springer-Verlag, New York, 2002. MR1876657
- [44] Zachary L. Scherr. *Rational Polynomial Pell Equations*. PhD thesis, University of Michigan, 2013.
- [45] Jeffrey Shallit. Diophantine equation, $\sigma(n) = 2^m$. *Mathematics Magazine*, **63** (1990), no. 2, 129.
- [46] Wacław Sierpiński, *Elementary theory of numbers*, Translated from Polish by A. Hulanicki. Monografie Matematyczne, Tom 42, Państwowe Wydawnictwo Naukowe, Warsaw, 1964. MR0175840
- [47] Joseph H. Silverman, *The arithmetic of elliptic curves*, Graduate Texts in Mathematics, vol. 106, Springer-Verlag, New York, 1986. MR817210
- [48] Joseph H. Silverman, *The arithmetic of dynamical systems*, Graduate Texts in Mathematics, vol. 241, Springer, New York, 2007. MR2316407
- [49] Joseph H. Silverman. *A Friendly Introduction to Number Theory*. Pearson, 2012.
- [50] Joseph H. Silverman and John Tate, *Rational points on elliptic curves*, Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1992. MR1171452
- [51] Simon Singh, *Fermat’s enigma: The epic quest to solve the world’s greatest mathematical problem*, Walker and Company, New York, 1997. With a foreword by John Lynch. MR1491363

- [52] William Stein and David Joyner. *SAGE: System for algebra and geometry experimentation*. Communications in Computer Algebra (SIGSAM Bulletin), July 2005. <http://www.sagemath.org>.
- [53] Ian Stewart and David Tall, *Algebraic number theory and Fermat's last theorem*, 3rd ed., A. K. Peters, Ltd., Natick, MA, 2002. MR1876804
- [54] The PARI Group, Bordeaux. *PARI/gp, version 2.3.2*, 2007. available from <http://pari.math.u-bordeaux.fr/>.
- [55] Andrew Wiles, *Modular elliptic curves and Fermat's last theorem*, Ann. of Math. (2) **141** (1995), no. 3, 443–551, DOI 10.2307/2118559. MR1333035

List of Algorithms

0.1	Trial Division	3
1.1	Division Algorithm	9
1.2	Divisibility Test	9
1.3	Euclidean Algorithm	16
1.4	Extended Euclidean Algorithm	18
1.5	Sieve of Eratosthenes	25
2.1	Inverses from the Euclidean Algorithm	46
2.2	Chinese Remainder Theorem	57
3.1	m th Roots	81
4.1	Encryption with Affine Ciphers	94
4.2	Diffie–Hellman Key Exchange	97
4.3	RSA Public Key	99
4.4	Digital Signature and Verification with RSA	100
4.5	Digital Signature and Verification with Hash Function	102
4.6	Secret Sharing with Mignotte Sequences	104
7.1	Continued Fraction Expansion	170
7.2	Best Rational Approximation	180
8.1	Fundamental Solution to Pell’s Equation	207
9.1	Torsion Subgroup of an Elliptic Curve	236
10.1	Determining Rational Periodic Points by Good Reduction	267
10.2	Determining Rational Preperiodic Points by Good Reduction	268

List of Notation

\mathbb{N}	The natural numbers	5
\mathbb{Z}	The integers	5
$a \mid b$	a divides b	6
$a \nmid b$	a does not divide b	6
\gcd	Greatest common divisor	11
lcm	Least common multiple	11
$\pi(N)$	Prime counting function	24
$f \sim g$	The functions f and g are asymptotic.	26
$a \equiv b \pmod{n}$	a is congruent to b modulo n	39
\bar{a}	Residue class	40
$\varphi(n)$	Euler phi function	49
$\lambda(n)$	Carmichael function	60
$\left(\frac{a}{p}\right)$	The Legendre symbol	67
$\left(\frac{a}{n}\right)$	The Jacobi symbol	75
$\mu(n)$	Möbius function	114
Φ_n	n th cyclotomic polynomial	119
$\sigma_k(n)$	Sum of k th powers of divisors of n	121
$\omega(n)$	Number of distinct prime divisors of n	122
$s(n)$	Aliquot sum of n	126
$[x]$	The largest integer less than x	128
$p(n)$	The number of partitions of n	130
$p(n, k)$	The number of partitions of n with at most k parts	133
\mathbb{Q}	The set of rational numbers	143

$\mathbb{Q}(\sqrt{d})$	Quadratic number field	145
$N(a)$	The norm of a	146
$\mathbb{Z}[\sqrt{d}]$	The set $\{a + b\sqrt{d} : a, b \in \mathbb{Z}\}$	155
$H(t)$	The height of the rational number t	159
$\{x\}$	The decimal part of the number x	164
$\{x\}$	The decimal part of x	164
$[a_0; a_1, a_2, a_3, \dots]$	A simple continued fraction	167
$g(k)$	Waring problem function	209
\mathbb{Z}_p	p -adic integers	217
\mathbb{Q}_p	p -adic numbers	217
$E(\mathbb{Q})$	Mordell–Weil group	237
$f^{\circ n}(x)$	The n th iterate of x by f	247
$\mathcal{O}_f(z)$	The (forward) orbit of z by f	248
$\Phi_n(f)$	The n th dynatomic polynomial of f	255
$\text{Res}(g, h)$	The resultant of g and h	259
$\text{Res}(f)$	The resultant of the rational function f	259
f^α	The conjugate of f by α	273
$h(a)$	The logarithmic height of a	274
$\hat{h}_f(z)$	The canonical height of z by f	274
$\mathbb{Z}[x]$	Polynomial with integer coefficients	275
$\mathbb{Q}[x]$	Polynomial with rational coefficients	275
$\mathbb{F}_p[x]$	Polynomial with coefficients modulo p	276
$ f $	The number $p^{\deg(f)}$	283
$I_d(x)$	The product of all monic irreducible polynomials of degree d in $\mathbb{F}_p[x]$	287
$N_p(d)$	The number of monic irreducible polynomials of degree d in $\mathbb{F}_p[x]$	288
T_n	The n th Chebyshev polynomial of the first kind	290
U_n	The n th Chebyshev polynomial of the second kind	290
$G_p(n)$	Waring problem function for polynomials modulo p	292

Index

- abundant number, 1, 140
- Adleman, Leonard, 99
- affine cipher, 93
 - frequency analysis, 94
 - using encryption key, 95
- algebraic integer, 148
 - irreducible, 149
 - prime, 149
 - see also* algebraic number
- algebraic number
 - algebraic integer, 148
 - characteristic polynomial, 144
 - degree, 144
 - fundamental unit, 203
 - Liouville's approximation theorem, 164
 - multiple factorizations, 146
 - norm, 146
 - Roth's approximation theorem, 166
 - unit, 149
- Algorithm, 3
 - Best Rational Approximation, 180
 - Chinese Remainder Theorem, 57
 - Continued Fraction Expansion, 170
 - Diffie–Hellman Key Exchange, 97
 - Digital Signature with Hash Function, 102
 - Digital Signature with RSA, 100
 - Divisibility test, 9
 - Division Algorithm, 9
 - Encryption with Affine Cipher, 94
 - Euclidean Algorithm, 16
 - Euler Factorization, 37
 - Extended Euclidean Algorithm, 18
 - Fermat Factorization, 36
 - Fundamental Solution to Pell's Equation, 207
 - Inverse from the Euclidean Algorithm, 46
 - Lucas–Lehmer Primality Test, 37
 - m th Roots, 81
 - Pollard rho, 37
 - Rational Periodic Points by Good Reduction, 267
 - Rational Preperiodic Points by Good Reduction, 268
 - RSA Public Key, 99
 - Secret Sharing with Mignotte Sequences, 104
 - Sieve of Eratosthenes, 25
 - Torsion Subgroup, 236
 - Trial Division, 3, 36
- aliquot sequence, 140
- aliquot sum, 126, 140
 - abundant number, 140
 - deficient number, 140
 - perfect number, 128
- almost perfect number, 140
- almost prime, 35
- amicable pair, 139
- approximation, best, *see* best approximation
- arithmetic axioms, 5
- arithmetic function, 109
 - Euler totient function, 109–113
 - Möbius function, 113–117
 - number of distinct prime divisors, 122

- number of distinct prime factors, 128, 140
- partition function, 130–134
- sum of divisors function, 122–125
- Arithmetic Mean–Geometric Mean, 175
- arithmetic progression, 32
 - primes in, *see* Dirichlet’s theorem
- asymmetric cipher, 98
- asymptotic functions, 26
- bad reduction, 260
- Bang–Zsigmondy Theorem, 29
- best approximation
 - first kind, 159
 - first kind for π , 162
 - second kind, 171
 - second kind are convergents, 178
- Blum, Manuel, 107
- Brahmagupta, 202
- Brouncker, William, 202
- canonical height, 274
- Cantor, Georg, 144, 158
- Carmichael function, 60, 63, 88, 89
- Carmichael number, 1, 48, 58, 61, 62, 135
- Catalan equation, 214
- character frequency, 94, 96
- characteristic polynomial, 144
- Chaum, David, 103
- Chebyshev polynomials, 290
- check digit, 102
- Chinese Remainder Theorem, 54
 - solving Diophantine equations, 191
- cipher, 92
 - affine, 93
 - asymmetric, 98
 - Diffie–Hellman key exchange, 97
 - discrete log problem, 97, 228
 - NIST key size recommendations, 100
 - RSA public key, 99, 108
 - shared secret, 104, 108
 - symmetric, 96
 - Vigenère, 95
- complete systems of residues, 40
- composite number, 7
- congruent modulo n , 39
- congruent number, 218, 238
 - t -congruent number, 246
 - generalizations, 246
- conjugation, 273
- continued fraction, 166, 167, 181
 - convergent, 171
 - finite, 170
 - for polynomials, 297
 - periodic, 185
 - solving quadratic equations, 186
- convergent, 171
 - as best approximations, 176
 - as solutions to Pell’s equation, 206, 297
 - error bound, 174
 - recursive formula for, 172
- counting function, 23, 160, 184
 - prime numbers, 24
- critical points, 273
- cubic reciprocity, 89
- CvHP hash function, 103
- cycle structure, 249
- cyclotomic polynomial, 117–121, 139, 144, 294
- Davis, Martin, 189
- decimal part of a number, 164, 168
- Dedekind, Julius, 150
- deficient number, 140
- degree
 - algebraic number, 144
 - polynomial, 275
 - rational function, 249
- density, 23
- descent, method of, 200
- Diffie, Whitfield, 97
- Diffie–Hellman key exchange, 97
 - with elliptic curves, 244
- digital signature, 100
- Diophantine approximation, 158–166
 - best rational approximation
 - algorithm, 179
 - by continued fractions, 171–181
 - Dirichlet’s theorem, 163, 175, 176, 204
 - Liouville number, 165
 - Liouville’s theorem, 164
 - Roth’s theorem, 166
 - simultaneous approximation, 185
- Diophantine equation
 - congruent numbers, 218
 - Fermat’s Last Theorem, *see* Fermat’s Last Theorem
 - Hasse principle, 190
 - Hensel’s Lemma, 193
 - method of descent, 200
 - modulo primes, 189–198
 - number of solutions mod p , 216

- Pell's equation, *see* Pell's equation
- Pell-like equations, 218
- Pythagorean triple, *see* Pythagorean triple
- solving using Chinese Remainder Theorem, 192
- Waring problem, *see* Waring problem
- Diophantus, 158, 187, 209, 223
- Dirichlet's theorem
 - arithmetic progression, 32, 191
 - Diophantine approximation, 163, 175, 176, 204
- Dirichlet, Johann, 163
- discrete log problem, 97
 - for elliptic curves, 228
- discriminant, 221, 270
- divisibility
 - for algebraic integers, 148
 - for integers, 6–9
 - for polynomials, 278
 - test modulo n , 41
- Division algorithm, 8, 39
 - for polynomials, 279
- divisor, 6
- dynamical system, 247
 - algorithm to compute all rational periodic points, 267
 - algorithm to compute all rational preperiodic points, 268
 - conjugation, 273
 - finitely many bad primes, 261
 - forward orbit, 248
 - good reduction, 260
 - Lagrange interpolation, 252
 - n th iterate, 247
 - period bound via good reduction, 266
 - periodic point, 248
 - Poonen's conjecture, 253
 - post-critically finite, 272
 - preperiodic point, 248
 - wandering point, 248
- dynamomic polynomial, 254–258
- Eisenstein's criteria, 277
 - generalization, 295
- elliptic curve, 221
 - addition of points, 224–228
 - algorithm to compute the torsion subgroup, 236
 - congruent numbers, 237, 240
 - Diffie–Hellman key exchange, 244
 - discriminant, 222, 242
 - doubling function, 250
 - for polynomials, 298
 - integer points, 230–244
 - Lattès map, 272
 - Lenstra factorization, 245
 - linearly independent points, 236
 - Mazur's theorem, 235, 253
 - Mordell curve, 223, 227, 230, 236, 241, 243
 - Mordell–Weil group, 237
 - Mordell–Weil Theorem, 237
 - Nagel–Lutz theorem, 234
 - order of a point, 229
 - over finite fields, 244
 - point at infinity, 224
 - point of finite order, 229
 - rank, 237
 - torsion point, 229–230
 - torsion subgroup, 235
- encryption key, 95
- Eratosthenes, 24
- Erdős, Paul, 135
- Euclid, 13, 22
- Euclidean algorithm, 15
 - applied to continued fractions, 168
 - efficiency of, 18–20
 - extended, 17, 18
 - for polynomials, 280
 - see also* extended Euclidean algorithm
- Euler factorization, 37
- Euler totient function, 49, 109–113
 - for polynomials, 284
 - formula for computing, 111
 - is multiplicative, 111
- Euler's criterion, 67
- Euler's formula, 49
 - for polynomials, 285
- Euler, Leonhard, 71
- Euler, Leonhard, 32, 130, 202, 223, 238
- experimental mathematics, 4
 - development process, 2
- extended Euclidean algorithm, 17
 - inverses from, 46
- factorial, 47
- factorization
 - Euler factorization, 37
 - Fermat factorization, 36
 - Lenstra elliptic curve, 245
 - multiple factorizations for algebraic numbers, 146

- Pollard rho factorization, 37
- trial division algorithm, 36
- unique for integers, 21
- unique for polynomials, 281
- Fermat equation, *see* Fermat's Last Theorem
- Fermat factorization, 36
- Fermat near miss, 213, 218
- Fermat number, 3, 33, 271
- Fermat pseudoprime, 48, 58, 62
 - see also* Carmichael number
- Fermat's Last Theorem, 1, 187, 200–202, 215
 - for polynomials, 289
 - generalized, 296
 - $n = 4$, 200
 - near miss, 218
- Fermat's Little Theorem, 46
 - Fermat pseudoprime, 58, 62
 - for polynomials, 286
 - primality test, 48, 61
- Fermat, Pierre de, 1, 200, 202, 223, 238
- Fibonacci numbers, 1, 18, 31, 58
 - convergents of the golden ratio, 181
 - in terms of the golden ratio, 31
 - prime divisors of, 61
 - worst-case for Euclidean algorithm, 18–20
- floor function, 168
- Ford, Kevin, 136
- formal periodic point, 257
- forward orbit, *see* orbit
- Fundamental Theorem of Algebra, 51, 54, 276
- fundamental unit, 203
- Gauss' Lemma, 276
- Gauss, Carl Frederick, 71, 190, 210
- Germain, Sophie, 2
- Goldbach conjecture, 135
- Goldbach partitions, 141
- Goldbach, Christian, 135
- golden ratio, 19, 31, 180
- good reduction, 260
- greatest common divisor, 11
 - as linear combination, 12, 280
 - computing, 13
 - polynomial, 280
- hash function, 101
 - check digit, 102
 - CvHP, 103
 - ISBN, 103
 - MD5, 101
 - SHA-1, SHA-2, SHA-3, 102
 - UPC, 102
- Hasse principle, 190
- Hasse, Helmut, 190
- height, 159
 - canonical, 274
 - logarithmic, 274
 - of a rational number, 159
 - points of bounded height, 184
- Hellman, Martin, 97
- Hensel lifting, *see* Hensel's Lemma
- Hensel's Lemma, 193, 194
- Heron triangle, 246
- Heron of Alexandria, 246
- Hilbert's 10th problem, 3
- Hilbert's 10th problem, 189
- Hilbert's number, 182
- Hilbert, David, 3, 189, 209
 - 10th problem, 4, 189
- integers
 - abundant number, 140
 - almost perfect, 140
 - almost prime, 35
 - composite, 7
 - deficient number, 140
 - density, 23
 - divisibility, 6–9
 - Division algorithm, 8
 - factorial, 47
 - Fermat number, *see* Fermat number
 - greatest common divisor, 11
 - least common multiple, 11
 - linear combination, 11
 - Mersenne prime, 34
 - partitions, 130–134
 - perfect number, 125–128
 - prime, 7
 - prime gap, 33
 - Prime Number Theorem, 28
 - relatively prime, 11
 - smooth number, 35
 - unique factorization, 21
 - well ordering property, 5
- International Standard Book Number (ISBN), 103
- inverse, 44
 - modulo n , 45
- irrational number, 157
- irreducible
 - algebraic integer, 149

- polynomial, 276
- Jacobi symbol, 75
- Jacobi, Carl, 75
- Julius Caesar, 91
- key exchange, Diffie–Hellman, 97
- Kummer, Ernst, 2, 150
- Lagrange interpolation, 252
- Lagrange polynomial, 252
- Lagrange’s Theorem, 84, 88
- Lagrange, Joseph Louis, 209
- Lamé’s Theorem, 19
- Lambert, Johann, 158
- Lattès map, 272
- lattice points, 184
- least common multiple, 11
- Legendre sieve, 128, 129
- Legendre symbol, 67
 - is multiplicative, 68
- Legendre, Adrien-Marie, 67, 71, 128, 210
- Lehmer, Derrick, 113
- Lenstra, Hendrik, 245
- Lind, Carl-Erik, 190
- Lindemann, Carl, 145
- linear combination, 11
 - greatest common divisor as, 12, 280
- linear congruence
 - multivariable, 62
 - solving in general, 52
 - solving with inverses, 50
 - system of, 54
 - see also* Chinese Remainder Theorem
- linear fractional transformation, 273
- Liouville function, 138, 141
- Liouville number, 145, 165
- Liouville’s theorem, 164
- Liouville, Joseph, 138, 141, 145
- logarithmic height, 274
- Lucas–Lehmer primality test, 37
- m th power residue, 76
- m th root modulo n , 76
- Mahler measure, 296
- Mandelbrot set, 250
- Matiyasevich, Yuri, 189
- Mazur’s theorem, 235, 253
- Mazur, Barry, 230
- MD5, 101
- Mersenne prime, 34, 127
 - relation with perfect numbers, 127
- Mersenne, Marin, 34
- Mignotte secret sharing, 104
- Mignotte sequence, 104
- Mignotte, Maurice, 104
- minimal period, 249
- Möbius function, 113–117, 255
 - is multiplicative, 114
- Möbius inversion formula, 116
- modular arithmetic, 41
 - for polynomials, 282
- modulus, 39
- monic polynomial, 148, 275
- Mordell curve, *see* elliptic curve
- Mordell, Louis, 227, 237
- Morton, Richard Patrick, 253
- Morton–Silverman uniform
 - boundedness conjecture, 253, 273
- multiplicative function, 110
- multiplicative order, 76
 - for polynomials, 286
- multiplier, 263
- n th iterate, 247
- Nagel–Lutz theorem, 234
- National Institute of Standards and Technology, 100, 102
- National Security Agency, 102
- Natural numbers, 5
- Newton map, 271
- nonlinear congruence, 62
- nontotient, 136
- norm, 146
- number field, *see* quadratic number field
- one-way function, 101
- orbit, 248
- p -adic number, 194, 198, 216
- parity problem, 130
- partition, 130–134
 - as power series coefficients, 131
 - Goldbach, 141
 - part, 130
 - prime partition, 141
 - Young diagram, 132
- Pell’s equation, 187, 202–208
 - for polynomials, 290
 - fundamental solution, 205
 - Pell-like equations, 218
 - solving with continued fractions, 206–208, 297
- Pell, John, 202

- perfect number, 1, 125–128
 - relation with Mersenne primes, 127
- periodic point, 248
 - determining by good reduction, 267
 - existence of, 272
 - formal periodic point, 257
 - minimal period, 249
 - multiplier, 263
 - period modulo p , 264
 - see also* dynatomic polynomial
- Pfitzmann, Birgit, 103
- pigeonhole principle, 163
- point at infinity, 224
- Polignac's Conjecture, 33
- Pollard rho factorization, 37
- polynomial, 275
 - content, 276
 - continued fraction, 297
 - degree, 275
 - Diophantine equations, 293
 - Diophantine equations, 288
 - Division algorithm, 279
 - Eisenstein's criteria, 277
 - Euclidean algorithm, 280
 - Euler totient function, 284
 - Euler's formula, 285
 - Fermat's Last Theorem, 289
 - Fermat's Little Theorem, 286
 - Fundamental Theorem of Algebra, 276
 - irreducible, 276
 - Mahler measure, 296
 - modular arithmetic for polynomials, 282
 - modulo irreducible polynomials, 288
 - monic, 148, 275
 - multiplicative order, 286
 - number of monic irreducible, 288
 - Pell's equation, 290
 - prime generating, 32
 - primitive, 276
 - Pythagorean triple, 289
 - Quadratic Reciprocity, 296
 - reducible, 276
 - root, 117, 275
 - unique factorization, 281
 - value, 282
 - Waring problem, 292, 298
- post-critically finite, 272
- preimage, 268
- preperiodic point, 248
 - cycle structure, 249
 - determining by good reduction, 268
- MS uniform boundedness conjecture, 253, 273
- tail, 249
- primality test, 24, 60
 - AKS, 61
 - Fermat's Little Theorem, 48, 61
 - Lucas, 61
 - Lucas–Lehmer, 37
 - Miller–Rabin, 61
 - pseudoprime, 61
 - Sieve of Eratosthenes, 24
 - trial division, 61
 - Wilson's theorem, 61
- prime gap, 33
 - Polignac's Conjecture, 33
 - Twin prime conjecture, 33
- prime number, 1, 7
 - algebraic integer, 149
 - consecutive primes, 33
 - generated by polynomials, 32
 - in arithmetic progression, 32
 - infinitely many, 22
 - Mersenne prime, 34
 - prime gap, 33
 - Sieve of Eratosthenes, 24, 128
 - twin primes, 33
- Prime Number Theorem, 28
- prime partition, 141
- prime splitting, *see* quadratic number field
- primitive polynomial, 276
- primitive prime divisor, 31, 32
- primitive root, 77
 - existence of, 85
- proper divisor, 126
- pseudoprime, 61
 - Fermat, 58, 62
 - see also* Carmichael number
- PSQL algorithm, 4
- Putnam, Hilary, 189
- pyramidal number, 243
- Pythagorean triple, 160, 187, 198–200
 - congruent number, 238
 - for polynomials, 289
- quadratic number field
 - basis, 145, 156
 - fundamental unit, 205
 - imaginary, 145
 - prime splitting, 151, 156
 - real, 145

- Quadratic Reciprocity, 69–74
 - for polynomials, 296
 - Supplemental Law, 69
- quadratic residue, 66
- quotient, 9
- Rabin coin flipping, 107
- Rabin, Michael, 107
- rational approximation, *see*
 - Diophantine approximation
- rational function, 249
 - bad reduction, 260
 - canonical height, 274
 - conjugation, 273
 - critical points, 273
 - degree, 249
 - dynamotic polynomial, 254–258
 - forward orbit, 248
 - good reduction, 260
 - Lattès map, 272
 - modulo p , 258
 - MS uniform boundedness conjecture, 253, 273
 - multiplier, 263
 - n th iterate, 247
 - Newton map, 271
 - post-critically finite, 272
 - preimage, 268
 - resultant, 259
- rational numbers, 143, 157
- Reichardt, Hans, 190
- relatively prime, 11
- remainder, 9
- remainder after division, 39
- residue class, 40
- resultant, 258
 - rational function, 259
- Rivest, Ronald, 99
- Robinson, Julia, 189
- root of unity, 69, 83
 - modulo n , 89
 - primitive, 117
 - see also* cyclotomic polynomial
- Roth's theorem, 166, 182
- Roth, Klaus, 166
- RSA public key, 99, 108
- Selmer, Ernst, 190
- Shamir, Adi, 99
- shared secret, 93, 108
- shift cipher, *see* affine cipher
- Siegel, Carl, 235
- Sieve of Eratosthenes, 24, 128
- sieve theory, 128
- Silverman, Joseph H., 253
- smooth number, 35
- sociable numbers, 140
- squarefree, 113
- sum of divisors function, 122–125, 139
 - as power series coefficients, 124
 - formula for computing, 123
 - is multiplicative, 123
- sum of two squares, 143, 152–153
- symmetric cipher, 96
- tail of a preperiodic point, 249
- Taylor, Richard, 200
- Thue, Axel, 223
- torsion point, 229
- torsion subgroup, 235
- totient function, *see* Euler totient function
- transcendental number, 144
 - Liouville number, 165
- Triangle Inequality, 176
- triangular number, 1, 215
- twin primes, 33
- unique factorization
 - integers, 21
 - polynomials, 281
- unit, 149
- Universal Product Code (UPC), 102
- van Heijst, Eugène, 103
- Vigenère cipher, 95
- wandering point, 248
- Waring problem
 - for integers, 208–212, 219
 - for polynomials, 292, 298
 - generalized four square problem, 216
 - modulo primes, 219
 - sum of four squares, 212
 - sum of four squares, 209
- Waring, Edward, 208
- Weil, André, 237
- well ordering property, 5, 8, 200, 205
- Wiles, Andrew, 2, 200
- Wilson's Theorem
 - for polynomials, 295
 - integer, 60
 - primality test, 61
- Young diagram, 132
- Young, Alfred, 132

Published Titles in This Series

- 31 **Benjamin Hutz**, An Experimental Introduction to Number Theory, 2018
- 30 **Steven J. Miller**, Mathematics of Optimization: How to do Things Faster, 2017
- 29 **Tom L. Lindstrøm**, Spaces, 2017
- 27 **Shahriar Shahriari**, Algebra in Action, 2017
- 26 **Tamara J. Lakins**, The Tools of Mathematical Reasoning, 2016
- 25 **Hossein Hosseini Giv**, Mathematical Analysis and Its Inherent Nature, 2016
- 24 **Helene Shapiro**, Linear Algebra and Matrices, 2015
- 23 **Sergei Ovchinnikov**, Number Systems, 2015
- 22 **Hugh L. Montgomery**, Early Fourier Analysis, 2014
- 21 **John M. Lee**, Axiomatic Geometry, 2013
- 20 **Paul J. Sally, Jr.**, Fundamentals of Mathematical Analysis, 2013
- 19 **R. Clark Robinson**, An Introduction to Dynamical Systems: Continuous and Discrete, Second Edition, 2012
- 18 **Joseph L. Taylor**, Foundations of Analysis, 2012
- 17 **Peter Duren**, Invitation to Classical Analysis, 2012
- 16 **Joseph L. Taylor**, Complex Variables, 2011
- 15 **Mark A. Pinsky**, Partial Differential Equations and Boundary-Value Problems with Applications, Third Edition, 1998
- 14 **Michael E. Taylor**, Introduction to Differential Equations, 2011
- 13 **Randall Pruim**, Foundations and Applications of Statistics, 2011
- 12 **John P. D'Angelo**, An Introduction to Complex Analysis and Geometry, 2010
- 11 **Mark R. Sepanski**, Algebra, 2010
- 10 **Sue E. Goodman**, Beginning Topology, 2005
- 9 **Ronald Solomon**, Abstract Algebra, 2003
- 8 **I. Martin Isaacs**, Geometry for College Students, 2001
- 7 **Victor Goodman and Joseph Stampfli**, The Mathematics of Finance, 2001
- 6 **Michael A. Bean**, Probability: The Science of Uncertainty, 2001
- 5 **Patrick M. Fitzpatrick**, Advanced Calculus, Second Edition, 2006
- 4 **Gerald B. Folland**, Fourier Analysis and Its Applications, 1992
- 3 **Bettina Richmond and Thomas Richmond**, A Discrete Transition to Advanced Mathematics, 2004
- 2 **David Kincaid and Ward Cheney**, Numerical Analysis: Mathematics of Scientific Computing, Third Edition, 2002
- 1 **Edward D. Gaughan**, Introduction to Analysis, Fifth Edition, 1998

This book presents material suitable for an undergraduate course in elementary number theory from a computational perspective. It seeks to not only introduce students to the standard topics in elementary number theory, such as prime factorization and modular arithmetic, but also to develop their ability to formulate and test precise conjectures from experimental data. Each topic is motivated by a question to be answered, followed by some experimental data, and, finally, the statement and proof of a theorem. There are numerous opportunities throughout the chapters and exercises for the students to engage in (guided) open-ended exploration. At the end of a course using this book, the students will understand how mathematics is developed from asking questions to gathering data to formulating and proving theorems.



© 2018 Janel Peyton Photography

The mathematical prerequisites for this book are few. Early chapters contain topics such as integer divisibility, modular arithmetic, and applications to cryptography, while later chapters contain more specialized topics, such as Diophantine approximation, number theory of dynamical systems, and number theory with polynomials. Students of all levels will be drawn in by the patterns and relationships of number theory uncovered through data driven exploration.

ISBN 978-1-4704-3097-9



9 781470 430979

AMSTEXT/31



For additional information
and updates on this book, visit

www.ams.org/bookpages/amstext-31



www.ams.org



This series was founded by the highly respected
mathematician and educator, Paul J. Sally, Jr.