Daniel A. Klain

Essentials of Number Theory

Daniel A. Klain

Essentials of Number Theory

Preliminary Edition
last updated March 22, 2020

Contents

Kein' Musik ist ja nicht auf Erden,
Die uns'rer verglichen kann werden.

Preface

This is an undergraduate level introduction to classical number theory, covering traditional topics (from discoveries of the ancient Greeks, to the work of Fermat, Euler, and Gauss), along with a few sections that outline newer applications of number theory made possible by 20th century computer science.

While familiarity with calculus and linear algebra may be helpful for reasons of mathematical maturity, most of the material in this book is accessible to readers having a solid background in high school level mathematics. Sections are kept sufficiently short and focused for single session of reading. The first 26 sections (omitting sections 13, 14, and 17 with no loss of continuity) form a core introduction to the subject, providing the basic tools needed for further study. Selections can be made from among the remaining sections (13, 14, 17, 27–33) for applications and further topics. The book is also designed with extracurricular readers in mind: a student using this textbook for a reading course can read every section in the order provided.

As with all mathematical studies, the exercises are of paramount importance. Section 35 contains hints, simple answers, and, in some cases, full solutions to selected exercises.

The first edition of any textbook is likely to want correction and refinement. I hope to find and correct errors and to fill some omissions for a more polished future edition. Comments are most welcome.

I am grateful to my students at UMass Lowell for their patience with earlier drafts of this book, as well as for their comments and corrections. I am also very grateful to Tanya Khovanova for carefully proofreading several sections and for making many detailed and helpful suggestions. Finally, I am grateful to Michael Artin, for introducing me to this subject, and to Glenn Stevens, for showing me that classical number theory is a deep, exciting, and *accessible* part of mathematics that every student of the sciences can enjoy.

Dan Klain
June 2017

1 What is number theory?

At first glance the term "number theory" seems mysteriously broad. Isn't all of mathematics about numbers? Is this just another name for mathematics in general? A more cautious reader might note that geometry and logic (for example) aren't really about numbers, even if numbers are sometimes used. But even leaving out these topics, the "study of numbers" still sounds overly broad. Indeed, the term *number theory* is traditional, and refers exclusively to the study of *whole numbers*; that is, the numbers we count with:

$$1, 2, 3, 4, \ldots$$

along with the daring addition of 0, and, when convenient, the negative integers. *Excluded* from consideration are fractions, real numbers, and complex numbers. Those abstractions, while called numbers in ordinary language, are traditionally studied in courses on *Analysis*.

While the whole numbers have their origins in counting, number theory is not about counting either. The study of advanced techniques in counting[1] is a field of mathematics all its own, called *Combinatorics*.

Number theory then is the pure study of whole numbers and their relations to one another, especially with regards to addition and multiplication, both of which will always transform whole numbers into whole numbers. For a sense of what this means, consider the following questions about whole numbers:

- Is the sum of two odd numbers even or odd? What about the product?
- If we divide n by 3, we have 2 left over. If we divide the same number n by 17, we have 9 leftover. What are the possible values for n?
- Can a power of two ever end in the digits "...324"?
- When can a positive integer n be written as a sum of two integer squares?
- Is the number $1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}$ ever a whole number if $n > 1$?
- Does the equation $12x - 57y = 39$ have integer solutions x and y? What if 39 is replaced with 38?

Number theory, as described so far, may seem a rather abstract topic to spend months (years?) studying. Indeed, because of its ostensible purity and great distance from industrial or scientific applications, number theory was once known as the "Queen of Mathematics."

[1] For example, if four married couples are seated at a round table, how many ways can they be arranged, alternating by gender, so that no one sits next to his or her spouse?

This is no longer the case. While still considered an exemplar of abstract mathematical elegance, number theory now provides concrete applications to information theory and computer science, including cryptography, data compression, error-correcting codes, and pseudorandom number generation, to name just a few examples.

All the same, the most compelling reason to study number theory is for its unique combination of simplicity and mystifying complexity, which provides a setting for mathematical beauty, surprise, and the sudden clarity that can be so thrilling to those who enjoy mathematics.

2 *Integers and arithmetic*

Our primary setting is the set $\mathbb{Z}$ of integers; that is, whole numbers, positive and negative:

$$\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, 4, 5, 6, \ldots\}.$$

It is also convenient to denote by $\mathbb{N}$ the set of natural numbers; that is, the positive integers:

$$\mathbb{N} = \{1, 2, 3, 4, 5, 6, 7, 8, \ldots\}.$$

We will assume familiarity with properties of the integers that the reader should recall from grade school. In particular, the integers have an ordering, and they are closed under addition, subtraction, and multiplication. Moreover, we will also assume that $\mathbb{N}$ is closed under addition and multiplication.

The following identities summarize the algebraic properties of the integers.

Theorem 2.1 (Basic properties of integer arithmetic)**.**
Let $a, b, c \in \mathbb{Z}$.

- $a + b = b + a$ (addition is commutative)
- $a + (b + c) = (a + b) + c$ (addition is associative)
- $a + 0 = a$ (zero is the additive identity)
- $a + (-a) = 0$ (every integer has an additive inverse)
- $ab = ba$ (multiplication is commutative)
- $a(bc) = (ab)c$ (multiplication is associative)
- $a \cdot 1 = a$ (1 is the multiplicative identity)
- $a(b + c) = ab + ac$ (distributive law)
- *If $ab = 0$ then either $a = 0$ or $b = 0$ (or both).* (integral domain)

The last property in the list above implies the following *Cancellation Law*.

Proposition 2.2. *Let $a, b, c \in \mathbb{Z}$, and suppose that $a \neq 0$. If $ab = ac$ then $b = c$.*

Proof. Since $ab = ac$, we have $ab - ac = 0$, so that $a(b - c) = 0$. Since $a \neq 0$, the last property in Theorem 2.1 implies that $b - c = 0$, so that $b = c$. ❑

Notice what we did *not* say in the proof above. We did not talk about "dividing both sides by a". Instead, the proof used properties of addition, subtraction,

and multiplication, without any direct reference to division. The reason for this circumlocution will become evident as the theory unfolds.

❧

For $a, b \in \mathbb{Z}$ we write $a < b$ if a is less than b. Denote $a \leq b$ if either a is less than b or $a = b$. The following identities summarize *some* of the order properties of the integers.

Theorem 2.3 (Order properties of $\mathbb{Z}$)**.**
Let $a, b, c \in \mathbb{Z}$.

- $a \leq a$ (reflexive property)
- *If $a \leq b$ and $b \leq a$ then $a = b$* (antisymmetric property)
- *If $a \leq b$ and $b \leq c$ then $a \leq c$* (transitive property)
- *If $a, b \in \mathbb{Z}$ then $a \leq b$ or $b \leq a$* (total ordering)
- *If $a \geq b$ then $a + c \geq b + c$.*
- *If $a \geq b$ and $c \geq 0$ then $ac \geq bc$.*
- *If $a \geq 0$ and $b \geq 0$ then $a + b, ab \geq 0$.* (closure of $\mathbb{N}$)

The assertions of Theorem 2.3 should already be familiar to the reader, and we will not belabor them. However, the next property of the integers, while *fundamental*, is not generally emphasized in grade school.

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

Note that $\mathbb{Z}$ does *not* satisfy the well-ordering principle. The set of all integers has no minimum. However, the well-ordering principle still holds for the set $\mathbb{N} \cup \{0\}$ of non-negative integers: If $S \subseteq \mathbb{N} \cup \{0\}$ is nonempty, either $0 \in S$, in which case 0 is the minimum, or $0 \notin S$, in which case $S \subseteq \mathbb{N}$, and the original well-ordering principle supplies a minimum for S.

The well-ordering principle enables us to use *mathematical induction* to prove facts about the integers.

Theorem 2.4 (The Principle of Mathematical Induction)**.** *Let $S \subseteq \mathbb{N}$ such that $1 \in S$, and such that, whenever $n \in S$, we have $n + 1 \in S$ as well.*

Then $S = \mathbb{N}$.

Proof. Denote by T the complement of S in $\mathbb{N}$; that is,

$$T = \mathbb{N} - S = \{n \in \mathbb{N} \mid n \notin S\}.$$

Suppose that $S \neq \mathbb{N}$. In this case the set T is not empty, so the well-ordering principle implies that T has a minimal element t. Since $1 \in S$, we know that $t \neq 1$, so that $t-1$ is a positive integer. Since t is the minimum of T, the integer $t-1 \notin T$, so that $t-1 \in S$. Set $n = t-1$. The hypotheses of the theorem assert that, since $n \in S$, we have $t = n+1 \in S$ as well. So $t \in S$, contradicting the fact that $t \in T$. It follows that T cannot have a minimal element, so that $T = \emptyset$ and $S = \mathbb{N}$. ❑

Theorem 2.4 is often used to prove that a property holds for all positive integers by performing the following two steps:

- Verify that the property holds for the number 1. (The Trivial Case)
- Prove that, if the property holds for a positive integer n, then it must also hold for $n+1$. (The Induction Step)

Theorem 2.4 then implies that the set S of positive integers that satisfy the given property is the entire set $\mathbb{N}$; in other words, the property is true for all positive integers.

Here an example involving the factorial function: $n! = 1 \cdot 2 \cdot 3 \cdots (n-1) \cdot n$.

Example: Prove that there are $n!$ ways to seat n people in a row of n seats.

> First observe that there is $1! = 1$ way to seat one person in one seat. Suppose that there are $k!$ ways to seat k people in k seats, for some particular integer k. To seat $k+1$ people in $k+1$ seats, we have $k+1$ choices of whom to place in the first seat. There are now $k!$ ways to seat the remaining k people (by the *induction hypothesis*). It follows that there are a total of
>
> $$(k+1) \cdot k! = (k+1)!$$
>
> ways to seat $k+1$ people. The Principle of Mathematical Induction now implies that the factorial formula gives the correct answer for all n.

❧

Induction is sometimes helpful for verifying identities.

Example: Prove that

$$1 \cdot 2 \; + \; 2 \cdot 3 \; + \; 3 \cdot 4 \; + \; \cdots \; + \; n \cdot (n+1) \;\; = \;\; \frac{n(n+1)(n+2)}{3} \qquad (2.1)$$

for all positive integers n.

> First, consider the case $n = 1$. In this case the left-hand-side of the identity is $1 \cdot 2 = 2$, while the right-hand-side is
>
> $$\frac{1(1+1)(1+2)}{3} \;\; = \;\; 2$$

as well.

Next, suppose the identity holds for some integer $n \geq 1$, and consider the situation for $n+1$. In this case the left-hand-side of the identity (2.1) becomes

$$\begin{aligned}
&1\cdot 2 + 2\cdot 3 + 3\cdot 4 + \cdots + n\cdot(n+1) + (n+1)\cdot(n+2) \\
&\quad = \Big(1\cdot 2 + 2\cdot 3 + 3\cdot 4 + \cdots + n\cdot(n+1)\Big) + (n+1)\cdot(n+2) \\
&\quad = \frac{n(n+1)(n+2)}{3} + (n+1)(n+2) \qquad \text{(by the induction assumption)} \\
&\quad = (n+1)(n+2)\left(\frac{n}{3}+1\right) \\
&\quad = \frac{(n+1)(n+2)(n+3)}{3}.
\end{aligned}$$

Meanwhile, substituting $n+1$ for n on the right-hand-side of the identity (2.1) yields

$$\frac{(n+1)((n+1)+1)((n+1)+2)}{3} = \frac{(n+1)(n+2)(n+3)}{3},$$

as well. This completes the induction step and the proof.

❧

Sometimes proofs by induction are more easily understood in the language of the well-ordering, and sometimes the opposite. When you have a choice, use the approach that seems simplest to you.

❧

Recall that the positive integers $\mathbb{N}$ are closed under addition and multiplication. This observation has the following elementary but important consequence.

Proposition 2.5. *If $a, b \in \mathbb{N}$ then $ab \geq a$, with equality if and only if $b = 1$.*

Proof. The proof is by induction with respect to b. Let $a \in \mathbb{N}$. Clearly $a \cdot 1 = a$. Moreover, since $a > 0$, we have $2a = a + a > a + 0$, so that $2a > a$. Suppose that $ab > a$ for some $b \in \mathbb{N}$. We then have

$$a(b+1) = ab + a > a + a > a,$$

as well. It follows that $ab > a$ for all $b \geq 2$. ❑

❧

While $\mathbb{Z}$ is closed under addition, subtraction, and multiplication, note that division is not always possible. For example, $5/3$ is not an integer. Ratios of

integers (which may not always be integers) form another important category of numbers: A real number x is said to be a *rational number* if x can be expressed in the form

$$x = \frac{a}{b} \tag{2.2}$$

where $a, b \in \mathbb{Z}$. Every integer is also a rational number (why?), but not every rational number is an integer. Moreover, the expression (2.2) is not unique, since

$$\frac{2}{5} = \frac{4}{10} = \frac{400}{1000} = \frac{-8}{-20},$$

for example. However, for any particular rational number x, there is a unique expression (2.2) for which the denominator b is positive and minimal (see Exercise 2.14). When b is chosen this way, the expression a/b is said to be in *lowest terms*.

❧

Exercise 2.1. Show that every integer is also a rational number.

Exercise 2.2. Use mathematical induction to prove that, for $n \in \mathbb{N}$, we have

$$1 + 2 + 3 + \cdots + (n-1) + n = \frac{n(n+1)}{2}.$$

Can you also prove this formula more directly, without using induction?

Exercise 2.3. Use mathematical induction to prove that, for $n \in \mathbb{N}$, we have

$$1 + 4 + 9 + \cdots + (n-1)^2 + n^2 = \frac{n(n+1)(2n+1)}{6}.$$

Exercise 2.4. Use mathematical induction to prove that, for $n \in \mathbb{N}$, we have

$$1 + 8 + 27 + \cdots + (n-1)^3 + n^3 = [1 + 2 + 3 + \cdots + (n-1) + n]^2.$$

Hint: Exercise 2.2 is helpful.

Exercise 2.5. Given $n \in \mathbb{N}$, guess a formula for the sum

$$1 + 3 + 5 + \cdots + (2n-1),$$

and use induction to prove that your guess is correct.

Exercise 2.6. Use mathematical induction to prove that, for $n \in \mathbb{N}$, we have

$$1! \cdot 1 + 2! \cdot 2 + 3! \cdot 3 + \cdots + n! \cdot n = (n+1)! - 1.$$

Exercise 2.7. Let $F_0 = 0$ and $F_1 = 1$. For $n \geq 2$ define

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

The sequence $\{F_n\}_{n=0}^{\infty}$ is called the *Fibonacci sequence*, and the values F_n are called the *Fibonacci numbers*.[1]

(a) Use induction to prove that, for $k \geq 1$,

$$F_0 + \cdots + F_k = F_{k+2} - 1.$$

(b) Use induction to prove that $F_k \leq 2^{k-1}$ whenever $k \geq 1$.

(c) Use induction to prove that $F_k \geq 2^{k/2}$ whenever $k \geq 6$.

[1] Named for the Italian mathematician Fibonacci (c. 1170 – c. 1250), also known as Leonardo Bonacci and Leonardo of Pisa, and author of the *Liber Abaci*.

Exercise 2.8. Use mathematical induction to prove that, for integers $n \geq 4$, we have $3^n > 4n^2$.

Exercise 2.9. Use mathematical induction to prove that, for integers $n \geq 6$, we have $3^n > 2n^3 + 1$.
Hint: $2n^3 + 1 < 3n^3$.

Exercise 2.10. Use mathematical induction to prove that, for integers $n \geq 4$, we have $n! > 2^n$.

Exercise 2.11. Prove that a set S of n distinct elements has 2^n subsets.
Hint: If S is non-empty, suppose $x \in S$. Use induction on the size of a set, along with the fact that every subset of S either contains x or does not contain x.

Exercise 2.12. Prove that every non-empty subset $S \subseteq \mathbb{Z}$ contains an element of smallest absolute value.

Exercise 2.13. Show that if the number $\frac{1}{2}$ were an integer then the well-ordering principle would be violated.

Exercise 2.14. Use the well-ordering principle to prove that every rational number has a *unique* expression of the form (2.2) in *lowest terms*.
Note: This exercise asks you to prove both existence and uniqueness.

Exercise 2.15. Prove the *Principle of Strong Induction:*

> Let $S \subseteq \mathbb{N}$ such that $1 \in S$, and such that, whenever $\{1, 2, 3, \ldots, n\} \subseteq S$, we have $n + 1 \in S$ as well. Then $S = \mathbb{N}$.

Exercise 2.16. Find an example of integers $a, b \in \mathbb{Z}$ such that $ab < a$. Can you find an example where a and b are both non-zero? Can you find an example where both $ab < a$ and $ab < b$?

Exercise 2.17. Suppose that $x, y \in \mathbb{R}$ and $x, y > 0$. Is it always true that $xy \geq x$?

Exercise 2.18. Use mathematical induction to prove:

(a) For integers $n \geq 6$, we have $n! < \left(\frac{n}{2}\right)^n$.
(b) For integers $n \geq 1$, we have $n! > \left(\frac{n}{3}\right)^n$.
Hint: Logarithms and a little calculus are helpful for this exercise.

3 Some useful algebraic identities

In this section we recall two fundamental algebraic identities that you may have seen before: the geometric sum formula and the binomial theorem. These identities are theorems of algebra, not of number theory, but will be useful to us later on.

❧

Theorem 3.1. *For all positive integers n and all x, y,*

$$x^n - y^n = (x^{n-1} + x^{n-2}y + x^{n-3}y^2 + \cdots + xy^{n-2} + y^{n-1})(x - y) \tag{3.1}$$

Proof. Multiply out the right side, and notice that everything cancels except the first and last terms. ❑

There are many special cases of Theorem 3.1 worth examining in detail. For example, when $n = 2$ we have

$$x^2 - y^2 = (x + y)(x - y),$$

the classic "difference of two squares."

If $n = 3$ we have

$$x^3 - y^3 = (x^2 + xy + y^2)(x - y).$$

Setting $y = -z$ in the previous expression yields

$$x^3 + z^3 = (x^2 - xz + z^2)(x + z).$$

❧

If $y = 1$ in (3.1), we obtain the *Geometric Sum Formula*:

$$1 + x + x^2 + \cdots + x^n = \frac{1 - x^{n+1}}{1 - x} \tag{3.2}$$

which holds provided that $x \neq 1$.

Notice that if $|x| < 1$ in (3.2) then $x^{n+1} \to 0$ as $n \to \infty$. It follows that, for $-1 < x < 1$,

$$1 + x + x^2 + x^3 + \cdots = \frac{1}{1-x}. \tag{3.3}$$

This infinite sum is called a *geometric series*.

❧

Before stating the next fundamental identity, recall that "*n factorial*", denoted $n!$ is defined by

$$n! = 1 \cdot 2 \cdot 3 \cdots (n-1) \cdot n,$$

for all positive integers n. By convention we also define $0! = 1$. The number $n!$ counts the number of ways to arrange n different objects in a row: n choices for the first position, then $(n-1)$ choices for the second position, and so on until all of the n objects are arranged.

Next, define the symbol $\binom{n}{k}$ by

$$\binom{n}{k} = \frac{n!}{k!(n-k)!},$$

where n and k are non-negative integers. For example,

$$\binom{8}{3} = \frac{8!}{3!5!} = \frac{8 \cdot 7 \cdot 6}{3 \cdot 2 \cdot 1} = 56.$$

Notice that

$$\binom{n}{k} = \binom{n}{n-k},$$

and that

$$\binom{n}{0} = \binom{n}{n} = 1 \quad \text{and} \quad \binom{n}{1} = \binom{n}{n-1} = n.$$

With such a complicated looking denominator, it isn't obvious that $\binom{n}{k}$ is an integer. One way to see this is to observe that $\binom{n}{k}$ must be a whole number because it *counts* something. Suppose a class has n students, and we must choose a team of k students to go on a mission. The teacher picks the team by choosing k students one by one, having them stand in a line. There are n choices for the first student, then $(n-1)$ choices for the second student, and so one, giving

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

ways to make this *ordered* set of choices. But the membership of the team is the *same* no matter how the k lucky students are lined up. Since there are $k!$ ways to line up (re-arrange) the lucky k students, the number of actual distinct teams is

$$\frac{n!}{(n-k)!k!} = \binom{n}{k},$$

which is, therefore, an integer. This counting argument will lead to an explanation of the following algebraic identity.

Theorem 3.2 (Binomial Theorem)**.** *For all positive integers n and all x, y,*

$$\begin{aligned}(x+y)^n &= x^n + nx^{n-1}y + \frac{n(n-1)}{2}x^{n-2}y^2 + \cdots + \binom{n}{k}x^k y^{n-k} + \cdots + nxy^{n-1} + y^n \\ &= \sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k}. \end{aligned} \tag{3.4}$$

Because of their appearance as coefficients in expansions of binomial powers $(x+y)^n$, the numbers $\binom{n}{k}$ are called *binomial coefficients*.

Proof. Observe that

$$(x+y)^n = (x+y)(x+y)\cdots(x+y).$$

When this is multiplied out, and like terms are collected, each term will be of the form $a_k x^k y^{n-k}$, for some integer a_k, since each term is formed by picking either an x or a y from each of the n copies of $(x+y)$ and multiplying these n choices together. The number a_k counts the number of times we choose exactly k 'x's and $n-k$ 'y's as we select from each factor. Think of each $(x+y)$ factor as a student, where x means pick the student for our team, and y means don't pick that student. It follows that a_k is the number of ways to pick a team of k students from an n student class, and that's why $a_k = \binom{n}{k}$. ❑

The following special case of the Binomial Theorem 3.2 will be useful later on.

Corollary 3.3. *For all positive integers n and all x,*

$$(1+x)^n = 1 + nx + \binom{n}{2}x^2 + \cdots + \binom{n}{k}x^k + \cdots + nx^{n-1} + x^n. \tag{3.5}$$

❧

Exercise 3.1. Use Theorem 3.1 to prove the geometric sum formula (3.2).

Exercise 3.2. Prove that, for positive integers n

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

Exercise 3.3. Prove that

$$1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \cdots = 2.$$

Exercise 3.4. Prove that

$$\frac{9}{10} + \frac{9}{10^2} + \frac{9}{10^3} + \cdots = 1.$$

What does this tell you about the decimal expansion $0.99999\ldots$?

Exercise 3.5. Express the repeating decimal expansion $0.127127127\ldots$ as a fraction by first rewriting this expansion as a geometric series and then summing the series using (3.3).

Exercise 3.6. Prove that, for positive integers $k < n$ and all $x \neq 1$,

$$x^k + x^{k+1} + \cdots + x^n = \frac{x^k - x^{n+1}}{1 - x}$$

Exercise 3.7. Prove that, for positive integers k and $|x| < 1$,

$$x^k + x^{k+1} + \cdots = \frac{x^k}{1 - x}$$

Exercise 3.8. Factor the polynomial $x^4 - y^4$ into 3 polynomial factors.

Exercise 3.9. Factor the polynomial $x^6 + y^6$ into 2 polynomial factors.

Exercise 3.10. Factor the polynomial $x^6 - y^6$ into 3 polynomial factors.

Exercise 3.11. Use the binomial theorem to prove that the sum of the binomial coefficients satisfies

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n} = 2^n.$$

Exercise 3.12. Use the binomial theorem to prove that the alternating sum of the binomial coefficients satisfies

$$\binom{n}{0} - \binom{n}{1} + \binom{n}{2} + - + \cdots + (-1)^n \binom{n}{n} = 0.$$

Exercise 3.13. Prove that

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}.$$

Exercise 3.14. Use mathematical induction to give another proof of the geometric sum formula (3.2).

Exercise 3.15. Use mathematical induction to give another proof of the Binomial Theorem 3.2.

Exercise 3.16. Prove that, if n is an odd positive integer, then $7^n + 4^n$ is divisible by 11 (that is, $7^n + 4^n = 11k$ for some integer k).

Exercise 3.17. Prove that, if n is an even positive integer, then $7^n - 4^n$ is divisible by 33 (that is, $7^n - 4^n = 33k$ for some integer k).
Hint: Exercise 3.16 is helpful.

4 Divisibility

Given integers a, b with $a \neq 0$, we say that "a divides b", denoted $a|b$, if there exists $k \in \mathbb{Z}$ such that $b = ka$. In this case we say that a is a *factor* or *divisor* of b, and that b is a *multiple* of a.

If a does not divide b, we sometimes write $a \nmid b$. For example, $3 \nmid 8$, since $\frac{8}{3}$ is not an integer. More precisely, $8 \neq 3k$ for every integer k.

Divisibility satisfies the following elementary properties.

Proposition 4.1. *Let $a, b, c, x, y \in \mathbb{Z}$.*

(i) If $a \neq 0$ then $a|a$.

(ii) If $a, b > 0$ and $a|b$ then $a \leq b$.

(iii) If $a|b$ and $b|a$ then $a = \pm b$.

(iv) If $a|b$ and $b|c$ then $a|c$.

(v) If $a|b$ and $a|c$ then $a|(bx + cy)$ for all integers x, y.

Parts (i), (iii), and (iv) imply that the divisibility relation $|$ defines a *partial order* on the positive integers.

Proof. Part (i) is left as an exercise. Part (ii) follows from Proposition 2.5. To prove part (iii), suppose that $a, b \in \mathbb{Z}$, where $a|b$ and $b|a$. It follows that, on taking absolute values, $|a|$ divides $|b|$ and that $|b|$ divides $|a|$. Since $|a|, |b| \in \mathbb{N}$, from Proposition 2.5 implies that both $|a| \leq |b|$ and $|b| \leq |a|$, whence $|a| = |b|$. In other words, $a = \pm b$.

To prove part (iv), go back to the definition of divisibility: If $a|b$ and $b|c$ then there exist integers m, n so that $b = ma$ and so that $c = nb$. Therefore,

$$c = nb = n(ma) = (nm)a,$$

so that $a|c$.

The proof of part (v) is similar, and is left as an exercise. ❑

❧

Recall from grade school that, while not every positive integer divides another, we can always attempt a division, provided we are willing to accept a "remainder." For example, while $7 \nmid 45$, we know that

$$45 = 7 \cdot 6 + 3,$$

where 3 is a remainder.

We could also have written

$$45 = 7 \cdot 5 + 10,$$

but this is somehow inferior, because we know that another 7 can still be extracted from the 10. The process of division with remainder is not complete unless the remainder is smaller than the divisor. This assertion is made more precise by the following theorem.

Theorem 4.2 (Division with remainder)**.** *Let $a, b \in \mathbb{N}$. There exist unique integers q and r such that*

$$b = aq + r,$$

where $0 \leq r < a$.

The value q is called the *quotient*, and r is called the *remainder*.

Proof. Consider the set S of all non-negative integers of the form $b - am$, where $m \in \mathbb{Z}$. If $m = 0$ then $b - am = b > 0$, so that $b \in S$, and S is not the empty set. It follows from the well-ordering principle that S has a minimum, which we will denote by r. Since $r \in S$, it follows that $r \geq 0$ and that

$$r = b - aq$$

for some particular integer q.

Suppose $r \geq a$. In this case we have

$$b - a(q+1) = b - aq - a = r - a \geq 0.$$

This means that $r' = b - a(q+1) \in S$. But $r' = r - a < r$, since $a > 0$. This contradicts our choice of r as the *minimum* of the set S. Therefore, $0 \leq r < a$.

It remains to show that r and q are unique. Suppose that $b = aq_1 + r_1$, where $0 \leq r_1 < a$. In this case, we have

$$aq + r \;=\; b \;=\; aq_1 + r_1,$$

so that

$$|a(q - q_1)| = |r_1 - r|.$$

Since $0 \leq r, r_1 < a$, we know that $|r_1 - r| < a$, so that $|a(q - q_1)| < a$. This can only happen if $q - q_1 = 0$, so that $q_1 = q$ and $r_1 = r$. ❑

The style of argument in the last part of this proof is often used to prove uniqueness in mathematics: To show that an object X with certain properties is unique, suppose that another object Y also has those same properties, and then use those properties to prove that $X = Y$.

❧

Ancient mathematicians believed that all numbers could be expressed as fractions, however, the Pythagorean school discovered that this is not the case. Consider a square region with total area 2. The edge of this square must have length s, where $s^2 = 2$. Suppose we try to express s as a ratio of integers, $s = a/b$. If this is possible at all, then the well-ordering principle implies that this can be done in lowest terms.

In this case, we have $a = sb$, so that $a^2 = s^2b^2 = 2b^2$. In other words, a^2 is even. This means that a must also be even (see Exercises 4.6 and 4.14), so that $a = 2m$ for some integer m. Hence, $2b^2 = a^2 = 4m^2$, which implies that $b^2 = 2m^2$, so that b is also even, where $b = 2n$ for some n. We conclude that

$$s = \frac{a}{b} = \frac{2m}{2n} = \frac{m}{n},$$

where $0 < n < b$, since $n|b$. This contradicts our choice of a/b being already in lowest terms. It follows that s cannot be expressed as a rational number. We typically denote s by the symbol $\sqrt{2}$, and we say that $\sqrt{2}$ is *irrational*.

❧

Exercise 4.1. Show that $a|0$ for all non-zero integers a.

Exercise 4.2. Show that $1|n$ for all $n \in \mathbb{Z}$.

Exercise 4.3. Show that $13|1001$.

Exercise 4.4. What is the highest power of 3 (e.g. $3^0, 3^1, 3^2, \ldots$) that divides 999?

Exercise 4.5. Find the remainders of the following numbers after division by 7:

$$73, \quad 4, \quad 1001, \quad -12, \quad 0$$

Exercise 4.6. An integer n is said to be *even* if n is divisible by 2; otherwise, n is said to be *odd*. The quality of being even or odd is called the *parity* of a number. Answer each of the following questions, and prove that your answer is correct in general.
(a) Is the sum of two even numbers even or odd?
(b) Is the sum of two odd numbers even or odd?
(c) Is the sum of an even and an odd number even or odd?
(d) What happens if addition is replaced with multiplication in part (a)-(c)?

Exercise 4.7. Suppose that $a, b \in \mathbb{Z}$ and that $a + b$ is odd. Prove that ab is even.

Exercise 4.8. Let k be an even positive integer. Is it ever possible to express 1 as a sum of reciprocals,

$$1 = \frac{1}{n_1} + \frac{1}{n_2} + \cdots + \frac{1}{n_k},$$

where every denominator n_i is an odd positive integer?

If not, prove it is impossible. If so, give an explicit example.

Exercise 4.9. A natural number will be called *awesome* if it can be represented in the form $a^b + b^a$, where a and b are natural numbers. For example, number 177 is awesome, because $177 = 2^7 + 7^2$. Is 2015 awesome?[1]

Exercise 4.10. List all numbers $n \in \mathbb{N}$ such that $n = 1 + 2 + \cdots + (n-1)$, and prove that your list is complete.

Exercise 4.11. Prove parts (i) and (v) of Proposition 4.1.

Exercise 4.12. State and prove an extended version of the Theorem 4.2 for all integers $a, b \in \mathbb{Z}$.

Exercise 4.13. Prove that, if $0 \leq r, r_1 < a$, then $|r_1 - r| < a$, as claimed near the end of the last proof.
Hint: Consider the two cases in which $r > r_1$ and $r \leq r_1$.

Exercise 4.14. Let a be an integer. Prove that a is even if and only if $4|a^2$.

Exercise 4.15. Prove that if $n \in \mathbb{Z}$ then $n^5 + n^4 + 1$ is always odd.

Exercise 4.16. Let $n, m \in \mathbb{Z}$. Suppose that the remainder of n after division by 3 is 2, and suppose that the same holds for m.
(a) What is remainder of $m + n$ after division by 3?
(b) What is remainder of mn after division by 3?

Exercise 4.17. **(a)** Suppose that m and n are odd integers. Prove that

$$x^2 + 2mx + 2n = 0$$

has no integer solutions.[†]
Hint: Think about divisibility by 2 and by 4. Consider the different cases that may arise.

(b) Now use part (a) to prove that this quadratic equation has no rational solutions.

Exercise 4.18. Let a be an integer.
(a) Prove that a is divisible by 3 if and only if a^2 is divisible by 9.
(b) Prove that $\sqrt{3}$ is irrational.

Exercise 4.19. Prove that $2^{\frac{1}{3}}$ is irrational.

Exercise 4.20. Recall the Fibonacci sequence $\{F_n\}$ of Exercise 2.7.
(a) Is the Fibonacci number F_{1000} even or odd? What about $F_{1000000}$?
(b) Is the Fibonacci number F_{1000} divisible by 3? If not, what is the remainder after division by 3? What about $F_{1000000}$?
Note: You should be able to answer these questions without any sort of computer assistance (or long computations).

[1] If you like awesome numbers, then you might also enjoy this delightful mathematics blog: `http://blog.tanyakhovanova.com`

[†] From the 1907 Eötvös Competition [25].

5 Representations of numbers

Here are some examples of natural numbers:

$$45 \qquad 6 \qquad 1635 \qquad 3003000 \qquad 11111.$$

What do these symbols mean exactly? Traditionally we represent numbers "base ten," meaning that the first ten non-negative integers are represented by the distinct symbols

$$0, 1, 2, 3, 4, 5, 6, 7, 8, 9,$$

after which we use multiple symbols to count higher powers of the number ten, which is itself rendered "10," meaning one ten and zero ones. For example, the symbol "1635" is a shorthand that represents the integer

$$5 \ + \ 3 \cdot 10 \ + \ 6 \cdot 10^2 \ + \ 1 \cdot 10^3.$$

The symbols $1, 6, 3, 5$ are called *digits*. Since the number has been rendered base 10 in this case, these particular symbols are *decimal digits*.[1]

Notice that if we use the division algorithm to divide the number 1635 by the number 10, the remainder is the "ones" digit 5. In other words,

$$1635 = 163 \cdot 10 \ + \ 5.$$

If we then repeat the procedure with the quotient 163, we find

$$163 = 16 \cdot 10 \ + \ 3,$$

and so on.

❧

More generally, the representation of an integer n in a base b is a shorthand for expressions of the form

$$n = d_0 + d_1 b + d_2 b^2 + \cdots d_k b^k,$$

where the values $d_0, d_1, \ldots, d_k$ are the *base b digits* and must each satisfy $0 \leq d_i \leq b - 1$.

This representation is obtained by repeated division with remainder by the value b. Theorem 4.2 guarantees that each digit d_i lies in the list $\{0, 1, \ldots, b-1\}$ of possible remainders, and that the resulting expansion is *unique*. This phenomenon is expressed by the following theorem.

[1] The word *decimal* is derived from the Latin word *decem* for the number 10. The word *digit* is derived from the Latin word *digitus* meaning finger.

Theorem 5.1. *Let $n, b \in \mathbb{N}$, where $b > 1$. The number n has a unique expansion of the form*

$$n = d_0 + d_1 b + \cdots + d_k b^k, \tag{5.1}$$

where $0 \leq d_i < b$ for each $i = 0, \ldots, k$.

In other words, every positive integer has *exactly one* way of being written down in base b notation.

❧

From a purely mathematical perspective, ten is not such a desirable base, since it is not divisible by 3 or 4. Indeed, the ancient Babylonians used a notation in base 60, which amends this flaw, at the expense of having many more distinct symbols to deal with. As far as I can tell the only reason most human civilizations have settled on base ten is that we have ten fingers to count with.

Computers store information in the form of switches that have one of two possible states, *on* or *off*. Consequently numbers are stored by computers in *binary* form; that is, base 2. So, for example, the binary number 10110111 represents the decimal number

$$\begin{aligned} &1 \cdot 2^0 \ + \ 1 \cdot 2^1 \ + \ 1 \cdot 2^2 \ + \ 0 \cdot 2^3 \ + \ 1 \cdot 2^4 \cdot \ + 1 \cdot 2^5 \cdot \ + 0 \cdot 2^6 \ + \ 1 \cdot 2^7 \\ &= \ 1 + 2 + 4 + 16 + 32 + 128 \quad = \quad 183. \end{aligned}$$

Binary digits are so important that the phrase "binary digit" has been condensed to the word "bit."

If there is any doubt how a numerical representation might be interpreted, we write $(d_k d_{k-1} \cdots d_1 d_0)_b$ to denote a number in base b notation. For example, we have

$$(10110111)_2 \ = \ (183)_{10}.$$

But this tedious notation is dropped when there is no ambiguity.

❧

The iterated division algorithm provides a method for converting one base to another. For example, to express 1635 in base 8 (octal), repeat divisions by 8 to obtain:

$$\begin{aligned} 1635 &= 8 \cdot 204 + 3 \\ 204 &= 8 \cdot 25 + 4 \\ 25 &= 8 \cdot 3 + 1 \\ 3 &= 8 \cdot 0 + 3 \end{aligned}$$

so that $(1635)_{10} = (3143)_8$. Note that the digits are written in *reverse* order of their appearance in the sequence of divisions.

Conversely,

$$3 + 4 \cdot 8 + 1 \cdot 8^2 + 3 \cdot 8^3 = 1635$$

once again.

❧

When a base b is a number greater than 10, new symbols must be added to represent the decimal numbers $10, 11, \ldots, (b-1)$ in the new base.

An important special case is base 16, which uses the *hexadecimal* digits:

$$0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F.$$

For example,

$$(18)_{10} = (12)_{16} \qquad (45)_{10} = (2D)_{16} \qquad (256)_{10} = (100)_{16}.$$

The most commonly used bases are 2, 8, 10, and 16. Base 8 digits are called *octal digits*. In this book you should assume that a number is expressed in base 10 unless otherwise noted.

❧

Exercise 5.1. Convert the following binary numbers to decimal:

(a) 1
(b) 11
(c) 110
(d) 1000
(e) 10101
(f) 11101101

Exercise 5.2. Convert the following decimal numbers to binary:

(a) 5
(b) 15
(c) 16
(d) 24
(e) 60
(f) 729

Exercise 5.3. Convert the binary numbers in Exercise 5.1 to octal and hexadecimal by grouping binary digits in 3s and 4s (respectively) and converting these groups into octal or hexadecimal digits (respectively). Why does this work?

Exercise 5.4. Convert the decimal numbers in Exercise 5.2 to octal and hexadecimal.

Exercise 5.5. Convert the following octal numbers to binary:
Hint: Each octal digit expands to a group of three bits. (Why?)

(a) 3
(b) 10
(c) 30
(d) 32
(e) 70
(f) 100

Exercise 5.6. Convert the following hexadecimal numbers to binary and to decimal:

(a) 9
(b) C
(c) 10
(d) 15
(e) $2A$
(f) ABC

Exercise 5.7. Let n be an integer.
(a) Prove that n is divisible by 2 if and only if its final decimal digit is even.
(b) Prove that n is divisible by 4 if and only if its final two decimal digits form a number that is divisible by 4.
(c) Prove that 2^n never ends with the digits '14'.
Hint: Notice that the final decimal digit of n is the remainder after division of n by the number 10.

Exercise 5.8. Suppose that $n \in \mathbb{N}$. What possible values can be found as the final decimal digit of n^2?

Exercise 5.9. Use the results of the previous exercise to explain why you know *at a glance* that the number

$$574189075420685700469527892875489843284092292096285751873$$

is not the square of an integer.

Exercise 5.10. Suppose that $n \in \mathbb{N}$ has 6 as its final decimal digit. Prove that all of the higher powers $n^2, n^3, n^4, \ldots$ also have 6 for a final decimal digit.

Exercise 5.11. Suppose a number x is given by the *binary* expansion

$$0.10011101110111\ldots,$$

where the pattern 0111 repeats forever. Rewrite this expansion using geometric series, and then sum the series using (3.3) to express x as a fraction using decimal digits.

Exercise 5.12. Express the decimal fraction $1/3$ as an infinite *binary* expansion.
Hint: Express the number 3 in binary, and perform a long division in base 2. Continue until a pattern appears.

6 Common divisors

An integer $c \neq 0$ is said to be a *common divisor* of integers a and b if $c|a$ and $c|b$.

Suppose that a and b are not both zero. In this case, Proposition 2.5 implies that, if $c|a$ and $c|b$, then $|c| \leq \max\{|a|, |b|\}$, so that there are a *finite* number of integers c that divide both a and b. Since finite set of integers always has a maximum, there is a unique *largest* integer $d > 0$ such that $d|a$ and $d|b$. We call this integer the *greatest common divisor* of a and b, and denote it by $\gcd(a, b)$.

Here are some elementary properties of the gcd.

Proposition 6.1. *Let $a, b \in \mathbb{Z}$, not both zero.*

- $\gcd(a, b) = \gcd(b, a)$.
- $\gcd(a, 1) = 1$.
- $\gcd(a, 0) = |a|$ *for all* $a \neq 0$.
- $\gcd(a, b) = |a|$ *if and only if* $a|b$.

The proof of this proposition is left as an exercise (see Exercise 6.5).

The following proposition will lead to an easy and fast way to compute the gcd.

Proposition 6.2. *If $a, b \in \mathbb{Z}$ then*

$$\gcd(a, ax + b) = \gcd(a, b)$$

for all integers $x \in \mathbb{Z}$.

Proof. Let $d = \gcd(a, b)$ and let $d' = \gcd(a, ax + b)$. Since $d|a$ and $d|b$, it follows from part (v) of Proposition 4.1 that $d|(ax + b)$. In other words, d is a common divisor of a and $ax + b$. Since d' is the greatest such (by definition), it follows that $d \leq d'$.

Similarly, since $d'|a$ and $d'|(ax + b)$ we have $d'|((ax + b) - ax)$, so that $d'|b$. Therefore, $d' \leq \gcd(a, b)$; that is, $d' \leq d$. It now follows that $d' = d$. ❑

Proposition 6.2 allows us to use division with remainder to efficiently compute the gcd. To compute $\gcd(a, b)$, use Theorem 4.2 to perform a division:

$$b = aq + r,$$

where $0 \leq r < a$. By Proposition 6.2 we have

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

Since $r < a$ we can repeat the procedure:

$$a = rq_1 + r_1,$$

where $0 \leq r_1 < r$. Again, by Proposition 6.2 we have $\gcd(a, b) = \gcd(r, a) = \gcd(r, r_1)$. Iterating this procedure yields a succession of remainders $r > r_1 > r_2 > \ldots \geq 0$, such that

$$\gcd(a, b) = \gcd(r, a) = \gcd(r_1, r) = \cdots = \gcd(r_k, r_{k-1}) = \cdots.$$

Since each remainder is non-negative and strictly smaller than its predecessor, this procedure must terminate; that is, eventually $r_{k+1} = 0$, so that $\gcd(a, b) = \gcd(0, r_k) = r_k$.

This method of computing the gcd is called *Euclid's Algorithm*.

Example:

To compute $\gcd(1029, 840)$, we iterate division with remainder to obtain

$$\begin{aligned} 1029 &= 840 \cdot 1 + 189 \\ 840 &= 189 \cdot 4 + 84 \\ 189 &= 84 \cdot 2 + 21 \\ 84 &= 21 \cdot 4 + 0 \end{aligned} \tag{6.1}$$

so that $\gcd(1029, 840) = 21$.

Sure enough, a simple computation reveals that $1029 = 21 \cdot 49$ and that $840 = 21 \cdot 40$. Notice that, after factoring out the gcd of 21 from the input numbers $(1029, 840)$, the new pair $(49, 40)$ has no positive common factors except 1. In grade school one might have said that

$$\frac{1029}{840} = \frac{49}{40},$$

where the second fraction is expressed in *lowest terms*.

Example:

To compute $\gcd(573, 46)$, we iterate division with remainder to obtain

$$\begin{aligned} 573 &= 46 \cdot 12 + 21 \\ 46 &= 21 \cdot 2 + 4 \\ 21 &= 4 \cdot 5 + 1 \\ 4 &= 1 \cdot 4 + 0 \end{aligned} \tag{6.2}$$

so that $\gcd(573, 46) = 1$. In other words, the fraction $\frac{573}{46}$ is already in lowest terms.

If the number 1 appears as a remainder (as in the preceding example), we may immediately conclude that the gcd $= 1$. In this case, we say that the original pair of numbers (such as 573 and 46) are *relatively prime* or *co-prime*.

❧

The greatest common divisor of three or more integers $a_1, a_2, \ldots, a_n$ is defined to be the largest integer d that divides every integer a_i.

The list is *mutually relatively prime* if $\gcd(a_1, \ldots, a_n) = 1$. A list of integers $a_1, a_2, \ldots, a_n$ is said to be *pairwise relatively prime* if $\gcd(a_i, a_j) = 1$ for every pair of numbers a_i, a_j in the list. You should think about why these two concepts are not the same.

❧

Given integers a and b, not both zero, and denoting $d = \gcd(a, b)$, we will reverse Euclid's algorithm to solve the equation

$$ax + by = d$$

where the solutions x and y are both **integers**. This is called a *linear Diophantine equation*. More generally, *Diophantine equations*[1] are equations that require integer-valued solutions.

To see how this works, consider the example of $\gcd(1029, 840)$ derived earlier. Reversing the equations (6.1) of division with remainder, we find that

$$\begin{aligned} 21 &= 189 + 84 \cdot (-2) \\ &= 189 + (840 + 189 \cdot (-4)) \cdot (-2) \;=\; 840 \cdot (-2) + 189 \cdot 9 \\ &= 840 \cdot (-2) + (1029 + 840 \cdot (-1)) \cdot 9 \;=\; 840 \cdot (-11) + 1029 \cdot 9 \end{aligned}$$

so that

$$840 \cdot (-11) + 1029 \cdot 9 \;=\; 21.$$

We have solved $804x + 1029y = 21$ with integer solutions $x = -11$ and $y = 9$. But what a mess! This tedious procedure will be simplified later on, but for the moment we give an alternative proof that this always works.

Theorem 6.3. *Let $a, b \in \mathbb{Z}$, not both zero. Let $d = \gcd(a, b)$. There exist integers x and y such that*

$$ax + by = d. \tag{6.3}$$

[1] Named after Diophantus of Alexandria, a 3rd century Greek mathematician and author of *Arithmetica*. In spite of his enormous influence, very little is known about the life of Diophantus [33, p. 50-51]

Proof. Let S be the set of all positive integers of the form $ax + by$, where $x, y \in \mathbb{Z}$. Since a and b are not both zero, S is not empty. (For example, setting $x = a$ and $y = b$ we have $ax + by = a^2 + b^2 > 0$, so that S contains this number.) Since S is a non-empty set of positive integers, the well-ordering principle asserts that S has a minimum element, which we will denote by e. Since $e \in S$, it follows that $e = ax_0 + by_0$ for some particular integers x_0 and y_0. It remains to show that e is the gcd.

Using division with remainder, write $a = eq + r$, where $0 \leq r < e$. Note that

$$r \;=\; a - eq \;=\; a - (ax_0 + by_0)q \;=\; a(1 - qx_0) + b(-qy_0).$$

If $r > 0$ then this formulation implies that $r \in S$. But this violates the minimality of e in S, since $r < e$. It follows that $r = 0$, so that $e|a$. Similarly, $e|b$. In other words, e is a common factor of a and b.

Let $d = \gcd(a, b)$. Since e is a common factor of a and b, we immediately know that $e \leq d$. However, since $d|a$ and $d|b$, it follows that $d|(ax_0 + by_0)$, so that $d|e$. Hence, $d \leq e$, so that $e = d$, and

$$\gcd(a, b) \;=\; d \;=\; e \;=\; ax_0 + by_0,$$

so $x = x_0$ and $y = y_0$ are integers that solve the equation (6.3). ❑

❧

The solution provided by Theorem 6.3 is not unique. For example, the equation

$$5x + 3y = 4$$

has integer solutions $x = 2$ and $y = -2$. However, $x = 8$ and $y = -12$ also solve this equation.

❧

Theorem 6.3 has many useful corollaries.

Corollary 6.4. *An integer k is a common factor of a and b if and only if $k|\gcd(a, b)$.*

Proof. Let $d = \gcd(a, b)$.

Suppose that $k|d$. Since $d|a$ we have $k|a$ by transitivity (Proposition 4.1, part (iv)). Similarly, $k|b$, so that k is a common factor of a and b.

For the converse, suppose that $k|a$ and $k|b$. By Theorem 6.3, there exist integers x, y so that $d = ax + by$. Since $k|a$ and $k|b$, it follows (again from Proposition 4.1) that $k|(ax + by)$, so that $k|d$. ❑

Corollary 6.5. *Let $a, b \in \mathbb{Z}$, not both zero, and let $k \in \mathbb{N}$. Then*

$$\gcd(ka, kb) = k \gcd(a, b).$$

Proof. Let $d = \gcd(a, b)$. Since $d|a$, it follows that $a = dx$ for some integer x, so that $ka = kdx$. In other words, $kd|ka$. Similarly, $kd|kb$. It follows that $kd \leq \gcd(ka, kb)$.

Meanwhile, since $k|ka$ and $k|kb$, we know that $k|\gcd(ka, kb)$ by Corollary 6.4. This means that $\gcd(ka, kb) = ke$ for some integer e.

Since $\gcd(ka, kb)|ka$, we have $ke|ka$, so that $e|a$. Similarly, $e|b$. Therefore, $e \leq d$, so that $\gcd(ka, kb) = ke \leq kd$.

On combining these two results, we have $\gcd(ka, kb) = kd$, and the corollary follows. ❑

❧

Exercise 6.1. Compute the following greatest common divisors:

(a) $\gcd(6, 15)$
(b) $\gcd(6, 14)$
(c) $\gcd(6, 13)$
(d) $\gcd(15, 10^{100})$
(e) $\gcd(2^{57}, 2^{43})$
(f) $\gcd(2^3 \cdot 3^2, 3^3 \cdot 2^2)$

Exercise 6.2. What is the $\gcd(0, 6)$? $\gcd(0, -6)$?

Exercise 6.3. Let n be a positive integer.
(a) What is the $\gcd(n, n + 1)$?
(b) What is the $\gcd(n, n + 2)$?

Exercise 6.4. Prove that $\gcd(a, b) > 0$ always.

Exercise 6.5. Prove Proposition 6.1.

Exercise 6.6. Use Euclid's algorithm to compute $\gcd(3864, 3335)$. You can do this without the help of an electronic device!

Exercise 6.7. What is the $\gcd(-84, -123)$?

Exercise 6.8. Recall the Fibonacci sequence $\{F_n\}$ of Exercise 2.7. Prove that if $n \geq 0$ then $\gcd(F_n, F_{n+1}) = 1$.

Exercise 6.9. Explain why the equation $6x + 3y = 4$ can have *no* integer solutions.

Exercise 6.10. Let $a, b \in \mathbb{Z}$, not both zero. Let $d = \gcd(a, b)$, where $a = da'$ and $b = db'$. Prove that $\gcd(a', b') = 1$.

Exercise 6.11. Let $a, b, c \in \mathbb{N}$. Prove that $\gcd(a, b, c) = \gcd(a, \gcd(b, c))$.

Exercise 6.12. Find $a, b, c \in \mathbb{N}$ that are pairwise relatively prime, but are not mutually relatively prime.

Exercise 6.13.
(a) Find positive integer solutions to the equation $21x - 15y = 3$.
(b) Explain why the equation $21x - 15y = 2$ has no integer solutions.
(c) Are there integer solutions to the equation $21x - 15y = 6$?

7 Relatively prime integers

Given integers a, b not both zero, we say that a and b are *relatively prime* (or *co-prime*) if $\gcd(a, b) = 1$.

The following special case of Theorem 6.3 provides a fundamental step in many of the proofs that follow.

Theorem 7.1. *The equation*

$$ax + by = 1$$

has integer solutions x and y if and only if a and b are relatively prime.

Proof. If a and b are relatively prime, then the result follows from Theorem 6.3.

For the converse, suppose that $ax + by = 1$, where a and b are integers, and let $d = \gcd(a, b)$. Since $d|a$ and $d|b$, it follows that $d|(ax + by)$, which means that $d|1$. It follows that $d = 1$ and that a and b are relatively prime. ❑

❧

The next result is fundamental, as is its method of proof.

Theorem 7.2. *If $\gcd(a, b) = 1$ and $a|bc$, then $a|c$.*

Proof. By Theorem 6.3, there exist integers x and y so that

$$ax + by = 1$$

It follows that

$$acx + bcy = c.$$

Since $a|bc$, it now follows that $a|(acx + bcy)$, so that $a|c$. ❑

The next result has a similar proof, left as an important exercise for the reader (see Exercise 7.1).

Proposition 7.3. *If $\gcd(a, b) = 1$ and $a|c$ and $b|c$, then $ab|c$.*

❧

For $a, b \in \mathbb{N}$, denote by $\text{lcm}(a, b)$ the *least common multiple* of a and b. When considering the least common multiple of arbitrary non-zero integers in $\mathbb{Z}$, only positive multiples are considered, so that $\text{lcm}(a, b)$ is never negative.

More generally, the least common multiple of three or more integers $a_1, a_2, \ldots, a_n$ is defined to be the smallest integer e that is divisible by every integer a_i.

Note that

$$\text{lcm}(ka, kb) = k \cdot \text{lcm}(a, b).$$

The proof of this identity follows in a simple manner from the definition of the lcm and is left as an exercise (see Exercise 7.11).

Proposition 7.4. *Let $a, b \in \mathbb{N}$. If $\gcd(a, b) = 1$ then $\text{lcm}(a, b) = ab$.*

Proof. Let $e = \text{lcm}(a, b)$. Since $\gcd(a, b) = 1$, there exist $x, y \in \mathbb{Z}$ such that $ax + by = 1$. It follows that

$$aex + bey = e.$$

Since $a|e$ and $b|e$, there are $k, l \in \mathbb{N}$ such that $e = ak$ and $e = bl$. Therefore,

$$a(bl)x + b(ak)y = e,$$

so that $ab|e$. In particular, $ab \leq e$.

Meanwhile, since e is the *least* common multiple, and since ab is (obviously!) a common multiple of a and b, it follows that $e \leq ab$. Hence, $e = ab$. ❑

The previous proposition generalizes in the following way to arbitrary pairs of positive integers.

Theorem 7.5. *If $a, b, \in \mathbb{N}$, then*

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

Proof. Let $d = \gcd(a, b)$, and let $a' = a/d$ and $b' = b/d$. It follows from Corollary 6.5 that a' and b' are relatively prime. By the previous proposition, $\text{lcm}(a', b') = a'b'$. It follows that

$$\text{lcm}(a, b) = \text{lcm}(da', db') = d \cdot \text{lcm}(a', b') = da'b',$$

so that

$$\gcd(a, b) \cdot \text{lcm}(a, b) = d^2 a'b' = da'db' = ab.$$

❑

Since Euclid's algorithm gives us an efficient way to compute $\gcd(a,b)$, Theorem 7.5 gives us a simple algorithm for computing $\text{lcm}(a,b)$ as well.

❧

Exercise 7.1. Prove Proposition 7.3.

Exercise 7.2. Find integers $a, b, c \in \mathbb{Z}$ such that $a|c$ and $b|c$ but $ab \nmid c$.

Exercise 7.3. Let $d = \gcd(a,b)$. Use Theorem 6.3 to give a short proof that $\frac{a}{d}$ and $\frac{b}{d}$ are relatively prime.

Exercise 7.4. Suppose that $x = \frac{a}{b}$, where a and b are positive integers.
(a) Prove that if the specific expression $\frac{a}{b}$ for x is in lowest terms (that is, b is the smallest positive integer denominator possible), then $\gcd(a,b) = 1$.
(b) Prove that if $\gcd(a,b) = 1$. the specific expression $\frac{a}{b}$ for x is in lowest terms.

Exercise 7.5. Suppose that $a, b \in \mathbb{N}$ and that $\gcd(a,b) = 1$. Prove that $\gcd(a+b, a-b)$ is either 1 or 2.

Exercise 7.6. Let a be an integer.
(a) Show that $\gcd(3a+7, 2a+5) = 1$
(b) What are possible values of $\gcd(5a+4, 2a+1)$?

Exercise 7.7. Suppose that $\gcd(a^2, b^2) = 1$. Use Theorem 7.1 to prove that $\gcd(a,b) = 1$.

Exercise 7.8. Suppose that $\gcd(a,b) = 1$.
(a) Use Theorem 7.2 to prove that $\gcd(a, b^2) = 1$.
Hint: Let $d = \gcd(a, b^2)$. What is $\gcd(d, b)$?
(b) Now use part (a) to prove that $\gcd(a^2, b^2) = 1$.

Exercise 7.9. Let $n \in \mathbb{N}$. Compute the following least common multiples:

(a) $\text{lcm}(12, 28)$
(b) $\text{lcm}(12, 29)$
(c) $\text{lcm}(132, 143)$
(d) $\text{lcm}(2^3 \cdot 5^6, 3^2 \cdot 5^3)$
(e) $\text{lcm}(n, 1)$
(f) $\text{lcm}(n, 2)$
(g) $\text{lcm}(n, n+1)$
(h) $\text{lcm}(n, n+2)$

Exercise 7.10. Suppose that $a, b \in \mathbb{N}$. Prove that $\text{lcm}(a,b) = a$ if and only if $b|a$.

Exercise 7.11. Suppose that $a, b, k \in \mathbb{N}$. Prove that $\text{lcm}(ka, kb) = k \cdot \text{lcm}(a,b)$.

Exercise 7.12. Prove that an integer m is a common multiple of a and b if and only if $\text{lcm}(a,b)|m$.
Hint: Division with remainder is helpful here.

Exercise 7.13. Let $a, b, c \in \mathbb{N}$. Prove that $\text{lcm}(a,b,c) = \text{lcm}(a, \text{lcm}(b,c))$.
Hint: The result of the previous exercise can be helpful here.

Exercise 7.14. Prove or disprove: If $\gcd(a,b,c) = 1$, then $\text{lcm}(a,b,c) = abc$.

Exercise 7.15. Suppose that $\gcd(n, 4) = 1$. If we divide n by 4, what possible remainders could occur?

Exercise 7.16. Suppose that $\gcd(n, 5) = 1$. If we divide n by 5, what possible remainders could occur?

Exercise 7.17. Suppose that m and n are odd integers, and let $d \in \mathbb{N}$. Prove that if 2^d divides $m^3 - n^3$ then 2^d divides $m - n$.[†]

Exercise 7.18. Find an integer such that $\gcd(n, 3) = 1$ and $\gcd(n, 4) = 1$, or explain why this is impossible.

Exercise 7.19. Find an integer such that $\gcd(n, 4) = 2$ and $\gcd(n, 6) = 3$, or explain why this is impossible.

Exercise 7.20. If $\gcd(n, 12) = 4$ and $\operatorname{lcm}(n, 12) = 84$, what is n?

[†]From the 1908 Eötvös Competition [25].

8 Solving linear Diophantine equations

For the case in which $\gcd(a, b) = 1$, the reversal of the identities generated by Euclid's algorithm allows us to find x and y so that

$$ax + by = 1.$$

However, reversing those identities can be tedious, and result in a tree of confusingly nested parentheses.

For example, to find integer solutions to $573x + 46y = 1$, our results from (6.2) yield

$$\begin{aligned} 1 &= 21 - 4 \cdot 5 \\ &= 21 - [46 - 21 \cdot 2] \cdot 5 \\ &= [573 - 46 \cdot 12] - [46 - [573 - 46 \cdot 12] \cdot 2] \cdot 5 \\ &= 573 \cdot 11 \ + \ 46 \cdot (-137). \end{aligned}$$

This is not a pleasant procedure.

A *continued fraction* makes this process much easier. Consider again the example earlier in which we used Euclid's Algorithm to show that $\gcd(573, 46) = 1$. Re-write the first division in (6.2) in the form

$$\frac{573}{46} \ = \ 12 + \frac{21}{46} \ = \ 12 + \frac{1}{\frac{46}{21}}.$$

Then write the next division of (6.2) as

$$\frac{46}{21} \ = \ 2 + \frac{4}{21} \ = \ 2 + \frac{1}{\frac{21}{4}}.$$

Putting the two together, we have

$$\frac{573}{46} \ = 12 + \frac{1}{2 + \frac{1}{\frac{21}{4}}}.$$

Continuing with the rest of (6.2), we have

$$\frac{573}{46} \ = 12 + \frac{1}{2 + \dfrac{1}{5 + \dfrac{1}{4}}} \tag{8.1}$$

To solve $573x + 46y = 1$, remove the last bottom fraction from the expression (8.1) above, and then simplify the continued fraction that remains:

$$12 + \cfrac{1}{2 + \cfrac{1}{5}} = \cdots = \frac{137}{11}.$$

The numerator and denominator of this last fraction will, with a suitable change of sign, yield values for x and y. In this example, we have $x = 11$ and $y = -137$, so that

$$573 \cdot 11 + 46 \cdot (-137) = 1.$$

The short explanation for why this works is that the continued fraction expansion provides a simple bookkeeping device for Euclid's Algorithm that makes unwinding the division identities much easier.[1] A more detailed explanation, including a careful *proof* that this trick always works, will be given in Section 32.

❧

Euclid's algorithm (and the continued fraction shorthand) can be used to solve more general linear Diophantine equations. Here are some more examples to illustrate the method.

Example: Find integers x and y such that $130x + 88y = 12$.

Using Euclid's algorithm we find that $\gcd(130, 88) = 2$ (an exercise for you). Dividing the given Diophantine equation by 2 yields the equivalent equation

$$65x + 44y = 6.$$

To find integer solutions x and y, compute the continued fraction for $\frac{65}{44}$ via Euclid's algorithm, so that

$$\frac{65}{44} = 1 + \cfrac{1}{2 + \cfrac{1}{10 + \cfrac{1}{2}}}. \tag{8.2}$$

Since

$$1 + \cfrac{1}{2 + \cfrac{1}{10}} = \frac{31}{21},$$

we see that $65(21) + 44(-31) = 1$. Multiplying by 6 yields

$$65(126) + 44(-186) = 6.$$

Example: Find integers x and y such that $130x + 78y = 12$.

[1] Notice the analogous role of the *partial quotients* 12, 2, 5, and 4 in the continued fraction (8.1) and in the display of Euclid's algorithm in (6.2).

Using Euclid's algorithm we find that $\gcd(130, 78) = 26$. This means that $26|(130x + 78y)$ whenever $x, y \in \mathbb{Z}$. Since $26 \nmid 12$, this Diophantine equation has *no solutions*.

The next theorem summarizes our ability to solve linear Diophantine equations.

Theorem 8.1. *Let $a, b \in \mathbb{Z}$, both non-zero. The linear Diophantine equation*

$$ax + by = e$$

has integer solutions x and y if and only if $\gcd(a, b)|e$.

If a particular solution (x_0, y_0) exists, then there are infinitely many solutions, all having the form

$$x = x_0 + \frac{kb}{\gcd(a, b)} \quad \text{and} \quad y = y_0 - \frac{ka}{\gcd(a, b)},$$

where $k \in \mathbb{Z}$.

Continuing with the example above, the complete set of *all* integer solutions to the equation

$$130x + 88y = 12$$

is given by

$$x = 126 + 44k \qquad \text{and} \qquad y = -186 - 65k \text{ for } k \in \mathbb{Z}.$$

Proof of Theorem 8.1. The first part of the Theorem (existence of solutions) follows from Theorem 6.3. There remains to describe all possible solutions. Suppose the pair (x_0, y_0) solves the equation, and let (x, y) be a second solution. We have

$$\begin{aligned} ax_0 + by_0 &= e \quad \text{as well as} \\ ax + by &= e. \end{aligned} \tag{8.3}$$

Subtracting, we have

$$a(x_0 - x) = b(y - y_0).$$

Let $d = \gcd(a, b)$, so that a/d and b/d are *integers*, and $\gcd(a/d, b/d) = 1$. We now have

$$(a/d)(x - x_0) = (b/d)(y_0 - y).$$

Since $\gcd(a/d, b/d) = 1$, it follows from Theorem 7.2 that (b/d) divides $x - x_0$, so that $x = x_0 + k(b/d)$ for some $k \in \mathbb{Z}$, whence $y = y_0 - k(a/d)$. Meanwhile, it is an immediate consequence of (8.3) that the pair

$$x = x_0 + k(b/d) \qquad y = y_0 - k(a/d)$$

solves $ax + by = e$ for every $k \in \mathbb{Z}$. ❑

❧

A re-interpretation of Euclid's algorithm as a sequence of matrix row operations provides yet another fast algorithm for solving linear Diophantine equations.[2]

In order to solve the equation

$$ax + by = \gcd(a, b),$$

begin with the 2×3 matrix

$$\left[\begin{array}{c|cc} a & 1 & 0 \\ b & 0 & 1 \end{array}\right]. \tag{8.4}$$

In accordance with Euclid's algorithm we will use repeated division with remainder on the *first column* entries to determine a sequence of elementary row operations that subtract an ***integer*** multiple of one row from another, continuing until the matrix has the form

$$\left[\begin{array}{c|cc} \gcd(a, b) & * & * \\ 0 & * & * \end{array}\right],$$

where the symbol $*$ represents whatever various numerical results appear in the remaining 4 matrix entries. The row operations needed will correspond exactly to the steps of Euclid's algorithm. An exchange of rows may be necessary at the final step.

It will turn out that the final reduced matrix will always have the form

$$\left[\begin{array}{c|cc} \gcd(a, b) & x & y \\ 0 & * & * \end{array}\right]$$

where $ax + by = \gcd(a, b)$.

The reason this always works will be explained below, but first we illustrate with a numerical example, using the matrix reduction method to solve the equation

$$573x + 46y = 1. \tag{8.5}$$

Begin with the matrix

$$\left[\begin{array}{c|cc} 573 & 1 & 0 \\ 46 & 0 & 1 \end{array}\right],$$

and use division with remainder on the first column entries to obtain

$$573 = 46 \cdot 12 + 21.$$

[2]Readers unfamiliar with matrix algebra and matrix row reduction can skip this discussion without loss of continuity. An introduction to matrix row reduction and its connection to solving linear equations can be found in [4] or in any modern linear algebra text.

Subtract 12 of the bottom row from the top row to obtain a new matrix:

$$\left[\begin{array}{cc|cc} 573 & & 1 & 0 \\ 46 & & 0 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{cc|cc} 21 & & 1 & -12 \\ 46 & & 0 & 1 \end{array}\right]$$

Continuing as in (6.2), we perform the following sequence of row operations on the matrix (8.4):

$$\begin{gathered} \left[\begin{array}{cc|cc} 573 & & 1 & 0 \\ 46 & & 0 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{cc|cc} 21 & & 1 & -12 \\ 46 & & 0 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{cc|cc} 21 & & 1 & -12 \\ 4 & & -2 & 25 \end{array}\right] \\ \longrightarrow \left[\begin{array}{cc|cc} 1 & & 11 & -137 \\ 4 & & -2 & 25 \end{array}\right] \longrightarrow \left[\begin{array}{cc|cc} 1 & & 11 & -137 \\ 0 & & -46 & 573 \end{array}\right] \end{gathered} \tag{8.6}$$

In first step we subtracted 12 times the bottom row from the top row. We then subtract 2 times the top row from the bottom row, and so on until the final form is achieved. The row multipliers 12, 2, 5, and 4 are precisely the quotients we found via Euclid's algorithm in (6.2), and also appear as partial quotients in the continued fraction (8.1).

From the top row of the final matrix in (8.6) we conclude that $x = 11$ and $y = -137$, so that

$$573 \cdot 11 \; + \; 46 \cdot (-137) \; = \; 1,$$

as we also discovered using the continued fraction (8.1). Notice that the final step of the row reduction was not really necessary to find x and y, but does reveal a striking symmetry: the bottom two entries of the right-hand square form a vector perpendicular to the vector $(573, 46)$. This is not a coincidence.

❧

Why does this matrix algorithm work?

In order to solve the equation

$$ax + by = \gcd(a, b),$$

begin with the matrix algebra identity

$$\begin{bmatrix} a & 1 & 0 \\ b & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{8.7}$$

This matrix identity is equivalent to the assertion that

$$\begin{aligned} au + 1v + 0w &= 0 \\ bu + 0v + 1w &= 0 \end{aligned} \tag{8.8}$$

has a solution where $u = -1$, $v = a$, and $w = b$. Recall that if we add or subtract one of these equations from another (any number of times), the solutions (u, v, w) to the system of equations *does not change*. Adding one equation to another in (8.8) corresponds in (8.7) to adding one row of the 2×3 matrix to another.

With this in mind, suppose that $a \geq b$, and perform a division with remainder to obtain $a = bq + r$, where $0 \leq r < b$. If we subtract the bottom row of the 2×3 matrix from the top row q times, the matrix identity (8.7) becomes

$$\begin{bmatrix} r & 1 & -q \\ b & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Continuing in this manner, we perform Euclid's algorithm on the entries of the first column of the 2×3 matrix until that matrix has the final form

$$\begin{bmatrix} \gcd(a,b) & x & y \\ 0 & x' & y' \end{bmatrix}$$

where x, x', y, and y' represent whatever numbers appear in the remaining entries after the process is complete. The matrix identity (8.7) now becomes

$$\begin{bmatrix} \gcd(a,b) & x & y \\ 0 & x' & y' \end{bmatrix} \begin{bmatrix} -1 \\ a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tag{8.9}$$

and the corresponding equations (8.8) become

$$\begin{aligned} -\gcd(a,b) + ax + by &= 0 \\ ax' + by' &= 0 \end{aligned} \quad \Longleftrightarrow \quad \begin{aligned} ax + by &= \gcd(a,b) \\ ax' + by' &= 0 \end{aligned}$$

In other words, the values (x, y) appearing in (8.9) solve the Diophantine equation $ax + by = \gcd(a, b)$.

❧

Exercise 8.1. Use Euclid's algorithm to show that $\gcd(130, 88) = 2$, as asserted in the example of this section.

Exercise 8.2. Use Euclid's algorithm on the pair $(65, 44)$ to verify that the continued fraction shown in the identity (8.2) is correct.

Exercise 8.3. Use Euclid's algorithm to show that $\gcd(130, 78) = 26$, as asserted in the example of this section.

Exercise 8.4. Find all solutions $x, y \in \mathbb{Z}$ such that $51x + 27y = 15$.

Exercise 8.5. Find all solutions $x, y \in \mathbb{Z}$ to $130x + 88y = 12006$, where x and y are both *positive*.

Exercise 8.6.
(a) Find all $x, y \in \mathbb{Z}$ such that $4x + 13y = 200$.
(b) Find all $x, y \in \mathbb{N}$ such that $4x + 13y = 200$.
(c) Find all $x, y \in \mathbb{N}$ such that $12x + 39y = 600$.
(d) Find all $x, y \in \mathbb{N}$ such that $12x + 39y = 400$.

Exercise 8.7. Find all integer solutions to the equation $19x + 30y = 1$.

Exercise 8.8. Find all *positive* integer solutions to the equation $19x + 30y = 800$.

Exercise 8.9. Find $a, b \in \mathbb{N}$ such that

$$\frac{a}{14} + \frac{b}{9} = \frac{101}{126}.$$

Exercise 8.10. A toy store sells small teddy bears for \$5 and large teddy bears for \$8. I walk into the store with n dollars, but find that no matter how many teddy bears I choose to buy, I will have some money left over. What is the largest possible value for n?

Exercise 8.11. Prove that there are no integer solutions x, y, z to the equation

$$15x - 27y + 42z = 904.$$

Exercise 8.12. Without finding actual specific values for x, y, z, prove that the equation

$$120x + 102y + 425z = 1999,$$

has integer solutions x, y, z.

Exercise 8.13. Find integer solutions x, y, z to the equation

$$12x + 70y + 65z = 1.$$

Exercise 8.14. A toy store sells small dolls for \$6, medium-sized dolls for \$10, and large dolls for \$15. Suppose that I have \$425 exactly.
(a) Show that if I spend all my money then I must have bought an odd number of large dolls.
(b) If I buy exactly 7 small dolls and as many of the others as I can, how much money will I have left over? (Assume I have spent as much money as possible.)
(c) Suppose instead that I walk into the store with n dollars, but find that no matter how many dolls I choose to buy, I will have some money left over. What is the largest possible value for n?

Exercise 8.15. A candy store sells jelly beans for 6 cents each and gum balls for 25 cents each. If I spend exactly 10 dollars in the store, what the largest total number of candies I could buy?

Exercise 8.16. Use the matrix row reduction method to solve the Diophantine equations:

(a) $22x + 17y = 1$. **(b)** $576x + 84y = 12$.

Exercise 8.17. Adapt the matrix row reduction method to solve the Diophantine equations:

(a) $25x + 9y = 4$. **(b)** $28x + 26y + 91z = 17$.

Hint: For part (b) set up and row reduce a suitable 3×4 matrix. It is permissible to multiply any row by a non-zero integer if this is helpful, but non-integer fractions should never appear at any step in your computations.

Exercise 8.18. Adapt the matrix row reduction method to solve the *system of simultaneous* Diophantine equations:

$$3x + 4y + 7z = 5$$
$$2x + y - 8z = 1$$

Hint: Set up and row reduce a suitable 3×5 matrix. It is permissible to multiply any row by a non-zero integer if this is helpful, but non-integer fractions should never appear at any step in your computations.

9 Prime numbers and unique factorization

A positive integer $n > 1$ is *prime* if it has no positive factors except itself and 1. The number 1 is called a *unit*.[1] All other positive integers are said to be *composite*.

For example, 2, 3, 5, and 7 are prime, while 4, 6, 8, and 9 are composite.

❧

A negative integer n is said to be prime (respectively unit, composite) if $|n|$ is prime (respectively unit, composite).

For example, -3 is prime, since $|-3| = 3$ is a prime positive integer.

Zero is a composite number.

❧

Proposition 9.1. *Every integer $n > 1$ is divisible by at least one prime number.*

Proof. We use the well-ordering principle. Let S be the set of natural numbers greater than 1 that have no prime factors. We will show that S is the empty set.

If S is not empty, then S has a minimal element m. Since $m|m$ and has no prime factors, m cannot be prime. Therefore, m is composite, so that $m = ab$, where $1 < a < m$ and $1 < b < m$. Since $a < m$ and m is the minimum of S, it follows that $a \notin S$. From the definition of S it then follows that a has a prime factor p. Since $p|a$ and $a|m$, we also have $p|m$, contradicting the fact that $m \in S$. Therefore, the set S has no such minimum and must be empty. ❑

One way to determine if a number n is prime is to try division by the numbers $2, 3, 4, \ldots, n-1$. If every division fails (leaves a non-zero remainder) then n has no factors besides itself and 1, so that n is prime. This approach certainly works, but is also quite tedious, since we must perform $n-2$ divisions. If $n > 10^{12}$, for example, this could take a long time, even on a computer. One easy improvement is to notice that if $d|n$ and $d \neq n$ then $d \leq \frac{n}{2}$. This halves the number of trials we need to make. Since a composite number must have a proper *prime* factor, it is also acceptable if we skip trial divisions by even numbers greater than 2. This reduces the number of necessary divisions to

[1] The number 1 is singled out in this way, because 1 is the *only* positive integer whose reciprocal is also an integer. For the same reason, the units in $\mathbb{Z}$ consist of the numbers 1 and -1.

approximately $\frac{n}{4}$. While additional similar considerations (avoiding successive multiples of 3 and 5, for example) reduce the necessary labor a bit more, the prospect of verifying primality for the number

$$n = 837514498321$$

is still daunting. For example, by testing divisibility only by 2, by 3, and by subsequent numbers of the form $6k+1$ and $6k+5$ (thereby avoiding even numbers and numbers divisible by 3) up to $\frac{n}{3}$, we are still required to test on the order of

$$\frac{837514498321}{9} \approx 10^{11}$$

possible cases!

The following theorem is a big help.

Theorem 9.2. *Let $n > 1$ be a positive integer. Either n is prime, or there exists a prime number $p \leq \sqrt{n}$ such that $p|n$.*

Proof. Suppose that n is composite, so that $n = ab$ where $1 < a \leq b < n$. If $a > \sqrt{n}$ then $b \geq a > \sqrt{n}$ so that

$$n = ab > \sqrt{n}\sqrt{n} = n,$$

a contradiction. It follows that $1 < a \leq \sqrt{n}$. Since a has a prime factor $p \leq a$, it follows that $p|n$ and $p \leq \sqrt{n}$. ❑

Theorem 9.2 implies that, if the number 837514498321 is composite, it has a prime factor

$$p \leq \lfloor \sqrt{837514498321} \rfloor = 915158.$$

A short program on a personal computer can now verify primality in a few seconds.

While Theorem 9.2 provides a dramatically improved algorithm for primality testing, modern cryptographic applications (which may use prime numbers that are hundreds of digits long) require even faster methods. We will return to this topic in Section 31.

❧

Proposition 9.1 is really only the beginning. In fact, we can prove a much stronger result.

Theorem 9.3 (The Fundamental Theorem of Arithmetic). *If $n > 1$ then there are unique primes $p_1 < \cdots < p_k$ and integers $a_1, \ldots, a_k > 0$ such that*

$$n = p_1^{a_1} \cdots p_k^{a_k}.$$

The important point is not only that n factors into primes, but that it does so in *exactly one way* (uniqueness). Before we prove Theorem 9.3, here is a fundamental Lemma, a special case of Theorem 7.2:

Lemma 9.4. *Let p be a prime integer. If $p|ab$ then $p|a$ or $p|b$.*

More generally, if $p|b_1 b_2 \cdots b_k$ then $p|b_i$ for some i.

Proof. Suppose that $p|ab$. Recall that the only divisors of p are 1 and p. If $p \nmid a$, then $\gcd(p, a) = 1$. It follows from Theorem 7.2 that $p|b$.

The more general case follows by induction on $k \geq 2$. ❑

We are now ready to prove the Fundamental Theorem of Arithmetic.

Proof of Theorem 9.3. Let S be the set of integers $n > 1$ that violate Theorem 9.3. If S is not empty, then S has a minimal element m.

If m is prime, then m is its own prime factorization, and can be factored no other way, which contradicts $m \in S$.

If m is composite, then $m = st$ for integers s, t, such that $1 < s < m$ and $1 < t < m$. By the minimality of m, both s and t satisfy Theorem 9.3 and are both products of primes. Therefore, $m = st$ is also a product of primes.

We have shown that every positive integer except 1 is a product of primes. The only way the minimal element $m \in S$ can violate the theorem is by having two or more *different* prime factorizations. Suppose this happens.[2] This means that

$$m = p_1^{a_1} \cdots p_k^{a_k} = q_1^{b_1} \cdots q_s^{b_s}, \tag{9.1}$$

where $p_1 < \cdots < p_k$ and $q_1 < \cdots < q_s$ are primes. Since p_1 is prime and $p_1|m$, it follows from Lemma 9.4 that $p_1|q_j$ for some j. Since p_1 and q_j are both prime, it follows that $p_1 = q_j$. Let $N = m/p_1$, so that

$$N = p_1^{a_1 - 1} \cdots p_k^{a_k} = q_1^{b_1} \cdots q_j^{b_j - 1} \cdots q_s^{b_s}.$$

Since $1 < N < m$, the minimality of m in S implies that these factorizations of N into prime powers must be identical, so that $k = s$, and each $p_i = q_i$, and each $a_i = b_i$. But this also implies that the factorizations of m in (9.1) are identical, contradicting our original choice. It follows that S is the empty set, so that every $n > 1$ has a unique prime factorization. ❑

If n is a non-zero integer and p is prime, we write

$$p^k || n$$

[2] Recall the general paradigm for proving uniqueness in mathematics: suppose that there are two, and prove that those two must be equal.

when $p^k|n$ and $p^{k+1} \nmid n$. In other words, this notation indicates that k is the highest power of p that divides n. For example, $5^1||40$ and $2^3||40$, since $40 = 8 \cdot 5 = 2^3 \cdot 5^1$.

❧

Theorem 9.3 provides an easy way to express the gcd and lcm of two numbers $n, m > 1$ once we have their factorizations into primes. Specifically, if

$$n = p_1^{a_1} \cdots p_k^{a_k} \quad \text{and} \quad m = p_1^{b_1} \cdots p_k^{b_k},$$

where each $a_i, b_i \geq 0$, then

$$\gcd(n, m) = p_1^{\min\{a_1, b_1\}} \cdots p_k^{\min\{a_k, b_k\}}$$

and

$$\operatorname{lcm}(m, n) = p_1^{\max\{a_1, b_1\}} \cdots p_k^{\max\{a_k, b_k\}}.$$

In practice, however, it is not computationally feasible to factor a large integer into prime powers. For this reason, the gcd (and lcm) are typically obtained using Euclid's algorithm, which does not require any information about the factorizations of n and m.

❧

Given that every integer n is a product of primes, it would make sense (from a number theoretic perspective) to enumerate these basic building blocks of the integers as best we can. The first observation along these lines, made by Euclid[3] over 2000 years ago, is that the list of primes is infinite.

Theorem 9.5. *There are infinitely many distinct prime integers.*

Proof. Suppose that there are only a finite number of prime integers, say

$$p_1 < p_2 < \cdots < p_m,$$

where m is the total number of primes. Let $N = (p_1 p_2 \cdots p_m) + 1$. Since $N > 1$ there is a prime p such that $p|N$, so that $N = pk$ for some integer $k \geq 1$. Since p is prime, we have $p = p_i$ for some p_i in the complete list above. It follows that $N = p_i k$, so that

$$1 \;=\; N - p_1 p_2 \cdots p_i \cdots p_m \;=\; p_i[k - (p_1 \cdots p_{i-1} p_{i+1} \cdots p_m)].$$

In other words, $p_i|1$, which is impossible. It follows that p cannot be on the list of primes above, and no such finite list is ever complete. ❑

[3] Euclid was a Greek mathematician active during the 3rd century BC.

Notice that every prime $p > 2$ is odd,[4] so that every such prime has the form $4n + 1$ or $4n + 3$. A small variation of the previous proof can be used to show that there are infinitely many primes of the form $4n + 3$, as well as other special forms (see Exercises 9.13 and 9.14). A slightly more complicated variation shows that there are also infinitely many primes of the form $4n + 1$ (see Exercise 22.15).

Dirichlet's theorem[5] asserts that, for every relatively prime pair $a, b \in \mathbb{Z}$, the arithmetic progression

$$\{a + tb \mid t = 0, 1, 2, 3, \ldots\}$$

contains an infinite number of primes. A proof of this deep and difficult theorem can be found in [3]. However, the special case of Dirichlet's Theorem in which $a = 1$ and b is a prime has an elementary proof (see Exercise 20.17).

❧

Here is a list of the positive integer primes less than 1000:

2	3	5	7	11	13	17	19	23	29
31	37	41	43	47	53	59	61	67	71
73	79	83	89	97	101	103	107	109	113
127	131	137	139	149	151	157	163	167	173
179	181	191	193	197	199	211	223	227	229
233	239	241	251	257	263	269	271	277	281
283	293	307	311	313	317	331	337	347	349
353	359	367	373	379	383	389	397	401	409
419	421	431	433	439	443	449	457	461	463
467	479	487	491	499	503	509	521	523	541
547	557	563	569	571	577	587	593	599	601
607	613	617	619	631	641	643	647	653	659
661	673	677	683	691	701	709	719	727	733
739	743	751	757	761	769	773	787	797	809
811	821	823	827	829	839	853	857	859	863
877	881	883	887	907	911	919	929	937	941
947	953	967	971	977	983	991	997		

Notice that primes often come in pairs $p, p + 2$. These are called *twin primes*. Examples include

$$\{3, 5\} \quad \{5, 7\} \quad \{11, 13\} \quad \{17, 19\} \quad \{41, 43\} \quad \{857, 859\}.$$

The *Twin Primes Conjecture* asserts that there are infinitely many twin prime pairs. Although there has been substantial progress on this question in recent

[4] It has been pointed out that, since 2 is the only prime number that is also an even number, this makes 2 a very odd prime indeed!

[5] Discovered by Peter Gustav Lejeune Dirichlet (1805–1859).

years,[6] the conjecture remains open.

Cousin primes are prime pairs of the form $p, p+4$, and *sexy primes* are prime pairs of the form $p, p+6$.

❧

Here is another simple assertion that, even after centuries of study, no one has been able to verify or disprove.

Goldbach's conjecture: Every even number greater than 2 is the sum of two primes.

Settle this conjecture, and become world-famous overnight!

❧

Mersenne primes are primes of the form $p = 2^n - 1$ for some integer n. Examples include

$$3 = 2^2 - 1 \qquad 7 = 2^3 - 1 \qquad 31 = 2^5 - 1 \qquad 127 = 2^7 - 1.$$

It is not difficult to show (using algebraic identities) that n must be prime in order for p to be prime (see Exercise 9.32). On the other hand, this provides no guarantee. For example,

$$2^{11} - 1 = 2047 = 23 \cdot 89.$$

The largest known Mersenne prime is

$$2^{77,232,917} - 1,$$

the 50th discovered so far, having 23,249,425 digits.[7] It is not known if there are infinitely many Mersenne primes.

❧

Fermat primes are primes of the form $p = 2^{2^n} + 1$ for some integer n. Examples include

$$3 = 2^{2^0} + 1 \qquad 5 = 2^{2^1} + 1 \qquad 17 = 2^{2^2} + 1 \qquad 257 = 2^{2^3} + 1 \qquad 65537 = 2^{2^4} + 1.$$

However,

$$2^{2^5} + 1 \;=\; 2^{32} + 1 \;=\; 4294967297 \;=\; 6700417 \cdot 641,$$

[6] In 2014 it was shown that there are infinitely many pairs of primes that differ by some positive integer N, where $N \leq 246$. For the twin primes conjecture to be true, this assertion must hold for $N = 2$.

[7] As of January 2018.

is not prime. In fact, no other Fermat primes have been found.

❧

Once the infinitude of primes has been established, it is natural to ask how the primes are distributed. For example, Bertrand's postulate (later proved by Chebyshev)[8] asserts that, if $n > 1$, then there is at least one prime number p such that $n < p < 2n$.

Legendre made a related conjecture: If $n > 1$ then there is at least one prime number p such that $n^2 < p < (n+1)^2$. This conjecture remains open!

More generally still, define $\pi(n)$ to be the number of positive prime integers $p \leq n$. For example, $\pi(10) = 4$, since there are 4 primes $p \leq 10$, namely $2, 3, 5, 7$. Unfortunately no simple formula has been found for $\pi(n)$. However, some asymptotic results are known. For example, it follows from Theorem 9.5 that $\lim_{n\to\infty} \pi(n) = \infty$. This leads to the question of *how fast* $\pi(n)$ grows. The *Prime Number Theorem* answers this question in part with the assertion that

$$\lim_{n\to\infty} \frac{\pi(n)}{\frac{n}{\ln n}} = 1.$$

In other words, for large values of n, we have the approximation:

$$\pi(n) \approx \frac{n}{\ln n}.$$

Said differently,

$$\frac{\pi(n)}{n} \approx \frac{1}{\ln n}$$

for large values of n. This means that, if n is a large positive integer, and if a random number is chosen from the list $\{1, 2, \ldots, n\}$, the probability of picking a prime number is about $\frac{1}{\ln n}$.

The proofs of these and other distribution theorems for primes can be found in [3]. A better understanding of how primes are distributed over the integers remains a principal goal of modern number theory research.

❧

Exercise 9.1. Prove that there are infinitely many composite numbers.

Exercise 9.2. Suppose that p is an odd prime. What remainder can occur if p is divided by 4? By 6? By 8?

Exercise 9.3. Prove that there is a positive integer N such that $N, N+1, \ldots, N+999$ are all composite.

Exercise 9.4. Express each of the following integers as a product of prime powers:

[8] Pafnuty Lvovich Chebyshev (1821–1894) was a Russian mathematician also known for his work in modern analysis and probability theory.

(a) 99
(b) 101
(c) 343
(d) 999
(e) 1331
(f) 1001
(g) 999999
(h) 2560000
(i) 10!

Exercise 9.5.
(a) Which integers $n > 0$ have exactly 2 positive divisors?
(b) Which integers $n > 0$ have exactly 3 positive divisors?
(c) Which integers $n > 0$ have exactly 4 positive divisors?

Exercise 9.6.
(a) If 3000! is written out in decimal digits, how many zeroes lie at the end?
(b) If $2^k || 3000!$ then what is k?

Recall that, for non-negative integers n and k the (n,k)-binomial coefficient is given by

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Exercise 9.7.
(a) Prove that, if p is prime and $1 \leq k \leq p-1$, then $\binom{p}{k}$ is divisible by p.
(b) Prove that, if $n > 2$ is even, then $\binom{n}{2}$ is *not* divisible by n.

Exercise 9.8. Prove that if $a^3 | b^2$ then $a | b$.

Exercise 9.9. Use Theorem 9.3 to give another proof that $\gcd(a,b) \cdot \operatorname{lcm}(a,b) = ab$.

Exercise 9.10. Suppose that $a, b \in \mathbb{N}$. When is $\operatorname{lcm}(a,b)$ a prime number?

Exercise 9.11. What are the possible values of $\operatorname{lcm}(n!, n+1)$?

Exercise 9.12. Suppose that $a, b \in \mathbb{N}$ and that $\gcd(a,b) = 1$.
What are the possible values of $\operatorname{lcm}(a+b, a-b)$?

Exercise 9.13. Prove that there are infinitely many prime numbers of the form $4n+3$.
Hint: Suppose there are finitely many. Omitting 3, multiply the rest of them together to obtain an integer N. Then consider the number $4N+3$. Can all of its prime factors be of the form $4n+1$?

Exercise 9.14. Prove that there are infinitely many prime numbers of the form $3n+2$.
Hint: Adapt the hint for Exercise 9.13.

Exercise 9.15. Prove that there are infinitely many prime numbers of the form $6n+5$.

Exercise 9.16. Prove that if p is prime then $\sqrt{p}$ is irrational.

A positive integer n is said to be a *perfect square* if $n = k^2$ for some integer k. An integer n is *square-free* if n is not divisible by any perfect square other than $1^2 = 1$.

Exercise 9.17. Prove that if $n > 1$ is square-free, then the number of positive divisors of n must be a power of 2.

Exercise 9.18. Suppose that n is a perfect square. Prove that if p is prime and $p^k || n$, then k is even.

Exercise 9.19. Use Bertrand's postulate to prove that, if $n > 1$, then $n!$ is never a perfect square.
Hint: Let p be the largest prime such that $p \leq n$. How many times does p divide $n!$?

Exercise 9.20. Suppose that $n \in \mathbb{N}$ has the property that whenever p is prime and $p|n$, we have $p^2|n$. Prove or disprove the assertion that n must be a perfect square.

Exercise 9.21. Suppose that a positive integer n is a ***not*** a perfect square. Prove that $\sqrt{n}$ is irrational.
Hint: First prove that some prime p divides n an odd number of times. Then adapt the proof that $\sqrt{2}$ is irrational, making use of p where necessary.

Exercise 9.22. Adapt the previous exercise to state and prove a theorem about the kth root of an integer, where $k \in \{3, 4, 5 \ldots\}$.

Exercise 9.23. Suppose that a and b are positive integers with $b > 1$, such that $\gcd(a, b) = 1$. Prove that $10^{\frac{a}{b}}$ is irrational.

Exercise 9.24. Prove that $\log_{10} 2$ is irrational.

Exercise 9.25. Suppose that $n > 1$ is an integer. Prove that $\log_{10} n$ is rational if and only if $n = 10^k$ for some positive integer k.

Exercise 9.26. Show that there are no triplets of twin primes $p, p+2, p+4$ except $3, 5, 7$.

Exercise 9.27. Find all triplets of primes p, q, r such that

$$r - q, \quad r - p, \quad q - p$$

are all primes.

Exercise 9.28. Show that there are no triplets of cousin primes $p, p+4, p+8$ except $3, 7, 11$.

Exercise 9.29. Show that there are no 5-tuples of sexy primes $p, p+6, p+12, p+18, p+24$ except $5, 11, 17, 23, 29$.

Exercise 9.30. Prove that Goldbach is correct if and only if every odd number greater than 5 is the sum of three primes.

Exercise 9.31. Prove that every integer greater than 11 is the sum of two distinct positive composite numbers.

Exercise 9.32. Suppose that $a, n \in \mathbb{N}$, where $n > 1$. Prove that if $a^n - 1$ is a prime number, then $a = 2$ and n is prime.
Hint: The geometric sum formula is helpful here.

Exercise 9.33. Prove that if k is not a power of 2 then $2^k + 1$ is composite.
Hint: Try factoring $x^k + 1$.

Exercise 9.34. Suppose that n is an integer.
(a) For what values of n is $n^2 - 1$ prime?
(b) For what values of n is $n^2 - 2n - 3$ prime?

Exercise 9.35. Prove that if $n > 1$ then $n^5 + n^4 + 1$ is composite.

Exercise 9.36. Prove that if $n > 1$ then $1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}$ is not an integer.
Hint: Count the twos.

10 Modular arithmetic

Modular arithmetic is an arithmetic of remainders. The idea is to replace integers with their remainders after division by some fixed value m. The next definition provides a useful way to express this idea.

Definition 10.1. Let $m > 1$ be an integer. For $a, b \in \mathbb{Z}$ we say that

$$b \equiv a \mod m$$

if $m|(b-a)$.

In plain English, we say: "b is congruent to a modulo m." The term *modulo* is Latin for *by the modulus* or *by the small measure.*

Observe that $a \equiv 0 \mod m$ if and only if $m|a$.

For example, we have

$$\begin{aligned} 100 &\equiv 0 \mod 4 \\ 19 &\equiv 5 \mod 7 \\ 100 &\equiv 9 \mod 13 \\ -10 &\equiv 2 \mod 3 \end{aligned}$$

❧

Notice that $n \equiv 0 \mod 2$ whenever n is even, and that $n \equiv 1 \mod 2$ whenever n is odd.

❧

If $m \neq 0$ and $b \in \mathbb{Z}$, then we can apply division with remainder (Theorem 4.2) to write

$$b = mq + r,$$

where $0 \leq r < m$. Moreover, Theorem 4.2 tells us that r is unique in the list $\{0, \ldots, m-1\}$. In the language of modular arithmetic, we can rewrite Theorem 4.2 as follows.

Theorem 10.2 (Division with remainder (modular arithmetic version))**.**
Let $m, b \in \mathbb{Z}$, where $m > 1$. There exists a unique $r \in \{0, 1, \ldots, m-1\}$ such that

$$b \equiv r \mod m.$$

For example, every integer n is congruent to $0, 1$, or 2 modulo 3, depending on its remainder after division by 3.

❧

Congruence mod m is an equivalence relation. More specifically, we have the following.

Proposition 10.3. *For $a, b, c \in \mathbb{Z}$ and an integer $m > 1$,*

- $a \equiv a \bmod m$.
- *If* $a \equiv b \bmod m$ *then* $b \equiv a \bmod m$.
- *If* $a \equiv b \bmod m$ *and* $b \equiv c \bmod m$ *then* $a \equiv c \bmod m$.

The proof is an exercise. (See Exercise 10.2.)

Congruence mod m is also compatible with addition and multiplication. This is expressed more precisely by the following theorem.

Theorem 10.4. *Let $a, a', b, b' \in \mathbb{Z}$, and let $m > 1$ be an integer. If $a \equiv a' \bmod m$ and $b \equiv b' \bmod m$, then*

- $a + b \equiv a' + b' \mod m$.
- $a - b \equiv a' - b' \mod m$.
- $ab \equiv a'b' \mod m$.

Note well the absence of division in Theorem 10.4. Division in a modulus is a more complicated matter, to be addressed in detail later on.

Proof of Theorem 10.4. The proof that $a \pm b \equiv a' \pm b' \bmod m$ is left as an exercise (see Exercise 10.3).

To prove that $ab \equiv a'b' \mod m$, observe that, since $a \equiv a'$ and $b \equiv b'$ we have $m|(a - a')$, and $m|(b - b')$. In other words, there exist integers k, l such that

$$a - a' = mk \quad \text{and } b - b' = ml.$$

It follows that

$$\begin{aligned} ab - a'b' &= ab - ab' + ab' - a'b' \\ &= a(b - b') + (a - a')b' \\ &= aml + mkb' = m(al + kb'), \end{aligned}$$

so that $m|(ab - a'b')$. ❑

Theorem 10.4 is a helpful tool for faster computation. For example, since $100 \equiv 0 \mod 4$, it follows that

$$73587459 \;=\; 735874 \cdot 100 + 59 \;\equiv\; 735874 \cdot 0 + 59 \;\equiv\; 59 \;\equiv\; 3 \mod 4.$$

Continuing with similar reasoning, we have

$$(73587459)^5 - 4781 \cdot 601 \;\equiv\; 3^5 - 1 \cdot 1 \;\equiv\; 243 - 1 \;\equiv\; 3 - 1 \;\equiv\; 2 \mod 4.$$

This sort of computation is much easier than explicitly computing $(73587459)^5$ and then dividing by 4, etc.

❧

Large exponents are more easily computed by the method of *repeated squares*. To illustrate, suppose we want to compute $3^{58} \bmod 19$. Observe that 58 has the binary expression $(111010)_2$; that is,

$$58 = 2 + 8 + 16 + 32.$$

By repeated squaring, we can quickly compile a list of exponentials to powers of 2 mod 19:

$$\begin{aligned} 3^2 &\equiv 9 \mod 19 \\ 3^4 &\equiv 81 \equiv 5 \mod 19 \\ 3^8 &\equiv 25 \equiv 6 \mod 19 \\ 3^{16} &\equiv 36 \equiv -2 \mod 19 \\ 3^{32} &\equiv 4 \mod 19 \end{aligned}$$

It now follows that

$$3^{58} \;\equiv\; 3^2 \cdot 3^8 \cdot 3^{16} \cdot 3^{32} \;\equiv\; 9 \cdot 6 \cdot (-2) \cdot 4 \;\equiv\; 5 \mod 19.$$

This method is much faster than performing 57 multiplications by 3.

❧

For $k = 0, 1, \ldots, m-1$, let

$$\langle k \rangle_m = \{n \in \mathbb{Z} \mid n \equiv k \mod m\}.$$

In other words, the sets $\langle 0 \rangle_m, \langle 1 \rangle_m, \ldots, \langle m-1 \rangle_m$ are the m equivalence classes of integers modulo m. Every integer $n \in \mathbb{Z}$ is an element of exactly one $\langle k \rangle_m$, determined by the remainder k upon division of n by m. For example,

$$\langle 3 \rangle_m = \{\ldots,\; 3-2m,\; 3-m,\; 3,\; 3+m,\; 3+2m,\; 3+3m,\; 3+4m, \ldots\}.$$

For a more specific example, if $m = 7$ then

$$\langle 3\rangle_7 = \{\ldots, -11, -4, 3, 10, 17, 24, 31, \ldots\}.$$

Denote by $\mathbb{Z}_m$ the set of equivalence classes of integers mod m. That is,

$$\mathbb{Z}_m = \{\langle 0\rangle_m, \langle 1\rangle_m, \ldots, \langle m-1\rangle_m\}.$$

We have seen that if $a \in \langle k\rangle_m$ and $b \in \langle l\rangle_m$ then

$$a + b \in \langle k + l\rangle_m \quad \text{and } ab \in \langle kl\rangle_m,$$

by Theorem 10.4. This allows us to define an arithmetic on the set $\mathbb{Z}_m$ itself, by defining

$$\langle k\rangle_m + \langle l\rangle_m = \langle k + l\rangle_m \quad \text{and} \quad \langle k\rangle_m \langle l\rangle_m = \langle kl\rangle_m.$$

This arithmetic on equivalence classes mod m is called *modular arithmetic*.

For simplicity of notation, when doing arithmetic mod m we dispense with the $\langle k\rangle_m$ notation and simply write "k mod m". To help remember that we are doing modular arithmetic (rather than ordinary integer arithmetic) we also typically replace the equal sign '$=$' with the equivalence symbol '$\equiv$'. For example, if we are working with arithmetic mod 7, the cumbersome notation

$$\langle 5\rangle_7 + \langle 4\rangle_7 \;=\; \langle 2\rangle_7$$

is replaced with the simpler expression

$$5 + 4 \;\equiv\; 2 \mod 7.$$

In practice, arithmetic mod m is the result of doing ordinary arithmetic of integers, along with the added property that the number m is equivalent to zero (along with whatever consequences that leads to).

❧

So far we have carefully avoided division in the context of modular arithmetic. Indeed, a naive approach to division in modular arithmetic can lead to trouble. For example,

$$2 \cdot 8 \equiv 2 \cdot 2 \mod 12$$

but

$$8 \not\equiv 2 \mod 12.$$

Moreover,

$$2 \cdot 6 \equiv 0 \mod 12,$$

while $2 \not\equiv 0 \mod 12$ and $6 \not\equiv 0 \mod 12$. In other words, it is possible for the product of non-zero values to be zero mod 12. Examples of this phenomenon occur in every composite modulus.

This disturbing state of affairs is mitigated in part by the following.

Theorem 10.5 (Cancellation Law).
If $ac \equiv bc \mod m$ *and* $\gcd(m,c) = 1$, *then* $a \equiv b \mod m$.

Proof. If $ac \equiv bc \mod m$, then $m|(ac - bc)$, so that $m|(a-b)c$. Since $\gcd(m,c) = 1$, it follows from Theorem 7.2 that $m|(a-b)$, so that $a \equiv b \mod m$. ❑

The cancellation law suggests that being relatively prime to the modulus m plays a role in division. To make this idea even more precise, let's reformulate a special case of Theorem 6.3 in the language of modular arithmetic.

Theorem 10.6. *The equation*

$$ax \equiv b \mod m$$

has a solution if and only if $\gcd(a,m)|b$.

Proof. Let $d = \gcd(a,m)$. The equation $ax \equiv b \mod m$ is equivalent to the assertion that

$$ax + my = b$$

for some integers x, y. These integers exist if and only if $d|b$, by Theorem 8.1. ❑

The next corollary emphasizes an important special case.

Corollary 10.7. *The equation*

$$ax \equiv 1 \mod m$$

has a solution if and only if $\gcd(a,m) = 1$.

In this case we denote $x = a^{-1}$ and say that a is a *unit* mod m, with inverse a^{-1}.

Corollary 10.8. *If* p *is prime, and if* $p \nmid n$, *then there exists* $a \in \{1, 2, \ldots, p-1\}$ *such that*

$$an \equiv 1 \mod p.$$

Corollary 10.8 is an immediate consequence of Corollary 10.7, since $\gcd(p,n) = 1$ whenever $p \nmid n$. It is worthy of special note, because it tells us that *every* non-zero value modulo a prime p is a unit mod p.

❧

Sometimes the easiest way to find inverses or to solve equations mod m is by astute guessing:

Example: $3^{-1} \equiv 2 \mod 5$, since $3 \cdot 2 \equiv 1 \mod 5$. Similarly $2^{-1} \equiv 3 \mod 5$.

Example: To solve $8x \equiv 3 \bmod 11$, first notice that $7 \cdot 8 = 56 \equiv 1 \bmod 11$. It follows that $x \equiv 7 \cdot 3 \equiv 21 \equiv 10 \bmod 11$.

Example: The equation $24x \equiv 2 \bmod 243$ has no solution, since $3|24$ and $3|243$, while $3 \nmid 2$.

However, sometimes it's best to use an algorithm:

Example: To solve $25x \equiv 2 \bmod 243$, first use Euclid's algorithm to find a, b such that $25a + 243b = 1$. This yields in the solution

$$25 \cdot 175 \ + \ 243 \cdot (-18) \ = \ 1,$$

so that

$$25 \cdot 175 \ \equiv \ 1 \mod 243.$$

It follows that $175 \equiv 25^{-1} \bmod 243$. We can now multiply both sides of the original equation by 175 to reveal that $x \equiv 2 \cdot 175 \equiv 350 \equiv 107$.

❧

If b is not a unit modulo m (that is, if b has no inverse mod m), then b is called a *zero divisor* mod m. The reason for this terminology is revealed in the next proposition.

Proposition 10.9. *Suppose that $m > 1$, and that $m \nmid b$. The following are equivalent.*

(i) *b is a zero divisor* mod m.

(ii) *b has no inverse* mod m.

(iii) *There exists $a \in \mathbb{Z}$, such that $m \nmid a$ and $ab \equiv 0$ mod m.*

Proof. We prove that (i) $\Leftrightarrow$ (ii) $\Leftrightarrow$ (iii).

The assertions of (i) and (ii) are equivalent by the definition of zero divisor.

Suppose (ii) holds, so that b has no inverse mod m. It then follows from Corollary 10.7 that $\gcd(b, m) = d > 1$. In other words, $m = ad$ and $b = sd$, where $d, a, s > 1$. Since $1 < a < m$, we have $a \not\equiv 0 \bmod m$. Moreover,

$$ab = asd = ms \equiv 0 \mod m,$$

so that (iii) follows.

Suppose that (iii) holds, so there exists $a \in \mathbb{Z}$, such that $m \nmid a$ and $ab \equiv 0 \bmod m$. If b has an inverse, then $bb^{-1} \equiv 1$, so that

$$0 \equiv 0 \cdot b^{-1} \equiv (ab)b^{-1} \equiv a(bb^{-1}) \equiv a \mod m,$$

contradicting $m \nmid a$. Therefore, b has no inverse, and (ii) follows. ❑

For example, 3 is a zero divisor mod 12, since $3 \cdot 4 \equiv 0 \bmod 12$, while $12 \nmid 4$.

The special case of when m is a prime is so important that we repeat (more or less) this special case with its own separate theorem.

Theorem 10.10. *Let p be prime. If $ab \equiv 0 \bmod p$, then either $a \equiv 0$ or $b \equiv 0 \bmod p$.*

In other words, there are no zero divisors in a prime modulus except 0.

❧

Exercise 10.1. Perform the following modular arithmetic computations. In each case, find the smallest positive integer that satisfies the given relation.

(a) $66577689 \mod 10$
(b) $66577689 \mod 100$
(c) $(4673)(536) + 77 \mod 3$
(d) $(-5553)(271) + 6^8 \mod 5$
(e) $6! \mod 3$
(f) $6! \mod 7$
(g) $6! \mod 10$
(h) $55799^{16} \mod 4$

Exercise 10.2. Use Definition 10.1 to prove Proposition 10.3.

Exercise 10.3. Prove the first two parts of Theorem 10.4: If $a \equiv a' \mod m$ and $b \equiv b' \mod m$, then

$$a + b \equiv a' + b' \mod m \quad \text{and} \qquad \text{and} \quad a - b \equiv a' - b' \mod m.$$

Exercise 10.4. Why don't we talk about mod 1?

Exercise 10.5. Explain why Theorem 10.10 follows from Proposition 10.9.

Exercise 10.6. Let $a, x, y, m \in \mathbb{Z}$, where $a \neq 0$ and $m > 1$.
(a) Prove that

$$ax \equiv ay \mod am \iff x \equiv y \mod m.$$

(b) Use part (a) to find all values of x mod 30 such that $12x \equiv 18 \bmod 30$.

Exercise 10.7. Prove that n is an odd number if and only if $n^2 \equiv 1 \bmod 4$.

Exercise 10.8. Prove that n is an odd number if and only if $n^2 \equiv 1 \bmod 8$.

Exercise 10.9. When is $n^2 \equiv 2 \bmod 3$?

Exercise 10.10. Prove that if n is an integer then $n^3 \equiv n \bmod 3$.

Exercise 10.11. Let $n \in \mathbb{Z}$.
(a) Prove that $n^3 \in \{0, 1, -1\}$ modulo 7.
(b) Prove that 70000000003 is not a perfect cube.
(c) Prove that 70000000003 can neither be written as a sum of 2 perfect cubes, nor as a difference of 2 perfect cubes.

Exercise 10.12. Prove there are no integers $x, y \in \mathbb{Z}$ such that $x^2 + 5y^3 = 2$.

Exercise 10.13. Find $a, m \in \mathbb{Z}$ so that $a^2 \equiv 0 \bmod m$, but $a \not\equiv 0 \bmod m$. Can you find an example where m is not a perfect square?

Exercise 10.14. Find $a, m \in \mathbb{Z}$ so that $a^3 \equiv 0 \mod m$, but $a^2 \not\equiv 0 \mod m$. Can you find an example where m is not a perfect cube?

Exercise 10.15. Let $n = 8^{800} + 5^{500} + 4^{400} + 3^{300}$.
(a) Prove that n is neither a perfect square nor a perfect cube.
(b) Can n be the kth power of an integer for any integer $k > 1$?

Exercise 10.16. By repeated squaring and reduction mod 23 compute the values of $5^2, 5^4, 5^8, 5^{16}$, $5^{32}, 5^{64} \mod 23$, and then use this information to compute $5^{83} \mod 23$.

Exercise 10.17. Suppose $n = a^2 + b^2$, where a and b are integers.
(a) What are the possible values of n mod 4?
(b) Can the number 20202043 ever be expressed as a sum of two integer squares? Why or why not?

Exercise 10.18. Suppose that $p \in \mathbb{N}$ is prime.
(a) Prove that if $n = 2^{p-1}(2^p - 1)$, then the decimal expression for n must end in the digit 6 or the digit 8.
(b) Prove that if $n = 2^{p-1}(2^p - 1)$, and if the decimal expression for n does not end in the digit 6, then it must end with the digits 28.

Exercise 10.19. Suppose that x and y are integers. Prove that, if $2x + 5y$ is divisible by 23, then $7x + 6y$ must also be divisible by 23.

Exercise 10.20. What are units mod 15? What are the zero divisors?

Exercise 10.21. For each of the following, solve for x if possible, or explain why no solution exists.

(a) $3x \equiv 1 \mod 16$
(b) $4x \equiv 3 \mod 19$
(c) $4x \equiv 3 \mod 18$
(d) $4x \equiv 2 \mod 18$
(e) $253x + 47 \equiv 900 \mod 7$
(f) $x^2 \equiv 3 \mod 13$
(g) $x^2 + 1 \equiv 0 \mod 2$
(h) $x^2 + 1 \equiv 0 \mod 3$

Exercise 10.22. Compute the first few powers of 10 mod 3. Use the results to explain why

$$8450326123897 \equiv 8 + 4 + 5 + 0 + 3 + 2 + 6 + 1 + 2 + 3 + 8 + 9 + 7 \mod 3.$$

What happens mod 9? mod 11?

Exercise 10.23. For each of the following, solve for x if possible, or explain why no solution exists.

(a) $100x \equiv 1 \mod 77$
(b) $25x \equiv 31 \mod 63$
(c) $18x \equiv 5 \mod 42$
(d) $18x \equiv 6 \mod 42$

Exercise 10.24. **(a)** Prove that, if p is a positive prime and $1 \leq k \leq p - 1$, then

$$\binom{p}{k} \equiv 0 \mod p.$$

(b) Show by example that (a) may no longer hold if p is not prime.
(c) Combine part (a) with the Binomial Theorem 3.2 to prove that, if p is prime, then

$$(x + y)^p \equiv x^p + y^p \mod p,$$

for all x and y.

Exercise 10.25. Prove that, if k is a positive integer, then

$$(2^k+1)^2 \equiv (2^k-1)^2 \equiv 1 \mod 2^{k+1}.$$

Exercise 10.26. Recall the Fibonacci sequence $\{F_n\}$ of Exercise 2.7. Suppose that $n \geq k \geq 3$.
(a) Prove that $F_n = F_k F_{n-k+1} + F_{k-1} F_{n-k}$.
(b) Combine the identity in part (a) with Exercise 6.8 to show that $F_n \equiv 0 \mod F_k$ if and only if $F_{n-k} \equiv 0 \mod F_k$.
(c) Use part (b) and division with remainder to prove that $F_k | F_n$ if and only if $k | n$.

Exercise 10.27. Can a power of two ever end in the digits "...324"?

Exercise 10.28. Prove that, if $n > 1$ is an integer, then $3^n + 1$ is not divisible 2^n.[†]

Exercise 10.29. For which $n \in \mathbb{N}$ is the number $n! + 5$ a perfect cube?

Exercise 10.30. For each of the following, solve for x if possible, or explain why no solution exists.

(a) $x^2 + x + 1 \equiv 0 \mod 2$
(b) $x^2 + x + 1 \equiv 0 \mod 3$
(c) $x^2 - x + 3 \equiv 0 \mod 5$
(d) $8x^2 - 6x + 43 \equiv 0 \mod 7$
(e) $x^2 - 1 \equiv 0 \mod 8$
(f) $9x^2 - 10 \equiv 0 \mod 8$

Exercise 10.31. For which integers n is $2^n + 5^n - 14$ a prime number?

Exercise 10.32. Let $m, n \in \mathbb{N}$, and suppose that $m > 2$.
(a) Prove that $3^m - 2^n \neq -1$.
(b) Prove that $3^m - 2^n \neq 1$.[*]

[†]From the 1911 Eötvös Competition [25].

[*]Parts (a) and (b) together can be stated as follows: the only instances of pairs $(2^n, 3^m)$ where 2^n and 3^m are within 1 of each other are the pairs $(1,1)$, $(2,1)$, $(2,3)$, $(4,3)$, and $(8,9)$. This theorem is attributed to Levi ben Gershon (1288–1344), also known as Gersonides, a medieval mathematician, astronomer, and Jewish philosopher. It is a special case of Catalan's conjecture: If $m, n \in \mathbb{N}$ with $m, n > 1$, then the equation $x^n - y^m = 1$ has no integer solutions besides $3^2 - 2^3 = 1$. Conjectured by Eugène Catalan (1814–1894) in 1844, this assertion was proved by Preda Mihăilescu in 2002 [18].

11 The Chinese remainder theorem

A farmer has a basket of apples. If he puts the apples into sacks of 8 apples each, there are 3 apples left over. If he puts the apples into crates of 25 apples each, there are 7 apples left over. How many apples does the farmer have?

A moment of thought reveals that this problem does not have a unique answer. On the other hand, there are certainly some stringent conditions being set on the number A of apples the farmer could have. In particular, we know that

$$A \equiv 3 \mod 8 \quad \text{and} \quad A \equiv 7 \mod 25.$$

Notice that $200 \equiv 0 \bmod 8$ as well as mod 25. This means that if A is a solution to the problem, then $A + 200$ is another solution, as is $A + 400$, $A + 600$, and so on. (If it were possible to have a negative number of apples, say by owing a debt of apples to the farmer next door, then $A - 200$ is also a solution.)

The previous observation suggests we look for an initial answer between 0 and 199. Enlightened trial and error leads to the possibility that the farmer has 107 apples. So maybe $A = 107$. If not, then at least we can say that

$$A \equiv 107 \mod 8 \quad \text{and} \quad A \equiv 107 \mod 25,$$

so that $A - 107$ is divisible by both 8 and 25. Since $\gcd(8, 25) = 1$, this means that $A - 107$ is divisible by $8 \cdot 25 = 200$. In other words, the number of apples A must be an integer such that

$$A \equiv 107 \mod 200.$$

In the absence of more information, this is the most we can know for certain about the number A.

This situation is summarized by the following theorem, historically attributed to the mathematicians of ancient China.

Theorem 11.1 (Chinese remainder theorem)**.** *If $m_1, \ldots, m_k$ are pairwise relatively prime,*[1] *and $a_1, \ldots, a_k \in \mathbb{Z}$, there is a unique value $x \mod (m_1 \cdots m_k)$ such that*

$$\begin{aligned} x &\equiv a_1 \mod m_1 \\ x &\equiv a_2 \mod m_2 \\ &\vdots \\ x &\equiv a_k \mod m_k \end{aligned} \tag{11.1}$$

simultaneously.

[1] That is, $\gcd(m_i, m_j) = 1$ for every $i \neq j$.

Proof. For $j = 1, \ldots, k$, denote

$$M_i = m_1 \cdots m_{i-1} m_{i+1} \cdots m_k;$$

that is, let M_i be the product of all the relevant moduli *except* m_i. Since the m_i are pairwise relatively prime, we have $\gcd(m_i, M_i) = 1$. It follows that there is an integer y_i such that

$$M_i y_i \equiv 1 \mod m_i,$$

for each i. To solve the system of equivalences (11.1), set

$$x = a_1 M_1 y_1 + a_2 M_2 y_2 + \cdots + a_k M_k y_k.$$

Since $m_1 | M_i$ for each $i > 1$, while $M_1 y_1 \equiv 1 \bmod m_1$, it follows that

$$x \equiv a_1(1) + 0 + \cdots + 0 \equiv a_1 \mod m_1,$$

and, similarly, $x \equiv a_i \bmod m_i$ for each i. We have shown that a solution x exists.

Next, suppose that z is another solution to the given system of equivalences. This implies that $z \equiv x \equiv a_i \bmod m_i$ for each i, so that

$$m_i | (z - x)$$

for each i. Since the m_i are pairwise relatively prime, it follows from Proposition 7.3 that the product

$$m_1 m_2 \cdots m_k | (z - x),$$

so that $z \equiv x \bmod m_1 \cdots m_k$. ❑

In practice there are two methods to solve a Chinese remainder problem, as the next example illustrates.

Example: Find all integers x such that

$$\begin{aligned} x &\equiv 3 \mod 4 \\ x &\equiv 2 \mod 5 \\ x &\equiv 10 \mod 13 \end{aligned}$$

are simultaneously satisfied.

Solution #1: Following the proof of the theorem, set $M = 4 \cdot 5 \cdot 13 = 260$, and set

$$M_1 = 5 \cdot 13 = 65 \qquad M_2 = 4 \cdot 13 = 52 \qquad M_3 = 4 \cdot 5 = 20.$$

We now solve for integers y_1, y_2, and y_3 so that

$$65y_1 \equiv 1 \mod 4 \qquad 52y_2 \equiv 1 \mod 5 \qquad 20y_3 \equiv 1 \mod 13.$$

After simplification in each modulus, these identities become

$$y_1 \equiv 1 \mod 4 \qquad 2y_2 \equiv 1 \mod 5 \qquad 7y_3 \equiv 1 \mod 13.$$

A moment of consideration yields the integer solutions

$$y_1 = 1, \qquad y_2 = 3, \qquad y_3 = 2,$$

so that

$$x \equiv 3 \cdot 65 \cdot 1 + 2 \cdot 52 \cdot 3 + 10 \cdot 20 \cdot 2 \equiv 907 \mod 260.$$

Reducing mod 260 then yields the solution

$$x \equiv 127 \mod 260.$$

Solution #2: An alternative approach[2] is to solve the problem one identity at a time. To begin, we have $x \equiv 3 \bmod 4$, so that

$$x = 3 + 4r$$

for some integer r. Next, we are told that $x \equiv 2 \bmod 5$, so that

$$3 + 4r \equiv 2 \mod 5,$$

which implies that

$$4r \equiv -1 \equiv 4 \mod 5,$$

so that $r \equiv 1 \bmod 5$. In other words, $r = 1 + 5s$ for some $s \in \mathbb{Z}$. Putting these results together, we have

$$x = 3 + 4r = 3 + 4(1 + 5s) = 7 + 20s.$$

Notice that we have now simultaneously solved the first two identities, since

$$7 \equiv 3 \mod 4 \quad \text{and} \quad 7 \equiv 2 \mod 5.$$

Continuing, the last condition on x tells us that

$$x = 7 + 20s \equiv 10 \mod 13,$$

so that

$$7s \equiv 3 \mod 13,$$

which implies that $s \equiv 6 \bmod 13$. In other words, $s = 6 + 13t$ for some $t \in \mathbb{Z}$. Combining this with the previous results, we obtain

$$x = 7 + 20s = 7 + 20(6 + 13t) = 127 + 260t,$$

so that

$$x \equiv 127 \mod 260.$$

[2] My experience has been that most students prefer this next method.

❧

The relatively prime condition on the moduli $m_1, \ldots, m_k$ is necessary to assure the existence of a solution. For example, there is *no* integer x such that

$$\begin{aligned} x &\equiv 3 \mod 4 \\ x &\equiv 2 \mod 6. \end{aligned}$$

Indeed, the first condition implies that x is odd, while the second implies that x is even. The discrepancy mod 2 is no coincidence, since $\gcd(4,6) = 2$, so that Theorem 11.1 cannot be applied.

On the other hand, the simultaneous system

$$\begin{aligned} x &\equiv 3 \mod 4 \\ x &\equiv 5 \mod 6. \end{aligned}$$

does have solutions mod 24, namely, $x \equiv 11$ and $x \equiv 23$. Notice that these two different solutions are congruent mod $12 = \operatorname{lcm}(4,6)$. These examples are special cases of the following.

Proposition 11.2. *Let $m_1, m_2 > 1$ be integers, and let $a_1, a_2 \in \mathbb{Z}$. Let $d = \gcd(m_1, m_2)$.*

If $d \mid (a_1 - a_2)$, then the equations

$$\begin{aligned} x &\equiv a_1 \mod m_1 \\ x &\equiv a_2 \mod m_2 \end{aligned} \tag{11.2}$$

have a unique simultaneous solution $x \mod \operatorname{lcm}(m_1, m_2)$.

If $d \nmid (a_1 - a_2)$ then (11.2) has no simultaneous solution.

❧

Exercise 11.1. If we divide n by 3, we have 2 left over. If we divide the same number n by 17, we have 9 leftover. What are the possible values for n?

Exercise 11.2. Find all integers x such that

$$\begin{aligned} x &\equiv 2 \mod 7 \\ x &\equiv 3 \mod 9 \end{aligned}$$

simultaneously hold.

Exercise 11.3. Find all integers x such that

$$\begin{aligned} x &\equiv 4 \mod 5 \\ x &\equiv 1 \mod 7 \\ x &\equiv 8 \mod 41 \end{aligned}$$

all simultaneously hold.

Exercise 11.4. Find the **largest negative integer** x satisfying the modular identities in Exercise 11.3.

Exercise 11.5. Find the **smallest positive integer** x such that

$$\begin{aligned} x &\equiv 1 \mod 2 \\ x &\equiv 2 \mod 3 \\ x &\equiv 3 \mod 5 \\ x &\equiv 5 \mod 7 \end{aligned}$$

all simultaneously hold.

Exercise 11.6. Explain why the equations

$$\begin{aligned} x &\equiv 9 \mod 15 \\ x &\equiv 14 \mod 21 \end{aligned}$$

have no simultaneous solution.

Exercise 11.7. Prove Proposition 11.2.

Exercise 11.8 (Exploratory Exercise)**.** How does Proposition 11.2 generalize to 3 or more modular arithmetic equations?

Exercise 11.9. Find all integers x such that

$$\begin{aligned} x &\equiv 8 \mod 15 \\ x &\equiv 14 \mod 21 \end{aligned}$$

simultaneously hold.

Exercise 11.10. Find all integers x such that

$$\begin{aligned} x &\equiv 1 \mod 5 \\ x^2 &\equiv 3 \mod 11 \end{aligned}$$

simultaneously hold.

Exercise 11.11. Find all integers x such that

$$\begin{aligned} x^2 - x + 3 &\equiv 0 \mod 5 \\ x^2 &\equiv 1 \mod 13 \end{aligned}$$

simultaneously hold.

Exercise 11.12. Find 3 consecutive integers that are each divisible by some perfect square greater than 1.

12 *Divisibility tests*

A divisibility test is a quick test to determine if an integer a is divisible by another integer m. The test should be easier to apply than the actual process of dividing m by a (with possible remainder), or we would just do that! A divisibility test may offer a shortcut to the yes/no question of divisibility, without necessarily supplying a quotient or a remainder. Sometimes, however, a test will also provide us with a value for $a \bmod m$.

The simplest example is one you undoubtedly know already: An integer is even (divisible by 2) iff* its last digit is even. To see why this holds, recall that a positive integer n has decimal digits $d_0, d_1, \ldots, d_k \in \{0, 1, 2, \ldots, 9\}$ when

$$n = 10^k d_k + 10^{k-1} d_{k-1} + \cdots + 10^2 d_2 + 10 d_1 + d_0, \tag{12.1}$$

where d_0 is the final digit. Since $10 \equiv 0 \bmod 2$, it immediately follows that $n \equiv d_0 \bmod 2$, so that the final digit determines whether n is even or odd.

It might seem like we are making too much out of an obvious fact. But the proof is useful in that it reveals immediate generalizations. For example, since $100 \equiv 0 \bmod 4$, a similar argument using the expansion (12.1) implies that a number is congruent mod 4 to the number formed by its last two digits. More generally, since $10^m = 2^m 5^m$, we have divisibility tests for all powers of 2 and 5:

- A number is divisible by 2^m iff the number formed by its last m decimal digits is divisible by 2^m.
- A number is divisible by 5^m iff the number formed by its last m decimal digits is divisible by 5^m.

For example, one can see at a glance that the number 229811340 is divisible by 2 and 4 (because $4|40$), but is not divisible by 8 (since $8 \nmid 340$). Similarly, this number is divisible by 5, but not by 25.

The expansion (12.1) is also used to derive divisibility tests for 3, 9, and 11, using the fact that $10 \equiv 1 \bmod 3$ and mod 9, while $10 \equiv -1 \bmod 11$.

- A number n is always congruent to the sum of its decimal digits mod 3 and mod 9.
- A number n is always congruent to the alternating sum of its decimal digits mod 11.

*"iff" is shorthand for "if and only if".

To understand the first assertion, observe that

$$10^k \equiv 1^k \equiv 1 \mod 3 \quad (\text{or mod } 9),$$

for all integer exponents k. Combining this with (12.1) yields

$$n \equiv d_k + d_{k-1} + \cdots + d_2 + d_1 + d_0 \mod 3,$$

and similarly mod 9.

For the mod 11 test, combine the fact that

$$10^k \equiv (-1)^k \mod 11,$$

with (12.1) to obtain

$$n \equiv d_0 - d_1 + d_2 - \cdots + (-1)^k d_k \mod 11.$$

Note that this alternating sum begins with the last digit d_0 in order to avoid a possible sign error.

For example, the number 229811340 satisfies

$$229811340 \equiv 2+2+9+8+1+1+3+4+0 = 30 \equiv 3+0 = 3 \mod 9,$$

while

$$229811340 \equiv 0-4+3-1+1-8+9-2+2 = 0 \mod 11.$$

Exercise 12.1. Find and prove a simple test for divisibility by 6.

Exercise 12.2. Use mental arithmetic (without writing on paper) to determine which of the following numbers is divisible by 3, by 4, by 6, by 9, and by 11.

(a) 41414
(b) 414414
(c) 41424
(d) 41404
(e) 10001
(f) 11001
(g) 11100
(h) 11010
(i) 4700000036

Exercise 12.3. Find and prove modular arithmetic shortcuts for mod 20, mod 250, and mod 1250, in analogy to the methods for mod 2, 4, and 5 described in this section.

Exercise 12.4. Find and prove modular arithmetic shortcuts for mod 99, mod 999, and mod 101, in analogy to the methods for mod 3, 9, and 11 described in this section.

Exercise 12.5. **(a)** Find a simple shortcut for reducing very large numbers mod 37.
Hint: What is $37 \cdot 27$?
(b) Use the method you found in part (a) to compute 229811340 mod 37.

Exercise 12.6. For $k = 1, 2, 3 \ldots$, let E_k denote the number whose decimal expansion consists of k ones; that is, $E_1 = 1$, $E_2 = 11$, and so on, so that

$$E_k = \underbrace{111\cdots1}_{k \text{ digits, all ones}}$$

(a) Prove that E_k is divisible by 11 if and only if k is even.
(b) Prove that E_k is divisible by 3 if and only if $3|k$.
(c) Prove that if E_k is prime then k is prime.
(d) Prove that if k is prime then E_k is sometimes prime and sometimes composite.

Exercise 12.7. Let n be a positive integer. Let a be the last digit of n expressed in decimal form. Let b be the number formed from n by removing the last digit. (For example, if $n = 3516$ then $a = 6$ and $b = 351$).

Prove that, for all $n \in \mathbb{N}$, we have $7|n$ if and only if $7|(b - 2a)$.

Exercise 12.8. Let n be a positive integer. Following the notation of the previous exercise, show that $n \equiv 3(b - 2a) \bmod 7$.

Exercise 12.9. Use your solution to the previous two exercises regarding mod 7 to find analogous tests for divisibility by mod $13, 17, 19$ and 23.

Exercise 12.10. Prove that a positive integer n is congruent mod 15 to the sum of the digits of its hexadecimal expansion.

Exercise 12.11. Let n be a positive integer.
(a) Invent a simple method for computing n mod 3 when n is given as a binary expansion.

(b) Use your answer to part (a) to compute the value of the ***binary*** numbers 1111001000 and 1011011101111 mod 3.

Exercise 12.12. Let n be a positive integer.
(a) Invent a simple method to compute n mod 7 when n is given as an octal expansion.

(b) Use your answer to part (a) to compute the value of the octal number $(4156332)_8$ mod 7.

(c) Use your answer to part (a) to compute the value mod 7 of the binary expansions from Exercise 12.11

Exercise 12.13. Let m be a 5-digit positive whole number whose decimal expansion has the form

$$m = 5aba4$$

where a and b are decimal digits.

Suppose that $33|m$ and that m is **not** divisible by 4.

What is $m = ?$

13 Checksums

Checksums and checkdigits are used to detect errors in transmission and copying.

When a stream of bits $(b_1, b_2, \ldots, b_n)$ is transmitted,[1] include an extra *parity bit*,

$$b_{n+1} \equiv b_1 + b_2 + \cdots + b_n \mod 2.$$

The recipient compares the mod 2 sum of first n bits received with the parity bit. If the results do not match, there are an odd number of errors in the transmission. This is an effective checksum when errors are expected to be uncommon and distributed randomly.

More generally, given an integer x expressed in binary, the *even parity bit* $e(x)$ is obtained by summing the bits of each x (that is, counting the 1's that appear in the binary expression of x), and returning the value of this sum modulo 2. So the $e(x) = 0$ if x has even number of 1s in its binary expression and $e(x) = 1$ if x has an odd number of 1s in its binary expression. For example, the (decimal) integer 23 has the representation

$$10111$$

in binary, so $e(23) \equiv 1 + 0 + 1 + 1 + 1 \equiv 0 \mod 2$.

In Section 29 even parity bits will play an important role in pseudorandom number generation.

❧

International Standard Book Number (ISBN) codes are used in the labelling of retail books to facilitate easy identification via barcode scanners (for example). ISBN codes include check-digits in order to detect read errors by scanners and copy errors by humans.

An ISBN-10 code is a sequence of values $a_1, \ldots, a_{10}, c$, where each a_i is a decimal digit, and where the symbol $c \in \{0, 1, \ldots, 9, X\}$ is a check-digit satisfying the rule

$$a_1 + 2a_2 + 3a_3 + \cdots + 9a_9 + 10c \equiv 0 \mod 11.$$

The value $c = X$ is used to represent the value of 10 mod 11. An ISBN-10 code is often printed with dashes, which are ignored in the checksum calculation.

[1] A bit is a single binary digit, either 0 or 1.

For example, the code 0-321-69394-5 is a valid ISBN-10 code, whereas 0-321-69934-5 contains an error. Similarly, the code 0-321-69399-X is a valid ISBN-10 code.

❧

An ISBN-13 code is a sequence of values $a_1, \ldots, a_{12}, c$, where each of the a_i are decimal digits, and where the number c is a check-digit satisfying the rule

$$c \equiv a_1 + 3a_2 + a_3 + 3a_4 + \cdots + a_{11} + 3a_{12} \mod 10,$$

where coefficients 1 and 3 alternate in the sum.

For example, the code 978-0-321-69394-9 is a valid ISBN-13 code, whereas 978-0-324-69394-9 contains an error.

❧

Another common example of a check digit is the final digit in a Universal Product Code (UPC). A UPC-A number is the 12-digit number seen at the base of the UPC barcode printed on many retail packages. To compute the check digit for the UPC-A coding system, denote the first 11 digits of the UPC by $a_1, a_2, \ldots, a_{11}$. The final check digit c is given by the formula

$$c \equiv -3(a_1 + a_3 + a_5 + a_7 + a_9 + a_{11}) - (a_2 + a_4 + a_6 + a_8 + a_{10}) \mod 10.$$

For example, the code 3-40034-21778-5 is a valid UPC-A code, whereas 3-40134-21778-5 contains an error.

❧

Exercise 13.1. Suppose we express the non-negative integers $0, 1, 2, 3, \ldots$ in binary notation:

$$0, 1, 10, 11, 100, 101, 110, 111, 1000, \ldots$$

and then compute the even parity bit $e(x)$ of each integer in this list. The result is an infinite sequence $\{e_n\}_{n=0}^{\infty}$ of bits.
(a) What are the even parity bits $e_0, e_1, \ldots, e_{16}$ of the first 17 non-negative integers? Do you see a pattern?
(b) Prove that the sequence $\{e_n\}$ never has 3 consecutive zeroes and never has 3 consecutive ones.

Exercise 13.2. For integers $x \geq 0$ (expressed in binary), prove or disprove:

(a) $e(2x) = e(x)$.
(b) $e(x+1) = e(x) + 1$.
(c) $e(2x+1) = e(x) + 1$.
(d) $e(2^x) = 1$.
(e) $e(2^x - 1) = 0$.
(f) $e(2^x + 1) = 0$.

Exercise 13.3. Show that the check-digit c of the ISBN-10 code satisfies

$$c \equiv \sum_{i=1}^{9} i a_i \mod 11.$$

Exercise 13.4. Show that a valid ISBN-10 code satisfies

$$10a_1 + 9a_2 + 8a_3 + \cdots + 2a_9 + c \equiv 0 \mod 11.$$

Exercise 13.5. What is the check-digit of an ISBN-10 code having the form 6-519-20700-□?

Exercise 13.6. What is the missing digit of the ISBN-10 code 3-888-1234□-3?

Exercise 13.7. Show that the check-digit of an ISBN-10 code will detect an error if any two digits in a valid code are exchanged, or if any single digit has the wrong value.

Exercise 13.8. What is the check-digit of an ISBN-13 code having the form 800-6-452-83813-□?

Exercise 13.9. What is the missing digit of the ISBN-13 code 020-1-666-22□31-7?

Exercise 13.10. Show that the check-digit of an ISBN-13 code will detect an error if exactly one digit has the wrong value.

Exercise 13.11. Show that the check-digit of an ISBN-13 code might *not* detect an error if two adjacent digits in a valid code are exchanged.

Exercise 13.12. What is the check-digit of a UPC-A code having the form 3-80882-20070-□?

Exercise 13.13. What is the missing digit of the UPC-A code 1-41414-1414□-8?

Exercise 13.14. Show that the check-digit of a UPC-A code will detect an error if exactly one digit has the wrong value.

Exercise 13.15. Show that the check-digit of a UPC-A code might *not* detect an error if two adjacent digits in a valid code are exchanged.

14 Pollard's Rho

This next application of modular arithmetic provides a way to factor large integers more efficiently than the trial division method of Section 9.

Given a positive integer n, how can we find its prime factorization? The first step is to determine whether n itself is prime. Methods for doing this efficiently will be described in Section 31. If we know that n is composite, the next step is to find a prime factor p of n. Writing $n = pn'$ we can then iterate by factoring n', until the original integer n is completely factored.

The difficult step is finding a prime p that divides n. The naive approach from Section 9 was to use trial division of n by the sequence of primes

$$2, 3, 5, 7, 11, 13, 17, 19, 23, \ldots$$

until a prime factor is found. By Theorem 9.2 a composite number n is divisible by some prime $p \leq \sqrt{n}$, so this trial division process will require at most $\sqrt{n}$ steps.

The method of trial division has obvious defects. First, it requires us to compile a complete list of primes $p \leq \sqrt{n}$. We can get around this problem by performing trial division by all integers (or, say, all odd integers after 2) up to $\sqrt{n}$, without worrying about which is prime. But this still leaves a procedure of great inefficiency. For example, if n is a 13 digit number, then $\sqrt{n}$ is on the order of 10^6, possibly requiring millions of trial divisions. If n has more than 24 digits, then this procedure may require trillions of trial divisions.

Faster methods of factoring integers have been developed. In the 1970s John Pollard discovered an especially simple and beautiful stochastic algorithm for finding prime factors of a large composite integer [22]. This algorithm, known as *Pollard's Rho*, is implemented as follows:

> Let $f(x)$ be a quadratic polynomial (such as $f(x) = x^2 + 1$).
> Set $x_0 = y_0 = 2$.
> For $i \geq 1$, let $x_i = f(x_{i-1}) \bmod n$, let $y_i = f(f(y_{i-1})) \bmod n$,
> and let $d_i = \gcd(x_i - y_i, n)$.
>
> - If $d_i = 1$, then continue to step $i + 1$.
> - If $d_i = n$, then the algorithm *fails* (see below).
> - If $1 < d_i < n$, then the algorithm returns $d = d_i$, a proper divisor of n.

Observe that, at each step of the sequence above, we have $y_i = x_{2i}$. Pollard's Rho computes $\gcd(x_i - x_{2i}, n)$ for $i = 1, 2, \ldots$ until either a proper factor of n appears,

or $x_{2i} \equiv x_i \bmod n$. In the latter case, the algorithm fails and should be restarted using a different choice of quadratic function $f(x)$, such as $f(x) = x^2 - 1$.

Pollard's Rho assumes that n is a composite number. A *primality test* (such as the Miller-Rabin test; see Section 31) should always be used before Pollard's Rho is applied, and primality should also be checked every time a factor of n has been isolated.

If the algorithm succeeds in returning a proper divisor d of n, we have $n = dn'$ for some integers $d, n' < n$. After testing d and n' for primality, we can now iterate the procedure on d and n' until n is factored completely into primes.

❧

While Pollard's Rho is usually used for large integers and implemented on a computer, the following step-by-step example with $n = 1219$ illustrates the remarkable efficiency of this algorithm. In this example we use $f(x) = x^2 + 1$. Keep in mind that all computations are done mod 1219. The results of each step appear in the following table.

i	x_i	y_i	$x_i - y_i$	$\gcd(x_i - y_i, 1219)$
0	2	2	0	1219
1	5	26	21	1
2	26	1205	1197	1
3	677	1021	344	1
4	1205	1021	1035	23 (Success!)
5	197			
6	1021			
7	197			
8	1021			

(14.1)

As seen in the table 14.1 above, the sequence x_i begins to cycle mod 1219 at x_5 with a period of 2. The cycling mod 23 is revealed the at step 4, where the gcd of 23 appears. It follows that $23|1219$. After one more integer division we find the prime factorization $1219 = 23 \cdot 53$.

In practice the algorithm would have terminated at $i = 4$ in 14.1. The values of x_i are shown here through $i = 8$ for illustrative purposes only. Notice that $y_i = x_{2i}$ for each i.

❧

Suppose that n is a composite integer whose smallest prime factor is p. In the analysis that follows we will show that Pollard's Rho typically finds p (or some other non-trivial factor of n) after approximately $\sqrt{p}$ steps. Since $p \leq \sqrt{n}$, Pollard's Rho can often be used to find a factor of a composite number n after

approximately $\sqrt[4]{n}$ steps, a dramatic improvement over earlier methods. For example, if n has 13 digits, then Pollard's Rho should find a factor of n after only few thousand steps, rather than the millions required by the method of trial division.

❧

While Pollard's Rho is easy to describe and to implement, it is not obvious *why* this algorithm finds a factor as quickly as it does. The following non-rigorous analysis assumes some knowledge of probability theory.

Suppose we select k items uniformly at random from a list of n distinct items *with replacement* (so that the same item may be selected multiple times). Let $\mathcal{P}$ denote probability that the k items selected are all different from one another; i.e., that no item is selected more than once. A classical result known as the *Birthday Problem* [28, p. 147] asserts that, if k is much smaller than n, then

$$\mathcal{P} \approx 1 - e^{-\frac{k^2}{2n}},$$

where $e = 2.718\ldots$ is Euler's famous constant.[1]

In the classic example, let $n = 365$, the number of days in a typical year, and let k be the number of randomly chosen people from a population. If $k \geq 23$ then $\mathcal{P} \geq 0.5$, so that there is a better than 50% chance that at least two people chosen have the same birthday. If $k \geq 58$ then this probability rises to almost 99%! More generally, if $k > 2\sqrt{n}$ then $\mathcal{P} \geq 0.86$, while if $k > 4\sqrt{n}$ then $\mathcal{P} > 0.999$.

Let us suppose (for the moment) that the sequence x_i generated by Pollard's Rho is a uniformly random sequence of integers when reduced mod p, where p is the smallest prime dividing n. The solution to the Birthday Problem asserts that, if we examine the sequence out to $i \geq 4\sqrt{p}$, then we are more than 99.9% likely to find a *repeated value* mod p.

Recall that $x_{i+1} = f(x_i)$. It follows that, whenever $x_i = x_j$, we also have $f(x_i) = f(x_j)$, so that $x_{i+1} = x_{j+1}$ and so on. In other words, once $x_i = x_{i+u}$ for some u, the sequence will be forever periodic with period length u.

With this in mind, let s denote the first index of x such that $x_s = x_{s+u}$. There is no reason to expect that Pollard's Rho will find s, since the algorithm only compares x_i with x_{2i}. Amazingly, however, this turns out to be good enough in practice.[2] To see why, notice that

$$\begin{aligned} x_i \equiv x_{2i} \quad &\text{iff} \;\; i \geq s \text{ and } 2i = i + uk \text{ for some } k \\ &\text{iff} \;\; i \geq s \text{ and } i = uk \text{ for some } k \\ &\text{iff} \;\; i \geq s \text{ and } i \equiv 0 \mod u \end{aligned}$$

[1] A better estimate is $\mathcal{P} \approx 1 - e^{-\frac{k(k-1)}{2n}}$, but either will serve our purposes here.

[2] This phenomenon is known as *Floyd cycle detection* and is used to improve the efficiency of many advanced factorization algorithms [16].

This means that Pollard's Rho will detect the cycle when $i \geq s$ and $i \equiv 0 \bmod u$. To see how soon this happens, use division with remainder:

$$s = uq + r,$$

where $0 \leq r < u$, and then set $i = s - r + u$. Evidently $i > s$ and

$$i \;=\; (s-r)+u \;=\; uq+u \;\equiv\; 0 \mod u.$$

It follows that Pollard's Rho finds a pair $x_{2i} \equiv x_i \bmod p$ no later than

$$i = s - r + u \leq s + u.$$

Since x_{s+u} is the first repeated value in the sequence $\{x_i \bmod p\}$, this will occur with moderate probability if $i \geq \sqrt{p}$, and with very high probability if $i \geq 4\sqrt{p}$.

Since a composite integer n always has a prime factor $p \leq \sqrt{n}$, Pollard's Rho is highly likely to find a proper factor of n after $4\sqrt[4]{n}$ steps, with a decent probability of success even after only $\sqrt[4]{n}$ steps. This is much faster than the $\sqrt{n}$ steps that may be needed using the method of trial division.

There are some obvious objections to this analysis. One objection is that we are comparing apples and oranges: Trial by division does one division at each step, while Pollard's Rho requires a gcd computation at each step (along with some quadratic computations). However, Euclid's algorithm provides a very efficient means of computing the gcd (on the order of $\log(n)$ divisions), rendering this point moot for larger values of n.

A more serious objection is that the solution to the Birthday Problem applies to uniformly random selections, while the sequence x_i is generated by a deterministic (though arguably pseudorandom) formula. While this turns out not to matter in practice, it renders the precise stochastic mechanism behind Pollard's Rho a bit mysterious. Indeed, this mechanism remains a topic of current research. In the meantime, a more detailed stochastic analysis of Pollard's Rho can be found in [6, p.426-432].

❧

More sophisticated methods of factoring integers have also been discovered, including *Fermat factorization*, the *continued fraction method*, the *quadratic sieve*, and various *number field sieves*. These algorithms lie beyond the scope of this book, but are explained in [6, 10, 30], for example.

There are still no integer factorization algorithms that run in polynomial time relative to the number of digits of n, although advances in quantum computing may change this someday [29].

❧

Exercise 14.1. Use Pollard's Rho to factor the number 111 by hand. Remember that all computations take place mod 111.

Exercise 14.2. Use Pollard's Rho to factor the number 1333, constructing a table with columns for i, x_i, y_i, $x_i - y_i$, and $\gcd(x_i - y_i, 1333)$ as in the example of 14.1. Remember that all computations take place mod 1333. This can be done by hand in a few minutes with a hand-held calculator. How many steps does your procedure require? How does this compare to factoring 1333 using trial by division?

Exercise 14.3. Use Pollard's Rho to factor the number 2279, as in Exercise 14.2. How many steps does your procedure require? How does this compare to factoring 2279 using trial by division? *Hint:* You should notice a definite pattern in the x_i after 7 or 8 steps.

Exercise 14.4. Pollard's Rho should only be applied to an integer n that is known to be composite. What happens of you apply Pollard's Rho to a prime?

Exercise 14.5 (Project Exercise)**.** Write a program in a computer language of your choice to factor integers using the method of trial division. Use your program to factor the following composite numbers

$$\begin{array}{l} 4056187 \\ 8165104431 \\ 64227559510527 \\ 171684933869539 \\ 82079243425409089 \end{array} \tag{14.2}$$

Note how long it takes your program to factor each number.[3]

Exercise 14.6 (Project Exercise)**.** Write a program in a computer language of your choice to find a proper factor of a composite integer using Pollard's Rho. Use your program to find proper factors of the numbers in the list (14.2) from the previous exercise. Note how long it takes your program to find a proper factor of each number. How do these times compare to those in Exercise 14.5?

Then use your program to find a proper factor of 12193263122374638001.

[3] If your program takes too long to factor 82079243425409089, then just make note of that and move on to Exercise 14.6.

15 $\mathbb{Z}_p$ and Fermat's theorem

Recall from Corollary 10.8 that, if p is prime, then every residue except 0 is a unit mod p. This fact has remarkable consequences, among them a fundamental property of exponentials discovered by Fermat.[1]

Theorem 15.1 (Fermat's Theorem). *If $p \nmid a$ then $a^{p-1} \equiv 1 \mod p$.*

Theorem 15.1 is sometimes known as *Fermat's Little Theorem*, in order to distinguish it from the more famous *Fermat's Last Theorem*.[2]

Proof. Consider the list of values

$$a, 2a, 3a, \ldots, (p-1)a \mod p.$$

Since $p \nmid a$, the cancellation law implies each of these values is non-zero and distinct mod p, for if $ar \equiv as$, then we can cancel a to obtain $r \equiv s \mod p$. It follows that the list above is the same as the list of values

$$1, 2, 3, \ldots, (p-1) \mod p,$$

listed in a possibly different order. Therefore, multiplying all of the numbers in the first list will give the same outcome mod p as multiplying all of the numbers in the second list. That is to say,

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

Applying the cancellation law once again, we have $a^{p-1} \equiv 1 \mod p$. ❑

Corollary 15.2.

- *If $a \in \mathbb{Z}$ then $a^p \equiv a \mod p$.*
- *If $p \nmid a$ then $a^{-1} \equiv a^{p-2} \mod p$.*

This corollary follows readily from Fermat's theorem (see Exercise 15.1).

❧

[1] Pierre de Fermat (1601–1665) was the founder of modern number theory. For a short history of his contributions, see [3, p. 5] or [33, p. 207-223].

[2] Fermat's Last Theorem addresses the very difficult Diophantine equation $a^n + b^n = c^n$. See the discussion of the equation (28.6) in Section 28.

Everyone knows that a number has at most two square roots. Except when it doesn't! Consider, for example, the equation

$$x^2 \equiv 1 \mod 8.$$

Checking the possible values $1, 2, 3, \ldots, 8$ reveals that this equation has **4** distinct solutions, namely, $x = 1, 3, 5, 7$.

This anomaly turns out to be a consequence of working in a composite modulus. The next proposition shows that quadratic equations are more well-behaved when the modulus is a prime number p.

Proposition 15.3. *If p is an odd prime, and $p \nmid a$, then the equation*

$$x^2 \equiv a \mod p$$

has either exactly two distinct roots or no solutions at all.

Before proving this proposition, consider the following question: Can it ever be the case that $a \equiv -a \mod m$? Of course this is true when $a \equiv 0$, but what if $a \not\equiv 0$? The answer is Yes. For example, $2 \equiv -2 \mod 4$. More generally, $a \equiv -a \mod m$ holds if and only if $2a \equiv 0 \mod m$. If m is even, say $m = 2k$, then this identity holds for $a \equiv k$ (or $-k$) mod m. On the other hand, if m is odd, then $\gcd(2, m) = 1$, so that

$$a \equiv -a \quad \Leftrightarrow \quad 2a \equiv 0 \quad \Leftrightarrow \quad a \equiv 0 \mod m,$$

since we can divide by 2 in this instance. In particular, if p is a odd prime number then

$$a \equiv -a \mod p \quad \Leftrightarrow \quad a \equiv 0 \mod p.$$

Proof of Proposition 15.3. If there no solutions (can you think of an example?), then we are done.

Suppose instead that there is a solution b, so that $b^2 \equiv a \mod p$. It follows that $(-b)^2 \equiv b^2 \equiv a$ as well, giving a second solution, since $b \not\equiv -b$ when p is odd.

Suppose that x is yet another (third?) solution. In this case we would have

$$x^2 \equiv a \equiv b^2 \mod p$$

so that

$$x^2 - b^2 \equiv 0 \mod p.$$

Factoring the difference of squares, we have

$$(x - b)(x + b) \equiv 0 \mod p,$$

so that either $x - b \equiv 0$ or $x + b \equiv 0 \mod p$, by Theorem 10.10. In other words, $x \equiv b$ or $x \equiv -b$; there are no other possibilities. ❑

Proposition 15.3 explains the following famous curiousity regarding factorials in a prime modulus.

Theorem 15.4 (Wilson's Theorem)**.** *If p is prime, then $(p-1)! \equiv -1 \mod p$.*

Proof. Consider the list of numbers $\{1, 2, \ldots, p-1\}$. Since no number on this list is divisible by p, each of them has an inverse mod p. In other words, for each number a in this list, there is a number a^{-1}, also in the list, such that $aa^{-1} \equiv 1 \mod p$.

But wait! Can any such number be its own inverse? Suppose $a \equiv a^{-1} \mod p$. In this case, multiply both sides by a to obtain $a^2 \equiv 1$. By Proposition 15.3 the only possible cases are $a \equiv 1$ and $a \equiv -1 \equiv p-1$. Therefore, for each $a \in \{2, \ldots, p-2\}$, we have $a \not\equiv a^{-1}$. Since both a and a^{-1} appear in that list, everything in this shorter list must cancel upon multiplication, so that

$$2 \cdot 3 \cdots (p-2) \equiv 1 \mod p.$$

It follows that

$$1 \cdot 2 \cdot 3 \cdots (p-2) \cdot (p-1) \equiv 1 \cdot (p-1) \equiv -1 \mod p.$$

❑

❧

Exercise 15.1. Use Theorem 15.1 to prove Corollary 15.2.

Exercise 15.2. Compute $5^{843} \mod 43$.

Exercise 15.3. Compute $2^{64} \mod 19$.

Exercise 15.4. Prove that, if $m > 4$ is composite, then $(m-1)! \equiv 0 \mod m$. What happens when $m = 4$?

Exercise 15.5. Prove that, if p is prime, then $(p-2)! \equiv 1 \mod p$.

Exercise 15.6. Compute $76! \mod 79$.

Exercise 15.7. Use Wilson's theorem to simplify $3 \cdot 4 \cdots 11 \mod 13$.

Exercise 15.8. For which values of k does the equation $x^2 \equiv k$ have a solution mod 4? mod 5? mod 3? mod 7?

Exercise 15.9. List *all* quadratic equations mod 2. Which have solutions mod 2? Which have no solutions?

Exercise 15.10. List *all* quadratic equations mod 3. Which have solutions mod 3? Which have no solutions?

Exercise 15.11. Suppose that n is a positive integer. Prove that

$$1^n + 2^n + 3^n + 4^n$$

is a multiple of 5 iff n is not a multiple of 4.[†]

Exercise 15.12. Prove that, if $n > 1$, then $\frac{3^n - 2^n}{n}$ is not an integer.
Hint: Consider various cases. What happens if $2|n$? If $3|n$? What happens if n is a prime greater than 3? A power of such a prime? A product of prime powers?

Exercise 15.13. For which values of $n \in \mathbb{N}$ is $\frac{5^n - 3^n}{n}$ an integer?

Exercise 15.14. Suppose that $b > a$ are positive integers. For which values of $n \in \mathbb{N}$ is $\frac{b^n - a^n}{n}$ an integer?

❧

Recall that, for non-negative integers n and k the (n,k)-binomial coefficient is given by

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \tag{15.1}$$

Exercise 15.15. Let p be a positive prime number. Use the results of Exercise 10.24[*] to give an induction proof of Fermat's Theorem, by showing that $x^p \equiv x \bmod p$ for all integers x.

Exercise 15.16 (Exploratory Exercise)**.** Recall that the binomial coefficients (15.1) form a famous patterns when arranged as Pascal's Triangle. What happens to the values in Pascal's triangle when binomial coefficients are computed mod 2? What happens modulo other primes? Or modulo composite numbers?

[†]From the 1901 Eötvös Competition [24].

[*]If you haven't done Exercise 10.24, try to do it now, but be sure to solve Exercise 10.24 *without* using Fermat's Theorem.

16 Units in $\mathbb{Z}_n$ and Euler's function

Recall that an integer k is a unit mod n if and only if $\gcd(k,n) = 1$. For $n > 1$ let U_n denote the set of units mod n. For example,

$$U_4 = \{1,3\}, \qquad U_5 = \{1,2,3,4\}, \qquad U_{12} = \{1,5,7,11\}.$$

Proposition 16.1. *If $u, v \in U_n$, then $uv \in U_n$ and $u^{-1} \in U_n$.*

Proposition 16.1 tells us that U_n is closed under multiplication and inverses. In other words, U_n is a group[1] under multiplication mod n.

Proof. If u is a unit mod n with inverse u^{-1}, then $uu^{-1} = 1$, so that $u = (u^{-1})^{-1}$. It follows that $u^{-1} \in U_n$ as well. If $u, v \in U_n$, with respective inverses u^{-1} and v^{-1}, then

$$(uv)(v^{-1}u^{-1}) \;=\; u(vv^{-1})u^{-1} \;=\; uu^{-1} \;=\; 1,$$

so that $uv \in U_n$. ❑

Denote by $\phi(n) = |U_n|$, the number of units mod n. By convention we also define $\phi(1) = 1$. The function ϕ is sometimes called "Euler's function"[2] or the "Euler Phi function".[3]

From the examples above we see that

$$\phi(4) = 2, \qquad \phi(5) = 4, \qquad \phi(12) = 4.$$

If we are careful to restrict our attention to the units mod n, we obtain the following generalization of Fermat's Theorem 15.1.

Theorem 16.2 (Euler's Theorem). *If $\gcd(a,n) = 1$ then $a^{\phi(n)} \equiv 1 \mod n$.*

The proof is almost the same as that of Fermat's Theorem.

Proof. Consider the list of all units:

$$u_1, u_2, \ldots, u_{\phi(n)} \mod n. \tag{16.1}$$

[1] You don't need to know group theory to enjoy classical number theory, but the two subjects are closely related. See, for example, [4].

[2] Leonhard Euler (1707–1783) was a Swiss mathematician. The name "Euler" is pronounced like the English word "oiler" and *not* like "yooler". [Euclid was Greek and lived more than 1000 years before Euler. These names are not related.]

[3] Some texts may refer to ϕ as the "totient."

Multiply all of these units by the value a, to obtain a new list:

$$au_1, au_2, \ldots, au_{\phi(n)} \mod n. \tag{16.2}$$

Since $\gcd(a,n) = 1$, each au_i is also a unit, by Proposition 16.1. Moreover, the cancellation law implies each of these values is distinct mod n, for if $ar \equiv as$, then we can cancel a to obtain $r \equiv s \mod n$. It follows that the lists (16.1) and (16.2) have the same values, possibly arranged in a different order. Therefore, multiplying all of the numbers in the first list will give the same outcome mod n as multiplying all of the numbers in the second list. That is to say,

$$u_1 u_2 \cdots u_{\phi(n)} \equiv a^{\phi(n)} u_1 u_2 \cdots u_{\phi(n)} \mod n.$$

Applying the cancellation law once again, we have $a^{\phi(n)} \equiv 1 \mod n$. ❑

❧

If p is prime then the units mod p are the non-zero values $\{1, 2, \ldots, p-1\} = U_p$, so that

$$\phi(p) = p - 1.$$

For this special case Euler's Theorem 16.2 simply re-states Fermat's Theorem 15.1.

❧

For $k \geq 2$, the units U_{p^k} consist of all values from $1, 2, \ldots, p^k - 1$ that are not divisible by p. The reader should use this observation to verify that

$$\phi(p^k) = p^k - p^{k-1}. \tag{16.3}$$

In order to compute $\phi(n)$ for more general composite numbers n, the following observation is helpful.

Proposition 16.3. *Let $m, n > 1$. The following are equivalent:*

- *k is a unit* mod *mn*
- *k is a unit* mod *m and a unit* mod *n*

Proof. Recall that k is unit mod mn iff $\gcd(k, mn) = 1$. It is left to the reader to show that $\gcd(k, mn) = 1$ iff both $\gcd(k, n) = 1$ and $\gcd(k, m) = 1$ (see Exercise 16.2). ❑

The next theorem makes the computation of $\phi(n)$ much easier, provided we know the factorization of n.

Theorem 16.4. *If* $\gcd(m,n) = 1$ *then* $\phi(mn) = \phi(m)\phi(n)$.

Proof. Suppose that $\gcd(m,n) = 1$. Consider the map

$$G : U_{mn} \to U_m \times U_n$$

given by

$$G(k) = (k \bmod m,\ k \bmod n).$$

We will show that the map G is a bijection,[4] so that $\phi(mn) = \phi(m)\phi(n)$.

First, observe that $G(k)$ is well-defined; that is, $G(k)$ is indeed a pair of units mod m and mod n by Proposition 16.3.

To show that G is one-to-one, suppose that $G(k_1) = G(k_2)$. This means that $k_1 \equiv k_2 \bmod m$ and that $k_1 \equiv k_2 \bmod n$. In other words, $m|(k_2 - k_1)$ and $n|(k_2 - k_1)$. Since $\gcd(m,n) = 1$, it follows from Proposition 7.3 that $mn|(k_2 - k_1)$, so that $k_1 \equiv k_2 \bmod mn$.

To show that G is onto, suppose that k_1 is a unit mod m and that k_2 is a unit mod n. Since $\gcd(m,n) = 1$, the Chinese remainder theorem implies that there exists a unique value k mod mn such that $k \equiv k_1 \bmod m$ and $k \equiv k_2 \bmod n$, so that $G(k) = (k_1, k_2)$.

Since G is a bijection between the finite sets U_{mn} and $U_m \times U_n$, these sets must have same number of elements, so that $\phi(mn) = \phi(m)\phi(n)$. ❑

The multiplicative property of ϕ given by Theorem 16.4 leads to a useful formula for $\phi(n)$ in the case where we are able to factor n into powers of primes.

Proposition 16.5. *If* $n = p_1^{a_1} \cdots p_k^{a_k}$, *then*

$$\phi(n) \;=\; (p_1^{a_1} - p_1^{a_1-1}) \cdots (p_k^{a_k} - p_k^{a_k-1}) \;=\; n\Big(1 - \frac{1}{p_1}\Big) \cdots \Big(1 - \frac{1}{p_k}\Big).$$

Proof. Combine Theorem 16.4 with the identity 16.3. ❑

The final expression in the identity of Proposition 16.5 enables us to compute $\phi(n)$ without knowing the powers a_i of the primes in the factorization of n.

❧

For $n \in \mathbb{N}$, the function ϕ satisfies the following identity:

$$\sum_{d|n} \phi(d) = n. \tag{16.4}$$

[4] A *bijection* is a function that is both one-to-one and onto; that is, an invertible function.

For example, the divisors of the number 10 are 1, 2, 5, and 10, while

$$\phi(1) + \phi(2) + \phi(5) + \phi(10) \;=\; 1 + 1 + 4 + 4 \;=\; 10.$$

To prove the identity 16.4, first verify the case of $n = p$, a prime. More generally, the left hand side of (16.4) gives a straightforward telescoping sum when n is a power of a prime. The general formula for $\phi(n)$ together with some algebra can be used to verify the remaining cases. (See Exercise 16.12.)

❧

Functions $f : \mathbb{N} \to \mathbb{N}$ are called *arithmetic*[5] functions. An arithmetic function is said to be *multiplicative* if $f(mn) = f(m)f(n)$ whenever $\gcd(m, n) = 1$.

Theorem 16.4 asserts that ϕ is a multiplicative function. Simpler examples of multiplicative functions include the functions of the form

$$f(n) = n^k,$$

where k is a positive integer constant.

Theorem 16.6. *If $f : \mathbb{N} \to \mathbb{N}$ is a multiplicative function, then*

$$g(n) = \sum_{d|n} f(d)$$

is also an multiplicative function.

Theorem 16.6 can be used to prove the identity 16.4, once the prime power case is verified.

The proof of Theorem 16.6 uses the following lemma.

Lemma 16.7. *If $\gcd(m, n) = 1$ and $d|mn$ then there are unique $a|m$ and $b|n$ such that $d = ab$.*

Proof. Let $a = \gcd(d, m)$, and let $b = \gcd(d, n)$. Clearly $a|m$ and $b|n$. Since $\gcd(m, n) = 1$, it follows that $\gcd(a, b) = 1$ as well.

Note that $a|d$ and $b|d$. Since $\gcd(a, b) = 1$, it follows that $ab|d$. Meanwhile, there exist integers x and y so that

$$a = dx + my$$

as well as integers w and z so that

$$b = dw + nz,$$

so that

$$ab = d^2xw + dnxz + dmyw + mnyz.$$

[5]Pronounced with an accent on the *third* syllable.

Since $d|mn$, it now follows that $d|ab$. Since $ab|d$ as well (shown earlier), we have $d = ab$.

To prove uniqueness, suppose that $d = a'b'$ where $a'|m$ and $b'|n$. Then $a'b' = ab$. Since $a'|m$ and $b|n$, we have $\gcd(a', b)|\gcd(m, n) = 1$, so that $\gcd(a', b) = 1$. It follows that $a'|a$. By a similar and symmetrical argument $a|a'$. It follows that $a = a'$ and $b = b'$. ❑

Exercise 16.13 asks you to use Lemma 16.7 to prove Theorem 16.6.

❧

Since the functions 1 and n are multiplicative, it follows from Theorem 16.6 that

$$\tau(n) = \sum_{d|n} 1 \quad \text{(the number of divisors of } n\text{)},$$

$$\sigma(n) = \sum_{d|n} n \quad \text{(the sum of the divisors of } n\text{)},$$

are both multiplicative functions.

❧

The *Möbius function* μ is defined by

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1 \\ (-1)^s & \text{if } n = p_1 p_2 \cdots p_s \text{ (i.e., if } n \text{ is square-free)} \\ 0 & \text{if } n \text{ is not square-free} \end{cases}$$

It is easy to verify from the definition above that μ is multiplicative. Moreover, Theorem 16.6 has striking consequences when applied to μ (see Exercises 16.27 and 16.28).

❧

Exercise 16.1. Find an example of units u, v mod m where $u + v \not\equiv 0$ but $u + v$ is still not a unit mod m.

Exercise 16.2. Let $m, n > 1$. Prove that $\gcd(k, mn) = 1$ iff both $\gcd(k, n) = 1$ and $\gcd(k, m) = 1$.

Exercise 16.3. Prove that if $n > 2$ then $\phi(n)$ is even.

Exercise 16.4. Prove that if $a|b$ then $\phi(a)|\phi(b)$.

Exercise 16.5. Suppose that n is odd and that $4 \nmid \phi(n)$. Prove that $n = p^k$ for some $k \in \mathbb{N}$ and some prime p such that $p \equiv 3$ mod 4.

Exercise 16.6. For which integers n is $\phi(n) = 2$? $\phi(n) = 4$? $\phi(n) = 6$?

Exercise 16.7. How are the values of $\phi(n)$ and $\phi(2n)$ related?

Exercise 16.8. How are the values of $\phi(n)$ and $\phi(n^2)$ related?

Exercise 16.9. Prove that, if n is composite and $\phi(n)|(n-1)$, then n is a square-free product of at least 3 distinct primes.

Exercise 16.10. Prove that $\phi(mn) > \phi(m)\phi(n)$ whenever $\gcd(m,n) > 1$.

Exercise 16.11. An integer n between 1 and 1000000 is chosen uniformly[6] at random. What is the probability that n is relatively prime to 1000000?

Exercise 16.12. **(a)** Use Proposition 16.5 to prove the identity (16.4) for the special case in which $n = p^k$, a power of a prime number.
(b) Prove the identity (16.4) for any integer n.
Hint: Suppose n is factored into powers of distinct primes. Apply part (a) to these powers, and then combine Theorem 16.4 with some algebra to finish the proof.

Exercise 16.13. Use Lemma 16.7 to prove Theorem 16.6.

Exercise 16.14. Use Theorems 16.4 and 16.6 to give another proof of the identity (16.4).

Exercise 16.15. Prove that, for all integers $k \geq 0$, the arithmetic function $f(n) = n^k$ is multiplicative.

Exercise 16.16. Prove directly (from the definitions) that the arithmetic functions τ, σ, and μ are all multiplicative functions, without using Theorem 16.6.

Exercise 16.17. Let $n = p_1^{a_1} \cdots p_k^{a_k}$ (where $p_1 < \cdots < p_k$ are prime). Find a simple formula for $\tau(n)$ in terms of the exponents a_i.

Exercise 16.18. Let $n = p_1^{a_1} \cdots p_k^{a_k}$ (where $p_1 < \cdots < p_k$ are prime). Prove that

$$\sum_{d|n} \tau(d) = \prod_{i=1}^{k} \binom{a_i + 2}{2}.$$

Exercise 16.19. Let $p \in \mathbb{N}$ be prime. Prove that

$$\sigma(p^k) = \frac{p^{k+1} - 1}{p - 1}.$$

Exercise 16.20. Let $n = 81000000$. Without use of any electronic device, compute $\tau(n)$, $\sigma(n)$, and $\mu(n)$.
Hint: Use the multiplicative property, along with the results of Exercises 16.17 and 16.19.

Exercise 16.21. For which integers n is $\tau(n) = 2$? $\tau(n) = 3$? $\tau(n) = 4$?

Exercise 16.22. For which integers n is $\tau(n)$ a power of 2?

Exercise 16.23. For which integers n is $\tau(n)$ odd?

Exercise 16.24. Prove that $\tau(mn) \leq \tau(m)\tau(n)$ for all $m, n \in \mathbb{N}$. When does equality hold?

Exercise 16.25. Prove that if $n > 2$ is square-free then $\sigma(n)$ is even.

[6] This means that each possible choice is equally likely.

Exercise 16.26. A positive integer n is said to be *perfect* if n is the sum of its proper divisors. For example, the number 6 is perfect, since $6 = 1+2+3$, while 12 is not perfect, since $1+2+3+4+6 = 16 \neq 12$.
(a) Show that the numbers 28 and 496 are perfect.
(b) Show that n is perfect if and only if $\sigma(n) = 2n$.
(c) Show that, if p and $2^p - 1$ are both prime, then $2^{p-1}(2^p - 1)$ is a perfect number.[7]
(d) Show that, if p is prime and $k \in \mathbb{N}$ then p^k is not perfect.[8]
Remark: Euler proved a partial converse of part (c); that is, every *even* perfect number has the form given in part (c).[9] Are there any *odd* perfect numbers? Nobody knows!

Exercise 16.27. Find a surprisingly simple formula for

$$\sum_{d|n} \mu(d).$$

Exercise 16.28. Suppose that f is a multiplicative function on $\mathbb{N}$ and that

$$g(n) = \sum_{d|n} f(d).$$

Use the result of Exercise 16.27 to prove the *Möbius Inversion Formula*:

$$f(n) = \sum_{d|n} \mu(n/d) g(d).$$

Exercise 16.29. Recall the function $\pi(n)$ that counts the number of positive prime integers less than or equal to n. Show that $\pi(n)$ is *not* a multiplicative function.

Exercise 16.30. Compute 5^{686} mod 18.

Exercise 16.31. What are the last two (least significant) decimal digits of 3^{1201}?

Exercise 16.32. What are the last two (least significant) decimal digits of 2^{1201}?

Exercise 16.33. What are the last two (least significant) decimal digits of $7^{80} + 80^7$?

❧

Knuth[15] introduced the following *up-arrow notation* for generating large integers. Define

$$a \uparrow b = a^{a^{\cdot^{\cdot^{a}}}}$$

where the value a appears b times in this exponential expression. For example,

$$a \uparrow 4 = a^{a^{a^a}} = a^{(a^{(a^a)})}.$$

Note carefully the order of operations: the exponentials are computed from the top down. For example,

$$3 \uparrow 4 = 3^{3^{3^3}} = 3^{3^{27}} = 3^{7625597484987}.$$

[7] Primes of the form $2^p - 1$ are the *Mersenne primes* described Section 9.
[8] This result can be extended further: an odd perfect number must have at least three distinct prime factors [9, p. 13].
[9] Euler's work on perfect numbers is surveyed in [9]. On a more elementary related note, see Exercise 10.18.

Next, define

$$a \uparrow\uparrow b = \underbrace{a \uparrow a \uparrow \cdots \uparrow a}_{b \text{ copies of } a} = a \uparrow (a \uparrow (\cdots \uparrow a) \cdots),$$

noting once again the order of operations (from right to left). For example,

$$2 \uparrow\uparrow 3 = 2 \uparrow 2 \uparrow 2 = 2 \uparrow (2 \uparrow 2) = 2 \uparrow 4 = 2^{2^{2^2}} = 65536.$$

More generally, for integers $k \geq 1$ define

$$a \uparrow^k b = a \underbrace{\uparrow\uparrow \cdots \uparrow} b = \underbrace{a \uparrow^{k-1} a \uparrow^{k-1} \cdots \uparrow^{k-1} a}_{b \text{ copies of } a}.$$

Exercise 16.34. Prove that $2 \uparrow^k 2 = 4$ for all k.

Exercise 16.35. Which is larger?
(a) $2 \uparrow\uparrow\uparrow 3$ or $4 \uparrow\uparrow 3$?
(b) $3 \uparrow\uparrow\uparrow 3$ or $4 \uparrow\uparrow 4$?

Exercise 16.36. Prove that, for $a, b \in \mathbb{N}$ and $b > 1$, we have $a \uparrow b = a^{a \uparrow (b-1)}$.

Exercise 16.37. Suppose that $a, b, n \in \mathbb{N}$. True or False?
(a) If $a \equiv b \bmod m$ then $a \uparrow n \equiv b \uparrow n \bmod m$.
(b) If $a \equiv b \bmod m$ then $n \uparrow a \equiv n \uparrow b \bmod m$.
(c) If $a \equiv b \bmod m$ then $a \uparrow a \equiv b \uparrow b \bmod m$.

Exercise 16.38. Compute:

(a) $3 \uparrow 6 \mod 4$
(b) $4 \uparrow 4 \mod 9$
(c) $2 \uparrow 5 \mod 5$
(d) $5 \uparrow\uparrow 5 \mod 10$
(e) $3 \uparrow\uparrow\uparrow 2 \mod 7$
(f) $10 \uparrow\uparrow\uparrow 10 \mod 12$

Exercise 16.39. For which positive integers n does $n \uparrow n$ have 3 as its final digit?

Exercise 16.40. For which positive integers n does $n \uparrow n$ have 6 as its final digit?

17 Elementary cryptography

Cryptography is the science of secret codes; that is, of expressing information in a way that is only meaningful to privileged viewers and can not be understood by eavesdroppers. Cryptography usually relies on the use of a shared secret or password (called the *key*), with which an encrypted message becomes easy to read. For viewers without the key, the encrypted message should be very difficult or impossible to read.

Cryptanalysis is the science of code breaking; that is, the decoding of cryptographic messages by cleverly guessing the key or by exploiting weaknesses in the code that allow one to read the message without knowing the key at all. Cryptanalysis involves a blend of number theory, statistics, and computer programming,[1] along with whatever hints are made available from context and social engineering.[2]

Information in its original form, readable by anyone, is called *plaintext*. Information in encrypted form, readable only by someone who has the key, is called *ciphertext*.

❧

A famous ancient example of cryptography is the Caesar cipher, attributed to Julius Caesar.[3] The example described here is adapted slightly to English text.[4] Given an alphabetic plaintext string:

```
MEET US NEXT MONDAY
```

shift each letter forward in the alphabet by 3 letters (wrapping around again if necessary). The resulting ciphertext is

```
PHHW XV QHAW PRQGDB.
```

In practice one would remove spaces, since these give obvious clues to the structure of the message. The ciphertext then becomes

```
PHHWXVQHAWPRQGDB.
```

[1] Useful for brute force attacks that try every possible key.

[2] That is, persuading insiders to reveal information via deception or duress.

[3] See Suetonius, *De Vita Caesarum, Divus Julius*, Section 56. An English translation can be found in [35, p. 150].

[4] Caesar wrote in Latin, whose classical alphabet conflated the symbols $I = J$ and $U = V$ and omitted the symbol W.

To decrypt a message, one simply reverses the process, shifting each letter of the ciphertext backwards by 3 letters, restoring the original MEETUSNEXTMONDAY.

This cryptosystem uses modular arithmetic. Assigning values

$$\mathtt{A} = 0,\ \mathtt{B} = 1,\ \mathtt{C} = 2,\ \ldots,\ \mathtt{Z} = 25, \tag{17.1}$$

the Caesar cipher acts on each plaintext character p to produce a ciphertext character c via the formula

$$c \equiv p + 3 \mod 26.$$

This is not a secure cryptosystem. If an eavesdropper knows that the system uses a shift of the alphabet, then a brute force attack will break the code easily, since there are only 25 possible shifts to choose from (keeping in mind that a shift of zero, or, equivalently, 26 letters, would leave the original text unencrypted).

A more secure approach might be to choose a random permutation[5] of the letters $\{\mathtt{A}, \mathtt{B}, \mathtt{C}, \ldots, \mathtt{Y}, \mathtt{Z}\}$ instead of simply a shift. Again assigning values as in (17.1) one might select a permutation such as

$$(1\ 4\ 25\ 17\ 18)\ (2\ 21)(3\ 14\ 0\ 9\ 10\ 13)(5\ 19\ 23\ 12\ 6)(7\ 20\ 11)(15\ 24\ 22\ 16\ 8)^*$$

This substitution cipher might seem more secure, since there are $26! \approx 4 \cdot 10^{26}$ permutations for a brute force attacker to try. However, this kind of encryption is highly vulnerable to *frequency analysis*. For example, it is well known that the most common letters in typical English text are $E, T, A, \ldots$ in that order. One might then suspect that the most common character in the ciphertext represents E, the next most common T, and so on. Variations on this sort of analysis will quickly break any cipher based on a simple permutation of the alphabet.

❧

During the Middle Ages, a variation of the Caesar cipher known as the *Vigenère cipher*[7] was regarded as far more secure. Instead of shifting each letter of the plaintext by 3 steps, choose a finite sequence of numbers, say 17, 4, 3, then proceed to

- shift the 1st letter of the plaintext by 17 steps,
- shift the 2nd letter of the plaintext by 4 steps,
- shift the 3rd letter of the plaintext by 3 steps,
- shift the 4th letter of the plaintext by 17 steps,
- shift the 5th letter of the plaintext by 4 steps,

[5] A permutation is a rearrangement.

*This cycle notation is read in the following way: $1 \to 4 \to 25 \to 17 \to 18 \to 1$, $2 \leftrightarrow 21$, etc.

[7] Named for Blaise de Vigenère (1523-1596)

- and so on, until the plaintext is exhausted.

Following the scheme above, in which we view the English alphabet as symbols for the integers mod 26, we can remember the key $(17, 4, 3)$ as the letter string RED.

Example: Let's encrypt the message "Rain expected tomorrow" using the password RED. A simple way to do this is to write out the plaintext with the password repeated underneath, and then perform letter by letter 'addition' mod 26 using the table in Figure 17.1:

```
RAINEXPECTEDTOMORROW
REDREDREDREDREDREDRE
--------------------
IELEIAGIFKIGKSPFVUFA
```

The ciphertext IELEIAGIFKIGKSPFVUFA is then sent to the recipient, who uses the secret password RED to decrypt by *subtracting* mod 26 (or, alternatively, by "encrypting" with the password JWX, since $(\mathtt{R}, \mathtt{E}, \mathtt{D}) + (\mathtt{J}, \mathtt{W}, \mathtt{X}) \equiv (\mathtt{A}, \mathtt{A}, \mathtt{A})$) to restore the original plaintext.

In this way, Vigenère encryption can be thought of as a form of vector addition. The plaintext is broken into short vectors of letters the same length as the password vector. The password vector is then added to each plaintext vector (by componentwise addition mod 26) to encrypt, and subtracted to decrypt. This point to view leads to more elaborate vector encryption schemes (see Exercise 17.11.)

The Vigenère cipher is more secure than the Caesar cipher in a number of ways. First, the keyspace is much larger, since there are 26^n possible choices for an n letter password. Frequency analysis is also more difficult, especially when the password is long relative to the length of the plaintext.

The primary weakness of the Vigenère cipher lies in the repetition of the password. This weakness can be mitigated or even eliminated altogether by replacing a repeated key pattern, such as REDREDRED..., with a non-repeating *stream* of password letters, either generated by some formula, or taken from a completely random stream of letters.

For example, the *Autokey cipher* is a Vigenère-style cipher that uses a key stream consisting of an initial password followed by the message itself:

```
RAINEXPECTEDTOMORROW
REDRAINEXPECTEDTOMOR
--------------------
IELEEFCIZIIFMSPHFDCN
```

Alternatively, one might use *pseudorandom number generator* (see Section 29) to produce a sequence of random-looking values mod 26 that would serve as a

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
B	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A
C	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
D	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
E	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
F	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
G	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F
H	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
I	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
J	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
K	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
L	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
M	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
N	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
O	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
P	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Q	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
R	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
S	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
T	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
U	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
V	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
W	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
X	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
Y	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Z	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y

Figure 17.1: A Vigenère table for the English alphabet (addition mod 26).

password stream.[8] The generating method (with initial values) would be shared by all members of the conversation and kept secret from everyone else.

Most secure of all is the *One-Time Pad*, which uses a truly random list of numbers (mod 26 for an alphabetic cipher, mod 2 for encrypting a bit stream) to encrypt a stream of plaintext. Each party must have a book or file containing the random number list to be used. The numbers must be truly random,[9] and any part of the list used for encryption by *either party* must *never be used again*. If these conditions are satisfied then the resulting ciphertext is provably secure: no eavesdropper can decrypt the ciphertext without illicit access to the plaintext or the random number pad.

Why is a one-time pad so secure? Suppose a 100 character message is picked up to be analyzed. Imagine a superfast computer that checks all 26^{100} possible passwords. When the computer finds legible English the message has been decrypted, right? Wrong! Because of the *truly random* nature of the key, the computer would have to consider every possible 100 character key, resulting in every possible 100 character English text. As a result, there would be no reason to conclude that `ATTACK AT DAWN...` is the correct decryption, rather than `ATTACK AT NOON...` or `RETURN TO BASE...`. Randomness means that you can never know when a guess is correct. Any lapse in randomness by the key will ruin this quality, making a legible (English text) guess more plausible. This is why the key pad must be truly random for perfect (theoretical) security.[10]

Outside of certain extreme diplomatic or military contexts, the one-time pad is regarded as too inefficient (read "expensive"), because of the massive random key streams that must be generated, shared, and maintained. A cost-effective cryptosystem is one that is reasonably effective using a short password or key, even if that key needs to be changed from time to time. At present, the most commercially popular cryptosystem of this kind is the *Advanced Encryption Standard* (AES). AES typically uses a 256 bit key and a complicated succession of transformations (including non-linear transformations that are, in theory, difficult to invert). A brute force decryption of AES ciphertext would require up to $2^{256} \approx 10^{77}$ trials, which would defeat the resources of most attackers. A detailed technical description of how AES operates lies beyond the scope of this book, but can be found in [34], for example.

❧

So far all of the cryptosystems described have been *symmetric key* cryptosystems.

[8] In modern applications both plaintext and ciphertext are stored and processed as bit streams, that is, using only the values 0 and 1. Consequently, modern encryption usually involves arithmetic mod 2 rather than mod 26.

[9] Finding a source of true randomness is another complicated issue, but let's assume for the moment that this ideal can be achieved.

[10] However, in real life applications, social engineering usually trumps even a theoretically perfect cryptosystem.

This means that knowledge of the *same* password is necessary for both encryption and decryption of a message. Symmetric cryptosystems are powerful and efficient, but leave users with the problem of how to share the password.

For example, suppose Alice and Bob want to transmit secret messages over a long distance. In order to use AES, or any of the other ciphers described above, they have to meet ahead of time to decide on a common secret password. They can't choose a password over their long distance communication link, because it may be monitored. It's a chicken-and-egg problem: they need a secret code in order to choose and share the password for their secret code.

This is an especially serious problem in the case where Alice and Bob have never met. And yet this is precisely the situation for a customer who wants to open an account with a bank or retail site over the Internet. How can you open an account and choose a secure password with a site that you have never done business with before, while having confidence that your initial transaction is protected from eavesdroppers?

The solution to this puzzle is a different kind of encryption, called *public key* encryption, which uses *different keys* for encryption and decryption. Moreover (and this is fundamental), knowing one of those two keys must ***not*** allow you to guess the other one!

Suppose for the moment that this kind of encryption is possible. Alice will have two keys for her own cryptosystem: e_A ('A' for Alice), which she shares with the world, say in a public database, and d_A which she keeps secret. If Bob wants to send her a private message, he sends her the message:

$$\Big[\text{Hi Alice, It's Bob. Let's talk privately using the AES key: } \texttt{3bH\%jl@4rY!A;rx(}\ \Big]_{e_A} \tag{17.2}$$

where the notation $[M]_{e_A}$ denotes a message M encrypted with a public key e_A. Alice receives this encrypted message and decrypts it with her private key d_A. She can then have a longer and more efficient exchange of information with Bob using symmetric encryption with the secret key Bob suggested. Meanwhile, anyone else who intercepts Bob's initial encrypted message can't read it: even though they know e_A (since it's public), they don't know d_A, so they can't decrypt the gibberish.

Public key encryption allows for *authentication* as well as secrecy. When Alice receives the message (17.2), she knows that no one else can read it. But she does *not* know if she is really talking to Bob. Anyone can write to her in this secret manner claiming to be Bob. How can she authenticate the speaker's identity? One way is for her to respond with a message encrypted using Bob's public key e_B and see if he is able to decrypt it. However, since authentication is so important in general, this is usually accomplished all at once with a *secure signature.* In other words, instead of the initial message (17.2), Bob would more likely send

something of the form:

$$\Big[\text{Hi Alice, Here's a message for you from Bob:}\quad \Big[\text{Let's talk privately using the AES key: } \texttt{3bH\%jl@4rY!A;rx(}\ \Big]_{d_B}\Big]_{e_A}$$

Notice that Bob encrypted part of the inner message with his own private *decryption* key d_B. When Alice decrypts this message with her decryption key d_A, she sees:

Hi Alice, Here's a message for you from Bob: <gibberish>

She then looks up Bob's public key e_B in the database available to everyone, and uses e_B to decrypt the gibberish. If this decryption attempt results in a legible message, then she can be reasonably certain it was Bob who sent the original message, since only Bob knows the key d_B that partners with the public key e_B.

❧

All of this is wonderful in theory, but how can it be implemented in practice? Does a cryptosystem exist that uses two keys in this way, so that knowledge of one key doesn't give away the other? This problem was open for some time, but since the mid-20th century several workable public key cryptosystems have been developed. The two most famous rely on classical number theoretic ideas and are described below.

Before we proceed, however, one important note: All currently known public key cryptosystems are *less efficient*[11] for encrypting large data files than traditional symmetric encryption (such as AES). Moreover, public key systems are regarded as *less secure*,[12] since their security relies on algorithmic properties of arithmetic (or abstract algebra) that are still not well understood. As a result, public key cryptography is usually *not* used to encrypt long exchanges of information. Instead, a public key system is used only to begin a conversation, in order to negotiate a secure shared secret key for AES or some other fast and secure symmetric cryptosystem. On other words, use public key cryptosystems for symmetric key exchange, and then exchange the bulk of all data with symmetric encryption.

❧

Our first example of public key encryption is the Diffie-Hellman key exchange using $\mathbb{Z}_p$ exponentiation:

Choose a large prime p, and an element $g \in \mathbb{Z}_p$ (preferably a primitive root).[13]

[11] That is, it takes even a fast computer significantly longer to encrypt and decrypt a file.

[12] That is, it may be easier for an eavesdropper to crack the code, especially when longer messages are encrypted, giving an attacker more data to work with.

[13] Primitive roots are discussed in Section 20.

This information is public.

Alice chooses a secret exponent $a \in \{1, \ldots, p-1\}$.
Bob chooses a secret exponent $b \in \{1, \ldots, p-1\}$.

Alice publishes g^a, and Bob publishes g^b, all computed mod p.
Alice and Bob can each compute $g^{ab} \bmod p$, but no one else can do this, so g^{ab} is their secret key.

The security of this approach depends on the assumption that, for a large prime p, knowledge of g^a and g^b mod p (and g and p) does not allow for easy computation of g^{ab}. In other words, discrete logarithms are difficult to compute within a feasible timeframe even on the fastest computers available.[14]

❧

Perhaps the most famous and widely used public key exchange protocol is RSA (named for its inventors: Ron Rivest, Adi Shamir, and Leonard Adelman).[15] This is a method of public key encryption whose security is built on the principle (still true given current technology) that it is very easy to multiply a pair of large (200+ digit) prime numbers together, while it is very difficult (even on the fastest computers) to factor a large (400+ digit) composite number having no small prime factors.

RSA is implemented as follows. Each person chooses two large prime integers $p \neq q$. Set $n = pq$ and choose a value e such that

$$\gcd\Big(e, (p-1)(q-1)\Big) = 1. \tag{17.3}$$

Note that $\phi(n) = (p-1)(q-1)$.

The pair (n, e) make up the *public key*.

Next, solve for d in the equation

$$ed \equiv 1 \mod (p-1)(q-1).$$

The existence of d is guaranteed by the gcd condition (17.3). The value d is *private key*, and all of the values

$$d, p, q, (p-1)(q-1)$$

are kept secret.

[14] In practical applications the values of p, a, and b, should be very large, with $p > 10^{300}$ and $a, b > 10^{100}$.

[15] The RSA algorithm was discovered even earlier by Clifford Cocks, but his discovery was kept secret by his employer, the British intelligence agency GCHQ, until 1998, many years after RSA had been successfully commercialized by R, S, and A.

In order to send a private message that only Bob can read, Alice looks up Bob's public key (n_B, e_B) in a public database. To send a plaintext value X to Bob with security, Alice sends him the value

$$C \equiv X^{e_B} \mod n_B.$$

When Bob receives the message C, he can decrypt it with his private key by computing

$$C^{d_B} \equiv \left(X^{e_B}\right)^{d_B} \equiv X^{e_B d_B} \equiv X \mod n_B,$$

where the final congruence follows from Euler's Theorem 16.2. He can then respond to Alice using her public key, and so on.

The security of this algorithm lies in the difficulty of factoring n. Because n is a product of very large primes, an eavesdropper cannot easily determine p and q, and therefore cannot compute the value of d, even though the value of e is public.[16]

Because of its inefficiency, and because overuse of an RSA key may reduce its security over time, RSA is usually used only for key exchange and authentication. Alice will send a short message to Bob encrypted with RSA, suggesting a shared password for using symmetric encryption for future transactions. Once the shared password is agreed upon, all future communications can be performed using a more efficient and more secure symmetric encryption method, such as AES.

❧

In order to implement RSA we must have a method of generating large prime numbers (having several hundred digits). Since such large numbers are difficult to factor, this motivates the question of how very large primes can be obtained. Efficient methods for *primality testing* will be addressed in Section 31.

❧

Exercise 17.1. Another simple cipher related to the Caesar cipher is ROT-13, which moves each alphabetic letter exactly 13 steps forward along the alphabet. The following message describes a cute feature of ROT-13:

```
QBVGGJVPRGBTRGGURBEVTVANYGRKGONPX
```

(a) What is the plaintext corresponding to the expression above?
(b) Explain how ROT-13 is actually a simple version of the Vigenère cipher. What is the password?

[16] Because effective attacks on RSA have been discovered in certain special cases, real-world implementations should take extra precautions into account. See, for example, [10, p. 164-170].

Exercise 17.2. Encrypt the message `ATTACK AT DAWN` using:
(a) Caesar cipher.
(b) Vigenère cipher with the password `UML`.

Exercise 17.3. A message is encrypted by a Vigenère cipher using a 5 letter password, but the password has been forgotten. How many possible passwords are there?

Exercise 17.4. The Atbash cipher[17] is a substitution cipher implemented as follows:

$$\texttt{A} \leftrightarrow \texttt{Z} \quad \texttt{B} \leftrightarrow \texttt{Y} \quad \texttt{C} \leftrightarrow \texttt{X} \quad \cdots \quad \texttt{M} \leftrightarrow \texttt{N}$$

Like ROT-13 (see Exercise 17.1), this cipher is really just a visual disguise and offers no real security.
(a) Decrypt the Atbash message: `RGSRMPGSVIVULIVRZN`.
(b) Number the English letters `A-Z` with respective integers 0-25. Using the corresponding arithmetic mod 26, find a formula for the Atbash character $A(x)$ given x mod 26.

Exercise 17.5. Philbert decides to make his encryption stronger by first encrypting his message with the Vigenère password `UBEL`, and then encrypting it *again* with the password `HBMS`. Explain why this double encryption scheme is no more secure than a single encryption.

Exercise 17.6. After realizing his error (see the previous exercise), Philbert decides to make his encryption stronger by first encrypting his message with the Vigenère password `UBEL`, and then encrypting it *again* with the password `HBFMS`. Show that this double encryption is indeed more secure than a single encryption using either of those separate passwords alone.

Exercise 17.7. Eve discovered that Harold was exchanging secret messages with someone using the Vigenère cipher. In his desk drawer she found the message:

```
KIAIASYSLUBUHRTCBNLXHZLOHCIDFGCMNXMGKINDTJRKOFWTFIFFKWCIRCHNLWACEQ
```

Eve thinks that Harold is having an affair, and she suspects that his lover may have begun the message with the words "Dear Harold, ..." Use this guess to figure out what the message says.[18]

Exercise 17.8. The ciphertext `FSSJCIVDFOIKCMHBTM` was sent using the autokey cipher with password `HEY`. What was the original plaintext?

Exercise 17.9. The ciphertext `VKTMTGJEYEVZIPGVCAAUIYWIK` is sent using the autokey cipher. An eavesdropper believes that the last words of the message are `PROVESNOTHING`. What is the original plaintext? What is the key?

Exercise 17.10. Billy encrypted a message to Gilly using a one-time pad of random values they have shared in advance. Unfortunately, Billy also sent a message yesterday using the same stream of values from the one-time pad.

Eve has recorded both encrypted messages. She negates (multiplies by -1) the stream from yesterday and then re-encrypts today's ciphertext message with the result. This leaves Eve with a new stream of characters. Why is this easier to ultimately decode (given sufficient computing power) than a correctly implemented one-time pad message?

[17] Atbash first appeared in Biblical and Talmudic texts and was originally performed by exchanging letters of the 22 character Hebrew alphabet.

[18] This method of cryptanalysis is called a *known plaintext attack*. A better cryptosystem would not allow Eve to guess the key using known plaintext.

Exercise 17.11. An *affine cipher* is a generalization of the Vigenère cipher that uses matrix multiplication as well as vector addition.[19]
(a) Consider the encryption that takes each plaintext character x and returns a cipher text character $c(x)$ via the function

$$c(x) \equiv 3x + 14 \mod 26$$

How is the word "hello" encrypted using this scheme?
(b) What is the formula for decrypting ciphertext characters from part (a)?
(c) Why is the alternative ciphertext function $c(x) \equiv 4x + 14 \mod 26$ a really bad idea?
(d) Consider the encryption scheme that takes breaks plaintext "hello there" into pairs "he ll ot he re" and then encrypts each pair xy, viewed as a vector with two coordinates, via the function

$$c(x,y) \equiv \begin{bmatrix} 2 & 19 \\ 7 & 11 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 12 \\ 5 \end{bmatrix} \mod 26$$

using matrix multiplication. How is the phrase "hello there" encrypted using this scheme?
(e) What is the secret key in part (d)?
(f) What is the formula for decrypting ciphertext character pairs from part (d)?
(g) Show that a matrix

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

is acceptable for use in an affine cipher if and only if $ad - bc \not\equiv 0 \mod 26$.

Exercise 17.12. Alice wants to communicate with Bob using a common secret password. The common password will be generated by a Diffie-Hellman key exchange using the prime $p = 97$ and base $g = 21$. Alice's secret exponent is $a = 46$. Bob's secret exponent is $b = 63$. What is their common secret password value?*

Exercise 17.13. Suppose $p = 17$ and $q = 23$ are used to construct an RSA public key with encryption exponent $e = 5$. What is the corresponding decryption exponent d?

Exercise 17.14. To implement a public/private key pair for RSA, Alice chooses two large distinct primes p and q, and then computes

$$n = pq = 17292864462617,$$

along with

$$(p-1)(q-1) = 17292856036560.$$

She then chooses an encryption exponent e_A (co-prime to $(p-1)(q-1)$) and computes the decryption exponent d_A by solving

$$e_A d_A \equiv 1 \mod 17292856036560.$$

By accident, however, she published as her public key the three values,

$$[n, e_A, (p-1)(q-1)] = [17292864462617, e_A, 17292856036560].$$

What are p and q? Show how your answer came from your knowledge of pq and $(p-1)(q-1)$.

Hint: Find formulas for p and q in terms of pq and $(p-1)(q-1)$. Then use a computer or calculator to compute p and q.

[19] An affine cipher that is implemented with matrix multiplication only, using no translation vector, is sometimes called a *Hill cipher*.

*The numbers are kept small so that this problem can be solved without a calculator. In a real life application the prime p and the secret exponents a and b would be very large integers.

18 Lagrange's root theorem

A function $f(x)$ is a *polynomial* of *degree* n, where $n \geq 1$, if $f(x)$ has the form

$$f(x) = c_n x^n + c_{n-1} x^n + \cdots + c_1 x + c_0,$$

where $c_n \neq 0$ (or, in the context of modular arithmetic, $c_n \not\equiv 0$). A constant function is said to be a polynomial of degree 0.

A common first step toward factoring a polynomial is to look for roots of that polynomial. The fact that roots of a polynomial yield linear factors is summarized as follows.

Theorem 18.1 (Factor Theorem)**.** *Suppose that $f(x)$ is a non-constant polynomial. If $f(r) = 0$ for some value r then*

$$f(x) = (x - r)g(x),$$

where $deg(g) = deg(f) - 1$.

Note the ambiguity in the statement of the previous theorem. We did not specify if f is a polynomial over the real numbers, or over complex numbers, or in some modulus. In fact, the theorem holds in all of these cases. Note carefully how each step of the following proof is valid whether the coefficients of f (and the value r) are real, or complex, or are values in some $\mathbb{Z}_m$ (where the symbols $=$ or $\equiv$ are used as required in each context).

Proof. Suppose that f is a polynomial of degree $n \geq 1$ and that $f(r) = 0$. The polynomial f has a expression of the form

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0,$$

where $a_n \neq 0$. Since $f(r) = 0$, we have

$$0 = a_n r^n + a_{n-1} r^{n-1} + \cdots + a_1 r + a_0.$$

Subtracting these two equations, we obtain

$$f(x) = a_n[x^n - r^n] + a_{n-1}[x^{n-1} - r^{n-1}] + \cdots + a_1[x - r].$$

For each k, it follows from the geometric sum formula (3.2) that

$$x^k - r^k = (x - r)(x^{k-1} + rx^{k-2} + \cdots + r^{k-2}x + r^{k-1}) = (x - r)g_k(x),$$

where we denote the longer factor of each $x^k - r^k$ by $g_k(x)$, a polynomial of degree $k - 1$. It now follows that

$$f(x) = (x - r)[a_n g_n(x) + a_{n-1} g_{n-1}(x) + \cdots + a_1] = (x - r)g(x),$$

where $g(x)$ is a polynomial of degree $n - 1$. ❑

In the next theorem we focus attention on the real (or complex) numbers only.

Theorem 18.2. *Suppose that $a_0, \ldots, a_n \in \mathbb{R}$ (or $\mathbb{C}$), where $a_n \neq 0$ and $n \geq 1$. Then the polynomial equation*

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0 \tag{18.1}$$

has at most n roots over $\mathbb{R}$ (or $\mathbb{C}$).

Proof. The theorem is trivial if $n = 1$, since the equation $a_1 x + a_0 = 0$ has exactly one root, namely, $x = -\frac{a_0}{a_1}$.

Suppose that $n \geq 2$ and that the theorem holds for polynomials of degree at most $n - 1$. If the equation (18.1) has *no* roots at all, we are done. If, instead, a root r exists, then the Factor Theorem 18.1 implies

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = (x - r) g(x),$$

where $g(x)$ has degree at most $n - 1$. By the induction assumption, $g(x)$ has at most $n - 1$ roots. Moreover, if s is any root of the equation (18.1), then

$$(s - r) g(s) = 0,$$

so that either $s = r$ or $g(s) = 0$ (or both). Therefore, there are at most n possible values for the root s. ❑

Notice that we never made explicit mention of $\mathbb{R}$ or $\mathbb{C}$ in the previous proof. The same proof would work just as well mod m, except in two steps: For the case of degree 1 it is necessary that a_1^{-1} exists mod m, and for the induction step we would need to know that

$$(s - r) g(s) \equiv 0 \quad \text{implies that } s \equiv r \text{ or } g(s) \equiv 0.$$

This implication would *not* always hold in a composite modulus. However, it *is* always true in a *prime* modulus. Therefore, the previous proof also verifies the following theorem.[1]

Theorem 18.3 (Lagrange's Root Theorem)**.** *Let p be a prime integer, and suppose that $a_0, \ldots, a_n \in \mathbb{Z}$, where $a_n \not\equiv 0 \bmod p$. Then the polynomial equation*

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \equiv 0 \mod p$$

has at most n roots mod p.

❧

[1] Named for Joseph-Louis Lagrange (1736–1813).

Exercise 18.1. Find all of the cube roots of 1 mod 4, as well as mod 5 and mod 7.

Exercise 18.2. Find all of the solutions to the equation $x^4 - 1 \equiv 0 \bmod 7$, as well as mod 13.

Exercise 18.3. Show that the values $1, 25, 31, 34$ are each solutions to the equation $x^4 - 1 \equiv 0 \bmod 39$. Are there any more solutions?

Exercise 18.4. Let $p \neq q$ be prime numbers. Prove that a quadratic equation has at most 4 distinct solutions mod pq.
Hint: The Chinese remainder theorem is helpful here.

Exercise 18.5. Factor $x^2 + x + 1 \bmod 3$.

Exercise 18.6. Factor $x^3 + 2x^2 + 4 \bmod 5$.

Exercise 18.7. Factor $x^4 - 1$ into linear factors mod 5. Any surprises? What about $x^6 - 1 \bmod 7$?

Exercise 18.8. Determine if each the following polynomials can be factored mod 2. If so, factor them into their irreducible[2] components.

(a) $x^2 + 1 \bmod 2$
(b) $x^2 + x + 1 \bmod 2$
(c) $x^3 + x^2 + x + 1 \bmod 2$
(d) $x^3 + x + 1 \bmod 2$
(e) $x^4 + 1 \bmod 2$
(f) $x^4 + x^2 + 1 \bmod 2$

[2] A polynomial is *irreducible* if it cannot be factored into polynomials of lower degree.

19 Polynomial equations and Hensel's lemma

In the Section 18 we considered polynomial equations mod p, where is p is prime. In this case, Lagrange's Theorem asserted that a polynomial of degree n has at most n roots. However, we have also seen that polynomials can have many more roots over composite moduli. For example, the equation

$$x^2 - 1 \equiv 0 \tag{19.1}$$

has 4 roots mod 8. Indeed, in spite of being a mere quadratic (degree 2) polynomial equation, there are 8 solutions to (19.1) mod 24, and 16 solutions mod 120.

❧

In order to characterize solutions to polynomial equations in a composite modulus, we first consider the case where the modulus is a prime power p^k.

Evidently, if a number is divisible by p^{k+1}, then it is also divisible by p^k. This observation implies the following proposition.

Proposition 19.1. *Let p be a positive prime integer, and let $k \in \mathbb{N}$. If $f(r) \equiv 0 \bmod p^{k+1}$ then $f(r) \equiv 0 \bmod p^k$.*

It follows that, in order to find roots of a polynomial equation $f(x) \equiv 0 \bmod p^{k+1}$, we should first find the roots mod p^k. This leads in turn to the following question: when does the converse of Proposition 19.1 hold? When does a root of $f(x) \equiv 0 \bmod p^k$ "lift" to a solution mod p^{k+1}? The answer, together with a lifting algorithm, is given by the next theorem.[1]

Theorem 19.2 (Hensel's Lemma)**.** *Let $f(x)$ be a polynomial of positive degree, and suppose that $f(r) \equiv 0 \bmod p^k$, so that $c = f(r)/p^k$ is an integer.*

- *If $f'(r) \not\equiv 0 \mod p$ then*

$$f(r + tp^k) \equiv 0 \mod p^{k+1}$$

iff

$$t \equiv -c[f'(r)]^{-1} \mod p.$$

[1] Named for German mathematician Kurt Hensel (1861–1941), who developed the theory of p-adic numbers.

- *If* $f'(r) \equiv 0 \mod p$ *and* $f(r) \equiv 0 \mod p^{k+1}$, *then*
$$f(r + tp^k) \equiv 0 \mod p^{k+1} \quad \textit{for all } t.$$
- *If* $f'(r) \equiv 0 \mod p$ *and* $f(r) \not\equiv 0 \mod p^{k+1}$, *then*
$$f(r + tp^k) \not\equiv 0 \mod p^{k+1} \quad \textit{for all } t.$$

Note that, in the first case of Hensel's Lemma above, the value of $c = f(r)/p^k$ is computed as an integer (not yet in a modulus), while $[f'(r)]^{-1}$ is computed mod p to produce a value from the list $0, 1, \ldots, p-1$. The value $r + tp^k$ is then computed mod p^{k+1}.

Proof. Let $f(x)$ be a polynomial of positive degree, and suppose that $f(r) \equiv 0 \mod p^k$, so that $c = f(r)/p^k$ is an integer.

For $t \in \mathbb{Z}$, Taylor's theorem asserts that

$$f(r + p^k t) \;=\; f(r) + f'(r)p^k t + \frac{f''(r)}{2!}p^{2k}t^2 + \text{ additional terms divisible by } p^{k+1}.$$

Note that, since f is a polynomial with integer coefficients, we have

$$n! | f^{(n)}(r)$$

for each of the derivatives $f^{(n)}$ (see Exercise 19.9), so that

$$f(r + p^k t) \;\equiv\; f(r) + f'(r)p^k t \mod p^{k+1}. \tag{19.2}$$

It follows that

$$f(r + p^k t) \;\equiv\; 0 \mod p^{k+1} \Leftrightarrow f(r) + f'(r)p^k t \equiv 0 \mod p^{k+1}.$$

If $f'(r) \not\equiv 0 \bmod p$, then $f'(r) \not\equiv 0 \bmod p^{k+1}$ as well, so that

$$\begin{aligned} f(r + p^k t) \;\equiv\; 0 \mod p^{k+1} &\Leftrightarrow f(r) + f'(r)p^k t \equiv 0 \mod p^{k+1}. \\ &\Leftrightarrow cp^k + f'(r)p^k t \equiv 0 \mod p^{k+1}. \\ &\Leftrightarrow c + f'(r)t \equiv 0 \mod p. \\ &\Leftrightarrow t \equiv -c[f'(r)]^{-1} \mod p. \end{aligned}$$

Meanwhile, if $f'(r) \equiv 0 \bmod p$, then $p | f'(r)$, so that $p^{k+1} | f'(r)p^k$. The identity (19.2) now becomes

$$f(r + p^k t) \;\equiv\; f(r) \mod p^{k+1},$$

so that $f(r + p^k t)$ is either *always* zero or *never* zero mod p^{k+1}, depending on the value of $f(r)$. ❑

Example: Find all solutions to the equation $x^2 \equiv 1 \bmod 8$ and mod 16.

The unique solution to $x^2 - 1 \equiv 0 \bmod 2$ is $x \equiv 1$. Setting $f(x) = x^2 - 1$, we have $f'(x) = 2x \equiv 0 \bmod 2$.

Since $f(1) \equiv 0 \bmod 4$, it follows from Hensel's Lemma (applying the case of the zero derivative) that $f(1) \equiv f(3) \equiv 0 \bmod 4$.

Since $f(1) \equiv f(3) \equiv 0 \bmod 8$, it follows similarly that the solution $x \equiv 1$ lifts to solutions $x \equiv 1$ and $1 + 4 \equiv 5$, while the solution $x \equiv 3$ lifts to solutions $x \equiv 3$ and $3 + 4 \equiv 7$, giving us four solutions $x \equiv 1, 3, 5, 7 \bmod 8$.

Since $f(1) \equiv f(7) \equiv 0 \bmod 16$, these roots mod 8 lift to solutions

$$1, \quad 1 + 8 = 9, \quad 7, \quad 7 + 8 = 15$$

mod 16. However, since $f(3) \equiv f(5) \not\equiv 0 \bmod 16$, these solution from mod 8 do not lift to solutions mod 16.

Example: Find all solutions to the equation $x^2 + x + 1 \equiv 0 \bmod 49$.

The solutions to $x^2 + x + 1 \equiv 0 \bmod 7$ are $x \equiv 2$ and $x \equiv 4$. Observe that

$$f(2) = 7 \quad \text{and} \quad f(4) = 21.$$

Since $f'(x) = 2x + 1$, we also have

$$f'(2) \equiv 5 \quad \text{and} \quad f'(4) = 9 \equiv 2 \mod 7,$$

so that

$$[f'(2)]^{-1} \equiv 3 \quad \text{and} \quad [f'(4)]^{-1} \equiv 4 \mod 7.$$

To lift the solution $x \equiv 2 \bmod 7$ to a solution mod 49, set $c = f(2)/7 = 1$ and compute $2 + 7t$, where

$$t \equiv -c[f'(2)]^{-1} \equiv -3 \equiv 4 \mod 7,$$

so that $x \equiv 2 + 7 \cdot 4 \equiv 30 \bmod 49$.

To lift the solution $x \equiv 4 \bmod 7$ to a solution mod 49, set $c = f(4)/7 = 3$ and compute $4 + 7t$, where

$$t \equiv -c[f'(4)]^{-1} \equiv -3 \cdot 4 \equiv -12 \equiv 2 \mod 7,$$

so that $x \equiv 4 + 7 \cdot 2 \equiv 18 \bmod 49$.

We conclude that the solutions to $x^2 + x + 1 \equiv 0 \bmod 49$ are $x \equiv 30$ and $x \equiv 18$.

❧

A combination of Hensel's Lemma 19.2 with the Chinese Remainder Theorem 11.1 yields an algorithm for solving a polynomial equation modulo m, provided we have the factorization of m.

To solve the equation $f(x) \equiv 0 \bmod m$, where m has a prime power factorization

$$m = p_1^{k_1} \cdots p_s^{k_s},$$

proceed with the following algorithm:

Step 1: Find solutions to $f(x) \equiv 0 \bmod p_i$ for each i.
If there are no solutions mod p_i for some i,
then there are no solutions mod m.
If there are solutions mod p_i for every i,
proceed to the next step.

Step 2: Use Hensel's Lemma 19.2 to lift the solutions mod p_i to solutions mod p_i^k, wherever possible.

Step 3: Use the Chinese Remainder Theorem 11.1 to combine values of x modulo each $p_i^{k_i}$ in order to obtain corresponding solutions mod m.

Example: Find all solutions to the equation $x^2 + x + 1 \equiv 0 \bmod 637$.

Since $637 = 49 \cdot 13$ we combine solutions to the equation mod 49 with solutions mod 13 using the Chinese remainder theorem.

Recall from the example above that the solutions mod 49 are $x \equiv 18$ or 30. By inspection we can determine that the solutions mod 13 are $x \equiv 6$ or 7. Applying the Chinese remainder theorem to the 4 solution combinations, we find that

$(x \bmod 13, x \bmod 49)$	$\Rightarrow$	$(x \bmod 637)$
$(6,\ 18)$	$\Rightarrow$	214
$(6,\ 30)$	$\Rightarrow$	422
$(7,\ 18)$	$\Rightarrow$	410
$(7,\ 30)$	$\Rightarrow$	618

so that $x \equiv 214$, 422, 410, or 618 mod 637.

❧

Exercise 19.1. Find all solutions to the equation $x^2 + x + 3 \equiv 0$

(a) mod 25.
(b) mod 125.
(c) mod 250.
(d) mod 375.

Exercise 19.2. Find all solutions to the equation $x^2 + x + 1 \equiv 0$

(a) mod 13.
(b) mod 27.
(c) mod 39.
(d) mod 169.

Exercise 19.3. Find all solutions to the equation $x^4 + x + 23 \equiv 0$
(a) mod 23.
(b) mod 25.
(c) mod 575.

Exercise 19.4. Find all solutions to the equation $x^2 - 4 \equiv 0$
(a) mod 999.
(b) mod 1001.

Exercise 19.5. Prove that there are 16 distinct solutions to the equation (19.1) mod 120. Your proof should explain the existence of these solutions without actually finding them.

Exercise 19.6. Find a general formula for the solutions to the equation $x^2 \equiv 1 \bmod 2^n$, for $n \geq 3$.

Exercise 19.7. Find a general formula for the solutions to the equation $x^3 \equiv 1 \bmod 3^n$, for $n \geq 2$.

Exercise 19.8. **(a)** Let p be an odd prime and $n \in \mathbb{N}$. Prove that $x \equiv \pm 1$ are the only solutions to the equation $x^2 \equiv 1 \bmod p^n$.
(b) Let $m > 1$ be an odd integer with prime factorization

$$m = p_1^{a_1} \cdots p_k^{a_k}.$$

Prove that the equation $x^2 \equiv 1 \bmod m$ has exactly 2^k distinct solutions.

Exercise 19.9. Let $f(x)$ be a polynomial function of x having integer coefficients, and let n be a positive integer. Here are some steps leading to a proof that $n! | f^{(n)}(r)$ for any integer r.
(a) For $m \in \mathbb{N}$, use that fact that $\binom{m+n}{n}$ is an integer to show that $n!$ divides the number

$$(m+1)(m+2)\cdots(m+n).$$

(b) Suppose that $g(x) = x^k$ for some integer $k \geq 0$. Show that $n! | g^{(n)}(0)$.
(c) Use part (b) to show that $n! | f^{(n)}(0)$.
(d) Apply part (c) to the function $h(x) = f(x+r)$ to complete the proof.

20 Primitive roots

Euler's Theorem 16.2 tells us that, if $u \in U_n$, then $u^{\phi(n)} \equiv 1 \bmod n$. However, it is often the case that $u^k \equiv 1$ for smaller exponents k. For example, it is always true that

$$(n-1)^2 \equiv (-1)^2 \equiv 1 \mod n.$$

For $u \in U_n$, define the *order* of u to be the smallest positive integer k such that $u^k \equiv 1 \mod n$. This value is denoted $\mathrm{ord}_n(u)$.

For example, if we compute the powers of 2 mod 7 we have

$$2^1 \equiv 2, \qquad 2^2 \equiv 4, \qquad 2^3 \equiv 1 \mod 7,$$

so that $\mathrm{ord}_7(2) = 3$.

Proposition 20.1. *Let $\alpha = \mathrm{ord}_n(u)$. For all m, $u^m \equiv 1 \bmod n$ if and only if $\alpha | m$.*

Proof. If $\alpha | m$ then $m = \alpha k$ for some integer k, so that

$$u^m \equiv u^{\alpha k} \equiv (u^\alpha)^k \equiv 1^k \equiv 1 \mod n.$$

To prove the converse, suppose that $u^m \equiv 1$. Write $m = \alpha q + r$, where $0 \leq r < \alpha$. Since $u^\alpha \equiv 1$, we have

$$1 \equiv u^m \equiv u^{\alpha q + r} \equiv (u^\alpha)^q u^r \equiv u^r \mod n.$$

If $0 < r < \alpha$, this violates the minimality of the order α. Therefore $r = 0$, and $\alpha | m$. ❑

By Euler's Theorem, $u^{\phi(n)} \equiv 1 \bmod n$ for every $u \in U_n$. It follows from the previous proposition that

$$\mathrm{ord}_n(u) | \phi(n)$$

for all $u \in U_n$.

In the discussion that follows we will often focus on arithmetic modulo a prime p. In this case we know that $\mathrm{ord}_p(u) | (p-1)$ for all units $u \in U_p$.

❧

If we take the powers of 2 mod 5 we have

$$2^1 \equiv 2 \qquad 2^2 \equiv 4 \qquad 2^3 \equiv 3 \qquad 2^4 \equiv 1 \mod 5$$

exhausting the units mod 5, so that $\text{ord}_5(2) = 4$. In other words, 2 generates the entire multiplicative group U_5, which turns out to be a cyclic group. A similar computation reveals that 3 generates the group U_7; that is, $\text{ord}_7(3) = 6 = \phi(7)$, so that every unit is a power of 3 mod 7.

Definition 20.2. A unit r mod n is a *primitive root* if

$$U_n = \{1, r, r^2, \ldots, r^{\phi(n)-1}\}.$$

In other words, r is a primitive root for n iff $\text{ord}_n(r) = \phi(n)$. In this case, the group U_n is *cyclic*, with r as a *generator*.

For example, we have seen that

$$U_5 = \{1, 2, 2^2, 2^3\}.$$

A few short computations also verify that

$$U_7 = \{1, 3, 3^2, 3^3, 3^4, 3^5\}.$$

Some moduli have primitive roots, and some do not. We will show (eventually) that every prime modulus p has at least one primitive root.

❧

It is not always easy to find a primitive root, when they exist at all. However, once we have found a primitive root r mod n, it is easy to find the others.

Proposition 20.3. *If* $\text{ord}_n(u) = \alpha$ *then* u^k *has order* α *if and only if* $\gcd(k, \alpha) = 1$.

Proof. Let $d = \gcd(k, \alpha)$, and let $\beta = \text{ord}_n(u^k)$. We need to show that $\beta = \alpha$ if and only if $d = 1$.

If $d > 1$ then $k = dx$ and $\alpha = dy$, where $x, y \in \mathbb{Z}$, and where $1 \leq y < \alpha$. In this case,

$$(u^k)^y \equiv u^{ky} \equiv u^{dxy} \equiv u^{\alpha x} \equiv (u^\alpha)^x \equiv 1^x \equiv 1 \mod p.$$

The minimality of the order β now implies that $\beta \leq y < \alpha$.

Suppose instead that $d = 1$. Since $\beta = \text{ord}_n(u^k)$, we have

$$u^{k\beta} = (u^k)^\beta = 1.$$

It follows from Proposition 20.1 that $\alpha | k\beta$. Since $d = 1$, the values α and k are relatively prime, so that $\alpha | \beta$.

Meanwhile,

$$(u^k)^\alpha \equiv (u^\alpha)^k \equiv 1^k \equiv 1 \mod p,$$

so that $\beta | \alpha$, again by Proposition 20.1. It now follows that $\beta = \alpha$. ❑

Corollary 20.4. *If $\mathbb{Z}_n$ has a primitive root r, then the primitive roots for $\mathbb{Z}_n$ are precisely those units r^k where $\gcd(k, \phi(n)) = 1$. In particular, there are $\phi(\phi(n))$ primitive roots mod n.*

Proof. If r is a primitive root mod n, then $\mathrm{ord}_n(r) = \phi(n)$. The previous proposition then implies that $\mathrm{ord}_n(r^k) = \phi(n)$ iff k is relatively prime to $\phi(n)$, giving $\phi(\phi(n))$ distinct cases. Since every $u \in U_n$ has the form r^k for some k (because r is primitive), this exhausts all possibilities for primitive roots mod n. ❑

Assuming that there is at least one primitive root modulo a prime p (to be shown in Section 21), it follows there are exactly $\phi(p-1) = \phi(\phi(p))$ primitive roots modulo p.

For example, since $\phi(\phi(13)) = \phi(12) = 4$, there are four primitive roots mod 13. The reader can verify that 2 is primitive mod 13. It follows from Corollary 20.4 that the primitive roots mod 13 are

$$2^1,\ 2^5,\ 2^7,\ 2^{11} \ \equiv\ 2,\ 6,\ 11,\ 7 \ \text{(respectively)} \mod 13.$$

❧

The following lemma is useful for generating elements of higher order, given elements of smaller order.

Lemma 20.5 (Multiplicative Lemma)**.** *Suppose that $\mathrm{ord}_n(a) = \alpha$ and $\mathrm{ord}_n(b) = \beta$. If $\gcd(\alpha, \beta) = 1$, then $\mathrm{ord}_n(ab) = \alpha\beta$.*

Proof. Let $\gamma = \mathrm{ord}_n(ab)$. Evidently

$$(ab)^{\alpha\beta} = a^{\alpha\beta}b^{\alpha\beta} = (a^\alpha)^\beta(b^\beta)^\alpha = 1^\beta 1^\alpha = 1,$$

so that $\gamma | \alpha\beta$, by Proposition 20.1. Meanwhile,

$$1 = 1^\alpha = ((ab)^\gamma)^\alpha = (ab)^{\alpha\gamma} = a^{\alpha\gamma}b^{\alpha\gamma} = b^{\alpha\gamma},$$

so that $\beta | \alpha\gamma$, again by Proposition 20.1. Since $\gcd(\alpha, \beta) = 1$, it follows that $\beta | \gamma$. By a similar and symmetrical argument, we also have $\alpha | \gamma$. Again, since $\gcd(\alpha, \beta) = 1$, we have $\alpha\beta | \gamma$. It now follows that $\gamma = \alpha\beta$. ❑

This lemma can be useful for finding primitive roots. For example, it is easy to see that 2 has order 5 mod 31, since $2^5 \equiv 32 \equiv 1 \mod 31$. And we always know that -1 has order 2 modulo an odd prime. If we can find an element c of order 3, the Multiplicative Lemma 20.5 implies that $-2c$ will have order 30, so it is primitive. Looking at a list of cubes:

$$1, 8, 27, 64, 125, 216, 343, \ldots$$

we see that $5^3 = 125 \equiv 1 \bmod 31$, so that $-10 \equiv 21$ is a primitive root mod 31. By Corollary 20.4, the complete list of primitive roots mod 31 will be congruent to some 21^k for k relatively prime to 30, that is, the values:

$$21, 21^7, 21^{11}, 21^{13}, 21^{17}, 21^{19}, 21^{23}, 21^{29} \mod 31$$

or (listed in the same order):

$$21,\ 11,\ 12,\ 22,\ 24,\ 13,\ 17,\ 3 \mod 31.$$

❧

In the exercises that follow you will see that 2 is often (though not always) a primitive root modulo an odd prime p. The question then arises, are there infinitely many primes p such that 2 is primitive mod p? The answer to this simple question is unknown. More generally, the Artin Conjecture[1] asserts that, if a is a positive integer that is not a perfect square, then there are infinitely many primes p such that a is primitive mod p.

❧

Exercise 20.1. What is $\mathrm{ord}_{23}(2)$? $\mathrm{ord}_{23}(3)$? $\mathrm{ord}_{23}(5)$?

Exercise 20.2. What is $\mathrm{ord}_{10}(13)$? What is $\mathrm{ord}_{13}(10)$?

Exercise 20.3.
(a) Is there a primitive root mod 10?
(b) Is there a primitive root mod 12?

Exercise 20.4. What are the orders of each of $1, 2, 3, 4, 5, 6$ mod 7? Are there any primitive roots?

Exercise 20.5. What are the orders of each of $1, 2, 4, 5, 7, 8$ mod 9? Are there any primitive roots? Why do we not also compute the orders of 3 and 6 mod 9?

Exercise 20.6. What are the orders of the four elements of U_8? Are there any primitive roots?

Exercise 20.7. What are the orders of the eight elements of U_{15}? Are there any primitive roots?

Exercise 20.8. Find all of the primitive roots mod 11.

Exercise 20.9. Find a primitive root mod 41.

Exercise 20.10. Let $u \in U_m$.
(a) Prove that $\mathrm{ord}_m(u^{-1}) = \mathrm{ord}_m(u)$.
(b) Prove that r is primitive mod m iff r^{-1} is primitive mod m.

Exercise 20.11. Suppose that p is an odd prime, and let r be a primitive root mod p.
(a) Prove that

$$r^{\frac{p-1}{2}} \equiv -1 \mod p.$$

(b) Prove that, if $p \equiv 1 \bmod 4$, then $-r$ is also a primitive root mod p.
(c) Is this still true when $p \equiv 3 \bmod 4$?

[1] Posed by Emil Artin (1898-1962) in 1927.

Exercise 20.12. Suppose that m is an integer with a primitive root. Let M denote the product of all the units mod m.
(a) Prove that $M \equiv -1 \bmod m$.
(b) Find an example of an integer m (lacking a primitive root) where the product $M \not\equiv -1 \bmod m$.
Remark: If m is prime this is equivalent to Wilson's theorem.

Exercise 20.13. Find positive integers n, a, b such that:
(a) $\text{ord}_n(ab) < \min\{\text{ord}_n(a), \text{ord}_n(b)\}$.
(b) $\text{ord}_n(ab) = \text{ord}_n(a) = \text{ord}_n(b)$.

Exercise 20.14. Suppose that $\text{ord}_n(a) = 4$ and $\text{ord}_n(b) = 6$. What is $\text{ord}_n(ab^2)$?

Exercise 20.15. Suppose that $\mathbb{Z}_m$ has a primitive root r, and let $k \in \mathbb{N}$. Prove that

$$\text{ord}_m(r^k) = \frac{\phi(m)}{\gcd(k, \phi(m))}.$$

Exercise 20.16. Suppose that $\mathbb{Z}_m$ has a primitive root, and let $\alpha \in \mathbb{N}$. Prove that, if $\alpha | \phi(m)$, then there are exactly $\phi(\alpha)$ units in U_m of order α.

Exercise 20.17. Suppose that p is a prime integer. The following steps lead to a proof that there are an infinite number of primes q such that $q \equiv 1 \bmod p$, a special case of Dirichlet's theorem (see Section 9).
To begin, suppose that there are only a finite number of such primes $q_1, \ldots, q_m$ (or none at all). Let

$$M = q_1 \cdots q_m p,$$

or set $M = p$ if no such q_i exist.
Next, let q be a prime factor of the number

$$M^{p-1} + M^{p-2} + \cdots + M + 1.$$

(a) Prove that $q \neq p$.
(b) Prove that $q \neq q_i$ for each i.
(c) Prove that $M^p \equiv 1 \bmod q$.
(d) Let $\alpha = \text{ord}_q(M)$. Prove that $\alpha = p$.
(e) Prove that $q \equiv 1 \bmod p$.
It follows from parts (b) and (e) that the finite list $q_1, \ldots, q_m$ can never be complete, so that there are an infinite number of primes q such that $q \equiv 1 \bmod p$.

❧

A rational number a/b is said to have a *finite* decimal expansion if its decimal expansion has only a finite number of non-zero digits. Examples include:

$$\frac{1}{2} = 0.5 \qquad \frac{3}{8} = 0.375 \qquad \frac{12}{5} = 2.4 \qquad \frac{8443}{10000} = 0.8443$$

A rational number is said to have a *purely periodic* decimal expansion if its decimal expansion consists of an infinitely repeating finite pattern of digits, such as:

$$\frac{1}{9} = 0.1111\ldots \qquad \frac{3}{11} = 0.272727\ldots \qquad \frac{2}{37} = 0.054054054\ldots \qquad \frac{8443}{9999} = 0.844384438443\ldots$$

The *period* is the length of the repeating pattern of digits. For example, the expansion for $\frac{2}{37}$ has period 3.

If the decimal expansion of a/b eventually repeats, it is simply called *repeating*, as in the case of:

$$\frac{3}{44} = 0.06818181... \qquad \frac{2003}{420} = 4.7690476190476190476... = 4.76\overline{904761}$$

Exercise 20.18. Let $m > 1$ be an integer.
(a) Prove that $\frac{1}{m}$ has a finite decimal expansion iff $m = 2^a 5^b$ for some integers a, b.
(b) Prove that, if $\frac{1}{m}$ has a purely periodic decimal expansion with period k, then

$$\frac{1}{m} = \frac{n}{10^k - 1}$$

for some integer n.
Hint: The formula (3.3) can be helpful here. See also Exercises 3.5 and 3.4.
(c) Prove that, if $\frac{1}{m}$ has a purely periodic decimal expansion, then $\gcd(m, 10) = 1$.
(d) Prove that, if $\gcd(m, 10) = 1$, then $\frac{1}{m}$ has a purely periodic decimal expansion with period equal to $\mathrm{ord}_m(10)$.

Exercise 20.19. Suppose that an integer $m > 1$ has prime factorization

$$m = p_1^{a_1} \cdots p_s^{a_s},$$

and that $\gcd(m, n) = 1$. What conditions on the factorization of m determine whether the decimal expansion of n/m is finite, purely periodic with period k, or periodic with period k after a (non-repeating) prefix of j digits?
Hint: Do Exercise 20.18 before attempting this one.

Exercise 20.20. Suppose that $m > 1$ and that $\gcd(n, m) = 1$. Prove that the period of the repeating part (if any) of the decimal expansion for $\frac{n}{m}$ must divide $\phi(m)$.
Hint: Do Exercise 20.18 before attempting this one.

In particular, the period of a repeating decimal is never greater than the Euler-ϕ value of its denominator when written as a fraction in lowest terms.

21 The existence of primitive roots

We now present a series of lemmas leading to a proof that, if p is prime, then $\mathbb{Z}_p$ has a primitive root.

It is a consequence of Fermat's Theorem 15.1 that the polynomial equation

$$x^{p-1} - 1 = 0$$

has at least $p-1$ distinct solutions mod p. By Lagrange's Theorem 18.3 this polynomial has at most $p-1$ roots, so it therefore has *exactly* the $p-1$ simple roots $1, 2, \ldots, p-1$ (and no repeated roots).

A similar argument yields the following.

Lemma 21.1. *If $x^{p-1} - 1 = g(x)h(x)$, where $\deg(g) = k$ and $\deg(h) = l$, then $g(x)$ has exactly k distinct roots* mod p*, and $h(x)$ has exactly l distinct roots* mod p*.*

Proof. To begin, notice that

$$p - 1 \;=\; \text{degree}(gh) \;=\; \text{degree}(g) + \text{degree}(h) \;=\; k + l.$$

If r is a root of $x^{p-1} - 1$, then $g(r)h(r) \equiv 0 \mod p$, so that either $g(r) \equiv 0$ or $h(r) \equiv 0 \mod p$. If g has fewer than k roots, then there are more than $l = p - 1 - k$ distinct roots of $x^{p-1} - 1$ remaining, all of which must then be roots of h. But h cannot have more than l roots. Therefore g must have exactly k roots, and similarly h must have exactly $p - 1 - k = l$ roots. ❑

The following algebraic identity is a variant of the geometric sum formula (3.2).

Lemma 21.2. *If $n = kl$, then*

$$x^n - 1 \;=\; (x^k)^l - 1 \;=\; (x^k - 1)(x^{k(l-1)} + x^{k(l-2)} + \cdots + x^k + 1).$$

Proof. Begin with the geometric sum identity:

$$u^l - 1 \;=\; (u - 1)(u^{l-1} + u^{l-2} + \cdots + u + 1).$$

The lemma follows after substituting $u = x^k$. ❑

We are now ready to prove the main theorem.

Theorem 21.3. *If p is prime, then $\mathbb{Z}_p$ has a primitive root.*

Proof. If $p = 2$ then 1 is primitive.

Suppose that p is an odd prime, and that $p-1$ has the prime power factorization

$$p - 1 = q_1^{a_1} \cdots q_k^{a_k},$$

where $q_1 < \cdots < q_k$.

By Lemma 21.2, we have

$$x^{p-1} - 1 \;=\; (x^{q_1^{a_1}} - 1)h(x),$$

where $h(x)$ is a polynomial. By Lemma 21.1 the factor $x^{q_1^{a_1}} - 1$ has exactly $q_1^{a_1}$ distinct roots mod p.

If s is a root of $x^{q_1^{a_1}} - 1$, then $s^{q_1^{a_1}} \equiv 1 \bmod p$. It follows from Proposition 20.1 that $\mathrm{ord}_p(s) | q_1^{a_1}$. Therefore, $\mathrm{ord}_p(s) = q_1^{b_1}$, for some $0 \leq b_1 \leq a_1$.

If every root of $x^{q_1^{a_1}} - 1$ has order strictly less than $q_1^{a_1}$, then every root of $x^{q_1^{a_1}} - 1$ is also a root of of $x^{q_1^{a_1-1}} - 1$. In other words, this polynomial of degree $q_1^{a_1-1}$ would have $q_1^{a_1}$ roots, which is impossible. It follows that at least one of the roots of $x^{q_1^{a_1}} - 1$ has order $q_1^{a_1}$. In other words, there exists an element $r_1 \in U_p$ having order $q_1^{a_1}$.

Repeating this argument for each q_i, we find, for each i, an element $r_i \in U_p$ of order $q_i^{a_i}$.

By the Multiplicative Lemma 20.5, the unit $r = r_1 \cdots r_k$ has order $q_1^{a_1} \cdots q_k^{a_k} = p - 1$, so that r is a primitive root mod p. ❑

❧

While we have proven the existence of primitive roots mod p, there remains the issue of how to find one.

Proposition 21.4. *Suppose that $\phi(p) = p - 1 = q_1^{a_1} \cdots q_s^{a_s}$, where $q_1 < \cdots < q_s$ are prime, and each $a_i > 0$. Then r is a primitive root* mod p *iff*

$$r^{\frac{p-1}{q_i}} \not\equiv 1 \mod p$$

for every q_i.

The proof of this proposition is left to the reader (see Exercise 21.1). Proposition 21.4 speeds the process of checking whether a value r is primitive. For

example, if $p = 31$ then $p - 1 = 30 = 2 \cdot 3 \cdot 5$. To determine if 3 is primitive mod 31 we need only check that

$$3^6 \not\equiv 1, \quad 3^{10} \not\equiv 1, \quad 3^{15} \not\equiv 1, \qquad \mod 31.$$

First, use repeated squaring to determine that

$$3^2 \equiv 9, \quad 3^4 \equiv -12, \quad 3^8 \equiv 20, \quad 3^{16} \equiv -3 \mod 31.$$

It is then easy to compute

$$3^6 \equiv 16, \quad 3^{10} \equiv 25, \quad 3^{15} \equiv -1 \mod 31$$

so that 3 must be primitive mod 31.

❧

We have shown that if p is an odd prime then $\mathbb{Z}_p$ has a primitive root. From here it is not difficult to show that $\mathbb{Z}_{p^2}$ has a primitive root as well. Evidently 3 is a primitive root mod 4. The remaining cases are addressed by the next proposition.

Proposition 21.5. *If p is an odd prime then $\mathbb{Z}_{p^2}$ has a primitive root.*

Proof. Let r be primitive mod p, and let $k = \mathrm{ord}_{p^2}(r)$. Since $r^k \equiv 1 \mod p^2$, we also have $r^k \equiv 1 \mod p$. Since r is primitive mod p, it follows that $(p-1)|k$, so that $k = s(p-1)$ for some integer s.

Meanwhile, $k|\phi(p^2)$; that is, $k|p(p-1)$. In other words,

$$s(p-1)|p(p-1)$$

so that $s|p$. This can only happen if $s = 1$ or $s = p$.

If $s = p$ then $\mathrm{ord}_{p^2}(r) = k = p(p-1) = \phi(p^2)$, so r is a primitive root mod p^2.

If $s = 1$, then $\mathrm{ord}_{p^2}(r) = k = (p-1)$, so $r^{p-1} \equiv 1 \mod p^2$. In this case, let $\tilde{r} = r + p$. Since $\tilde{r} \equiv r \mod p$, the value $\tilde{r}$ is also primitive mod p. By the previous argument, $\mathrm{ord}_{p^2}(\tilde{r}) = \tilde{s}(p-1)$ where $\tilde{s} \in \{1, p\}$. Moreover,

$$\begin{aligned}
\tilde{r}^{p-1} &\equiv (r+p)^{p-1} \\
&\equiv r^{p-1} + (p-1)r^{p-2}p + \text{terms divisible by } p^2 \\
&\equiv 1 + (p^2 - p)r^{p-2} \mod p^2 \\
&\equiv 1 - pr^{p-2} \mod p^2
\end{aligned}$$

If $\tilde{s} = 1$, then

$$1 \equiv \tilde{r}^{p-1} \equiv 1 - pr^{p-2} \mod p^2,$$

so that $pr^{p-2} \equiv 0 \mod p^2$. This implies that $p|r$, which contradicts our choice of r. Therefore, $\tilde{s} = p$, and $\tilde{r}$ is a primitive root mod p^2. ❑

The gist of the previous proof is the following: To find a primitive root mod p^2, start with a primitive root r mod p. Either r is also primitive mod p^2, or $r+p$ will be primitive mod p^2. This argument can be generalized to show that primitive roots exist for $\mathbb{Z}_n$ under the following conditions.

Theorem 21.6. *A primitive root exists for $\mathbb{Z}_n$ if and only if $n = 2, 4, p^e$ or $2p^e$, where p is an odd prime, and e is a positive integer.*

Proof. The cases of 2 and 4 are easily verified directly directly.

If p is an odd prime, Proposition 21.5 implies that there exists a primitive root r mod p^2 which is also primitive mod p. More generally, suppose that r is primitive mod $p, p^2, \ldots, p^e$ for some $e \geq 2$. We will show that r is also primitive mod p^{e+1}.

To this end, let $k = \text{ord}_{p^{e+1}}(r)$. Since $r^k \equiv 1 \mod p^{e+1}$, we also have $r^k \equiv 1 \mod p^e$. Since r is primitive mod p^e, it follows that $p^{e-1}(p-1)|k$, so that $k = sp^{e-1}(p-1)$ for some integer s.

Meanwhile, $k|\phi(p^{e+1})$; that is, $k|p^e(p-1)$. In other words,

$$sp^{e-1}(p-1)|p^e(p-1)$$

so that $s|p$. This can only happen if $s = 1$ or $s = p$.

Suppose that $s = 1$, so that

$$r^{p^{e-1}(p-1)} \equiv 1 \mod p^{e+1}. \tag{21.1}$$

By Euler's Theorem we also have

$$r^{p^{e-2}(p-1)} \equiv 1 \mod p^{e-1},$$

so that

$$r^{p^{e-2}(p-1)} = 1 + p^{e-1}t,$$

for some integer t. It follows from (21.1) that

$$1 \equiv r^{p^{e-1}(p-1)} = (1+p^{e-1}t)^p \equiv 1 + p^e t \mod p^{e+1},$$

so that $p^e t \equiv 0 \mod p^{e+1}$. In other words, $p^{e+1}|p^e t$, so that $p|t$. Writing $t = pm$, we now have

$$r^{p^{e-2}(p-1)} \equiv 1 + p^{e-1}t \equiv 1 + p^{e-1}pm \equiv 1 \mod p^e,$$

violating the assumption that r is primitive mod p^e.

It follows that $s \neq 1$. This means that $s = p$, so that $\text{ord}_{p^{e+1}}(r) = p^e(p-1)$. In other words, r is primitive mod p^{e+1} as well.

It now follows by induction that, if p is an odd prime, then there exists an integer r that is primitive mod p^e for all integer exponents $e \geq 1$. The proof of the rest of this theorem involves verifying a few additional cases, outlined by Exercises 21.5-21.9. ❑

❧

Exercise 21.1. Prove Proposition 21.4.

Exercise 21.2. Suppose that $p > 3$ is prime. Prove that the product of all of the primitive roots mod p is congruent to 1 mod p.
Hint: Part (b) of Exercise 20.10 is helpful.

Exercise 21.3. For which numbers n are there exactly two primitive roots?

Exercise 21.4. For which numbers n are there an odd number of primitive roots?

Exercise 21.5. **(a)** Suppose that $p \neq q$ are odd primes. Prove that pq has no primitive root.
Hint: Show that, for all $u \in U_{pq}$, we have $u^s \equiv 1 \bmod pq$, where $s = \operatorname{lcm}(p-1, q-1)$. Then show that $s < \phi(pq)$.
(b) Generalize part (a) show that $\mathbb{Z}_n$ has no primitive roots if n is divisible by two distinct odd primes.

Exercise 21.6. **(a)** Verify directly that $\mathbb{Z}_8$ has no primitive roots.
(b) Use part (a) as the basis for an an induction argument (with respect to the exponent e) to prove that $\mathbb{Z}_{2^e}$ has no primitive roots for $e \geq 3$.

Exercise 21.7. **(a)** Prove that, if p is an odd prime, then $\mathbb{Z}_{4p}$ has no primitive roots.
Hint: Suppose that r is primitive mod $4p$. Show that r must also be primitive mod p, and show that $r^2 \equiv 1 \bmod 4$. Then use these results to contradict the primitive property mod $4p$.

(b) Generalize part (a) show that $\mathbb{Z}_n$ has no primitive roots if n is divisible by $4p$ when p is an odd prime.

Exercise 21.8. Prove that, if p is an odd prime and e is a positive integer, then $\mathbb{Z}_{2p^e}$ has a primitive root.
Hint: The proof of Theorem 21.6 above yields a primitive root r for $\mathbb{Z}_{p^e}$. Show that either r or $r + p^e$ is an odd integer. Then prove that the odd choice is also primitive for $\mathbb{Z}_{2p^e}$.

Exercise 21.9. Let $k \geq 3$. Show that if $u \in U_{2^k}$, then $u^{(2^{k-2})} \equiv 1 \bmod 2^k$.
Hint: First verify the case of $k = 3$, and then proceed by induction with respect to k.

Exercise 21.10. Suppose that p is prime.
(a) For $n|(p-1)$ define

$$g(n) = \sum_{u^n \equiv 1 \bmod p} u$$

where each $u \in U_p$. Show that $g(1) = 1$ and that $g(n) = 0$ for $n > 1$.

(b) For $n|(p-1)$ define

$$f(n) = \sum_{\operatorname{ord}_p(u) = n} u$$

where each $u \in U_p$. Show that

$$g(n) = \sum_{d|n} f(d).$$

(c) Use part (b) and Möbius inversion (see Exercise 16.28) to prove that the sum of all of the primitive roots mod p is congruent to $\mu(p-1) \bmod p$.

22 *Quadratic residues*

In school level mathematics we first learn about the basic arithmetic operations $(+, -, \times, \div)$, and then continue with square roots, quadratic equations, and beyond. The development of modular arithmetic proceeds in a similar way. How do we solve quadratic equations modulo m? When do solutions exist?

❧

A unit $u \in U_n$ is called a *quadratic residue* if there exists $b \in U_n$ such that $b^2 \equiv u \mod n$. If unit u is not a quadratic residue, it is called a *quadratic non-residue*.[1]

Quadratic residues are the analogues of "perfect squares" among the units of a modulus.

For example, the squared units 1^2, 2^2, 3^2, 4^2, 5^2, $6^2 \mod 7$ yield respective values 1, 4, 2, 2, 4, 1 (in corresponding order). It follows that 1, 2, 4 are quadratic residues mod 7, while 3, 5, 6 are quadratic non-residues mod 7. We omit 0 from consideration, because 0 is not a unit.

A similar computation reveals that 1 is the only quadratic residue mod 8, while the quadratic residues mod 9 consist of $\{1, 4, 7\}$.

If $u \in U_n$ is a quadratic residue, and if $b^2 \equiv u \mod n$, then we say that b is a *square root* of $u \mod n$.

❧

In the discussion that follows we will focus on quadratic residues mod p, where p is an odd prime. In this case each quadratic residue u has exactly two distinct square roots. To see this, note that $b^2 \equiv u \mod p$ iff b is a root of the polynomial equation

$$x^2 - u \equiv 0 \mod p.$$

Since p is prime, this quadratic polynomial can have at most 2 roots. If b is one root, then $(-b)^2 \equiv b^2 \equiv u$, so that $-b$ is another. Moreover, $b \not\equiv -b \mod p$, since p is an odd prime, so b and $-b$ are distinct and are the only square roots of $u \mod p$.

Proposition 22.1. *If p is an odd prime, then there are exactly $\frac{p-1}{2}$ quadratic residues and exactly $\frac{p-1}{2}$ quadratic non-residues* mod p .

[1] It would make more sense to call them "non-quadratic residues," since they certainly are residues (i.e., remainders after division), but not quadratic (square). However, tradition trumps rational nomenclature in this instance.

Proof. Squaring the units $\{\pm 1, \pm 2, \ldots, \pm\frac{p-1}{2}\}$ gives us $\frac{p-1}{2}$ quadratic residues, which are distinct from one another, since each quadratic residue can have at most two square roots. This exhausts the units that we can square mod p, so there can be no other quadratic residues mod p. Since there are $p-1$ units in total, the remaining $\frac{p-1}{2}$ units are the quadratic non-residues. ❑

More generally, it is not difficult to prove the following (see Exercise 22.3).

Proposition 22.2. *If r is primitive* mod p*, then r^k is a quadratic residue iff k is even.*

❧

The next theorem describes a fundamental test for whether a given unit $u \in U_p$ is a quadratic residue.

Theorem 22.3 (The Euler Criterion). *Let p be an odd prime. If $u \in \mathbb{Z}$ and $p \nmid u$, then*

$$u^{\frac{p-1}{2}} \equiv \begin{cases} 1 & \text{if } u \text{ is a quadratic residue mod } p. \\ -1 & \text{otherwise.} \end{cases}$$

Proof. Suppose that $u \in U_p$. By Fermat's Theorem 15.1,

$$\left(u^{\frac{p-1}{2}}\right)^2 \equiv u^{p-1} \equiv 1 \mod p.$$

Since p is prime, it then follows from Lagrange's Root Theorem 18.3 (or, alternatively, Proposition 15.3) that

$$u^{\frac{p-1}{2}} \equiv \pm 1 \mod p. \tag{22.1}$$

If $u \equiv b^2$, then

$$u^{\frac{p-1}{2}} \equiv b^{p-1} \equiv 1 \mod p,$$

so that each of the $\frac{p-1}{2}$ quadratic residues is a root of the polynomial equation

$$x^{\frac{p-1}{2}} - 1 \equiv 0 \mod p.$$

Since this polynomial can have no more than $\frac{p-1}{2}$ roots, the remaining units (the quadratic non-residues) cannot satisfy this equation. By (22.1), the only remaining possible value for $u^{\frac{p-1}{2}}$ is -1. ❑

Corollary 22.4. *If p is an odd prime, then -1 is a quadratic residue* mod p *iff $p \equiv 1$* mod 4.

Proof. By the Euler criterion, -1 is a quadratic residue mod p iff $(-1)^{\frac{p-1}{2}} = 1$. But this holds iff $\frac{p-1}{2}$ is even; that is, iff $p \equiv 1 \bmod 4$. ❑

❧

If p is an odd prime and $p \nmid u$, denote

$$\left(\frac{u}{p}\right) = u^{\frac{p-1}{2}} \mod p.$$

The expression on the left-hand-side is called the *Legendre symbol*[2] for the pair (u, p). The Euler criterion tells us that

$$\left(\frac{u}{p}\right) \equiv \begin{cases} 1 & \text{if } u \text{ is a quadratic residue mod } p. \\ -1 & \text{otherwise.} \end{cases}$$

Legendre symbols satisfy the following properties.

Proposition 22.5. *Let p be an odd prime. For all $a, b \in U_p$,*

(i) *If $a \equiv b \bmod p$ then $\left(\frac{a}{p}\right) = \left(\frac{b}{p}\right)$.*

(ii) $\left(\frac{a^2}{p}\right) = 1$.

(iii) $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$.

(iv) $\left(\frac{-1}{p}\right) = \begin{cases} 1 & p \equiv 1 \mod 4. \\ -1 & p \equiv 3 \mod 4. \end{cases}$

Proof. Property (i) is immediate, while (ii) and (iii) follow from the Euler Criterion. Property (iv) is a re-statement of Corollary 22.4. ❑

Note: Contrary to appearances, the expression inside the parentheses of a Legendre symbol should *never* be viewed as a fraction.

❧

The Euler criterion provides a quick answer (in the form of Corollary 22.4) to the question of whether -1 is a quadratic residue mod p. We would like to derive similar conditions for other values. When is 2 a quadratic residue mod p? What about 3 and 5? More generally, given two odd primes p and q, how are their quadratic characters related?

[2] Named for Adrien-Marie Legendre (1752–1833).

Before stating the next theorem, we require some preliminary notation. Suppose that $u \in U_p$, and consider the list of values

$$u, 2u, 3u, \ldots, \left(\frac{p-1}{2}\right) u. \tag{22.2}$$

By the cancellation law, each of these values is distinct mod p. Reduce these values mod p so that each value lies within the range $(-\frac{p}{2}, \frac{p}{2})$. Suppose that s of these values are negative, and t of them are positive, so that $s + t = \frac{p-1}{2}$. The following result then holds.

Lemma 22.6 (Gauss's Lemma).
Following the notation introduced above, we have

$$\left(\frac{u}{p}\right) = (-1)^s.$$

Proof. First, note once again that the values in the list (22.2) are non-zero and distinct mod p, since u is a unit. We will show that, after reduction mod p to the range $(-\frac{p}{2}, \frac{p}{2})$, these numbers have distinct *absolute* values.

Suppose that, after reduction to the interval $(-\frac{p}{2}, \frac{p}{2})$, we obtain both of the values k and $-k$, where $k \in (0, \frac{p}{2})$. This means there are values au and bu from the list (22.2), such that $au \equiv k$ and $bu \equiv -k \mod p$. Adding these equations we find that

$$(a+b)u \;\equiv\; k + (-k) \;\equiv\; 0 \mod p,$$

so that $p|(a+b)$. However, the coefficients a and b also satisfy $a, b \in (0, \frac{p}{2})$. This means that

$$0 < a + b < p,$$

so that p cannot divide $a+b$, giving us a contradiction. Thus, if k appears on the reduced list, $-k$ cannot also appear.

Since there are $\frac{p-1}{2}$ integer values in the interval $(0, \frac{p}{2})$, and $\frac{p-1}{2}$ values in the reduced list (22.2), it follows that either the positive or the negative of each integer in the list $1, 2, \ldots, \frac{p-1}{2}$ must appear exactly once in the list (22.2) after reduction mod p to the interval $(-\frac{p}{2}, \frac{p}{2})$.

If we now multiply the numbers in (22.2) together, it follows that

$$u \cdot 2u \cdot 3u \cdots \left(\frac{p-1}{2}\right) u \;\equiv\; (-1)^s \left(\frac{p-1}{2}\right)! \mod p,$$

where the sign is determined by the number s of values in the negative half of the interval $(-\frac{p}{2}, \frac{p}{2})$. Applying the cancellation law to the values $2, 3, \ldots, \frac{p-1}{2}$ on both sides, we have

$$u^{\frac{p-1}{2}} \;\equiv\; (-1)^s \mod p.$$

The Lemma now follows from the Euler criterion. ❑

Gauss's Lemma 22.6 yields the following simple rule for determining when 2 is a quadratic residue.

Corollary 22.7. *If p is an odd prime, then*

$$\left(\frac{2}{p}\right) = \begin{cases} 1 & p \equiv \pm 1 \mod 8. \\ -1 & p \equiv \pm 3 \mod 8. \end{cases}$$

Proof. In each case set $u = 2$ in Gauss's Lemma 22.6, and consider the list $2, 4, 6, \ldots, p-1$.

If $p \equiv 1 \mod 8$, then $p = 8m+1$, so that $\frac{p-1}{2} = 4m$. In this case $0 < 2k \leq \frac{p-1}{2}$ iff $0 < k \leq 2m$. It follows that $s = t = 2m$, so that $\left(\frac{2}{p}\right) = (-1)^s = (-1)^{2m} = 1$.

If $p \equiv 3 \mod 8$, the $p = 8m+3$, so that $\frac{p-1}{2} = 4m+1$ is odd. In this case the largest even value from the list that appears in $(0, \frac{p}{2})$ is $\frac{p-3}{2} = 4m = 2t$. This means that $t = 2m$, so that $s = 2m+1$, an odd number. Hence, $\left(\frac{2}{p}\right) = (-1)^s = (-1)^{2m+1} = -1$.

The remaining cases are similar. ❑

❧

Exercise 22.1. List the quadratic residues in each of:

(a) mod 3
(b) mod 5
(c) mod 12
(d) mod 15
(e) mod 17
(f) mod 25

Exercise 22.2. Evaluate the following Legendre symbols:

(a) $\left(\frac{5}{3}\right)$
(b) $\left(\frac{-7}{13}\right)$
(c) $\left(\frac{7}{13}\right)$
(d) $\left(\frac{360}{7}\right)$
(e) $\left(\frac{-360}{7}\right)$
(f) $\left(\frac{384}{5}\right)$
(g) $\left(\frac{63}{13}\right)$
(h) $\left(\frac{36}{2003}\right)$
(i) $\left(\frac{32}{79}\right)$

Exercise 22.3. Prove Proposition 22.2.

Exercise 22.4. Use the Euler Criterion to prove part (iii) of Proposition 22.5.

Exercise 22.5. Suppose that p is prime and that u has odd order mod p. Prove that u is a quadratic residue mod p.

Exercise 22.6. Suppose that n is a composite integer, and let $u \in U_n$.
(a) If u is a quadratic residue mod n, does it follow that $u^{\frac{\phi(n)}{2}} \equiv 1 \mod n$?
(b) If $u^{\frac{\phi(n)}{2}} \equiv 1 \mod n$, does it follow that u is a quadratic residue mod n?
(c) Examine the proof of the Euler Criterion, and determine which step(s) fail if the prime p is replaced with a composite n (and $p-1$ replaced by $\phi(n)$).

Exercise 22.7. Suppose that $m > 1$ is a composite integer with a primitive root r. Prove that r^k is a quadratic residue mod m iff k is even.

Exercise 22.8. Let p be a prime such that $p \equiv 1 \bmod 4$, and suppose that u is a quadratic non-residue mod p.
(a) Prove that $\frac{p-1}{4}$ is an integer.
(b) Let $v = u^{\frac{p-1}{4}}$, and prove that $v^2 \equiv -1 \bmod p$.

Exercise 22.9. Let p be a prime such that $p \equiv 3 \bmod 4$, and suppose that u is a quadratic residue mod p.
(a) Prove that $\frac{p+1}{4}$ is an integer.
(b) Let $v = u^{\frac{p+1}{4}}$, and prove that $v^2 \equiv u \bmod p$.

Exercise 22.10. Let p be a prime, and let c be the smallest positive quadratic non-residue mod p. Prove that c is a prime number.

Exercise 22.11. Check the remaining cases in the proof of Corollary 22.7.

Exercise 22.12. Suppose that $a, b > 1$ are relatively prime integers. Prove that u is a quadratic residue mod ab iff u is a quadratic residue both mod a and mod b.

Exercise 22.13. What happens in Exercise 22.12 if a and b are not relatively prime?

Exercise 22.14. Use the quadratic formula and quadratic residue theory to determine which of these equations have solutions.

(a) $x^2 - x + 1 \mod 17$
(b) $x^2 - x + 1 \mod 31$
(c) $2x^2 - 3x + 4 \mod 11$
(d) $2x^2 - 3x + 4 \mod 23$

Exercise 22.15. Prove that there are infinitely many prime numbers of the form $4n + 1$.
Hint: Suppose there are finitely many. Multiply them together to obtain an integer N. Use Corollary 22.4 to show that the number $4N^2 + 1$ must have a prime factor of the form $4n + 1$.

Exercise 22.16. Prove that, if n is an integer, then $4n^2 + 4$ cannot be divisible by 127.

Exercise 22.17. Let p be an odd prime such that $p \equiv 2 \bmod 3$.
Prove that, for every $a \in \mathbb{Z}$, there exists $b \in \mathbb{Z}$ such that $a \equiv b^3 \bmod p$.
In particular, every unit mod p is a *cubic residue*.
Hint: Combine Fermat's theorem with the fact that $\gcd(3, p-1) = 1$ to find an actual formula for the cube root.

Exercise 22.18. Let p be an odd prime such that $p \equiv 1 \bmod 3$. Let $u \in U_p$.
Prove that there exists $b \in \mathbb{Z}$ such that $u \equiv b^3 \bmod p$ iff

$$u^{\frac{p-1}{3}} \equiv 1 \mod p.$$

Hint: Let r be a primitive root mod p. Since $p \nmid u$ we know that $u \equiv r^k$ for a unique value $k \in \{0, \ldots, p-2\}$. What happens if $3|k$? What happens if $3 \nmid k$?

Exercise 22.19. Suppose that $n, m > 1$ and that $n|m$. True or False?
(a) If u is a quadratic residue mod m, then u is a quadratic residue mod n.
(b) If u is a quadratic non-residue mod m, then u is a non-quadratic residue mod n.

Exercise 22.20. Can you find an integer $u > 1$ such that u is a quadratic residue in every modulus m such that $\gcd(u, m) = 1$?

23 The law of quadratic reciprocity

Gauss's Lemma is the stepping stone to the following remarkable symmetry law satisfied by quadratic residues.

Theorem 23.1 (The Law of Quadratic Reciprocity).
Let $p \neq q$ be positive odd primes.

- *If* $p \equiv q \equiv 3 \bmod 4$ *then* $\left(\frac{p}{q}\right) = -\left(\frac{q}{p}\right)$.
- *Otherwise* $\left(\frac{p}{q}\right) = \left(\frac{q}{p}\right)$.

This theorem has the following equivalent formulation:

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}} \tag{23.1}$$

Although the quadratic reciprocity law was formulated as a conjecture by Euler, its first proof is due to Gauss,[1] presented in his famous number theory text, the *Disquisitiones Arithmeticae*.[2]

A proof of Theorem 23.1 is deferred to the next section. Before that, we take a moment to illustrate how quadratic reciprocity makes it much easier to compute Legendre symbols, allowing us to avoid computing the exponentials indicated by the Euler Criterion.

[1] Carl Friedrich Gauss (1777–1855).
[2] Gauss was 21 years old when he wrote this enormously influential textbook. It is available in English translation; see [11].

Example:

$$\begin{aligned}
\left(\frac{59}{71}\right) &= -\left(\frac{71}{59}\right) \text{ by quadratic reciprocity, since } 59 \equiv 71 \equiv 3 \bmod 4,\\
&= -\left(\frac{12}{59}\right) \quad \text{since } 71 \equiv 12 \bmod 59,\\
&= -\left(\frac{4}{59}\right)\left(\frac{3}{59}\right) \quad \text{since } 12 = 3 \cdot 4,\\
&= -\left(\frac{3}{59}\right) \quad \text{since } 4 = 2^2,\\
&= \left(\frac{59}{3}\right) \quad \text{by quadratic reciprocity, since } 59 \equiv 3 \bmod 4,\\
&= \left(\frac{2}{3}\right) \quad \text{since } 59 \equiv 2 \bmod 3,\\
&= -1 \quad \text{by inspection. } \❑
\end{aligned}$$

Remember that the numerator of a Legendre symbol may be composite, but the denominator *must always be an odd prime*. In particular, a composite numerator must be factored *before* quadratic reciprocity can be applied.

Example:

$$\begin{aligned}
\left(\frac{90}{149}\right) &= \left(\frac{9}{149}\right)\left(\frac{2}{149}\right)\left(\frac{5}{149}\right)\\
&= 1 \cdot (-1) \cdot \left(\frac{5}{149}\right) \quad \text{since } 9 = 3^2 \text{ and } 149 \equiv 5 \bmod 8,\\
&= -\left(\frac{149}{5}\right) \quad \text{by quadratic reciprocity, since } 5 \equiv 1 \bmod 4,\\
&= -\left(\frac{4}{5}\right) \quad \text{since } 149 \equiv 4 \bmod 5,\\
&= -1 \quad \text{since } 4 = 2^2. \ \❑
\end{aligned}$$

In order to use Theorem 23.1 to compute Legendre symbols, one needs to determine that the numerator is prime, or to factor the numerator if it is not prime. When such factoring is not feasible (such as when the number is very large), one must instead return to the exponential formula given by the Euler Criterion.[3]

❧

Exercise 23.1. Show that the identity (23.1) is equivalent to Theorem 23.1.

[3] We will find a partial remedy to this difficulty in Section 26 on Jacobi symbols.

Exercise 23.2. Let p, q_1, q_2 be distinct odd primes. Use Theorem 23.1 to show that

$$\left(\frac{p}{q_1}\right) = \left(\frac{p}{q_2}\right)$$

whenever $q_1 \equiv q_2 \bmod 4p$.

Exercise 23.3. Use Theorem 23.1 to show that

$$\left(\frac{3}{p}\right) = \begin{cases} 1 & p \equiv \pm 1 \mod 12. \\ -1 & p \equiv \pm 5 \mod 12. \end{cases}$$

Exercise 23.4. **(a)** Use Theorem 23.1 to derive a quick and easy rule (in analogy to the previous exercise) for determining $\left(\frac{5}{p}\right)$ for odd primes $p \neq 5$.
(b) For which primes p is 10 a quadratic residue?

Exercise 23.5. Use Theorem 23.1 to derive a simple rule (in analogy to the previous exercise) for determining $\left(\frac{7}{p}\right)$ for odd primes $p \neq 7$.

Exercise 23.6. Compute the following Legendre symbols:

(a) $\left(\frac{7}{101}\right)$ **(c)** $\left(\frac{20}{61}\right)$ **(e)** $\left(\frac{1875}{2003}\right)$ **(g)** $\left(\frac{919}{683}\right)$ **(i)** $\left(\frac{7}{601}\right)$
(b) $\left(\frac{15}{127}\right)$ **(d)** $\left(\frac{72}{83}\right)$ **(f)** $\left(\frac{-490}{41}\right)$ **(h)** $\left(\frac{173}{401}\right)$ **(j)** $\left(\frac{11}{9901}\right)$

where you may accept as given that $401, 601, 683, 919, 2003$ and 9901 are prime numbers.

Exercise 23.7. **(a)** What is the smallest positive prime p such that 2 and 3 are *both* quadratic residues mod p?
(b) What is the smallest positive prime p such that 2, 3, and 5 are *all* quadratic residues mod p?
(c) What is the smallest quadratic non-residue modulo the prime p of part (b)?
Hint: Corollary 22.7 and the results of Exercises 23.3 and 23.4 make this exercise a lot easier. Brute force computations of every case are not necessary.

Exercise 23.8. Suppose that p is prime such that $3|(p-1)$, and let $m = \frac{p-1}{3}$. Prove that m is a quadratic residue mod p.

Exercise 23.9. Find a prime p such that 2, 3, 5, and 7 are *all* quadratic non-residues mod p.
Hint: Combine Corollary 22.7 and the results of Exercises 23.3, 23.4, and 23.5 with the Chinese remainder theorem.

Exercise 23.10. Use the quadratic formula and quadratic residue theory to determine which of the following quadratic equations have solutions. (But do not try to find the solutions unless you have a lot of spare time.)

(a) $x^2 - x + 1 \mod 53$
(b) $x^2 - x + 1 \mod 73$
(c) $2x^2 - 3x + 4 \mod 101$
(d) $2x^2 - 3x + 4 \mod 307$

Exercise 23.11. Given that

$$a = 1111111111111111111 \quad \text{and} \quad b = 11111111111111111111111$$

are prime integers (consisting respectively of 19 and 23 catenated decimal '1's), what is $\left(\frac{a}{b}\right)$?

24 *Proof of quadratic reciprocity*

We now give a proof Theorem 23.1, by proving the equivalent assertion that

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}}, \tag{24.1}$$

for odd primes $p \neq q$.

Proof of Quadratic Reciprocity. To begin, we apply Gauss's Lemma 22.6 to the Legendre symbol $\left(\frac{p}{q}\right)$. Specifically, consider the list of integers

$$p, 2p, \ldots, \left(\tfrac{q-1}{2}\right)p.$$

Reduce each number in this list modulo q, so that the resulting value lies in the interval $(-\frac{q}{2}, \frac{q}{2})$. In order to apply Gauss's Lemma, we need to count the number μ of negative values in the reduced list. In other words, we need to determine

$$\mu \;=\; \#\{x \mid xp \equiv s \text{ for some } s \in \left(-\tfrac{q}{2}, 0\right)\},$$

where $x \in \{1, 2, \ldots, \frac{q-1}{2}\}$. Gauss's Lemma then asserts that $\left(\frac{p}{q}\right) = (-1)^{\mu}$.

In order to count the number μ, note that

$$xp \equiv s \in \left(-\tfrac{q}{2}, 0\right) \mod q$$

if and only if

$$xp - yq \in \left(-\tfrac{q}{2}, 0\right) \tag{24.2}$$

for some $x, y \in \mathbb{Z}$.

If indeed (24.2) holds, then

$$\begin{array}{lrcll} & -\frac{q}{2} & < & xp - yq & < 0 \\ \text{iff} & -\frac{q}{2} - xp & < & -yq & < -xp \\ \text{iff} & xp & < & yq & < \frac{q}{2} + xp \end{array}$$

which means that

$$0 < y < \tfrac{1}{2} + \tfrac{xp}{q} < \tfrac{1}{2} + \tfrac{p}{2}$$

since $x < \frac{q}{2}$. Since p is odd and y is an integer, this means that

$$1 \leq y \leq \frac{p-1}{2}.$$

It now follows that (24.2) holds if and only if the line

$$xp - yq = s$$

passes through a point (x, y) in the region $\square OABD$ of Figure 24.1, where x and y are both integers and where $s \in \{-1, -2, \ldots, -\frac{q-1}{2}\}$.

Points in the plane having integer coordinates are called *lattice points*. In order to determine μ, we need to determine how many lines of the form $xp - yq = s$ pass through a lattice point in the region $\square OABD$.

Notice that if $xp - yq = s$ and $\bar{x}p - \bar{y}q = s$ for some pair of lattice points (x, y) and $(\bar{x}, \bar{y})$ then

$$xp - yq \;=\; s \;=\; \bar{x}p - \bar{y}q$$

so that $(x - \bar{x})p = (y - \bar{y})q$. This implies that $q|(x - \bar{x})$. Since $0 < x, \bar{x} < \frac{q}{2}$, it follows that $x = \bar{x}$ and, similarly, $y = \bar{y}$. In other words, each line of the form $xp - yq = s$ passes through *at most one* lattice point in the region $\square OABD$. This means that μ is simply the *total number of lattice points* in the interior[1] of the region $\square OABD$.

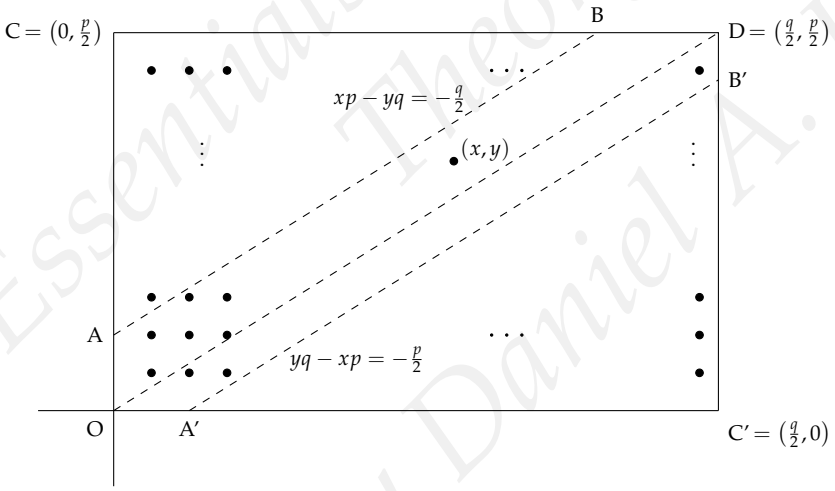

Figure 24.1: Counting (lines through) lattice points.

Similarly, after exchanging the roles of the primes p and q, we can apply Gauss's Lemma 22.6 to the Legendre symbol $\left(\frac{p}{q}\right)$, to discover that

$$\left(\frac{q}{p}\right) = (-1)^{\eta},$$

where (by a symmetrical argument) η is the number of lattice points in the interior of the region $\square OA'B'D$. It now follows that

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) \;=\; (-1)^{\mu+\eta},$$

[1] That is, excluding the boundary.

where $\mu + \eta$ is the total number of lattice points in the interior of the hexagonal region $\overline{OABDB'A'O}$.

Meanwhile, suppose that (x, y) is a lattice point in the interior of triangle $\triangle ABC$. This is true iff x and y satisfy the inequalities

$$0 < x < \frac{q}{2} \qquad 0 < y < \frac{p}{2} \qquad xp - qy < -\frac{q}{2}$$

all hold. But these inequalities hold iff

$$0 < \bar{x} < \frac{q}{2} \qquad 0 < \bar{y} < \frac{p}{2} \qquad q\bar{y} - \bar{x}p < -\frac{p}{2}$$

where

$$\bar{x} = \frac{p+1}{2} - x \qquad \text{and} \qquad \bar{y} = \frac{p+1}{2} - y.$$

The reader should check this algebra carefully. See Exercise 24.2.

It now follows that the lattice point (x, y) lies in the interior of the upper triangle $\triangle ABC$ iff the lattice point $(\bar{x}, \bar{y})$ lies in the interior of the lower triangle $\triangle A'B'C'$. Since the transforms $x \to \bar{x}$ and $y \to \bar{y}$ are each inverses of themselves (that is, $\bar{\bar{x}} = x$ and $\bar{\bar{y}} = y$), this means that $\triangle ABC$ and $\triangle A'B'C'$ contain the *same* number of interior lattice points. Call this number λ.

If now follows that there are a total of $\mu + \eta + 2\lambda$ lattice points in the interior of the rectangle $\square OCDC'$. Since this rectangle has $\frac{p-1}{2} \cdot \frac{q-1}{2}$ lattice points in its interior, we have

$$\mu + \eta + 2\lambda \;=\; \frac{p-1}{2} \cdot \frac{q-1}{2},$$

so that

$$\mu + \eta \;\equiv\; \frac{p-1}{2} \cdot \frac{q-1}{2} \mod 2,$$

which implies that

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) \;=\; (-1)^{\mu+\eta} \;=\; (-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}}.$$

❑

❧

While conjectured by Euler, the Law of Quadratic Reciprocity was first proven by Gauss, who ultimated published six different proofs over the course of his mathematical career. Gauss's third proof was later simplified by Eisenstein,[2] whose lattice point counting argument is the proof most commonly given in

[2] Ferdinand Gotthold Max Eisenstein (1823–1852).

modern introductory number theory texts, including this one. An exhaustive history of reciprocity laws can be found in [17].

❧

Exercise 24.1. Prove that $\bar{\bar{x}} = x$ and $\bar{\bar{y}} = y$.

Exercise 24.2. Verify the assertion (used in the proof of Theorem 23.1) that x and y satisfy the inequalities

$$0 < x < \frac{q}{2} \qquad 0 < y < \frac{p}{2} \qquad xp - qy < -\frac{q}{2}$$

if and only if $\bar{x}$ and $\bar{y}$ satisfy the inequalities

$$0 < \bar{x} < \frac{q}{2} \qquad 0 < \bar{y} < \frac{p}{2} \qquad q\bar{y} - \bar{x}p < -\frac{p}{2}$$

where

$$\bar{x} = \frac{p+1}{2} - x \qquad \text{and} \qquad \bar{y} = \frac{p+1}{2} - y.$$

Exercise 24.3. Prove that there are no lattice points on the line segment $\overline{OD}$ of Figure 24.1 except the origin O.

25 Quadratic residues over composite moduli

So far we have focused on the quadratic character of a unit mod p, where p is prime. Having developed the tools for addressing that special case, we now turn to the question of when a unit is a quadratic residue in a general modulus $m > 1$. This turns out to be straightforward, provided that m can be factored into powers of primes. The next three propositions explain.

Proposition 25.1. *Suppose that $m, n > 1$ and that $\gcd(m, n) = 1$. The following are equivalent.*

- *u is a quadratic residue* mod *mn.*
- *u is a quadratic residue* mod *m* ***and*** *u is a quadratic residue* mod *n.*

For example, a unit u is a quadratic residue mod $1001 = 7 \cdot 11 \cdot 13$ if and only if u is a quadratic residue in each of the moduli 7, 11, and 13.

Proof. If u is a quadratic residue mod mn, there exists $b \in \mathbb{Z}$ such that $u \equiv b^2$ mod mn. This means that $mn|(u - b^2)$, so that $m|(u - b^2)$ and $n|(u - b^2)$. In other words, $u \equiv b^2$ mod m and $u \equiv b^2$ mod n, so that u is a quadratic residue in both of these moduli.

For the converse, suppose that u is a quadratic residue mod m and mod n. This means that $u \equiv b^2$ mod m, and $u \equiv c^2$ mod n, for some integers b, c.

Since $\gcd(m, n) = 1$, the Chinese Remainder Theorem 11.1 provides an integer a such that $a \equiv b$ mod m and $a \equiv c$ mod n. This means that

$$a^2 \equiv b^2 \equiv u \mod m \quad \text{and} \quad a^2 \equiv c^2 \equiv u \mod n.$$

In other words, $m|(a^2 - u)$ and $n|(a^2 - u)$. Since $\gcd(m, n) = 1$, we may conclude that $mn|(a^2 - u)$, so that $u \equiv a^2$ mod mn, and a is a quadratic residue mod mn. ❑

Proposition 25.1 implies that, to determine the quadratic character of an integer mod 3645, it is sufficient to look at the separate cases mod 5 and mod 729. However, our methods so far do not help us with this question mod $729 = 3^6$, which is neither prime nor factorable into relatively prime proper divisors. To complete the story we still need to address the case of prime powers.

Proposition 25.2. *Let p be an odd prime, and suppose $p \nmid u$. Then u is a quadratic residue* mod *p^s, for $s \geq 2$, iff u is a quadratic residue* mod *p.*

If follows, for example, that a unit u is a quadratic residue mod $3645 = 3^6 \cdot 5$ if and only if u is a quadratic residue in each of the moduli 3 and 5, which is usually easy to determine.

Proof. If u is a quadratic residue mod p^s, then $p^s | (u - r^2)$ for some integer r. It follows that $p | (u - r^2)$, so that u is a quadratic residue mod p.

We will prove the converse by induction on s. Suppose that u is a quadratic residue mod p^s for some $s \geq 1$. This means that $u \equiv r^2 \bmod p^s$ for some integer r, so that

$$r^2 - u = p^s k \tag{25.1}$$

for some integer k.

We will show that u is a quadratic residue mod p^{s+1} as well. To see this, let

$$\bar{r} = r + tp^s, \tag{25.2}$$

where t is an integer to be determined. We have

$$\begin{aligned} \bar{r}^2 - u &\equiv (r + tp^s)^2 - u \mod p^{s+1} \\ &\equiv r^2 + 2rtp^s + t^2 p^{2s} - u \mod p^{s+1} \\ &\equiv p^s(k + 2rt) \mod p^{s+1}. \end{aligned}$$

It follows that $\bar{r}^2 - u \equiv 0 \bmod p^{s+1}$ iff

$$k + 2rt \equiv 0 \mod p,$$

which, in turn, holds iff there is an integer t such that

$$t \equiv -k(2r)^{-1} \mod p. \tag{25.3}$$

Such a value of t exists, because p is odd and $r \in U_p$.

We have shown that if u is a quadratic residue mod p^s, then u is a quadratic residue mod p^{s+1}. Therefore, if u is a quadratic residue mod p, then u is also a quadratic residue modulo every positive power of p. ❑

The proof above describes an algorithm for finding a square root of u mod p^{s+1}, if we are given a square root mod p^s. Combining the equations (25.1), (25.2), and (25.3) yields:

Corollary 25.3. *Let p be an odd prime, and let u be a quadratic residue* mod p. *If $r^2 \equiv u \bmod p^s$, then*

$$[r - v(r^2 - u)]^2 \equiv u \mod p^{s+1},$$

where $v \equiv (2r)^{-1} \bmod p$.

In other words, if r is a square root of u mod p^s, then $r - v(r^2 - u)$ is a square root of u mod p^{s+1}, where $v \equiv (2r)^{-1}$ is first computed mod p.

Example: Since $11 \equiv 1 \bmod 5$, it follows that 11 is a quadratic residue modulo every positive power of 5. To find the square root of 11 mod 25, first observe that $1^2 \equiv 1 \equiv 11 \bmod 5$, and compute $v \equiv (2 \cdot 1)^{-1} \equiv 3 \bmod 5$.

It follows that

$$1 - 3(1^2 - 11) \equiv 31 \equiv 6 \mod 25,$$

so that $6^2 \equiv 11 \bmod 25$.

Continuing, we can find a square root for 11 mod 125 by "lifting" our solution mod 25 to

$$6 - 3(6^2 - 11) \equiv 56 \mod 125,$$

so that $56^2 \equiv 11 \bmod 125$.

❧

The proof of Proposition 25.2 uses the fact that p is odd, so that 2^{-1} exists mod p. We can use a similar argument to determine the quadratic character of integers modulo 2^s, but some adjustment is needed.

To begin, recall that every odd number u is a quadratic residue mod 2, while u is a quadratic residue mod 4 iff $u \equiv 1 \bmod 4$. The remaining powers of 2 are addressed by the next proposition.

Proposition 25.4. *An odd number u is a quadratic residue* mod 2^s*, for $s \geq 3$, iff $u \equiv 1$* mod 8*.*

Recall that $u \equiv 1 \bmod 8$ is precisely the condition under which an odd number is a quadratic residue mod 8.

Proof. Suppose that u is a quadratic residue mod 2^s, where $s \geq 3$. Since $8 | 2^s$, it follows that u is also a quadratic residue mod 8, so that $u \equiv 1 \bmod 8$.

To prove the converse, suppose that u is a quadratic residue mod 2^s for some $s \geq 3$. This means that $u \equiv r^2 \bmod 2^s$ for some odd integer r.

If $u \equiv r^2 \bmod 2^{s+1}$, then u is also a quadratic residue mod 2^{s+1}.

If $u \not\equiv r^2 \bmod 2^{s+1}$, then $r^2 - u = 2^s k$ for some integer k.

Setting

$$\bar{r} = r + k2^{s-1},$$

we have

$$\begin{aligned}\bar{r}^2 - u &\equiv (r + k2^{s-1})^2 - u \mod 2^{s+1} \\ &\equiv r^2 + rk2^s + k^2 2^{2s-2} - u \mod 2^{s+1} \\ &\equiv 2^s k(1+r) \mod 2^{s+1} \quad \text{(since } s \geq 3\text{, so that } 2s-2 \geq s+1\text{)} \\ &\equiv 0 \mod 2^{s+1} \quad \text{(since } 1+r \text{ is even).}\end{aligned}$$

It follows that $\bar{r}^2 \equiv u \bmod 2^{s+1}$, so that u is a quadratic residue mod 2^{s+1}.

We have shown that if u is a quadratic residue mod 2^s, for some $s \geq 3$, then u is also a quadratic residue mod p^{s+1}. Therefore, if u is a quadratic residue mod 8, then u is also a quadratic residue modulo 2^s, for all $s \geq 3$. ❑

The following theorem now summarizes the situation for any modulus m other than a pure power of 2.

Theorem 25.5. *Suppose that m has the form*

$$m = 2^t p_1^{s_1} \cdots p_k^{s_k}, \tag{25.4}$$

where $p_1 < \cdots < p_k$ are odd primes, $t \geq 0$, and each $s_i > 0$. If $\gcd(u, m) = 1$, then u is a quadratic residue mod *m iff u is a quadratic residue modulo each of the p_i and*

- *u is odd if $t = 1$.*
- *$u \equiv 1 \mod 4$ if $t = 2$.*
- *$u \equiv 1 \mod 8$ if $t \geq 3$.*

Proof. This theorem is a straightforward consequence of Propositions 25.1, 25.2, and 25.4. ❑

❧

Example: To determine if 13 has a square root mod 1377, begin by factoring $1377 = 3^4 \cdot 17$. Since $13 \equiv 1 \bmod 3$, and

$$\left(\frac{13}{17}\right) = \left(\frac{17}{13}\right) = \left(\frac{4}{13}\right) = 1,$$

it follows that 13 is indeed a quadratic residue mod 1377.

To find a square root, start with $2^2 \equiv 4 \equiv 13 \bmod 9$. Compute $(2 \cdot 2)^{-1} \equiv 1 \bmod 3$, and then apply the formula from Corollary 25.3 to obtain a square root

$$2 - 1(2^2 - 13) \equiv 11 \mod 27.$$

Applying the lifting formula once again yields

$$11 - 1(11^2 - 13) \equiv 65 \mod 81.$$

so that $65^2 \equiv 13 \mod 81$.

Meanwhile, we can verify by trial and error that $8^2 \equiv 13 \mod 17$. Apply the Chinese Remainder Theorem 11.1 to the identities

$$r \equiv 8 \mod 17$$
$$r \equiv 65 \mod 81$$

to conclude that $r \equiv 875$ is a square root of 13 mod 1377.

❧

Recall that the quadratic residues mod m are the ***units*** mod m that are congruent to squared integers. What about squares that are not units?

An integer n is a *perfect square* mod m if $n \equiv c^2 \mod m$ for some integer c. Evidently a perfect square is a quadratic residue iff it is also a unit mod m.

The perfect squares modulo a prime p are just the quadratic residues, along with the extra value $0 \equiv 0^2$. Perfect squares in an odd prime power modulus p^s are described as follows.

Proposition 25.6. *Let p be an odd prime, and let s be a positive integer.*

An integer n is a perfect square mod p^s *iff $n \equiv p^e b \mod p^s$, where $e \geq 0$ is even and b is a quadratic residue* mod p.

Proof. If $p \nmid n$, then n is a unit mod p^s, so n is a perfect square mod p^s iff n is a quadratic residue mod p^s (by the definition of quadratic residue). This holds in turn iff n is a quadratic residue mod p, by Proposition 25.2.

If $p^s | n$, then $n \equiv 0 \equiv p^{2s} \mod p^s$, so the theorem holds in this case.

Suppose $p|n$, but $p^s \nmid n$. Assume n is reduced mod p^s, and write $n = p^e b$, where $1 \leq e \leq s-1$ and where $p \nmid b$.

If e is even and b is a quadratic residue mod p, then $e = 2t$ for some integer t, while b is a quadratic residue mod p^s. Therefore b and $p^e = p^{2t}$ are both perfect squares mod p^s, so that $n \equiv p^e b$ is a perfect square mod p^s.

For the converse, suppose that $n \equiv p^e b$ is a perfect square mod p^s, where $p \nmid b$. This means that

$$p^e b = c^2 + r p^s,$$

for some integer c. Therefore, $c^2 = p^e(b - r p^{s-e}) = p^e d$ for some integer d, so that

$$b = d + r p^{s-e}.$$

Since $p \nmid b$, it follows that $p \nmid d$, so that p divides c^2 exactly e times, and e must be an even integer. This also implies that d is a perfect integer square, say $d = g^2$. We now have

$$b = g^2 + rp^{s-e} \equiv g^2 \mod p^{s-e}.$$

Since $p \nmid b$, it follows that b is a quadratic residue mod p^{s-e} and therefore also mod p. ❑

Note that every integer is a perfect square mod 2, while the perfect squares mod 4 are $\{0, 1\}$. Perfect squares modulo higher powers of 2 are described as follows.

Proposition 25.7. *Let $s \geq 3$ be a positive integer. An integer n is a perfect square* mod 2^s *iff $n \equiv 2^e b \bmod 2^s$, where $e \geq 0$ is even, b is odd, and*

- $b = 1$ *if* $e \geq s - 2$,
- $b \equiv 1 \mod 8$ *if* $e \leq s - 3$.

The proof is similar to that of Proposition 25.6, keeping in mind the special cases characterizing quadratic residues mod 2^s given by Proposition 25.4.

The Chinese remainder theorem can now be used to prove the following more general result, characterizing perfect squares in a composite modulus m.

Theorem 25.8. *Suppose that*

$$m = p_1^{s_1} \cdots p_k^{s_k},$$

where $p_1 < \cdots < p_k$ are primes and each $s_i > 0$.

An integer n is a perfect square mod m *iff n is a perfect square* mod $p_i^{s_i}$ *for each i.*

The proof of Theorem 25.8 is left as an exercise (see Exercise 25.16).

❧

Exercise 25.1. Is 6 a quadratic residue mod 25? mod 35? mod 75? mod 95?

Exercise 25.2. Is 17 a quadratic residue mod 64? mod 65? mod 169? mod 338?

Exercise 25.3. Continue the example given after Corollary 25.3 to find the square roots of 11 mod 625 and mod 3125.

Exercise 25.4. Use Corollary 25.3 to find a square root of 13 mod 289, starting with the fact that $8^2 \equiv 13 \bmod 17$.

Exercise 25.5. Let $m > 1$ be an integer. Describe a simple condition for when -1 is a quadratic residue mod m.

Exercise 25.6. Let $m > 1$ be an odd integer. Describe a simple condition for when 2 is a quadratic residue mod m.

Exercise 25.7. Use Theorem 25.5 and the results of Exercise 23.3 to determine when 3 is a quadratic residue mod m, where $3 \nmid m$ and m has the factorization (25.4).

Exercise 25.8. Suppose that $\gcd(u, 10) = 1$. Describe a simple condition, in the spirit of Exercise 23.3, for when u is a quadratic residue mod 10^k, where $k \geq 3$. What if $k = 1$ or $k = 2$?

Exercise 25.9. **(a)** Develop an algorithm in the spirit of Corollary 25.3 for finding the square root of an odd number mod 2^{k+1} given a square root mod 2^k.
(b) Starting with the observation that $1^2 \equiv 1 \equiv 41 \bmod 8$, use your algorithm from part (a) to find a square root of 41 mod 64.

Exercise 25.10. Suppose that $p \neq q$ are odd primes. Show that a quadratic equation

$$x^2 + ax + b \equiv 0 \mod pq$$

has a solution iff it has solutions mod p and mod q.

Exercise 25.11. Suppose that p is an odd prime.
(a) Show that a quadratic equation

$$x^2 + ax + b \equiv 0 \mod p^2$$

has a solution iff it has a solution mod p.
(b) If r is a root of this equation mod p, what is the corresponding solution mod p^2?

Exercise 25.12. Read the proof of Proposition 25.2 carefully, and use the notation of that proof to explain why the formula of Corollary 25.3 is correct.

Exercise 25.13. What are the perfect squares mod 3? mod 9? mod 15? mod 27?

Exercise 25.14. Let p be an odd prime, and let k be a positive integer. Prove that an integer n is a perfect square mod p^2 iff either $n \equiv 0 \bmod p^2$ or n is a quadratic residue mod p. Does this still hold if $p = 2$?

Exercise 25.15. Give a detailed proof of Proposition 25.7. Use the proof of Proposition 25.6 as a model, and be mindful of the special conditions imposed by Proposition 25.4.

Exercise 25.16. Use the Chinese Remainder Theorem 11.1 to prove Theorem 25.8.

26 Jacobi symbols

Recall that, when $p \nmid a$, the Euler criterion asserts that

$$\left(\frac{a}{p}\right) = a^{\frac{p-1}{2}}. \tag{26.1}$$

Gauss's Lemma and the Law of Quadratic Reciprocity gave us more efficient means for computing Legendre symbols. However, there remains the difficulty of computing the value of (26.1) when p is a large prime and a has large composite value mod p. In this case we cannot apply quadratic reciprocity until a has been factored, and this may not be feasible for large values of a. It was Jacobi's insight to extend the notion of Legendre symbol to a larger context that makes these computations possible without requiring factorizations.[1]

We will see in Section 31 that Jacobi's generalization of the Legendre symbol also points to a practical and efficient algorithm for primality testing.

❧

To begin, suppose that p is an odd prime and that $a \in \mathbb{Z}$. Extend the definition of *Legendre symbol* for the pair (a, p) as follows

$$\left(\frac{a}{p}\right) = \begin{cases} 0 & \text{if } p|a, \\ 1 & \text{if } a \text{ is a quadratic residue mod } p, \\ -1 & \text{otherwise.} \end{cases}$$

Note that we have added the case where $p|a$, in which case the Legendre symbol is set equal to zero.

Suppose instead that $a \in \mathbb{Z}$ and $n > 1$ is an odd integer, *including the case where n is composite*. By the unique factorization theorem, n has a unique expression as a product of powers of distinct odd primes:

$$n = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}.$$

Define the *Jacobi symbol* for the pair (a, n) by

$$\left(\frac{a}{n}\right) = \left(\frac{a}{p_1}\right)^{e_1} \left(\frac{a}{p_2}\right)^{e_2} \cdots \left(\frac{a}{p_k}\right)^{e_k} \tag{26.2}$$

where each factor in the product above is a Legendre symbol. Observe that Jacobi symbols can take values $0, 1$ or -1.

[1] Carl Gustav Jacob Jacobi (1804–1851).

Note that if n is an odd prime, then the Jacobi symbol is simply the usual Legendre symbol.

Euler's criterion does *not* typically tell us the value of the Jacobi symbol directly. However, if $\gcd(a,n) = 1$, then Euler's criterion can be applied to each prime p_i in the factorization of n to compute each factor in the expression (26.2). The downside of this approach is that one needs to know the factorization of the number n.

Fortunately, there are easier ways to compute Jacobi symbols. The following elementary properties follow easily from similar properties of Legendre symbols.

Proposition 26.1. *Let $a, b \in \mathbb{Z}$, and let $m, n > 1$ be odd integers.*

(i) *If $a \equiv b \mod n$ then $\left(\frac{a}{n}\right) = \left(\frac{b}{n}\right)$.*

(ii) $\left(\frac{ab}{n}\right) = \left(\frac{a}{n}\right)\left(\frac{b}{n}\right)$.

(iii) $\left(\frac{a}{mn}\right) = \left(\frac{a}{m}\right)\left(\frac{a}{n}\right)$.

(iv) $\gcd(a,n) \neq 1$ *if and only if* $\left(\frac{a}{n}\right) = 0$.

(v) *If* $\gcd(a,n) = 1$ *then* $\left(\frac{a}{n^2}\right) = 1$.

(vi) *If* $\gcd(a,n) = 1$ *then* $\left(\frac{a^2}{n}\right) = 1$.

The proof is left as an exercise (see Exercise 26.3).

The converse of the property (vi) is **false** in general: It is possible that $\left(\frac{a}{n}\right) = 1$ even though a is *not* a quadratic residue mod n. For example, 2 is not a quadratic residue mod 15, even though

$$\left(\frac{2}{15}\right) = \left(\frac{2}{3}\right)\left(\frac{2}{5}\right) = (-1)(-1) = 1.$$

Moreover, notice that

$$2^{\frac{15-1}{2}} \equiv 2^7 \equiv 128 \equiv 8 \mod 15,$$

so that

$$2^{\frac{15-1}{2}} \neq \left(\frac{2}{15}\right).$$

In other words, a naive generalization of the Euler criterion may fail to hold when n is composite (see also Exercise 22.6).

❧

The next proposition shows that two important formulas for Legendre symbols generalize without change to Jacobi symbols.

Proposition 26.2. *Let $n > 1$ be an odd integer. Then*

$$\left(\frac{-1}{n}\right) = \begin{cases} 1 & \text{if } n \equiv 1 \mod 4 \\ -1 & \text{if } n \equiv 3 \mod 4 \end{cases} \quad \textit{and} \quad \left(\frac{2}{n}\right) = (-1)^{\frac{n^2-1}{8}}.$$

Recall that $(-1)^{\frac{n^2-1}{8}} = 1$ iff $n \equiv \pm 1 \bmod 8$. It follows that

$$\left(\frac{2}{n}\right) = \begin{cases} 1 & \text{if } n \equiv \pm 1 \mod 8. \\ -1 & \text{if } n \equiv \pm 3 \mod 8. \end{cases}$$

Proof. Another way of stating the first part is to assert that, if n is odd, then $\left(\frac{-1}{n}\right) \equiv n \mod 4$. This follows from the Euler criterion when n is prime. It then follows for more general n by the denominator product rule (iii) of Proposition 26.1.

To prove the second statement, observe that

$$(nm)^2 - 1 \;=\; n^2m^2 - n^2 + n^2 - 1 \;=\; n^2(m^2-1) + (n^2-1). \tag{26.3}$$

Since n is odd, we have $(-1)^{n^2} = -1$, so that

$$\begin{aligned} (-1)^{\frac{(nm)^2-1}{8}} &= (-1)^{\frac{n^2(m^2-1)+(n^2-1)}{8}} \\ &= \left((-1)^{n^2}\right)^{\frac{m^2-1}{8}} (-1)^{\frac{n^2-1}{8}} \\ &= (-1)^{\frac{m^2-1}{8}} (-1)^{\frac{n^2-1}{8}}. \end{aligned}$$

Suppose that

$$\left(\frac{2}{s}\right) \neq (-1)^{\frac{s^2-1}{8}} \tag{26.4}$$

for some odd integer $s > 1$, and let s be the smallest for which this occurs. We know that the proposition holds for primes, so s must be composite; that is, $s = nm$, where $n, m > 1$ are odd and strictly smaller than s. By the minimality of s, the proposition holds for n and m, so that

$$\left(\frac{2}{s}\right) = \left(\frac{2}{nm}\right) = \left(\frac{2}{n}\right)\left(\frac{2}{m}\right) = (-1)^{\frac{n^2-1}{8}}(-1)^{\frac{m^2-1}{8}} = (-1)^{\frac{(nm)^2-1}{8}} = (-1)^{\frac{s^2-1}{8}},$$

contradicting (26.4). The proposition now follows. ❑

The Law of Quadratic Reciprocity also generalizes to Jacobi symbols.

Theorem 26.3. *If $m, n > 1$ are odd integers, then*

$$\left(\frac{m}{n}\right) = \left(\frac{n}{m}\right)(-1)^{\frac{m-1}{2}\frac{n-1}{2}} = \begin{cases} \left(\frac{n}{m}\right) & \text{if } n \equiv 1 \mod 4 \quad \text{or if } m \equiv 1 \mod 4 \\ -\left(\frac{n}{m}\right) & \text{if } n \equiv m \equiv 3 \mod 4 \end{cases}$$

Note that this statement of quadratic reciprocity is presented a little differently from the original version for primes, in order to account for the case in which either Jacobi symbol is zero.

Proof. If $\gcd(m, n) \neq 1$ then both sides of the identity are zero. The more interesting case, in which $\gcd(m, n) = 1$, is left as an exercise (see Exercise 26.5). ❑

Theorem 26.3 and the propositions that precede it enable us to compute the Jacobi symbol $\left(\frac{a}{n}\right)$ without knowing the prime factorizations of a or n. In particular, it can be computationally feasible to compute $\left(\frac{a}{n}\right)$, even if the numbers a and n are too large for us to factor in a reasonable amount of time.

❧

Exercise 26.1. Evaluate the following Jacobi symbols using Propositions 26.1 and 26.2, but without recourse to Theorem 26.3:

(a) $\left(\frac{12}{45}\right)$ **(d)** $\left(\frac{7}{363}\right)$ **(g)** $\left(\frac{70}{125}\right)$

(b) $\left(\frac{13}{45}\right)$ **(e)** $\left(\frac{-7}{363}\right)$ **(h)** $\left(\frac{36}{77}\right)$

(c) $\left(\frac{7}{27}\right)$ **(f)** $\left(\frac{363}{7}\right)$ **(i)** $\left(\frac{-77}{15}\right)$

Exercise 26.2. Evaluate the following Jacobi symbols using the generalized quadratic reciprocity (Theorem 26.3) whenever it applies:

(a) $\left(\frac{7}{363}\right)$ **(c)** $\left(\frac{777}{363}\right)$ **(e)** $\left(\frac{1234}{4321}\right)$

(b) $\left(\frac{700}{363}\right)$ **(d)** $\left(\frac{7777}{363}\right)$ **(f)** $\left(\frac{20002}{10003}\right)$

Exercise 26.3. Prove Proposition 26.1.

Exercise 26.4. Show that if $\left(\frac{a}{n}\right) = -1$ then a is a quadratic **non**-residue mod n.

Exercise 26.5. Finish the proof of Theorem 26.3.
Here are some hints: Use the fact that quadratic reciprocity holds for odd primes. First, suppose that m is an odd prime and let n be the smallest odd composite number for which the theorem fails. Use an identity similar to (26.3) from the proof of Proposition 26.2 to derive a contradiction. This verifies the theorem when m is an odd prime and n is any odd number. Now repeat the process with a similar argument focusing on the general case for m.

Exercise 26.6. Evaluate the Jacobi symbol

$$\left(\frac{17292864462617}{17292864462677}\right)$$

without the help of an electronic device.

Exercise 26.7. Suppose that $n > 1$ is an even integer. Determine the exact conditions for when the Jacobi symbol

$$\left(\frac{n!}{n+1}\right)$$

is equal to 0, 1, or -1.
Hint: You might find Wilson's theorem helpful.

Exercise 26.8. If

$$a = 111111111111111111 \quad \text{and} \quad b = 11111111111111111111$$

(consisting respectively of 19 and 23 catenated decimal '1's), what is $\left(\frac{a}{b}\right)$?
Hint: While this is the same as Exercise 23.11, observe that the Jacobi symbol is easier to evaluate and that we no longer need to know in advance that a and b are prime numbers.

❧

Evidently if $n = k^2$ for some integer k then n is a quadratic residue in every modulus m (such that $\gcd(n, m) = 1$). The next exercise deals with the converse.

Exercise 26.9. Show that, if a positive integer $n \in \mathbb{Z}$ is a quadratic residue modulo every prime p (such that $p \nmid m$), then n is the square of an integer.
Here are some hints.
(a) Suppose n is not the square of an integer. Let n be minimal in this regard, and show that n is square-free, so that $n = p_1 p_2 \cdots p_s$ for some primes $p_1 < p_2 < \cdots < p_s$.
(b) Suppose that n is odd. Let c be a quadratic non-residue mod p_1. Show that there exists a positive integer m such that

$$\begin{aligned} m &\equiv 1 \mod 4 \\ m &\equiv c \mod p_1 \\ m &\equiv 1 \mod p_i \text{ for } i > 1. \end{aligned}$$

Then show that $\left(\frac{n}{m}\right) = -1$.
(c) Suppose that n is even, so that $n = 2p_1 p_2 \cdots p_s$. Show that there exists a positive integer m such that

$$\begin{aligned} m &\equiv 5 \mod 8 \\ m &\equiv 1 \mod p_i \text{ for each } i. \end{aligned}$$

Then show that $\left(\frac{n}{m}\right) = -1$.
(d) Use parts (a)-(c) to show that n cannot be a quadratic residue in every modulus.

Exercise 26.10. Let p be an odd prime. Prove that there exists an odd prime q such that q is a quadratic non-residue mod p.
Hint: See Exercise 22.10.

Exercise 26.11. Let p be an odd prime. Prove that there exists an odd prime q such that q is a quadratic residue mod p.
Hint: Use Exercise 20.17.

27 *Computing square roots mod p*

Suppose that p is an odd prime and that a is a quadratic residue mod p. How do we find an integer b such that $b^2 \equiv a \mod p$?

This is very easy when $p \equiv 3 \mod 4$. Since a is a quadratic residue mod p, Euler's criterion tells us that

$$a^{\frac{p-1}{2}} \equiv 1 \mod p,$$

so that

$$a^{\frac{p+1}{2}} \equiv a \mod p.$$

Since $p \equiv 3 \mod 4$, the expression $\frac{p+1}{4}$ is an integer. Set $b \equiv a^{\frac{p+1}{4}}$. We then have

$$b^2 \equiv a^{\frac{p+1}{2}} \equiv a \mod p. \tag{27.1}$$

❧

On the other hand, if $p \equiv 1 \mod 4$, then either $p \equiv 1 \mod 8$ or $p \equiv 5 \mod 8$.

Suppose $p \equiv 5 \mod 8$. Let c any quadratic non-residue[1] mod p, so that

$$c^{\frac{p-1}{2}} \equiv -1 \mod p.$$

If a is a quadratic residue, then

$$a^{\frac{p-1}{2}} \equiv 1 \mod p,$$

so that

$$a^{\frac{p-1}{4}} \equiv \pm 1 \mod p.$$

Since $8|(p+3)$, we set $b \equiv a^{\frac{p+3}{8}}$. We then have

$$b^2 \equiv a^{\frac{p+3}{4}} \equiv a^{\frac{p-1}{4}} \cdot a \equiv \pm a.$$

If $b^2 \equiv a$ we are done. If $b^2 \equiv -a$, then

$$(bc^{\frac{p-1}{4}})^2 \equiv -ac^{\frac{p-1}{2}} \equiv a.$$

❧

[1] For example, since $p \equiv 5 \mod 8$, we can set $c = 2$.

If $p \equiv 1 \bmod 8$, then either $p \equiv 1 \bmod 16$ or $p \equiv 9 \bmod 16$.

When $p \equiv 9 \bmod 16$ a slightly more complicated version of the previous argument yields a square root for the quadratic residue a.

When $p \equiv 1 \bmod 16$ one turns to the state of affairs mod 32, and so on. The result of this cascade of cases is the *Tonelli-Shanks algorithm*,[2] which will produce a square root for any quadratic residue modulo an odd prime p. Because this algorithm operates via a sequence of cases of indeterminate (but always finite) length, there is no resulting closed formula for the square root. It is an algorithm more suited to computer programming than to mathematical analysis. Instead, we will consider an alternative approach.

❧

Suppose once again that we seek the square roots of a quadratic residue a mod p, where $p \equiv 1 \bmod 4$.

Cipolla's algorithm offers an answer in two steps. To begin, find a value t such that $t^2 - a$ is a quadratic *non-residue*. In other words, we need to find t so that

$$\left(\frac{t^2 - a}{p}\right) = -1. \tag{27.2}$$

The value of t is found by trial and error. It can be shown that there is approximately a 50% chance of finding an acceptable value of t by guessing at random, so one would expect to find an acceptable value for t after a few guesses.[3] (See also Exercise 27.17.)

Having verified (27.2), introduce a symbol α into $\mathbb{Z}_p$ arithmetic with the rule that $\alpha^2 \equiv t^2 - a \bmod p$. Note that $\alpha \notin \mathbb{Z}_p$, since $t^2 - a$ is not a quadratic residue. Instead we have introduced a new element, in analogy to the imaginary $i = \sqrt{-1}$ used in the theory of complex numbers. Following this analogy, we continue doing mod p arithmetic in the larger ring

$$\mathbb{Z}_p[\alpha] = \{m + n\alpha \mid m, n \in \mathbb{Z}_p\},$$

always keeping in mind the rule that $\alpha^2 \equiv t^2 - a \bmod p$.

[2] Also known as the RESSOL algorithm. For a more detailed treatment, see [6, p. 32] or [20, p. 110].

[3] Although this might seem obvious at first glance, since half the units are quadratic non-residues, the reasons behind this are actually more subtle, since we are translating a only by quadratic residues t^2.

Example: Computing in $\mathbb{Z}_p[\alpha]$ we have

$$\begin{aligned}\frac{1+\alpha}{3-\alpha} \equiv \frac{1+\alpha}{3-\alpha}\cdot\frac{3+\alpha}{3+\alpha} &\equiv \frac{(1+\alpha)(3+\alpha)}{9-\alpha^2}\\ &\equiv \frac{3+4\alpha+\alpha^2}{9-\alpha^2}\\ &\equiv \frac{3+4\alpha+(t^2-a)}{9-(t^2-a)}\\ &\equiv [3+(t^2-a)][9-(t^2-a)]^{-1}+4[9-(t^2-a)]^{-1}\alpha,\end{aligned}$$

where $[3+(t^2-a)][9-(t^2-a)]^{-1}$ and $4[9-(t^2-a)]^{-1}$ are integers mod p. Note that the inverse of $9-(t^2-a)$ exists mod p because t^2-a is a quadratic non-residue, so that $t^2-a \not\equiv 9$.

The following proposition is helpful when doing arithmetic in $\mathbb{Z}_p[\alpha]$.

Proposition 27.1. *Suppose that $m, n \in \mathbb{Z}$.*

(i) $m+n\alpha \equiv 0$ *in* $\mathbb{Z}_p[\alpha]$ iff $m \equiv n \equiv 0 \bmod p$.

(ii) $(m+n\alpha)(r+s\alpha) \equiv 0$ *in* $\mathbb{Z}_p[\alpha]$
iff *either* $m \equiv n \equiv 0 \bmod p$, *or* $r \equiv s \equiv 0 \bmod p$, *or both.*

Proof. To prove (i), suppose that $m+n\alpha \equiv 0$ in $\mathbb{Z}_p[\alpha]$. If $n \not\equiv 0 \bmod p$, then n^{-1} exists mod p, so that $\alpha \equiv -n^{-1}m \in \mathbb{Z}_p$. This contradicts the assumption that $\alpha \notin \mathbb{Z}_p$ (because t^2-a is a quadratic non-residue). It follows that $n \equiv 0 \bmod p$, so that $m \equiv 0$ as well.

Conversely, if $n \equiv m \equiv 0 \bmod p$, then $m+n\alpha \equiv 0+0\alpha \equiv 0$ in $\mathbb{Z}_p[\alpha]$.

The proof of part (ii) is left as an exercise (see Exercise 27.1). ❑

We are now ready to find the square root of a in $\mathbb{Z}_p$. To do this, we set

$$b = (t+\alpha)^{\frac{p+1}{2}}, \tag{27.3}$$

and simplify the resulting algebra.

Theorem 27.2. *After simplification of the expression (27.3) we have*

- $b \in \mathbb{Z}_p$, *and*
- $b^2 \equiv a \mod p$.

Proof. We are given that $\alpha^2 \equiv t^2 - a$ in $\mathbb{Z}_p$. Since $t^2 - a$ is a quadratic non-residue mod p, the Euler criterion implies that

$$\alpha^{p-1} \equiv (t^2 - a)^{\frac{p-1}{2}} \equiv -1 \mod p.$$

It follows that $\alpha^p \equiv -\alpha$.

Observe that

$$(t+\alpha)^p = \sum_{k=0}^{p} \binom{p}{k} t^{p-k}\alpha^k \equiv t^p + \alpha^p,$$

since $\binom{p}{k} \equiv 0 \mod p$ for $1 \leq k \leq p-1$ (see Exercise 10.24). Moreover, Fermat's Theorem implies that $t^p \equiv t$, since t is a unit in $\mathbb{Z}_p$. Since we found that $\alpha^p \equiv -\alpha$, it now follows that

$$(t+\alpha)^p \equiv t^p + \alpha^p \equiv t - \alpha.$$

We now have

$$(t+\alpha)^{p+1} \equiv (t+\alpha)^p(t+\alpha) \equiv (t-\alpha)(t+\alpha) \equiv t^2 - \alpha^2 \equiv a,$$

It follows from the definition (27.3) that

$$b^2 \equiv (t+\alpha)^{p+1} \equiv a,$$

so that b is a square root of a in the ring $\mathbb{Z}_p[\alpha]$. It remains to show that b is actually a value in $\mathbb{Z}_p$.

We were given that a is a quadratic residue mod p. It follows that $a \equiv r^2$ for some $r \in \mathbb{Z}_p$. The previous argument implies that $b^2 \equiv r^2$ in $\mathbb{Z}_p[\alpha]$, so that

$$(b-r)(b+r) \equiv b^2 - r^2 \equiv 0$$

in $\mathbb{Z}_p[\alpha]$. It follows from Proposition 27.1 that either $b \equiv r$ or $b \equiv -r$. In particular, $b \in \mathbb{Z}_p$. ❑

Example: Let's find $\sqrt{2}$ mod 17. Since $17 \equiv 1 \mod 8$, we know that $\left(\frac{2}{17}\right) = 1$, so that $\sqrt{2}$ exists in $\mathbb{Z}_{17}$.

Setting $t = 3$ we have $t^2 - 2 = 7$. Since

$$\left(\frac{7}{17}\right) = \left(\frac{17}{7}\right) = \left(\frac{3}{7}\right) = -1,$$

this is an acceptable choice for t. It follows from (27.3) that

$$\sqrt{2} = (3+\alpha)^9,$$

where $\alpha^2 \equiv 7$. Of course, this still has to be simplified, since we want an answer in $\mathbb{Z}_{17}$.

To this end, observe that

$$(3+\alpha)^2 \;\equiv\; 9+6\alpha+\alpha^2 \;\equiv\; 9+6\alpha+7 \;\equiv\; 16+6\alpha \;\equiv\; 6\alpha-1.$$

After repeated squaring, similar computations yield

$$(3+\alpha)^4 \equiv 5\alpha-2 \quad \text{and} \quad (3+\alpha)^8 \equiv 9-3\alpha.$$

Hence,

$$(3+\alpha)^9 \;\equiv\; (9-3\alpha)(3+\alpha) \;\equiv\; 6,$$

so that $\sqrt{2} \equiv \pm 6 \bmod 17$.

Considering how small the modulus was in this example, it would have made more sense to make (at most) 8 consecutive guesses to find the answer. Cipolla's algorithm, while elegant in theory, can be cumbersome in practice.

❧

Exercise 27.1. Prove the second assertion of Proposition 27.1; that is,

$$(m+n\alpha)(r+s\alpha) \equiv 0$$

in $\mathbb{Z}_p[\alpha]$ iff $m \equiv n \equiv 0 \bmod p$ or $r \equiv s \equiv 0 \bmod p$ (or both).
Hint: Part (i) of Proposition 27.1 is helpful.

Exercise 27.2. Perform the following computations in $\mathbb{Z}_p[\alpha]$. Use the notation of this section, where $\alpha^2 \equiv t^2 - a \bmod p$. In every case, your final answer should be of the form $m+n\alpha$, where $m, n \in \mathbb{Z}_p$.

(a) α^3
(b) α^{-1}
(c) $(\alpha-a)^p$
(d) $(t+\alpha)^2$

Exercise 27.3. Perform the following computations in $\mathbb{Z}_{97}[\alpha]$, where $\alpha^2 \equiv 5$. In every case, your final answer should be of the form $m+n\alpha$, where $m, n \in \mathbb{Z}_{97}$.

(a) $(1+\alpha)(2+\alpha)$
(b) α^{-1}
(c) $\frac{10\alpha}{1+\alpha}$
(d) $(\alpha+100)^{98}$
(e) $\alpha^2 + \frac{1}{\alpha^2}$
(f) $\alpha^3 + \frac{1}{\alpha^3}$

Exercise 27.4. **(a)** When is α^n an element of $\mathbb{Z}_p$?
(b) Find a formula for the values $r, s \in \mathbb{Z}_p$ such that $\alpha^n = r + s\alpha$.

Exercise 27.5. Let $V = \mathbb{Z}_p \times \mathbb{Z}_p$ be the set of vectors (x, y) with coordinates $x, y \in \mathbb{Z}_p$. Define addition of vectors in the usual way, and define multiplication by the rule

$$(x,y)(u,v) \;=\; (xu + yv(t^2-a),\; xv+yu),$$

where t and a are defined as in the description of Cipolla's algorithm. Show that the function $f : V \to \mathbb{Z}_p[\alpha]$ defined by $f(x,y) = x + y\alpha$ is a one-to-one and onto function, such that

$$f(\vec{v}+\vec{w}) = f(\vec{v}) + f(\vec{w}) \quad \text{and} \quad f(\vec{v}\vec{w}) = f(\vec{v})f(\vec{w}),$$

for all $\vec{v}, \vec{w} \in V$.

Exercise 27.6. Let p be a positive integer prime.
(a) Suppose that $p \equiv 5 \bmod 8$. Prove that $\sqrt{-1} \equiv \pm 2^{\frac{p-1}{4}}$.
(b) Suppose that $p \equiv 5 \bmod 12$. Prove that $\sqrt{-1} \equiv \pm 3^{\frac{p-1}{4}}$.

Exercise 27.7. Let p be an odd prime integer, and let a be a quadratic **non-residue** mod p. Suppose that someone mistakenly believed that a had a square root mod p, and proceeded to use Cipolla's algorithm to generate the number

$$b = (t + \alpha)^{\frac{p+1}{2}},$$

as directed by that algorithm. What happens? What does b look like? How is b^2 related to a?

Exercise 27.8. Prove that $x^2 - \alpha \equiv 0$ has no solution $x \in \mathbb{Z}_p[\alpha]$.

Exercise 27.9. Let p be an odd prime integer of the form $p = 4m + 1$ for some integer m. Let a be a quadratic residue mod p. Choose t so that $t^2 - a$ is a quadratic non-residue mod p, and let

$$b = (t + \alpha)^{\frac{p+1}{2}},$$

as directed by Cipolla's algorithm, where $\alpha^2 = t^2 - a$ in $\mathbb{Z}_p[\alpha]$. Prove that

$$b \equiv \sum_{s=0}^{m} \binom{2m+1}{2s} t^{2(m-s)+1} (t^2 - a)^s \mod p,$$

and that

$$\sum_{s=0}^{m} \binom{2m+1}{2s+1} t^{2(m-s)} (t^2 - a)^s \equiv 0 \mod p.$$

Hint: Apply the binomial theorem to the identity $b = (t + \alpha)^{\frac{p+1}{2}}$, and recall what we proved about the value of b in Theorem 27.2 above.

Exercise 27.10. Use the formula from Exercise 27.9 to verify our computation of $\sqrt{2} \in \mathbb{Z}_{17}$ in the example earlier.

Exercise 27.11. Use Cipolla's algorithm to compute $\sqrt{10}$ mod 37.

Exercise 27.12. Recompute $\sqrt{10}$ mod 37 using the Tonelli-Shanks method. (See also Exercise 27.11.)

Exercise 27.13. Use the Tonelli-Shanks algorithm (outlined at the start of this section) to compute the following:

(a) $\sqrt{20} \mod 151$
(b) $\sqrt{2} \mod 71$
(c) $\sqrt{-3} \mod 103$
(d) $\sqrt{6} \mod 101$
(e) $\sqrt{5} \mod 61$
(f) $\sqrt{2} \mod 127$

Exercise 27.14. With regard to part (f) of Exercise 27.13, find a much easier way to compute $\sqrt{2}$ mod 127, using the fact that $2^7 = 128$.

Exercise 27.15. Use the results of Exercise 27.13(b) and the worked example in this section to compute $\sqrt{2}$ mod 1207.
Hint: $1207 = (71)(17)$.

Exercise 27.16. Let $m > 1$ be an integer, possibly composite. Suppose that u is a unit mod m of *odd* order $b = \mathrm{ord}_m(u)$. Find a simple formula for the square root of u mod m in terms of u and b.
Hint: Look for an analogy with the Tonelli-Shanks algorithm in the special case of $p \equiv 3 \bmod 4$.

Exercise 27.17. Here are some steps leading to a proof that a suitable value for t in Cipolla's algorithm always exists.

To begin, recall that half of the units mod p are quadratic non-residues. Let $a \in U_p$ be a quadratic residue, and suppose that c is a quadratic non-residue mod p.

(a) Prove that the set $\{c + ka \mid k = 0, \ldots, p-1\}$ exhausts *all* values mod p. In other words, prove that

$$\{c + ka \mod p \mid k = 0, \ldots, p-1\} = \{0, \ldots, p-1\}$$

(b) Prove that $c + ka$ is a quadratic residue mod p for some value of k.

(c) Let k be the smallest non-negative integer such that $c + ka$ is a quadratic residue mod p. Since we chose c to be a quadratic non-residue, we know that $k \neq 0$. Prove that there exists t such that $t^2 = c + ka \mod p$.

(d) Show that $t^2 - a$ is a quadratic non-residue mod p, as required for Cipolla's algorithm.

Exercise 27.18. Let p be an odd prime such that $p \equiv 2 \mod 3$. Combine Fermat's Theorem with the fact that $\gcd(3, p-1) = 1$ to find a formula for the cube root of any value mod p.

Exercise 27.19. Let p be an odd prime such that $p \equiv 7 \mod 9$, and suppose that u is a cubic residue mod p. Derive a simple formula for the cube root of u mod p in analogy to the square root formula (27.1).

28 Sums of squares

Possibly the most celebrated theorem of ancient mathematics is the Pythagorean theorem relating the sides of a right triangle: If a triangle has edge lengths $a \leq b \leq c$, where the two smaller edges meet at a right angle, then $a^2 + b^2 = c^2$. A proof of this theorem can be inferred from Figure 28.1.

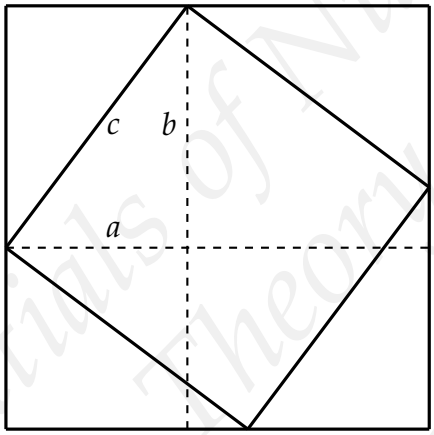

Figure 28.1: A "proof without words" of the Pythagorean theorem.

The Pythagoreans were especially interested in the cases where a, b, and c are integers. A sequence of three non-negative integers (a, b, c) is called a *Pythagorean triple* if $a^2 + b^2 = c^2$. The reader can verify the following examples of Pythagorean triples:

$$\begin{array}{ccc} 3^2+4^2=5^2 & 5^2+12^2=13^2 & 8^2+15^2=17^2 \\ 6^2+8^2=10^2 & 10^2+24^2=26^2 & 16^2+30^2=34^2 \\ 9^2+12^2=15^2 & 15^2+36^2=39^2 & 24^2+45^2=51^2 \\ \vdots & \vdots & \vdots \end{array} \tag{28.1}$$

A Pythagorean triple is *primitive* if $\gcd(a, b) = 1$. The triples in the top row of (28.1) are all primitive, while each column consists of triples that are multiples of the corresponding top row entry.

A simple way to generate Pythagorean triples arises from the algebraic identities:

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

After subtracting these identities, we have $(x+y)^2-(x-y)^2=4xy$, so that

$$(x-y)^2+4xy=(x+y)^2.$$

Setting $x=X^2$ and $y=Y^2$, we obtain *Euclid's formula:*

$$(X^2-Y^2)^2+(2XY)^2=(X^2+Y^2)^2. \tag{28.2}$$

The top row entries of (28.1) are generated by setting (X,Y) equal to $(1,2)$, $(2,3)$ and $(1,4)$, respectively. While Euclid's formula (28.2) is not quite sufficient to generate *all* Pythagorean triples, a small adjustment will do the trick.

Theorem 28.1. *All Pythagorean triples have the form*

$$Z^2(X^2-Y^2)^2+(2XYZ)^2=Z^2(X^2+Y^2)^2, \tag{28.3}$$

where $X,Y,Z\in\mathbb{Z}$.

Proof. Multiplying both sides of (28.2) by the integer Z^2 verifies that (28.3) is a valid algebraic identity, yielding a Pythagorean triple for all integer values of X, Y, and Z.

To prove that every Pythagorean triple has the form (28.3), suppose first that (a,b,c) is a ***primitive*** Pythagorean triple.

Since $\gcd(a,b)=1$, the integers a and b are not both even. If a and b are both odd, then

$$c^2\equiv a^2+b^2\equiv 1+1\equiv 2 \mod 4,$$

which is impossible, since every integer square is either 0 or 1 mod 4. It follows that a and b have opposite parity, so that c is odd.

Suppose, without loss of generality, that a is odd and b is even. We now have

$$b^2=c^2-a^2=(c-a)(c+a). \tag{28.4}$$

Since a and c are odd, both $c-a$ and $c+a$ are even, so that

$$\begin{aligned} c+a&=2s\\ c-a&=2t \end{aligned} \tag{28.5}$$

for some integers s and t.

If $k|s$ and $k|t$, then $k|(c+a)$ and $k|(c-a)$, so that $k|[(c+a)+(c-a)]=2c$ and $k|[(c+a)-(c-a)]=2a$. Since $\gcd(2a,2c)=2\gcd(a,c)=2$, it follows that $k|2$. But if $k=2$ then the identities (28.5) imply that $c+a\equiv c-a\equiv 0 \mod 4$, so that $2c\equiv 0 \mod 4$. Since c is odd, this is impossible. Therefore, $k\neq 2$, and $\gcd(s,t)=1$.

On combining (28.4) and (28.5) we have

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

Since $\gcd(s,t) = 1$, it follows from the Fundamental Theorem 9.3 that s and t are both integer squares. In other words,

$$s = X^2 \quad \text{and} \quad t = Y^2$$

for some integers X and Y. We now have

$$a = s - t = X^2 - Y^2$$
$$c = s + t = X^2 + Y^2$$

while $b^2 = 4st = (2XY)^2$, so that (a,b,c) conforms to (28.2).

More generally, if (a,b,c) is a Pythagorean triple with common factor Z, then

$$(a,b,c) = (Za', Zb', Zc') = Z(a',b',c'),$$

where (a',b',c') is primitive, so that

$$(a,b,c) = Z(X^2 - Y^2, 2XY, X^2 + Y^2) = (Z(X^2 - Y^2), 2XYZ, Z(X^2 + Y^2)),$$

for some integers X and Y. ❑

❧

Euclid's formula (28.3) for Pythagorean triples is described in the *Arithmetica*, an ancient Greek book of mathematics by Diophantus. Inspired by this presentation, Fermat was led to consider the more general equation

$$a^n + b^n = c^n, \tag{28.6}$$

for integers $n > 2$. He conjectured that, if $n > 2$, the equation (28.6) ***never*** holds for integers $a,b,c > 0$. Fermat famously wrote in the margin of his copy of *Arithmetica* that he had found a *demonstrationem mirabilem sane* (truly wonderful proof) of this conjecture, but that the margin was too small for him to explain his proof in detail.

While Fermat later published a proof that $a^4 + b^4 = c^4$ has no positive integer solutions, the details of his *demonstrationem mirabilem* were never revealed.[1] As a result, his more general conjecture, now known as *Fermat's Last Theorem*, remained open for over 400 years. During the centuries that followed many mathematicians verified Fermat's conjecture for various specific values of n, without

[1] That Fermat never published his *demonstrationem mirabilem* suggests that, as often happens, an apparently great idea when first conceived turned out to contain limitations or errors when reviewed later in careful detail. Take a lesson from this, and don't be discouraged when it happens to you.

solving the problem completely. It was not until 1994 that Fermat's Last Theorem was finally proved by Andrew Wiles for all $n > 2$.

A history of this famous conjecture and its proof can be found in [33, p. 210]. For an elementary presentation of Fermat's proof of the special case $n = 4$, see [21, p.72-74]. The proof by Wiles that turned Fermat's Last Theorem from a conjecture into a genuine theorem appears in [36].

❧

While the Pythagoreans were motivated by the geometry of right triangles, number theorists became interested in the more general question of when a positive integer n can be expressed as a sum of two integer squares. For example, we have:

$$\begin{array}{ll} 0^2+1^2=1 & 2^2+2^2=8 \\ 1^2+1^2=2 & 0^2+3^2=9 \\ 0^2+2^2=4 & 1^2+3^2=10 \\ 2^2+1^2=5 & 2^2+3^2=13 \end{array}$$

and so on. Evidently the integers 3, 6, 7, 11, and 12 cannot be expressed as sums of two integer squares.

Recall that if $a \in \mathbb{Z}$, then $a^2 \equiv 0$ or $a^2 \equiv 1 \bmod 4$. It follows that if $n = a^2 + b^2$, then

$$n \in \{0, 1, 2\} \mod 4.$$

In particular, if $n \equiv 3 \bmod 4$, then n is a not a sum of two integer squares. This explains some of the exceptions listed above. More generally, we have the following.

Proposition 28.2. *Suppose that p is a prime factor of n, that $p \equiv 3 \bmod 4$, and that p divides n an odd number of times. Then n cannot be expressed as a sum of two integer squares.*

Proof. If the proposition is false, let n be the smallest counterexample, where $n = a^2 + b^2$, for $a, b, \in \mathbb{Z}$.

Since $p|n$, we have $a^2 + b^2 \equiv 0 \bmod p$, so that

$$b^2 \equiv -a^2 \mod p.$$

If $p \nmid a$, then a and b are both units mod p, and

$$(ba^{-1})^2 \equiv -1 \mod p,$$

so that -1 is a quadratic residue mod p. This contradicts Corollary 22.4, since $p \equiv 3 \not\equiv 1 \bmod 4$.

It follows that $p|a$ and $p|b$, so that $p^2|n$. Write

$$n = p^2\tilde{n} \qquad a = p\tilde{a} \qquad b = p\tilde{b}.$$

Dividing by p^2, we have

$$\tilde{n}^2 = \tilde{a}^2 + \tilde{b}^2,$$

where $\tilde{n} < n$. Since p divides n an odd number of times, p also divides $\tilde{n}$ an odd number of times. Since n is the smallest counterexample to the proposition, we have a contradiction. Therefore, there are no counterexamples, and the proposition is true. ❑

❧

Proposition 28.2 explains why 3, 6, 7, 11, 12, and 30003 are not sums of integer squares. It turns out that the condition of Proposition 28.2 is the only obstruction. If n isn't divisible an odd number of times by any prime of the form $4k+3$, then n can indeed be expressed as a sum of squares. This remarkable fact, first proved by Fermat, is summarized by the next theorem.

Theorem 28.3 (Sums of Squares)**.** *Let $n \in \mathbb{N}$, and suppose that $n = s^2N$, where $s, N \in \mathbb{N}$ and N is square-free. Then n can be expressed as a sum of two integer squares iff every odd prime factor of N is congruent to* 1 mod 4*.*

In other words, the converse of Proposition 28.2 is true: n is a sum of squares iff every prime factor $p \equiv 3 \bmod 4$ divides n an even number of times.

A pair of lemmas will be helpful for the proof of Theorem 28.3. The first lemma, an algebraic identity, will allow us to address the situation for composite numbers.

Lemma 28.4. *For all a, b, c, d, we have*

$$(a^2 + b^2)(c^2 + d^2) = (ac - bd)^2 + (ad + bc)^2.$$

Lemma 28.4 is easily verified by simplifying the algebra.[2] This lemma implies that if n is a sum of squares and m is a sum of squares then nm is also a sum of squares.

The most challenging step in a proof of Theorem 28.3 is to show that every prime $p \equiv 1 \bmod 4$ is a sum of squares. The following lemma takes care of this.

[2]Simplifying the algebra verifies the lemma, but is not particularly edifying. Readers familiar with complex numbers should note that if $z = a + bi$ and $w = c + di$, then Lemma 28.4 is equivalent to the complex conjugation identity: $z\bar{z}w\bar{w} = zw\overline{zw}$.

Lemma 28.5. *If $p \in \mathbb{N}$ is prime and $p \equiv 1 \bmod 4$, then there exist $a, b \in \mathbb{N}$ such that $p = a^2 + b^2$.*

The following proof is due to Fermat.[3]

Proof. Suppose that $p \equiv 1 \bmod 4$ is a positive prime. By Proposition 22.4, the integer -1 is a quadratic residue mod p. It follows that $Y^2 \equiv -1 \bmod p$ for some integer u, so that

$$Y^2 + 1^2 = kp$$

for some integer k.

More generally, let s be the ***smallest*** positive integer such that

$$a^2 + b^2 = sp \tag{28.7}$$

for some $a, b \in \mathbb{Z}$. Note that

$$a^2 + b^2 \equiv 0 \mod p, \tag{28.8}$$

so that either of a or b can be replaced with $p - a$ or $p - b$ without altering (28.8). Since s is minimal, it now follows that $0 < a < \frac{p}{2}$ and $0 < b < \frac{p}{2}$, so that

$$sp = a^2 + b^2 < \frac{p^2}{2}.$$

Therefore, $0 < s < \frac{p}{2}$. Moreover, the minimality of s implies that $\gcd(a, s) = \gcd(b, s) = 1$.

We will prove the Lemma by showing that $s = 1$.

Suppose that $s > 1$. In this case we can reduce the identity (28.7) mod s, so that $a^2 + b^2 \equiv 0 \bmod s$. Reducing a and b mod s, there exist units $c \equiv a$ and $d \equiv b \bmod s$ such that $|c|, |d| < \frac{s}{2}$. It follows that

$$c^2 + d^2 \equiv a^2 + b^2 \equiv 0 \mod s,$$

so that $c^2 + d^2 = st$ for some integer $t > 0$. Moreover,

$$st = c^2 + d^2 < \frac{s^2}{4} + \frac{s^2}{4} = \frac{s^2}{2},$$

so that $0 < t < \frac{s}{2}$. In particular, note that $t < s$.

Applying Lemma 28.4, we obtain

$$s^2 pt = (a^2 + b^2)(c^2 + d^2) = (ad - bc)^2 + (ac + bd)^2.$$

[3] The description of Fermat's proof given here is essentially that given in Davenport's beautiful little book [8, p. 105-106].

Since

$$ad - bc \equiv ab - ba \equiv 0 \quad \text{and} \quad ac + bd \equiv a^2 + b^2 \equiv 0 \mod s,$$

it follows that $ad - bc = As$ and $ac + bd = Bs$ for some integers A and B. Therefore,

$$s^2pt = (As)^2 + (Bs)^2,$$

so that $A^2 + B^2 = pt$. Since $t < s$, this violates the minimality of s.

It follows that $s = 1$, so that $a^2 + b^2 = p$. ❑

We are ready to prove the full characterization for sums of two integer squares.

Proof of Theorem 28.3. Let $n \in \mathbb{N}$, and suppose that $n = s^2N$, where N is square-free.

If N has a prime factor $p \equiv 3 \mod 4$, then p divides n an odd number of times (since every prime divides s^2 an *even* number of times), so n cannot be expressed as a sum of two integer squares, by Proposition 28.2.

Suppose instead that every prime factor p of N satisfies $p \not\equiv 3 \mod 4$. Note that $2 = 1^2 + 1^2$ is a sum of squares, while every prime $p \equiv 1 \mod 4$ is a sum of squares by Lemma 28.5. It then follows from Lemma 28.4 that any finite product of these primes is also expressible as a sum of two integer squares. Therefore, there exist integers $a, b \in \mathbb{Z}$ such that $N = a^2 + b^2$, so that

$$n \;=\; s^2N \;=\; (sa)^2 + (sb)^2$$

is also expressible as a sum of two integer squares. ❑

❧

Fermat's theorem on sums of squares was later generalized. While some numbers (such as 7) cannot even be expressed as a sum of 3 integer squares, Lagrange proved that *every* positive integer can be expressed as a sum of 4 integer squares. Several different proofs of Lagrange's Four Square Theorem can be found in [13]. An elementary proof is described in [8, p. 111-114].

An exact description of the 3 square case was later given by Legendre. It is fairly elementary to show that a sum of 3 squares can never have the form $4^s(8k + 7)$ (see Exercise 28.15). Legendre proved that all other positive integers can indeed be expressed as sums of 3 squares. Unfortunately all known proofs of Legendre's theorem are rather difficult. A proof is given by Gauss in Section 291 of [11, p. 336-338]. For a modern proof, see [2].

❧

Exercise 28.1. Describe in detail the proof of the Pythagorean theorem implicit in Figure 28.1.

Exercise 28.2. Suppose that (a, b, c) is a primitive Pythagorean triple. Prove that:
(a) Exactly one of the integers a and b is divisible by 3.
(b) Exactly one of the integers a, b, or c is divisible by 5.

Exercise 28.3. Suppose that $a, b, c \in \mathbb{Z}$ and $a^2 + b^2 = c^2$. Prove that at least one of the values a and b is divisible by 4. Can you prove it without using Euclid's formula?

Exercise 28.4. Suppose that $a, b, c, d \in \mathbb{Z}$ and $a^2 + b^2 + c^2 = d^2$.
(a) Prove that, if d is even, then ***all*** of the values a, b, c are even.
(b) Prove that, if d is odd, then exactly one of the values a, b, c is odd.

Exercise 28.5. Suppose that $a, b, c, d, e \in \mathbb{Z}$ and $a^2 + b^2 + c^2 + d^2 = e^2$.
(a) Prove that, if e is odd, then exactly one of the values a, b, c, d is odd.
(b) Prove that, if e is even, all of the values a, b, c, d have the same parity.

Exercise 28.6. Prove that the Pythagorean triple given by Euclid's formula (28.2) is primitive iff the integers X and Y are relatively prime and of opposite parity.

Exercise 28.7. Prove that if k is an odd integer, then there is a Pythagorean triple of the form (k, b, c).

Exercise 28.8. Prove Lemma 28.4.

Exercise 28.9. Which of the following integers can be expressed as sums of two integer squares? (Answer Yes or No for each value.)

(a) 15
(b) 18
(c) 38
(d) 98
(e) 125
(f) 686
(g) 1100
(h) 12100
(i) 10!

Exercise 28.10. Suppose that $n = a^2 + b^2$, where $a, b \in \mathbb{Z}$. Find a formula for integers A and B such that $2n = A^2 + B^2$.

Exercise 28.11. **(a)** Express 10 as a sum of squares.
(b) Use part (a) and Lemma 28.4 to express 10000 as a sum of squares.

Exercise 28.12. Differences of squares are much easier to work with than sums of squares.
(a) Prove that every odd integer can be expressed a difference of two integer squares.
(b) Prove that an even integer m is a difference of two integer squares iff $4|m$.
Hint: What happens when you simplify $(a+1)^2 - a^2$? What about $(a+2)^2 - a^2$?

Exercise 28.13. Let p be an odd prime. Prove that:
(a) If $p = a^2 + 2b^2$ for some integers a and b, then $p \equiv 1$ or $p \equiv 3 \bmod 8$.
(b) If $p > 3$ and $p = a^2 + 3b^2$ for some integers a and b, then $p \equiv 1 \bmod 3$.
Note: Fermat proved that the converses of (a) and (b) are also true.

Exercise 28.14. **(a)** Show that 65 can be written as a sum of squares in two different ways.
(b) Use part (a) to show that there are two non-congruent right triangles with side lengths a, b, and 65, where $a < b < 65$.

Exercise 28.15. Suppose that $n = 4^s(8k + 7)$, where k and s are non-negative integers. Prove that n cannot be expressed as a sum of 3 integer squares.
Hint: First, consider the case where $s = 0$ and n is odd. Then show that if $n = a^2 + b^2 + c^2$ is a sum of 3 squares and if n is even, then so are a, b, and c. What happens next?

29 *Pseudorandom numbers*

Think of a list of 20 random digits between 0 and 9. Quickly now! Write them down. Now go make a cup of coffee, then come back and look again at your numbers. Do they really look random to you? Or was there an unintended pattern to your choices? A careful answer to this question might involve fancy statistical analysis, taking us outside the scope of number theory. But it is clear that sequences like

$$0, 4, 7, 0, 4, 7, 0, 4, 7, 0, 4, 7, \ldots$$
$$1, 2, 3, 4, 5, 6, 7, 8, 9, 8, 7, 6, \ldots$$
$$0, 1, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, \ldots$$

do not look random; indeed you can certainly guess what is likely to happen after the '...' in each case. On the other hand, the pattern of the sequence

$$3, 2, 7, 5, 3, 0, 1, 8, 1, 9, 8, 2, 6, \ldots$$

is harder to discern.[1]

While true randomness can only be (arguably) obtained from physical phenomena (such as coin flipping, dice rolls, radioactive emissions, Brownian motion of particles, ambient noise), it is often sufficient in applications to use *pseudorandom numbers*; that is, sequences of numbers that do have a predictable pattern, but *appear* random from the perspective of simple statistical tests. The advantage to pseudorandom numbers is that they are easy to generate in large quantities and at high speed on a computer. The disadvantage is that the underlying pattern (even if unknown) may undermine the application of those numbers, whether they are used to test scientific models (Monte-Carlo methods) or to provide cryptographic security.

A *pseudorandom number generator* (or PRNG) is based on the following principle: Beginning with a *seed* number s_0, an iterative procedure then transforms the seed s_0 into a new number s_1. The value s_1 is now the new seed, returned to the same procedure to produce s_2 and so on. The iterative procedure typically involves a simple algebraic expression, computed in a certain fixed modulus. The pseudorandom sequence $s_0, s_1, s_2, \ldots$ is then usually modified in some way to suit its application: for example, only the last digit (or bit) of each number s_i might be used. If the numbers s_i are expressed in binary, the even parity

[1] Least significant decimal digit (i.e., the last digit) of the iterated sequence $x_{n+1} = 7x_n + 1 \mod 23$ seeded with the number $x_0 = 3$.

bit of each s_i may be taken.[2] Bit sampling of a pseudorandom stream offers a simulation of a sequence of coin flips, which can then be easily manipulated to form random integers in any finite range.

In the discussion that follows we will examine some famous methods of generating pseudorandom numbers and examine briefly the utility and security of each method.

Our first example is due to John von Neumann,[3] who (in the era before modern computing) would generate pseudo-random numbers using the following 'middle square' method: Start with a (random?) initial 4-digit number s_0 as the seed.[4] To produce a new number in the sequence, square the current number to produce a string of 8 digits (padding to the left with zeroes if necessary). Then pick out the middle 4 digits.

For example, starting with $s_0 = 2013$, we obtain the sequence

$$2013, 4226, 8590, 7881, 1101, 2122, 5028, 2807, 8792, 2992, 9520, 6304, \ldots$$

which doesn't look too bad at first glance. On the other hand, starting with $s_0 = 1969$, we obtain the less appealing sequence

$$1969, 8769, 8953, 1562, 4398, 3424, 7237, 3741, 9950, 25, 6, 0, 0, 0, \ldots$$

revealing immediately a serious defect of this PRNG: If a number less than 10 ever appears, the sequence will consist of only zeroes after that.

The next proposition illustrates a defect of most PRNGs.

Proposition 29.1. *Regardless of the seed, the middle square algorithm produces a periodic sequence. That is, there exist k, M such that $a_{m+k} = a_m$ for all $m > M$.*

The proof is left as an exercise (see Exercise 29.4).

❧

While the middle square method can be tweaked in various ways to avoid some of its substantial defects, we turn instead to a more popular method, the *Linear Congruential Generator*.

The idea is to use a simple arithmetic iteration to quickly generate a sequence of numbers, and to perform this arithmetic in a modulus that prevents the numbers

[2] Recall from Section 13 that the *even parity bit* of a positive integer n is obtained by summing the bits of n (counting the 1's that appear in the binary expression of n) and returning the value of this sum modulo 2.

[3] John von Neumann (1903–1957) was a Hungarian-American mathematician and physicist, and a founder of modern computer science.

[4] Do you see something circular about this? How do we choose a random seed? This chicken-and-egg paradox is an issue with every pseudorandom number generator and is part of the 'pseudo' in pseudorandomness.

from becoming too large, while also, one hopes, disguising any pattern that the numbers form. Sampling only a few bits of each output number will enhance the "pseudorandom" effect.

For the Linear Congruential Generator (LCG) we use the simplest possible algebraic iteration, a linear function. Choose constants a (the multiplier), b (the increment), m (the modulus), and x_0 (the seed), so that $\gcd(a, m) = 1$. The elements of the pseudorandom sequence are given by

$$x_{n+1} \equiv ax_n + b \mod m \qquad \text{for } n \geq 0. \tag{29.1}$$

If $b = 0$, the LCG is said to be *purely multiplicative*.

An LCG is necessarily periodic, with period at most m. For computational convenience, the value of m in most implementations is $m = 2^{32}$ or $m = 2^{64}$. These moduli are typically used because computers store numbers in binary, at which point a computation mod 2^{32} (resp. 2^{64}) is simply a truncation to the last 32 (resp. 64) bits.

Remark: If $\gcd(a - 1, m) = 1$, then one should choose a seed x_0 such that $\gcd(x_0 - b(1 - a)^{-1}, m) = 1$. To understand why, see Exercise 29.12.

Example: Let $x_0 = 38$ and $x_{n+1} = 4x_n + 7 \mod 81$.

This results in the mod 81 sequence

$$36, 70, 44, 21, 10, 47, 33, 58, 77, 72, 52, 53, 57, 73, 56, 69, 40, \ldots$$

having least significant bits[5]

$$0, 0, 0, 1, 0, 1, 1, 0, 1, 0, 0, 1, 1, 1, 0, 1, 0, \ldots$$

and even parity bits

$$0, 1, 1, 1, 0, 1, 0, 0, 0, 0, 1, 0, 0, 1, 1, 1, 0, \ldots$$

Example: Let $x_0 = 23$ and $x_{n+1} = 11x_n + 4 \mod 100$. This results in the mod 100 sequence

$$23, 57, 31, 45, 99, 93, 27, 1, 15, 69, 63, 97, 71, 85, 39, 33, 67, \ldots$$

having least significant bits

$$1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, \ldots \text{ (not very helpful!)}$$

and even parity bits

$$0, 0, 1, 0, 0, 1, 0, 1, 0, 1, 0, 1, 0, 0, 0, 0, 1, \ldots$$

[5]That is, the last digit of each number when expressed in binary.

While the even parity bit sequence looks better, the fact that this example only generates odd numbers is a serious flaw, since at most half of the values mod 100 can ever appear. Measures can be taken to avoid losing so much of the modulus in this way.

Different choices of a, b and m produce different quality generators. For example, if a positive integer k divides all three of the constants a, b, m, it will divide every value x_n for $n \geq 1$. Moreover, if b and m are both even, then parity of x_n is constant for $n \geq 1$. For these and other reasons, one should be sure that $\gcd(a, m) = 1$ and $\gcd(b, m) = 1$.

Following the notation of (29.1), conditions for maximal period length are given by the following theorem.

Theorem 29.2 (Hull-Dobell)**.** *The period of an LCG is equal to its modulus m if and only if the following conditions hold:*

- $\gcd(b, m) = 1$.
- $a \equiv 1 \bmod p$ *for every prime* p *such that* $p|m$.
- $a \equiv 1 \bmod 4$ *if* $4|m$.

The proof of this theorem lies beyond the scope of these notes.[6] Note that the first example above satisfies the optimality conditions of the Hull-Dobell theorem, whereas the second example does not.

Example: Setting $a = 45$, $b = 229$, $m = 128$, with seed $s = 28$ yields the mod 128 sequence

81 34 95 24 29 126 11 84 41 26 119 80 117 118 35 12 1 18 15 8 77 110
59 68 89 10 39 64 37 102 83 124 49 2 63 120 125 94 107 52 9 122 87 48
85 86 3 108 97 114 111 104 45 78 27 36 57 106 7 32 5 70 51 92 17 98
31 88 93 62 75 20 105 90 55 16 53 54 99 76 65 82 79 72 13 46 123 4 25
74 103 0 101 38 19 60 113 66 127 56 61 30 43 116 73 58 23 112 21 22
67 44 33 50 47 40 109 14 91 100 121 42 71 96 69 6 115 28 81 ...

repeating for the first time at the 129th step. While the least significant bitstream of this sequence is the very uninteresting 1010101010..., since the numbers above alternate even by odd, the resulting even parity bitstream is

1000001111001110100101100001100111000111011000001001000000110100011111
0000110001011010011110011000111000100111110110111111010111...

❧

[6] See, for example, [16, p. 17]. See also Exercises 29.9, 29.10, and 29.11 for some hints on how to prove parts of Theorem 29.2.

A special case of LCG, the *Lehmer generator*, uses $b = 0$. In this case one should choose $m = p^k$ for some prime number p, and choose a to be a primitive root (or element of high order) mod m, in order to have a large period. It is also vital that $p \nmid x_0$ (the reader should think about why this is important).

Example: Let $a = 47$, $b = 0$, $m = 81$, and $x_0 = 28$. This results in the mod 81 sequence

20 49 35 25 41 64 11 31 80 34 59 19 2 13 44 43 77 55 74 76
8 52 14 10 65 58 53 61 32 46 56 40 17 70 50 1 47 22 62 79
68 37 38 4 26 7 5 73 29 67 71 16 23 28 20 ...

repeating for the first time at the 55th step, because $\phi(81) = 54$ and 47 is primitive mod 81. Replacing 47 with a non-primitive multiplier would shorten the period of the resulting sequence. The sequence above results in a least significant bitstream of

0111101100110101110000001011000010011001010001111110 10 ...

and an even parity bitstream of

0111111100111110011111100001101001111111011111010101 01 ...

❧

The *Blum-Blum-Shub* (BBS) algorithm can be used to generate a *cryptographically secure* stream of pseudorandom bits [5] (also [34, p. 336]). The idea is to produce a stream of pseudorandom bits using a private key, so that an eavesdropper that intercepts any number of bits cannot use that data to predict the next bit (or discern the overall pattern in the sequence). Since LCG's are highly insecure in this regard, alternative methods are needed for secure pseudorandomness.

To implement Blum-Blum-Shub, choose a modulus $m = pq$, where p and q are two large distinct primes such that $p \equiv q \equiv 3 \bmod 4$. Then choose a seed x_0 that is relatively prime to m. The numbers p, q, and m remain private and are not shared.

The elements of the pseudorandom sequence are now given by

$$x_{n+1} \equiv x_n^2 \mod m \qquad \text{for } n \geq 0.$$

The actual output of the algorithm is *not* the numbers x_n, but the sequence of even parity bits of these numbers (or, alternatively, the sequence of bits given by $x_n \bmod 2$).

The condition that $p \equiv q \equiv 3 \bmod 4$ helps to guarantee a longer period. To this end, one should also choose p and q so that $\gcd(\phi(p-1), \phi(q-1))$ is a small integer relative to m.

❧

If you're interested in pseudorandom numbers be sure to look up the *Mersenne twister*. This PRNG is increasingly common in modern software applications, being more effectively pseudorandom than LCGs (and their relatives), while still computationally feasible. Unfortunately the algorithm for the Mersenne twister is outside the scope of these notes, but there are many sources and discussions available on the internet.

❧

Exercise 29.1. Show that, if a number less than 100 ever appears in the middle square algorithm, then future numbers in the sequence will also be less than 100, and reduce to a sequence of zeroes shortly thereafter.

Exercise 29.2. Show that, if a number divisible by 100 ever appears in the middle square algorithm, then future numbers in the sequence will also be divisible by 100.

Exercise 29.3. The next few questions involve the middle square algorithm.
(a) Show that, if a number of the form '$a100$' ever appears in the middle square algorithm, where a is a decimal digit, then future numbers in the sequence will form a periodic pattern of 4 repeating numbers.
(b) Show that, if a number of the form '$a500$' ever appears, where a is a decimal digit, then future numbers in the sequence become constant.
(c) Show that, if a number of the form '$a600$' ever appears, where a is a decimal digit, then future numbers in the sequence have the same form.
(d) Show that, if a number of the form '$ab00$' ever appears, where a is a decimal digit and b is an even digit, then future numbers in the sequence eventually have the same form as in part (c).
(e) Show that, if a number of the form '$ab00$' ever appears, where a is a decimal digit and b is odd digit, then future numbers in the sequence eventually have the same form as in part (a).

Exercise 29.4. Prove Proposition 29.1.

Exercise 29.5. Consider the LCG defined by $x_n \equiv 6x_{n-1} + 9 \bmod 25$, where $x_0 = 7$.
(a) What is x_5?
(b) What is the length of the period of this LCG?

Exercise 29.6. Consider the LCG defined by $x_n \equiv 5x_{n-1} + 3 \bmod 8$, where $x_0 = 4$.
(a) What is the length of the period of this LCG?
(b) What is the least significant bit sequence generated by this LCG?
(c) What is the even parity bit sequence generated by this LCG?

Exercise 29.7. Consider the LCG defined by $x_n \equiv 3x_{n-1} + 5 \bmod 8$, where $x_0 = 4$.
(a) What is the length of the period of this LCG?
(b) What is the least significant bit sequence generated by this LCG?
(c) What is the even parity bit sequence generated by this LCG?

Exercise 29.8. Consider the Lehmer generator defined by $x_n \equiv 2x_{n-1} \bmod 27$, where $x_0 = 7$.
(a) What is the length of the period of this LCG?
(b) What is the least significant bit sequence generated by this Lehmer generator?
(c) What is the even parity bit sequence generated by this Lehmer generator?

Exercise 29.9. Consider the LCG defined by $x_{n+1} \equiv ax_n + b \bmod m$, where $4|m$.
(a) Show that if a is even then x_n is constant mod 4 for all $n \geq 2$.
(b) Show that if $a \equiv 3 \bmod 4$ then $x_{n+2} \equiv x_n \bmod 4$ for all $n \geq 0$.

Exercise 29.10. Consider the LCG defined by $x_{n+1} \equiv ax_n + b \bmod m$, where $p|m$ and $p|b$ for some odd prime p. Show that, for $n \geq 1$, either p *always* divides x_n or *never* divides x_n.

Exercise 29.11. Consider the LCG defined by $x_{n+1} \equiv ax_n + b \bmod m$, where $p|m$. Show that, if $a \not\equiv 1 \bmod p$, then $x_{n+k} \equiv x_n \bmod p$, where $k = \mathrm{ord}_p(a)$.
Hint: Show that, if $a \not\equiv 1 \bmod p$ and $k = \mathrm{ord}_p(a)$, then $1 + a + \cdots + a^{k-1} \equiv 0 \bmod p$.

Exercise 29.12. Consider the LCG defined by $x_{n+1} \equiv ax_n + b \bmod m$, where $\gcd(a-1, m) = 1$.
(a) What happens if $x_0 \equiv b(1-a)^{-1} \bmod m$?
(b) What happens if $\gcd(x_0 - b(1-a)^{-1}, m) \neq 1$?

Exercise 29.13. Show that, if $\gcd(a-1, m) = 1$, then the nth term generated by an LCG with parameters a, b, m and seed x_0 is given by the formula

$$x_n \equiv a^n x_0 + b(a-1)^{-1}(a^n - 1) \mod m.$$

Exercise 29.14. Suppose you are given a sequence of numbers $x_0, x_1, x_2, \ldots, x_n$, where $n > 4$. You are told these numbers come from an LCG mod m (where m is given), but that the values of a, b have been forgotten. How can you recover them?

Exercise 29.15. Suppose that a sequence of numbers x_i is generated via the Blum-Blum-Shub algorithm. Prove that

$$x_n \equiv \left(x_0^{2^n \bmod \mathrm{lcm}(p-1, q-1)}\right) \mod m.$$

Why doesn't this formula undermine the security of the Blum-Blum-Shub algorithm?

Exercise 29.16. Prove that if distinct primes $p \equiv q \equiv 3 \bmod 4$ then every quadratic residue mod $m = pq$ has at least one square root that is itself also a quadratic residue mod m.

Exercise 29.17 (Exploratory Exercise)**.** Suppose Bob chooses a secret pair of large primes p and q and then broadcasts a pseudorandom sequence of large integers $\{x_n\}$ generated by the Blum-Blum-Shub algorithm. Suppose that Eve records some of the integers x_n (say, 100 of them in succession). How can Eve use the data $x_k, x_{k+1}, \ldots, x_{k+99}$ to more easily guess the secret modulus m?
Remark: The point of this exercise is that, when using Blum-Blum-Shub, the public output should consists of one bit from each of the x_i. The full values of the x_i should remain private.

Exercise 29.18 (Exploratory Exercise)**.** Invent a PRNG that is not periodic. Given a suitable seed, the generator should yield a sequence of decimal digits as long as needed. Each subsequent digit in the sequence should yield a number that depends on the previous entries, but does *not* require any seeds other than the initial seed (nor any other 'external' sources of randomness).
Remark: Your invented PRNG will probably not be computationally efficient. That's ok: it's one of the points of this exercise.

30 Elementary primality testing

Given $n \in \mathbb{Z}$, how can we determine if n is prime without actually factoring n?

If n is composite, then n must have a prime factor $p \leq \sqrt{n}$. So one way to test for primality is to attempt division of n by p for all primes $p \leq \sqrt{n}$. But this is impractical for large n, even on a fast computer. There are better approaches.

❧

Fermat's theorem says that, if p is prime and $\gcd(a, p) = 1$, then $a^p \equiv a \bmod p$. Suppose that n is a large odd number. It follows from Fermat's theorem that, if $2^p \not\equiv 2 \bmod n$, then n cannot be prime.

On the other hand, if $2^n \equiv 2 \bmod n$, we still don't know whether or not n is prime. If such a number n is composite, n is said to be a *pseudoprime to base 2*. For example, $341 = 31 \cdot 11$ is a pseudoprime to base 2.

Pseudoprimes to other bases are defined analogously. A composite number n is a pseudoprime to base a if $\gcd(a, n) = 1$ and $a^n \equiv a \bmod n$. Unfortunately there are pseudoprimes for every base. In fact, some numbers are pseudoprimes for every base at once!

An composite integer is a *Carmichael number* if $a^n \equiv a \mod n$ for all integers a. In particular, a Carmichael number is a pseudoprime to every base.

The smallest Carmichael number is 561. Here is a condition for finding more of them.

Theorem 30.1. *A composite positive integer n is a Carmichael number if and only if n is square-free, and if, for each prime $p|n$, we have $(p-1)|(n-1)$.*

For example, the number $561 = 3 \cdot 11 \cdot 17$ is square-free. Moreover, since $560 = 2^4 \cdot 5 \cdot 7$ is divisible by $2, 10$, and 16, it follows that 561 is a Carmichael number.

On the other hand, while $341 = 11 \cdot 31$ is also square-free, 340 is not divisible by 30, so the theorem implies that 341 is not a Carmichael number. Indeed, 341 is not even a pseudoprime to base 3.

Proof of Theorem 30.1. To begin, suppose that n is a Carmichael number. If a prime p divides n, write $n = p^k Q$, where $\gcd(p, Q) = 1$. Let r be a primitive root for U_{p^k}. By the Chinese remainder theorem, there exists $s \in U_n$ such that $s \equiv 1 \bmod Q$ and $s \equiv r \bmod p^k$. Since n is a Carmichael number and s is a unit mod n, we have

$$s^{n-1} \equiv 1 \mod n,$$

so that

$$s^{n-1} \;=\; 1 + tn \;=\; 1 + tp^k Q,$$

for some integer t. Modulo p^k this becomes

$$r^{n-1} \equiv 1 \mod p^k.$$

Since r is primitive mod p^k, the divisibility theorem for order implies that $\phi(p^k)|(n-1)$, so that

$$p^{k-1}(p-1)|(n-1). \tag{30.1}$$

In particular, $(p-1)|(n-1)$ for each prime p dividing n. Moreover, if $k > 1$, then (30.1) implies that $p|(n-1)$. Since $p|n$, this contradicts the assumption that p is prime. It follows that $k = 1$, so that n is square-free.

Conversely, suppose that n is a square-free composite number, and that each prime p dividing n satisfies $(p-1)|(n-1)$. In this case we can write $n = p_1 \cdots p_k$ for primes $p_1 < \cdots < p_k$. Let $a \in \mathbb{Z}$. By Fermat, we have $a^{p_i} \equiv a \mod p_i$ for each i.

Suppose that $\gcd(a, n) = 1$. For each i we have $n - 1 = (p_i - 1)b_i$ for some b_i. It follows that

$$a^{n-1} \;=\; a^{(p_i-1)b_i} \;=\; (a^{p_i-1})^{b_i} \;\equiv\; 1 \mod p_i.$$

The Chinese remainder theorem then implies $a^{n-1} \equiv 1 \mod n$, so that $a^n \equiv a \mod n$.

By a similar argument, $p_j^n \equiv p_j \mod p_i$ for each $i \neq j$, while obviously $p_j^n \equiv 0 \equiv p_j \mod p_j$ as well. Again the Chinese remainder theorem implies that $p_j^n \equiv p_j \mod n$ for each j.

More generally, if $\gcd(a, n) \neq 1$ then $a = de$, where $\gcd(e, n) = 1$ and $d = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$. Therefore,

$$a^n = d^n e^n \equiv de \equiv a \mod n$$

by the preceding cases.

It follows in every case that n is a Carmichael number. ❑

It can be shown that there are infinitely many Carmichael numbers, although the proof is quite difficult (see [1]).

The existence of Carmichael numbers suggests that a better approach is needed for efficient primality testing.

❧

Exercise 30.1. Show that 91 is a pseudoprime to base 3, but not to base 2.

Exercise 30.2. Show that 341 is a pseudoprime to base 2, but not to base 3.

Exercise 30.3. Show that 2701 is a pseudoprime to base 2 and to base 3, but not to base 5.

Exercise 30.4. Show that 9017 is composite by showing that Fermat's theorem is violated when considering powers of 2. (Use a calculator or computer for this exercise.)

Exercise 30.5. Show that, if n is a pseudoprime to base 2, so is $2^n - 1$. It follows that there are infinitely many pseudoprimes to base 2.

Exercise 30.6. Use Theorem 30.1 to prove that 1105 and 2465 are both Carmichael numbers.

Exercise 30.7.
(a) Prove that $7 \cdot 11 \cdot 13 \cdot 41$ is Carmichael.

(b) Prove that $3 \cdot 5 \cdot 47 \cdot 89$ is Carmichael.

(c) Prove that $11 \cdot 13 \cdot 17 \cdot 31$ is Carmichael.

(d) Prove that $7 \cdot 11 \cdot 13 \cdot 31$ is **not** Carmichael.

31 Advanced primality testing

Recall that if p is an odd prime and $p \nmid a$ then Euler's criterion tells us that

$$\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}} \mod p.$$

The right hand expression is easy to compute (by repeated squaring, for example) even for large numbers, provided a computer is available.

On the other hand, we saw in the Section 26 that, if n is composite, then it is *possible* that

$$\left(\frac{a}{n}\right) \not\equiv a^{\frac{n-1}{2}} \mod n. \tag{31.1}$$

This provides a primality test: Given a large positive odd number n, choose a smaller positive integer a at random, and compute $\left(\frac{a}{n}\right)$ using the methods of the Section 26. If $\left(\frac{a}{n}\right) \not\equiv a^{\frac{n-1}{2}} \mod n$ then we know *for certain* that n is composite.

On the other hand, if $\left(\frac{a}{n}\right) \equiv a^{\frac{n-1}{2}} \mod n$, then n may or may not be prime. We still do not know. While this sounds like a fatal flaw for this proposed test, we will show that, if n is composite, then (31.1) will hold for at least $1/2$ of all possible choices of a. This means that a false positive (making n look prime when it is not) occurs with probability at most $1/2$. If we perform the test using 10 values of a, each chosen independently[1] and uniformly[2] at random, then the probability that a composite n will be falsely labelled prime every time is at most $(0.5)^{10} \approx 0.001$. If we run the test 100 times, the probability of falsely claiming n is prime goes down to less than 10^{-30}. On a typical computer this test allows one to determine primality for a very large number (too large to factor) with vanishingly small probability of error.

The test outlined above is called the *Solovay-Strassen Test*.[3] In order to justify these probabilistic assertions, we need to prove the following theorem.

Theorem 31.1. *Let $n > 1$ be an odd composite integer. There are at least $\frac{n}{2}$ positive integers $a \in \{1, 2, \ldots, n\}$ such that the Solovay-Strassen test applied to a and n will reveal that n is composite.*

In other words, there is at most a 50% chance that the Solovay-Strassen test will falsely claim that n is prime.

[1] The outcome of one choice should have no effect on how the other choices are made.
[2] Each value should be equally likely to be chosen.
[3] First described by Robert Solovay and Volker Strassen in 1977. See [31].

Before we prove Theorem 31.1, we will need a couple of lemmas. We begin with a lemma that guarantees the inequality (31.1) holds *at least once* when n is composite. Denote by U_n the group of units mod n.

Lemma 31.2. *If $n > 1$ is an odd composite integer, there exists $a \in U_n$ such that the inequality (31.1) holds.*

Proof. First, suppose that $n = p^e m$ for some prime p and $e \geq 2$, where $p \nmid m$.

Let r be a primitive root mod p^e, and let $a = r^2$. Evidently,

$$\left(\frac{a}{n}\right) = \left(\frac{r^2}{n}\right) = 1.$$

If $a^{\frac{n-1}{2}} \equiv 1 \mod n$ then $r^{n-1} \equiv 1 \mod p^e$. Since r is primitive for p^e, this implies that $\phi(p^e)|(n-1)$. In other words,

$$p^{e-1}(p-1)|(n-1),$$

so that $p|(n-1)$. Since $p|n$ this is a contradiction. Therefore, the inequality (31.1) holds for a and n in this case.

Next, suppose instead that $n = p_1 \cdots p_k$ where $p_1 < \cdots < p_k$ are prime and $k > 1$. Suppose also that

$$\left(\frac{a}{n}\right) \equiv a^{\frac{n-1}{2}} \mod n \qquad \text{for all } a \in U_n. \tag{31.2}$$

Since a is a unit mod n, this means that $a^{\frac{n-1}{2}} \equiv \left(\frac{a}{n}\right) \equiv \pm 1 \mod n$.

If $a \in U_n$, and if $a^{\frac{n-1}{2}} \equiv -1 \mod n$, then $a^{\frac{n-1}{2}} \equiv -1 \mod p_1$. Use the Chinese Remainder Theorem to obtain $c \in U_n$ so that

$$c \equiv a \mod p_1 \quad \text{and} \quad c \equiv 1 \mod p_s \text{ for } s > 1.$$

We then have

$$c^{\frac{n-1}{2}} \equiv -1 \mod p_1 \quad \text{and} \quad c^{\frac{n-1}{2}} \equiv 1 \mod p_s \text{ for } s > 1.$$

It follows that $c^{\frac{n-1}{2}} \not\equiv \pm 1 \mod n$, so that

$$\pm 1 \equiv \left(\frac{c}{n}\right) \not\equiv c^{\frac{n-1}{2}} \mod n$$

This contradicts our assumption (31.2). Hence,

$$a^{\frac{n-1}{2}} \equiv 1 \mod n$$

for all $a \in U_n$. It follows from (31.2) that

$$\left(\frac{a}{n}\right) = 1 \qquad \text{for all } a \in U_n. \tag{31.3}$$

Meanwhile, let d be a quadratic non-residue mod p_1, and use the Chinese Remainder Theorem to obtain $a \in U_n$ so that

$$a \equiv d \mod p_1 \quad \text{and} \quad a \equiv 1 \mod p_s \text{ for } s > 1.$$

We then have

$$\left(\frac{a}{n}\right) = \left(\frac{a}{p_1}\right)\left(\frac{1}{p_2}\right)\cdots\left(\frac{1}{p_s}\right) = -1,$$

contradicting (31.3). It follows that (31.2) must be false, so that the inequality (31.1) holds for some $a \in U_n$. ❑

The next lemma addresses *how often* (31.1) holds when $a \in U_n$.

Lemma 31.3. *Let $n > 1$ be an odd composite integer. There are at least $\frac{\phi(n)}{2}$ values $a \in U_n$ such that the inequality (31.1) holds.*

Proof. Suppose that

$$\left(\frac{b_i}{n}\right) \equiv b_i^{\frac{n-1}{2}} \mod n$$

for some values $b_1, \ldots, b_k \in U_n$.

By Lemma 31.2, there is an $a \in U_n$ such that (31.1) holds. For each b_i, we then have

$$\left(\frac{ab_i}{n}\right) \equiv \left(\frac{a}{n}\right)\left(\frac{b_i}{n}\right) \equiv \left(\frac{a}{n}\right) b_i^{\frac{n-1}{2}} \not\equiv a^{\frac{n-1}{2}} b_i^{\frac{n-1}{2}} \equiv (ab_i)^{\frac{n-1}{2}} \mod n$$

In other words, for every $b \in U_n$ violating (31.1) there is at least one distinct value $ab \in U_n$ satisfying (31.1). Therefore, at least half of the elements of U_n satisfy (31.1). ❑

The previous lemmas together lead to Theorem 31.1, stated earlier.

Proof of Theorem 31.1. Lemma 31.3 asserts that at most $\frac{\phi(n)}{2}$ values of $a \in U_n$ will give false positives for primality when $\gcd(a, n) = 1$.

Meanwhile, if $\gcd(a, n) \neq 1$ (which can be quickly determined by Euclid's algorithm) then n is immediately revealed to be composite. It follows that at most $\frac{\phi(n)}{2}$ values of $a \in \mathbb{Z}_n$ will give false positives for primality.

Since $\phi(n) \leq n - 1$ for all odd $n > 1$, the Solovay-Strassen test gives a false positive for primality in fewer than $\frac{n-1}{2}$ cases, and therefore gives a correct answer at least 50% of the time, when the value a is chosen uniformly at random from the set $\{1, \ldots, n-1\}$. ❑

❧

As elegant as the Solovay-Strassen test is, most modern primality testing software uses a different algorithm, known to be even faster. Like the Solovay-Strassen test, the *Miller-Rabin test* is also probabilistic, but gives false positives at most 1/4 of the time, resulting in fewer trials needed for a given tolerance of error. The Miller-Rabin test also uses quadratic residue theory and is even easier to describe than Solovay-Strassen. Unfortunately, the proof that Miller-Rabin is accurate with the claimed probability is more difficult, and beyond the scope of these notes. We will describe the test, and refer to reader to any of [10, 23, 27] for a proof of accuracy. The test is named for Gary Miller [19], who discovered an early version of the test, and Michael Rabin [23], who proved that the test determines primality with at least 75% accuracy (Theorem 31.5 below).

Lemma 31.4. *Let p be an odd prime integer. Suppose $p-1 = 2^e b$, where b is odd. If $p \nmid a$ then either*

$$a^b \equiv 1 \mod p,$$

or

$$a^{2^s b} \equiv -1 \mod p$$

for some non-negative integer $s < e$.

Proof. By Fermat's Theorem 15.1,

$$a^{2^e b} = a^{p-1} \equiv 1 \mod p.$$

Let t be the smallest non-negative integer such that

$$a^{2^t b} \equiv 1 \mod p. \tag{31.4}$$

If $t = 0$, then $a^b \equiv 1 \mod p$.

If $t > 0$, then set $x = a^{2^{t-1} b}$, so that

$$x^2 \equiv a^{2^t b} \equiv 1 \mod p.$$

Since p is prime, it follows that $x \equiv \pm 1 \bmod p$. Since t was chosen to be minimal so that (31.4) holds, we must have $x \equiv -1$, that is,

$$a^{2^{t-1} b} \equiv x \equiv -1 \mod p.$$

Set $s = t - 1$ to complete the proof of the lemma. ❑

Since Lemma 31.4 holds for all odd primes, this gives us another primality test: Given a large odd integer n, choose a random integer a such that $2 \leq a \leq n-1$. If $\gcd(a, n) \neq 1$, then n is composite. Otherwise, test to see if Lemma 31.4 holds.

If Lemma 31.4 fails to hold, then n is certainly composite. On the other hand, if Lemma 31.4 is true for this choice a, then we don't know if n is composite or prime. As with the Solovay-Strassen test, the Miller-Rabin test is always correct if it reports "composite," but may give a false answer if it reports "prime."

Theorem 31.5. *Let $n > 1$ be an odd composite integer. There are at least $\frac{3\phi(n)}{4}$ units $a \in U_n$ such that the Lemma 31.4 is violated for the pair (a, n).*

Since Lemma 31.4 is also violated whenever $\gcd(a, n) \neq 1$, a false positive (making n look prime when it is not) occurs with probability at most $1/4$. If we perform the test using 10 values of a, each chosen independently and uniformly at random, then the probability that a composite n will be falsely labelled prime every time is at most $(0.25)^{10} \approx 0.000001$. If we run the test 100 times, the probability of falsely claiming n is prime goes down to less than 10^{-60}. For a proof of Theorem 31.5, see any of [10, 23, 27].

In practice, the Miller-Rabin test is implemented as follows. Given a large odd number n, factor the powers of 2 out of $n - 1$ to express $n - 1 = 2^e b$, where b is odd. Then compute $a^b \mod n$. If

$$a^b \equiv \pm 1 \mod n,$$

the test says "probably prime."

Otherwise, if

$$a^b \not\equiv \pm 1 \mod n,$$

then compute

$$a^{2b}, a^{4b}, a^{8b}, \ldots, a^{2^{e-1}b} \mod n.$$

If -1 appears on this list at any point, then n is "probably prime."
If 1 appears on this list *before* -1 has appeared,[4] or if ± 1 never appear at all, then n is certainly composite.

❧

Exercise 31.1. Suppose we run the Solovay-Strassen test just once to determine if the number 15 is prime. What is the exact probability that the test is correct?

Exercise 31.2. Suppose we run the Miller-Rabin test just once to determine if the number 15 is prime. What is the exact probability that the test is correct?

Exercise 31.3 (Project Exercise)**.** Write a program in a computer language of your choice to determine the primality of a given integer n, by testing for divisibility by 2 and by all subsequent odd numbers up to $\sqrt{n}$. Use your program to determine which of following numbers are prime:

2003, 65537, 90901, 90109, 10909, 4115181073

[4] You should think about why -1 will never appear *after* 1 has appeared.

Exercise 31.4 (Project Exercise)**.** Write a program in a computer language of your choice to implement the Miller-Rabin test on a personal computer. Use your program to determine which of following numbers are prime:

4115181073
293417457040049
10000000000000001
1111111111111111111
862720373833956440063030449

Your program should run the Miller-Rabin test at least 50 times per trial to assure a high probability of accuracy.

If you did the previous exercise, how do the speeds of the algorithms compare when applied to the same numbers?

Exercise 31.5 (Project Exercise)**.** Write a program in a computer language of your choice to implement the Solovay-Strassen test on a personal computer. Use your program to determine which of following numbers are prime:

4115181073
293417457040049
10000000000000001
1111111111111111111
862720373833956440063030449

Your program should run the Solovay-Strassen test at least 100 times per trial to assure a high probability of accuracy.

If you did the previous two exercises, how do the speeds of the algorithms compare when applied to the same numbers?

Exercise 31.6. If you did either of Exercises 31.4 or 31.5, then write a program that draws on your previous written primality test to find the smallest prime number p such that $p > 10^{20}$.

Exercise 31.7 (Project Exercise)**.** Write a program in a computer language of your choice to find the prime factorization of an integer using the Miller-Rabin test to determine primality and Pollard's Rho (see Section 14) to find proper factors of composites. Use your program to find the prime factorizations of the numbers

12193263122374638001
3777907654122664049668 7
19643484842775033859113 89
13168757201679012362139921 8107

32 *Continued fractions*

In this section we explore some basic properties of continued fractions, including their relation to Euclid's algorithm and fundamental properties of relatively prime pairs.

❧

We begin with an alternative view of Euclid's algorithm. Consider the rational number

$$\frac{53}{16}.$$

Using elementary arithmetic we typically write

$$\frac{53}{16} = 3 + \frac{5}{16}. \tag{32.1}$$

This is the same as using division with remainder to write

$$53 = 16 \cdot 3 + 5.$$

In order to find $\gcd(53, 16)$, we iterate division with remainder as follows:

$$\begin{aligned} 53 &= 16 \cdot 3 + 5 \\ 16 &= 5 \cdot 3 + 1. \end{aligned}$$

This procedure, called Euclid's algorithm, implies that $\gcd(53, 16) = 1$. Moreover, we can fold the preceding equations together:

$$1 = 16 - 5 \cdot 3 = 16 - (53 - 16 \cdot 3) \cdot 3 = 16 \cdot 10 - 53 \cdot 3,$$

thereby obtaining an integer solution to the linear equation

$$16x + 53y = 1,$$

namely, $x = 10$ and $y = -3$.

Let's rewrite this procedure in another way. The equation $16 = 5 \cdot 3 + 1$ is equivalent to the equation

$$\frac{16}{5} = 3 + \frac{1}{5}.$$

On combining this with (32.1) we have

$$\frac{53}{16} = 3 + \frac{1}{\frac{16}{5}} = 3 + \frac{1}{3 + \frac{1}{5}}. \tag{32.2}$$

The final expression in (32.2) is called the *simple* or *regular continued fraction* expansion for the number $\frac{53}{16}$.

Similarly,

$$\frac{72}{25} = 2 + \frac{22}{25} = \cdots = 2 + \cfrac{1}{1 + \cfrac{1}{7 + \cfrac{1}{3}}}. \tag{32.3}$$

For shorthand we will write

$$[a_0, a_1, \ldots, a_n] = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\cdots + \cfrac{1}{a_n}}}. \tag{32.4}$$

Note that the continued fraction $[a_0, a_1, \ldots, a_n]$ always has the value 1 in each numerator of the display (32.4).

For example,

$$\begin{aligned} [2] &= 2 \\ [2,1] &= 2 + \frac{1}{1} = 3 \\ [2,1,7] &= 2 + \cfrac{1}{1 + \cfrac{1}{7}} = \frac{23}{8} \\ [2,1,7,3] &= \frac{72}{25}, \end{aligned}$$

as in (32.3).

The continued fractions $[a_0], [a_0, a_1], [a_0, a_1, a_2], \ldots$ leading up to $[a_0, \ldots, a_n]$ are called the *convergents* to $[a_0, \ldots, a_n]$. The numbers $a_0, a_1, a_2 \ldots$ are called the *partial quotients* (or simply the *quotients*) of the continued fraction $[a_0, \ldots, a_n]$.

Comparing the last two convergents of the continued fraction for $\frac{72}{25}$ we find that

$$\frac{72}{25} - \frac{23}{8} = \frac{72 \cdot 8 - 25 \cdot 23}{200} = \frac{1}{200}.$$

In particular,

$$72 \cdot 8 - 25 \cdot 23 = 1,$$

solving $72x + 25y = 1$ with the integers $x = 8$ and $y = -23$.

We will show later that this always holds: if $\gcd(a, b) = 1$ then the numerator and denominator of the next-to-last convergent to $\frac{a}{b}$ solve the linear equation $ax + by = 1$ (up to a possible change of sign).

Any rational number $\frac{a}{b}$ can be expanded into a regular continued fraction of finite length. This follows from the fact that Euclid's algorithm always terminates after a finite number of steps. What about regular continued fraction expansions for *irrational numbers*? Let x be any positive real number. Then x must lie between two natural numbers, $a_0 \leq x < a_0 + 1$; that is,

$$x = a_0 + r_0,$$

where $0 \leq r_0 < 1$. If $r_0 = 0$ then x is an integer and we are finished. If $0 < r_0 < 1$ then $\frac{1}{r_0} > 1$, so that

$$\frac{1}{r_0} = a_1 + r_1,$$

for some integer a_1 and some real $0 \leq r_1 < 1$. Hence,

$$x = a_0 + \frac{1}{\frac{1}{r_0}} = a_0 + \frac{1}{a_1 + r_1}.$$

We can continue this way, at each step defining a_k to be the integer part of the reciprocal $1/r_{k-1}$ of the previous remainder r_{k-1}. This procedure will terminate (yielding a finite continued fraction for x) if and only if x is rational.

❧

Using the preceding algorithm we can show that the simple continued fraction expansion for a positive rational number x is *unique*. For suppose that

$$x = [a_0, a_1, \ldots, a_n] = [b_0, b_1, \ldots, b_k]$$

with $a_n \neq 1$ and $b_k \neq 1$. Since the integer (whole number) part of

$$x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\cdots + \cfrac{1}{a_n}}}$$

is a_0, it follows that $x = a_0 + r_0$ with $0 \leq r_0 < 1$. But the same argument implies that the integer part of x is b_0. Therefore, $a_0 = b_0$. After subtracting $a_0 = b_0$ from both continued fractions and then taking reciprocals we obtain

$$[a_1, \ldots, a_n] = [b_1, \ldots, b_k].$$

A similar argument implies that $a_1 = b_1$ and so on, leading to the conclusion that $n = k$ and $a_i = b_i$ for $i = 1, \ldots, n$. Note however that we need to assume

that $a_n \neq 1$ and $b_k \neq 1$ for this argument to work at the final step, since

$$\begin{aligned}[a_0, a_1, \ldots, a_{n-1}, 1] &= a_0 + \cfrac{1}{a_1 + \cfrac{1}{\cdots + \cfrac{1}{a_{n-1} + \cfrac{1}{a_n}}}} \\ &= a_0 + \cfrac{1}{a_1 + \cfrac{1}{\cdots + \cfrac{1}{a_{n-1} + 1}}} \;=\; [a_0, a_1, \ldots, a_{n-1} + 1].\end{aligned}$$

❧

When the numbers a_i are integers, the value of $[a_0, a_1, \ldots, a_n]$ is always a rational number; that is, it can always be expressed as the ratio a/b of two relatively prime integers a and b. For integers $k = 0, 1, \ldots n$ denote

$$\frac{p_k}{q_k} = [a_0, a_1, \ldots, a_k],$$

also called the *kth convergent* of $[a_0, a_1, \ldots, a_n]$.

Computing the first few convergents we find that

$$\begin{aligned}[a_0] &= a_0, \\ [a_0, a_1] &= a_0 + \frac{1}{a_1} = \frac{a_0 a_1 + 1}{a_1}, \\ [a_0, a_1, a_2] &= a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2}} = \frac{a_0 a_1 a_2 + a_2 + a_0}{a_1 a_2 + 1}.\end{aligned}$$

so that, for example,

$$\begin{array}{lll} p_0 = a_0, & p_1 = a_0 a_1 + 1, & p_2 = a_0 a_1 a_2 + a_2 + a_0, \\ q_0 = 1, & q_1 = a_1, & q_2 = a_1 a_2 + 1. \end{array}$$

For notational convenience we will also formally define $p_{-1} = 1$ and $q_{-1} = 0$.

The most important properties of regular continued fractions are encoded in the following theorem.

Theorem 32.1. *For all $a_0, a_1, \ldots, a_n$,*

$$\begin{bmatrix} a_0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_1 & 1 \\ 1 & 0 \end{bmatrix} \cdots \begin{bmatrix} a_n & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{bmatrix} \tag{32.5}$$

where $\gcd(p_n, q_n) = \gcd(p_{n-1}, q_{n-1}) = 1$.

It will follow from its proof that the matrix identity (32.5) is still true for any complex values of a_k for which the continued fraction makes sense; that is, provided no zero denominators are ever generated in the continued fraction expansion. However, if the values a_k are not integers, then the values p_k and q_k may not be integers either.

Proof. The proof is by induction on the number of matrices in the product (32.5). To begin, we have

$$\begin{bmatrix} a_0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} p_0 & p_{-1} \\ q_0 & q_{-1} \end{bmatrix}$$

by the definition of p_k and q_k above. Suppose the theorem is true for the product of k matrices, so that, for example,

$$\begin{bmatrix} a_0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_1 & 1 \\ 1 & 0 \end{bmatrix} \cdots \begin{bmatrix} a_{k-1} & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} p_{k-1} & p_{k-2} \\ q_{k-1} & q_{k-2} \end{bmatrix}.$$

Denote

$$\begin{bmatrix} a_1 & 1 \\ 1 & 0 \end{bmatrix} \cdots \begin{bmatrix} a_k & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} p & \tilde{p} \\ q & \tilde{q} \end{bmatrix}.$$

Since this is also a product of only k matrices, our induction assumption implies that

$$\frac{p}{q} = [a_1, \ldots, a_k] \quad \text{and} \quad \frac{\tilde{p}}{\tilde{q}} = [a_1, \ldots, a_{k-1}],$$

where $\gcd(p, q) = \gcd(\tilde{p}, \tilde{q}) = 1$. It follows that

$$[a_0, a_1, \ldots, a_k] = a_0 + \frac{1}{[a_1, \ldots, a_k]} = a_0 + \frac{q}{p} = \frac{a_0 p + q}{p}$$

and

$$[a_0, a_1, \ldots, a_{k-1}] = a_0 + \frac{1}{[a_1, \ldots, a_{k-1}]} = a_0 + \frac{\tilde{q}}{\tilde{p}} = \frac{a_0 \tilde{p} + \tilde{q}}{\tilde{p}}.$$

where $\gcd(a_0 p + q, p) = \gcd(p, q) = 1$ and $\gcd(a_0 \tilde{p} + \tilde{q}, \tilde{p}) = \gcd(\tilde{p}, \tilde{q}) = 1$.

Meanwhile,

$$\begin{bmatrix} a_0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_1 & 1 \\ 1 & 0 \end{bmatrix} \cdots \begin{bmatrix} a_k & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} a_0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} p & \tilde{p} \\ q & \tilde{q} \end{bmatrix} = \begin{bmatrix} a_0 p + q & a_0 \tilde{p} + \tilde{q} \\ p & \tilde{p} \end{bmatrix}.$$

The theorem now follows by induction on k. ❑

Corollary 32.2. *For integers $n \geq 2$,*

$$p_n = a_n p_{n-1} + p_{n-2} \tag{32.6}$$

$$q_n = a_n q_{n-1} + q_{n-2} \tag{32.7}$$

$$p_n q_{n-1} - p_{n-1} q_n = (-1)^{n+1} \tag{32.8}$$

$$p_n q_{n-2} - p_{n-2} q_n = (-1)^n a_n. \tag{32.9}$$

Proof. By Theorem 32.1,

$$\begin{aligned}\begin{bmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{bmatrix} &= \begin{bmatrix} a_0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_1 & 1 \\ 1 & 0 \end{bmatrix} \cdots \begin{bmatrix} a_{n-1} & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_n & 1 \\ 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} p_{n-1} & p_{n-2} \\ q_{n-1} & q_{n-2} \end{bmatrix} \begin{bmatrix} a_n & 1 \\ 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} a_n p_{n-1} + p_{n-2} & p_{n-1} \\ a_n q_{n-1} + q_{n-2} & q_{n-1} \end{bmatrix}.\end{aligned}$$

The identities (32.6) and (32.7) now follow. To prove (32.8) take the determinant of both sides of (32.5). To prove (32.9), use (32.6) and (32.7), along with the basic properties of determinants, to obtain

$$\begin{aligned} p_n q_{n-2} - p_{n-2} q_n &= \det \begin{bmatrix} p_n & p_{n-2} \\ q_n & q_{n-2} \end{bmatrix} \\ &= \det \begin{bmatrix} a_n p_{n-1} + p_{n-2} & p_{n-2} \\ a_n q_{n-1} + q_{n-2} & q_{n-2} \end{bmatrix} \quad \text{(by identities (32.6) and (32.7))} \\ &= \det \begin{bmatrix} a_n p_{n-1} & p_{n-2} \\ a_n q_{n-1} & q_{n-2} \end{bmatrix} \\ &= a_n (p_{n-1} q_{n-2} - p_{n-2} q_{n-1}) \\ &= a_n (-1)^n. \end{aligned}$$

For the third identity above we use the fact that subtracting one column from another does not change the determinant, while the last identity follows from (32.8). ❑

The identity (32.8) of Corollary 32.2 implies that the continued fraction formulation for Euclid's algorithm always provides a solution to the Diophantine equation $ax + by = 1$ when a and b are relatively prime positive integers.

Corollary 32.3. *If* $\gcd(a, b) = 1$ *where* $0 < b < a$ *then the Diophantine equation*

$$ax + by = 1$$

is solved by setting

$$x = (-1)^{n+1} q_{n-1} \quad \text{and} \quad y = (-1)^n p_{n-1},$$

where $\frac{p_{n-1}}{q_{n-1}}$ *is the next-to-last convergent to the continued fraction for* $\frac{a}{b}$.

Proof. Following the notation in the statement of the corollary, we have $a = p_n$ and $b = q_n$. The corollary now follows from the identity (32.8). ❑

❧

The recursion identities (32.6) and (32.7) provide a useful shorthand for more rapid computations of the convergents to a continued fraction. Consider the regular continued fraction

$$[a_0, a_1, a_2, a_3, a_4] \;=\; [2, 4, 2, 3, 6].$$

For notational convenience we begin with

$$\begin{array}{ll} p_{-1} = 1 & p_{-2} = 0 \\ q_{-1} = 0 & p_{-2} = 1. \end{array}$$

Since $a_0 = 2$, we have

$$\begin{aligned} p_0 = 2p_{-1} + p_{-2} \;&=\; 2 \cdot 1 + 0 \;=\; 2 \\ q_0 = 2q_{-1} + q_{-2} \;&=\; 2 \cdot 0 + 1 \;=\; 1 \end{aligned}$$

so that

$$\begin{array}{lll} p_0 = 2 & p_{-1} = 1 & p_{-2} = 0 \\ q_0 = 1 & q_{-1} = 0 & p_{-2} = 1. \end{array}$$

Since $a_1 = 4$, we have

$$\begin{aligned} p_1 = 4p_0 + p_{-1} \;&=\; 4 \cdot 2 + 1 \;=\; 9 \\ q_1 = 4q_0 + q_{-1} \;&=\; 4 \cdot 1 + 0 \;=\; 4 \end{aligned}$$

so that

$$\begin{array}{llll} p_1 = 9 & p_0 = 2 & p_{-1} = 1 & p_{-2} = 0 \\ q_1 = 4 & q_0 = 1 & q_{-1} = 0 & p_{-2} = 1. \end{array}$$

All of this is easier to read if we drop the p's and q's from the notation and make a table:

6	3	2	4	2		a_{n+1}
		9	2	1	0	p_n
		4	1	0	1	q_n

Since the next partial quotient is $a_2 = 2$, we multiply the numbers in its column by 2 and add this to the previous column, to obtain the next column in the table:

6	3	2	4	2		a_{n+1}
	20	9	2	1	0	p_n
	9	4	1	0	1	q_n

Continuing in this manner, we complete the table:

	6	3	2	4	2		a_{n+1}
434	69	20	9	2	1	0	p_n
195	31	9	4	1	0	1	q_n

(32.10)

It follows that

$$[2,4,2,3,6] \;=\; \frac{434}{195}.$$

❧

Theorem 32.1 and Corollary 32.2 have many implications regarding the behavior of convergents to a continued fraction.

To begin, the difference between successive convergents is

$$\frac{p_n}{q_n} - \frac{p_{n-1}}{q_{n-1}} = \frac{p_n q_{n-1} - p_{n-1} q_n}{q_n q_{n-1}} = \frac{(-1)^{n+1}}{q_n q_{n-1}}, \tag{32.11}$$

where the last identity (in which the numerator is simplified) follows from the determinant identity (32.8). Since the values of a_n, and therefore of q_n, are positive integers, we obtain the following:

Proposition 32.4. *For $n \geq 1$,*

$$\frac{p_n}{q_n} > \frac{p_{n-1}}{q_{n-1}} \quad \text{if } n \text{ is odd,}$$
$$\frac{p_n}{q_n} < \frac{p_{n-1}}{q_{n-1}} \quad \text{if } n \text{ is even.}$$

Skipping two steps instead, we also find that

$$\frac{p_n}{q_n} - \frac{p_{n-2}}{q_{n-2}} \;=\; \frac{p_n q_{n-2} - p_{n-2} q_n}{q_n q_{n-2}} \;=\; \frac{(-1)^n a_n}{q_n q_{n-2}},$$

where the last identity (in which the numerator is simplified) follows from the determinant identity (32.9). In this case, we have the following:

Proposition 32.5. *For $n \geq 1$,*

$$\frac{p_0}{q_0} < \frac{p_2}{q_2} < \cdots < \frac{p_{2n}}{q_{2n}},$$

and

$$\frac{p_1}{q_1} > \frac{p_3}{q_3} > \cdots > \frac{p_{2n+1}}{q_{2n+1}}.$$

In other words, the even numbered convergents form an increasing sequence, while the odd number convergents are decreasing.

Propositions 32.4 and 32.5 are combined as follows.

Proposition 32.6. *Every even numbered convergent is less than every odd numbered convergent. Moreover, for $n \geq 1$,*

$$\frac{p_0}{q_0} < \frac{p_2}{q_2} < \cdots < \frac{p_{2n}}{q_{2n}} < \frac{p_{2n+1}}{q_{2n+1}} < \cdots \frac{p_3}{q_3} < \frac{p_1}{q_1}. \tag{32.12}$$

To see how quickly the convergents are "converging", take absolute values in (32.11). The *distance* between successive convergents is then given by

$$\left| \frac{p_n}{q_n} - \frac{p_{n-1}}{q_{n-1}} \right| = \left| \frac{p_n q_{n-1} - p_{n-1} q_n}{q_n q_{n-1}} \right| = \frac{1}{q_n q_{n-1}}, \tag{32.13}$$

where the last identity (in which the numerator is simplified) follows from the determinant identity (32.8).

Since every $a_i \geq 1$ for $i > 1$, the recursion relation (32.7) implies that

$$q_{n+1} \geq q_n + q_{n-1},$$

for $n \geq 1$. Since every denominator $q_i \geq 1$, it follows that $q_{n+1} > q_n$. In fact, it follows that, for $n \geq 0$,

$$q_n \geq F_n,$$

where $F_0, F_1, F_2, \ldots$ denotes the Fibonacci sequence:

$$F_0 = 0, \quad F_1 = 1, \quad F_{n+1} = F_n + F_{n-1} \quad \text{for } n \geq 2.$$

It is a simple exercise in mathematical induction (see Exercise 2.7) to show that, for $n \geq 6$,

$$F_n \geq 2^{\frac{n}{2}}.$$

Combining these observations with (32.13) yields

$$\left| \frac{p_n}{q_n} - \frac{p_{n-1}}{q_{n-1}} \right| = \frac{1}{q_n q_{n-1}} < \frac{1}{q_{n-1}^2} \leq \frac{1}{F_{n-1}^2} \leq \frac{1}{2^{n-1}}, \tag{32.14}$$

for $n \geq 7$.

We see that convergence is exponentially fast, since the inequalities of (32.14) imply that each step in the continued fraction reduces the distance between successive convergents by at least $1/2$. And this would be the slowest case: typically some values of a_n exceed 1, so that q_n is substantially greater than the Fibonacci number F_n, causing the convergence to proceed even more quickly. The larger the value of a_n, the faster the regular continued fraction will converge to its final value.

❧

Exercise 32.1. Express the continued fraction $[1,2,3,4,5]$ as an ordinary fraction.

Exercise 32.2. **(a)** Express the number $\frac{229}{72}$ as a regular continued fraction.
(b) Use the results of part (a) to find $x, y \in \mathbb{Z}$ such that $229x + 72y = 1$.

Exercise 32.3. Which is bigger $[2,3,7]$ or $[2,5,7]$? Which is bigger $[2,3,7]$ or $[2,3,8]$? What happens in general?

Exercise 32.4. Express the continued fractions

$$[1],\ \ [1,1],\ \ [1,1,1],\ \ [1,1,1,1],\ \ [1,1,1,1,1],\ \ [1,1,1,1,1,1]$$

as ordinary fractions. Does the resulting pattern of numerators and denominators remind you of anything? Use Corollary 32.2 to explain what you have found.

Exercise 32.5. Suppose that a real number x satisfies $x = [1,3,x]$. Find the value of x.

Exercise 32.6. Let $m \in \mathbb{N}$. Consider the sequence

$$s_0 = 1, \quad s_1 = m, \quad s_k = 4s_{k-1} + s_{k-2} \text{ for } k \geq 2.$$

Prove that, for $k \geq 2$,

$$\frac{s_{k+1}}{s_k} = [\underbrace{4,\ldots,4}_{k \text{ copies}},m].$$

Exercise 32.7. How does the formula for $\frac{s_{k+1}}{s_k}$ in Exercise 32.6 change when $s_0 > 1$?

33 *Infinite continued fractions*

If x is irrational, then x cannot be expressed as a fraction with integer numerator and denominator. It follows that, for irrational x, the procedure for constructing a simple continued fraction for x continues forever, and we obtain an infinite continued fraction:

$$[a_0, a_1, a_2, \ldots] = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cdots}}.$$

But does this expression make any sense? Suppose we view it as the limit of the sequence:

$$[a_0],\ [a_0, a_1],\ [a_0, a_1, a_2],\ \ldots.$$

In other words, define

$$[a_0, a_1, a_2, \ldots] \ = \ \lim_{n\to\infty} [a_0, a_1, \ldots, a_n].$$

Does this limit exist, and if so, does the sequence converge to our original irrational number x? We will see later that the answers to both of these questions is yes.

❧

Consider the following example. Let

$$x \ = \ [1, 1, 1, 1, 1, \ldots] \ = \ 1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cdots}}}. \tag{33.1}$$

Notice that the tail end of the expansion (33.1) looks just like the whole expansion. Suppose that the limit of this infinite continued fraction exists, and denote the value of this limit by x. In this case, we have

$$x = 1 + \frac{1}{x}$$

so that

$$x^2 - x - 1 = 0$$

or

$$x = \frac{1 + \sqrt{5}}{2} = 1.618\ldots.$$

Indeed, if we examine the convergents to the expansion (33.1) we see that

$$\begin{aligned}
[1] &= 1 \\
[1,1] &= 1 + \frac{1}{1} = 2 \\
[1,1,1] &= 1 + \frac{1}{1+\frac{1}{1}} = \frac{3}{2} = 1.5 \\
[1,1,1,1] &= \frac{5}{3} = 1.666\ldots \\
[1,1,1,1,1] &= \frac{8}{5} = 1.6,
\end{aligned}$$

The appearance of the Fibonacci numbers in the numerators and denominators of this sequence is no coincidence. They appear as a consequence of the recursion relations given in Corollary 32.2.

Notice also that convergents seem to alternate sides of the limit $1.618\ldots$, first below, then above, and so on, closing in on the limit from each side. This oscillatory behavior is consistent with Propositions 32.4, 32.5, and 32.6.

❧

To go other way, let's try applying a "Euclid-style" algorithm to the irrational number $x = \sqrt{3}$. At each stage we extract the whole number part, leaving a remainder r such that $0 < r < 1$. The reciprocal will then satisfy $1 < \frac{1}{r}$, so that we can repeat ad infinitum. Since the number $\sqrt{3}$ is irrational, the procedure can never terminate, but we might nonetheless discern a pattern.

Begin by observing that $1 < \sqrt{3} < 2$. The first step then yields

$$\sqrt{3} = 1 + (\sqrt{3} - 1).$$

Here $a_0 = 1$ and $r_0 = \sqrt{3} - 1$, where $0 < r_0 < 1$. Taking the reciprocal of the remainder r_0 gives a number greater than 1, namely,

$$\frac{1}{r_0} = \frac{1}{\sqrt{3}-1} = \frac{1}{\sqrt{3}-1} \cdot \frac{\sqrt{3}+1}{\sqrt{3}+1} = \frac{\sqrt{3}+1}{\sqrt{3}^2 - 1^2} = \frac{\sqrt{3}+1}{2}.$$

Since $1 < \sqrt{3} < 2$, we have $1 < \frac{\sqrt{3}+1}{2} < \frac{3}{2}$, so that the next partial quotient is also 1, and

$$\frac{\sqrt{3}+1}{2} = 1 + \left(\frac{\sqrt{3}+1}{2} - 1 \right) = 1 + \frac{\sqrt{3}-1}{2}.$$

In other words, $a_1 = 1$ and $r_1 = \frac{\sqrt{3}-1}{2}$. Taking reciprocals again yields

$$\frac{1}{r_1} = \frac{2}{\sqrt{3}-1} = \frac{2(\sqrt{3}+1)}{\sqrt{3}^2 - 1^2} = \sqrt{3}+1 = 2 + (\sqrt{3}-1),$$

so that $a_2 = 2$ and $r_2 = \sqrt{3} - 1$. Since $r_2 = r_0$, the pattern of reciprocals and quotients will repeat, so that

$$a_3 = 1,\ a_4 = 2,\ a_5 = 1,\ a_6 = 2,\ \ldots,$$

suggesting that

$$\sqrt{3} \ = \ [1, 1, 2, 1, 2, 1, 2, \ldots],$$

if this limit exists. (This will be shown further ahead.)

❧

A continued fraction of the form $[b_1, \ldots, b_k, \overline{a_1, \ldots, a_n}]$ is said to be *periodic*. In this notation the pattern of partial quotients beneath the overline is repeated infinitely often.

A continued fraction of the form $[\overline{a_1, \ldots, a_n}]$ is said to be *purely periodic*.

❧

Suppose we are given that the limiting value $x = [1, \overline{1, 2}]$ exists. Let us determine a numerical value for x. We have

$$x = 1 + \frac{1}{s}$$

where $s = [\overline{1, 2}]$ is *purely periodic*. To compute s, we use the periodicity to observe that

$$s \ = \ 1 + \frac{1}{2 + \dfrac{1}{s}} \ = \ 1 + \frac{s}{2s+1} \ = \ \frac{3s+1}{2s+1}.$$

so that $2s^2 + s = 3s + 1$ and, therefore, $2s^2 - 2s - 1 = 0$. The quadratic formula yields

$$s \ = \ \frac{2 \pm \sqrt{12}}{4} \ = \ \frac{1 \pm \sqrt{3}}{2}.$$

Since we know that $s > 0$, it follows that $s = \frac{1+\sqrt{3}}{2}$, so that

$$x \ = \ 1 + \frac{1}{s} \ = \ 1 + \frac{2}{\sqrt{3}+1} \ = \ 1 + \frac{2(\sqrt{3}-1)}{\sqrt{3}^2 - 1^2} \ = \ \sqrt{3},$$

as expected.

❧

The preceding examples suggest that if a periodic infinite continued fraction with integer entries converges to a value x, then x must be a *quadratic surd*; that is

$$x = \frac{a + b\sqrt{n}}{c}, \tag{33.2}$$

where a, b, c, n are integers, and $\sqrt{n}$ is irrational. Indeed, if x is purely periodic; say,

$$x = [\overline{a_0, \ldots, a_m}],$$

where $a_i \in \mathbb{N}$, then

$$x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\cdots + \cfrac{1}{a_m + \cfrac{1}{x}}}}.$$

Unwinding the algebra will result in a quadratic equation with integer coefficients having a solution of the form (33.2). If x has the more general form

$$x = [b_0, \ldots, b_k, \overline{a_0, \ldots, a_m}],$$

where $a_i, b_i \in \mathbb{N}$, then a similar argument applies. It is less obvious that the converse also holds.

Theorem 33.1 (Lagrange)**.** *A real number x has a representation as a periodic continued fraction if and only if x is a quadratic surd.*

The proof of this theorem is a bit involved. See, for example, any of [13, 14, 27].

❧

Given a sequence of integers $a_0, a_1, a_2, \ldots$, where $a_0 \geq 0$ and $a_n \geq 1$ for $n \geq 1$, we defined the infinite regular continued fraction $[a_0, a_1, a_2, \ldots]$ by

$$[a_0, a_1, a_2, \ldots] = \lim_{n\to\infty} [a_0, a_1, a_2, \ldots, a_n]. \tag{33.3}$$

But how do we know this limit exists? It follows from Proposition 32.6 that, if $k, l \geq n \geq 7$, then

$$\left|\frac{p_k}{q_k} - \frac{p_l}{q_l}\right| < \left|\frac{p_n}{q_n} - \frac{p_{n-1}}{q_{n-1}}\right| \leq \frac{1}{2^{n-1}},$$

where the last inequality follows from (32.14). Since $\lim_{n\to\infty} \frac{1}{2^{n-1}} = 0$, it follows that the convergents to an infinite regular continued fraction $[a_0, a_1, a_2, \ldots]$ form a Cauchy sequence, and therefore *always* converge to some limit.

Let's make this argument even more precise. Proposition 32.5 implies that the even numbered convergents form an increasing sequence. But Proposition 32.6

asserts that this entire sequence of even numbered convergents $\frac{p_{2n}}{q_{2n}}$ is bounded above by any odd numbered convergent; that is,

$$\frac{p_0}{q_0} < \frac{p_2}{q_2} < \frac{p_4}{q_4} < \cdots < \frac{p_1}{q_1}.$$

A bounded increasing sequence of real numbers must converge to some limit, call it $L = \lim_{n\to\infty} \frac{p_{2n}}{q_{2n}}$. Similarly, the odd numbered convergents form a bounded *decreasing* sequence converging to some value M. Since every even convergent is smaller than every odd convergent, we have $L \leq M$. In fact, for all $n > 0$,

$$\frac{p_{2n}}{q_{2n}} \leq L \leq M \leq \frac{p_{2n+1}}{q_{2n+1}}. \tag{33.4}$$

It follows that

$$|L - M| \leq \left| \frac{p_n}{q_n} - \frac{p_{n-1}}{q_{n-1}} \right| \leq \frac{1}{2^{n-1}},$$

for all $n \geq 7$. Therefore $L = M$, and the limit (33.3) is precisely this value.

Once again, convergence is exponentially fast, since the inequalities of (32.14) imply that each successive convergent (eventually) reduces the distance to the limit by at least $1/2$. As noted earlier, larger values of a_n will lead to much faster convergence.

Notice from (33.4) that the limit $L = [a_0, a_1, a_2, \dots]$ always lies in between successive convergents. This leads to the following important observation.

Proposition 33.2. *Given a sequence of integers $a_0, a_1, a_2, \dots$, where $a_0 \geq 0$ and $a_n \geq 1$ for $n \geq 1$, let $L = [a_0, a_1, a_2, \dots]$. For $n \geq 0$,*

$$\left| \frac{p_n}{q_n} - L \right| \leq \frac{1}{q_n q_{n+1}}.$$

Proof. Suppose that n is even, where $n = 2m$. It follows from (33.4) that

$$\frac{p_{2m}}{q_{2m}} \leq L \leq \frac{p_{2m+1}}{q_{2m+1}}$$

so that

$$0 \leq L - \frac{p_{2m}}{q_{2m}} \leq \frac{p_{2m+1}}{q_{2m+1}} - \frac{p_{2m}}{q_{2m}}.$$

Therefore,

$$\left| L - \frac{p_{2m}}{q_{2m}} \right| \leq \left| \frac{p_{2m+1}}{q_{2m+1}} - \frac{p_{2m}}{q_{2m}} \right| = \left| \frac{p_{n+1}}{q_{n+1}} - \frac{p_n}{q_n} \right| = \frac{1}{q_n q_{n+1}},$$

by the identity (32.8) of Corollary 32.2.

If n is odd, set $n = 2m + 1$, compare to $2m + 2$, and proceed similarly. ❑

❧

Some additional references to the theory of continued fractions are listed in the bibliography. The book by Rosen [27] gives an elementary introduction. Hardy and Wright [13] provide a more terse but thorough treatment. Stark emphasizes a geometric interpretation in [32]. For an advanced treatment of continued fractions see [14] and [26]. Some remarkable computational methods are described in an unpublished (but widely available) paper by Gosper [12]. Continued fraction generators can be found on the web. See, for example, [37].

❧

Exercise 33.1. **(a)** Suppose that m is a positive integer. Set $\alpha = [m, m, m, m, \ldots]$, and solve a quadratic equation to find a formula for the value of α in terms of m.

(b) Use your answer to part (a) to find continued fraction expansions for $\sqrt{2}$ and $\sqrt{5}$.

Exercise 33.2. **(a)** Suppose that m, n are positive integers. Set $\alpha = [m, n, m, n, \ldots]$, and solve a quadratic equation to find a formula for the value of α in terms of m and n.

(b) Use your answer to part (a) to find the continued fraction expansion for $\sqrt{15}$.

Exercise 33.3. Use a table of the form (32.10) to compute the first 5 convergents $\frac{p_k}{q_k}$ for the infinite continued fraction $\alpha = [1, 2, 3, 4, \ldots]$.

Exercise 33.4. Find the infinite continued fraction expansion for $\sqrt{7}$.

Exercise 33.5. Evaluate the continued fraction $x = [1, 6, \overline{1, 8}]$.

Exercise 33.6. Let $\delta = \sqrt[3]{2}$. Using a computer or hand-held calculator, find the integer part of δ, subtract it off, and take the reciprocal of the remainder. Repeat this procedure to find the first 8 partial quotients a_i of δ, so that $\delta = [a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, \ldots]$.

Exercise 33.7. The simple continued fraction expansion for $e = 2.718281828459...$ is given by

$$e = [2, 1, 2, 1, 1, 4, 1, 1, 6, 1, 1, 8, \ldots], \tag{33.5}$$

continuing with the pattern $1, 1, 2n$ for all $n \geq 2$.*
(a) Use a table of the form (32.10) to compute the first 10 convergents $\frac{p_k}{q_k}$ for the infinite continued fraction of (33.5).
(b) How far must you go in part (a) to find a fraction that correctly approximates e to two decimal places? To four decimal places?

Exercise 33.8. The simple continued fraction expansion for π is given by

$$\pi = [3, 7, 15, 1, 292, 1, 1, 1, 2, \ldots], \tag{33.6}$$

continuing with no discernible pattern.
(a) Use a table of the form (32.10) to compute the first 5 convergents $\frac{p_k}{q_k}$ for the infinite continued fraction of (33.6).
(b) How far must you go in part (a) to find a fraction that correctly approximates π to two decimal places? To four decimal places?

*A short proof of (33.5) can be found in [7].

34 Recommended reading

Here are some suggestions for further reading and some useful references for more specialized topics.

- Tom Apostol, *Introduction to Analytic Number Theory,* Springer, New York, 1976.
- William Dunham, *Euler: The Master of Us All,* Mathematical Association of America, New York, 1999.
- Paul Garrett, *Making, Breaking Codes: An Introduction to Cryptology*, Prentice Hall, Upper Saddle River, NJ, 2001.
- G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers, 6th ed.,* Oxford University Press, New York, 2006.
- Neil Koblitz, *A Course in Number Theory and Cryptography, 2nd ed.,* Springer, New York, 1994.
- C. Stanley Ogilvy and John T. Anderson, *Excursions in Number Theory*, Dover Publications, New York, 1988.
- Kenneth Rosen, *Elementary Number Theory and Its Applications, 5th ed.,* Addison-Wesley, New York, 2005.
- Simon Singh, *The Code Book,* Doubleday, New York, 1999.
- Douglas R. Stinson, *Cryptography, Theory and Practice, 3rd ed.,* Chapman and Hall/CRC, Boca Raton, FL, 2006.

The book by Ogilvy and Anderson is easy reading and a lot of fun for beginners. Rosen's book is a more traditional course textbook. Hardy and Wright is a classic general reference.

Dunham's book about Euler is a layman's introduction to different aspects of Euler's work, devoting different chapters to different mathematical topics. Some of Euler's contributions to number theory are described, with an emphasis on insights to be found in Euler's original arguments.

The book by Koblitz emphasizes connections between number theory and modern cryptography, while assuming some familiarity with basic abstract algebra (groups, rings, and fields). Garrett's book has a similar focus, but offers a more elementary and self-contained presentation. Simon Singh's book provides a history of cryptography, written for the general public, and requiring minimal mathematical background. Stinson's book is more practical and offers

specific technical details and cryptographic algorithms, as well as material on cryptanalysis (codebreaking).

Apostol's book moves beyond classical number theory to the modern analytic theory. This book is an especially suitable text for a second course in number theory.

35 Answers, hints, and solutions to selected exercises

Exercise 2.13: Use the fact that $\mathbb{N}$ is closed under multiplication.

Exercise 2.17: No.

❧

Exercise 3.5: $\frac{127}{999}$.

Exercise 3.9: Use the fact that $x^6 = (x^2)^3$ and $y^6 = (y^2)^3$.

Exercise 3.13: How many ways can a teacher choose a team of k students from a class with n students? How many ways if we know Harold was chosen? How many ways if we know Harold was *not* chosen?

Exercise 3.17: Since n is even, write $n = 2m$ for some integer m.

If m is odd , then $7^n - 4^n = 7^{2m} - 4^{2m} = (7^m - 4^m)(7^m + 4^m)$. Use algebraic factoring to show that $7^m - 4^m$ is divisible by $7 - 4 = 3$, and that $7^m + 4^m$ is divisible by 11.

If m is even, factor out 2 from n as many times as possible, so that $n = 2^s t$ where t is odd, and then use a similar argument.

❧

Exercise 4.1: This follows from the fact that $a \cdot 0 = 0$.

Exercise 4.3: $1001 = 7 \cdot 11 \cdot 13$.
This is a handy fact for estimation. Since $1001 \approx 1000$, we have

$$\frac{1}{13} \approx \frac{77}{1000} = 0.077,$$

and, similarly, $\frac{1}{11} \approx 0.91$ and $\frac{1}{7} \approx 0.143$.

Exercise 4.4: $999 = 27 \cdot 37 = 3^3 \cdot 37$.
This is another handy fact for estimation. Since $999 \approx 1000$, we have

$$\frac{1}{37} \approx \frac{27}{1000} = 0.027,$$

and, similarly, $\frac{1}{27} \approx 0.037$.

Exercise 4.5: $3, 4, 0, 2, 0$.

Exercise 4.10: If $n = 3$ we have $3 = 1 + 2$. If $n > 3$ then

$$1 + 2 + 3 + \cdots + (n-1) > 1 + (n-1) = n,$$

so the identity never holds again for $n > 3$.

Exercise 4.14: If a is even, then $a = 2k$ for some integer k, so that $a^2 = 4k^2$ is divisible by 4. If a is not even, then $a = 2k + 1$ for some integer k, so that $a^2 = 4k^2 + 4k + 1$ is not divisible by 4; indeed, it has a remainder of 1 after division by 4.

Exercise 4.16: The answer is 1 both in part **(a)** and **(b)**. To see this, use the fact that $n = 3a + 2$ and $m = 3b + 2$ for some integers a and b.

Exercise 4.19: Suppose that $2^{\frac{1}{3}} = \frac{a}{b}$, where this fraction is expressed in lowest terms. This means that $a = 2^{\frac{1}{3}} b$, so that $a^3 = 2b^3$. What can you now say about the parity of a? Of b?

❧

Exercise 5.1: **(a)** 1; **(b)** 3; **(c)** 6; **(d)** 8; **(e)** 21; **(f)** 237.

Exercise 5.2: **(a)** 101; **(b)** 1111; **(c)** 10000; **(d)** 11000; **(e)** 1111000; **(f)** 1011011001.

Exercise 5.3: The octal answers are **(a)** 1 **(b)** 3; **(c)** 6; **(d)** 10; **(e)** 25; **(f)** 355.
The hexadecimal answers are **(a)** 1 **(b)** 3; **(c)** 6; **(d)** 8; **(e)** 15; **(f)** *ED*.

Exercise 5.8: The possible final digits are $\{0, 1, 4, 5, 6, 9\}$. To see this, observe that if n ends with 0, then $n = 10k$ for some k, so that $n^2 = 100k$ also ends with 0. If n ends with 1, then $n = 10k + 1$ for some k, so that $n^2 = 100k^2 + 10k + 1$ also ends with 1. If n ends with 2, then $n = 10k + 2$ for some k, so that $n^2 = 100k^2 + 40k + 4 = 10(10k^2 + 4k) + 4$ ends with 4. Continue by similar checking the cases in which n ends with 3, 4, and so on, up to 9.

Exercise 5.11: $\frac{37}{60}$.

Exercise 5.12: $\left(\frac{1}{3}\right)_{10} = \left(\frac{1}{11}\right)_2 = 0.01010101\ldots$ in base 2.

❧

Exercise 6.1: **(a)** 3 **(b)** 2; **(c)** 1; **(d)** 5; **(e)** 2^{43}; **(f)** $2^2 \cdot 3^2$.

Exercise 6.2: The answer to both questions is 6.

Exercise 6.3: **(a)** $\gcd(n, n+1) = \gcd(n, 1) = 1$.
(b) $\gcd(n, n+2) = \gcd(n, 2) = \begin{cases} 1 & \text{if } n \text{ is odd.} \\ 2 & \text{if } n \text{ is even.} \end{cases}$

Exercise 6.6: The answer is 23.

Exercise 6.7: It's the same as the $\gcd(84, 123)$. Now finish the problem.

Exercise 6.9: $6x + 3y$ is always divisible by 3. But 4 is not divisible by 3.

Exercise 6.13: **(a)** One solution is $x = 3$, $y = 4$.
(b) This one is similar to Exercise 6.9.
(c) Yes. Can you find some?

❧

Exercise 7.9: **(a)** 84 **(b)** $12 \cdot 29$; **(c)** $11 \cdot 12 \cdot 13$; **(d)** $2^3 \cdot 3^2 \cdot 5^6$; **(e)** n; **(f)** n if n is even, $2n$ if n is odd; **(g)** $n(n+1)$; **(h)** $n(n+2)$ if n is odd, $\frac{n(n+2)}{2}$ if n is even.

Exercise 7.11: This can be proved using Theorem 7.5 and Corollary 6.5. Alternatively, you can prove it more directly: Let $e = \operatorname{lcm}(a, b)$. Since $k|\operatorname{lcm}(ka, kb)$, we can write $\operatorname{lcm}(ka, kb) = ke'$ for some integer e'. Since $ka|ke'$, we have $a|e'$ and, similarly, $b|e'$. The minimality of the lcm implies that $e \leq e'$. Meanwhile, since $a|e$ and $b|e$, we have $ka|ke$ and $kb|ke$. Minimality again implies that $ke' \leq ke$ and $e' \leq e$ as well.

Exercise 7.14: The assertion is false.

Exercise 7.15: 1 or 3.

Exercise 7.16: 1, 2, 3, or 4.

Exercise 7.17: What happens when you factor the expression $m^3 - n^3$?

Exercise 7.20: $n = 28$.

❧

Exercise 8.5: $(x, y) \in \{(91, 2), (47, 67), (3, 132)\}$.

Exercise 8.6: **(a)** $x = 11 + 13t$ and $y = 12 - 4t$ where $t \in \mathbb{Z}$; **(b)** $(x, y) \in \{(11, 12), (24, 8), (37, 4)\}$; **(c)** Same as (b); **(d)** No integer solutions.

Exercise 8.10: 27.

Exercise 8.12: Since $\gcd(120, 102) = 6$, that $120a + 102b = 6$ has integer solutions a and b. Since $\gcd(6, 425) = 1$, that $6c + 425d = 1$ has integer solutions c and d. Use these facts to finish the proof.

Exercise 8.14: **(a)** Use the equation $6S + 10M + 15L = 425$ to solve this part; **(b)** 3; **(c)** 29.

Exercise 8.15: 154 candies (4 gum balls and 150 jelly beans).

Exercise 8.16: **(a)**

$$\left[\begin{array}{c|cc} 22 & 1 & 0 \\ 17 & 0 & 1 \end{array}\right] \to \left[\begin{array}{c|cc} 5 & 1 & -1 \\ 17 & 0 & 1 \end{array}\right]$$
$$\to \left[\begin{array}{c|cc} 5 & 1 & -1 \\ 2 & -3 & 4 \end{array}\right] \to \left[\begin{array}{c|cc} 1 & 7 & -9 \\ 2 & -3 & 4 \end{array}\right]$$
$$\to \left[\begin{array}{c|cc} 1 & 7 & -9 \\ 0 & -17 & 22 \end{array}\right]$$

so that $x = 7$ and $y = -9$ solve the equation.
(b)

$$\left[\begin{array}{c|cc} 576 & 1 & 0 \\ 84 & 0 & 1 \end{array}\right] \to \cdots \to \left[\begin{array}{c|cc} 12 & -1 & 7 \\ 0 & 7 & 6 \end{array}\right]$$

so that $x = -1$ and $y = 7$ solve the equation.

Exercise 8.17: **(a)**

$$\left[\begin{array}{c|cc} 25 & 1 & 0 \\ 9 & 0 & 1 \end{array}\right] \to \cdots$$
$$\to \left[\begin{array}{c|cc} 1 & 4 & -11 \\ 0 & -9 & 25 \end{array}\right] \to \left[\begin{array}{c|cc} 4 & 16 & -44 \\ 0 & -9 & 25 \end{array}\right]$$

so that $25 \cdot 16 + 9 \cdot (-44) = 4$.
(b) Row reduce the following matrix to the suggested final form:

$$\left[\begin{array}{c|ccc} 28 & 1 & 0 & 0 \\ 26 & 0 & 1 & 0 \\ 91 & 0 & 0 & 1 \end{array}\right] \to \cdots \to \left[\begin{array}{c|ccc} 1 & x & y & z \\ 0 & * & * & * \\ 0 & * & * & * \end{array}\right]$$

where the symbol $*$ denotes whatever various integers appear elsewhere in the matrix. The resulting values of x, y, and z should solve the equation $28x + 26y + 91z = 1$. Then multiply both sides by 17 to solve the original problem. There are ***many*** different solutions, but you can easily check if yours is correct.

Exercise 8.18: Row reduce the following matrix to the suggested final form:

$$\left[\begin{array}{cc|ccc} 3 & 2 & 1 & 0 & 0 \\ 4 & 1 & 0 & 1 & 0 \\ 7 & -8 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{cc|ccc} 5 & 1 & x & y & z \\ * & * & * & * & * \\ * & * & * & * & * \end{array}\right]$$

where the symbol $*$ denotes whatever various integers appear elsewhere in the matrix. The resulting values of x, y, and z should solve both equations simultaneously.

❧

Exercise 9.5: **(a)** Primes; **(b)** Integers of the form p^2, where p is prime; **(c)** Integers of the form p^3 or pq, where $p \neq q$ are prime.

Exercise 9.6: **(a)** 748 zeroes; **(b)** $k = 2993$.

Exercise 9.10: Either $a = 1$ and b is prime, or vice versa, or $a = b$ is prime.

Exercise 9.11: What happens if $n + 1$ is prime? What happens if $n + 1$ is composite? Watch out for special cases!

Exercise 9.12: Exercise 7.5 is helpful here.

Exercise 9.20: The assertion is false. Find a counterexample.

Exercise 9.26: What are the remainders after division by 3?

Exercise 9.34: **(a)** $n = \pm 2$; **(b)** $n \in \{0, 2, 4, 6\}$.

Exercise 9.35: Factor n^5+n^4+1 algebraically. To see how, let $n = 0,1,2,3,4$, and look for patterns in how the values of n^5+n^4+1 factor (as integers) in order to guess an algebraic factor. If you can guess one factor, use long division of polynomials to find the other.

❧

Exercise 10.1: **(a)** 9; **(b)** 89; **(c)** 0; **(d)** 3; **(e)** 0; **(f)** 6; **(g)** 0; **(h)** 1.

Exercise 10.4: It would be boring.

Exercise 10.6: **(a)** $ax \equiv ay \mod am$ iff $am|(ax-ay)$ iff $am = (ax-ay)k$ for some integer k. Since $a \neq 0$, this holds iff $m = (x-y)k$ iff $m|(x-y)$ iff $x \equiv y \mod am$.
(b) $12x \equiv 18 \mod 30$ iff $2x \equiv 3 \mod 5$ iff $x \equiv 4 \mod 5$. This means that x has the form $x = 4+5t \mod 30$, where $t \in \mathbb{Z}$, so that $x \in \{4,9,14,19,24,29\} \mod 30$.

Exercise 10.7: If n is even, then $n = 2k$ for some integer k. If n is odd, then $n = 2k+1$ for some integer k. What does this tell us (in each case) about $n^2 \mod 4$?

Exercise 10.9: Never.

Exercise 10.12: What happens mod 5? Does this equation have any solutions mod 5?

Exercise 10.15: What happens mod 4?

Exercise 10.20: $\{1,2,4,7,8,11,13,14\}$ are the units mod 15. The zero divisors are $\{0,3,5,6,9,10,12\}$.

Exercise 10.23: **(a)** $x \equiv 67 \mod 77$ **(b)** $x \equiv 34 \mod 63$ **(c)** No solutions. **(d)** $x \equiv 5 \mod 7$, so that $x \in \{5,12,19,26,33,40\} \mod 42$.

Exercise 10.27: No.

Exercise 10.28: What happens if $n = 2$? If $n > 2$, then what happens mod 8?

Exercise 10.30: When the modulus m is small, there are only a few cases to check in order to solve any polynomial equation.

Exercise 10.32: **(a)** If n is even, try algebraic factoring. If n is odd, consider the situation mod 3.
(b) Prove that $n \geq 4$, and use an argument mod 16 to show that $4|m$. Then show the equation is impossible mod 5.

❧

Exercise 11.3: $x \equiv 869 \mod 1435$

Exercise 11.4: -566

Exercise 11.5: 173

Exercise 11.9: $x \equiv 98 \mod 105$.

Exercise 11.10: Since $x^2 \equiv 3 \mod 11$, we have $x \equiv 5 \mod 11$ or $x \equiv 6 \mod 11$. Combine each of these cases with the condition $x \equiv 1 \mod 5$ to find exactly two distinct solutions mod 55.

Exercise 11.11: Each quadratic equation has two distinct solutions in its respective modulus. Apply the Chinese remainder theorem to each pairing to find 4 distinct solutions mod 65.

❧

Exercise 12.5: Since $37 \cdot 27 = 999$, it follows that $1000 \equiv 1 \mod 37$. Use this to show that any large number is congruent to the sum of its groupings into blocks of 3 decimal digits.

Exercise 12.10: Use the fact that $16 \equiv 1 \mod 15$.

❧

Exercise 13.2: **(a)** True; **(b)** False; **(c)** True; **(d)** True; **(e)** False; **(f)** False.

Exercise 13.5: 4

Exercise 13.6: 5

Exercise 13.8: 8

Exercise 13.9: 8

Exercise 13.12: 6

Exercise 13.13: 9

❧

Exercise 15.2: **39**

Exercise 15.3: **17**

Exercise 15.6: Wilson's theorem makes this problem much easier.

Exercise 15.11: Check the cases of $n \in \{0,1,2,3\}$. Then use Fermat's theorem to finish the proof.

Exercise 15.12: The case where n is prime follows immediately from Fermat's Theorem. If $n = p^k$, where p is prime, then a similar argument also works. If $n = pq$ where $p < q$ are prime, then $\gcd(p-1,q) = 1$, so that $(p-1)x + qy = 1$ for some integers x, y. Combine this observation with Fermat's Theorem to verify the case of $n = pq$, and then generalize this approach to solve the problem.

Exercise 15.14: Do Exercises 15.12 and 15.13 first.

❧

Exercise 16.4: Use the final expression in the identity of Proposition 16.5.

Exercise 16.6: $\phi(3) = \phi(4) = \phi(6) = 2$;
$\phi(5) = \phi(8) = \phi(10) = \phi(12) = 4$;
$\phi(7) = \phi(9) = \phi(14) = \phi(18) = 6$.

Exercise 16.7: What happens when n is odd? When n is even?

Exercise 16.11: The probability is 0.4; that is, a 40% chance.

Exercise 16.19: The geometric sum formula (3.2) is helpful.

Exercise 16.31: What is $\phi(100)$?

Exercise 16.32: Euler's theorem doesn't work for this problem. (Why not?)

Exercise 16.33: The answer is 01. (Be careful with your use of Euler's theorem in this exercise.)

Exercise 16.38: **(a)** 3; **(b)** 4; **(c)** 1; **(d)** 5; **(e)** 6; **(f)** 4.

❧

Exercise 17.1:
(a) `DOITTWICETOGETTHEORIGINALTEXTBACK`;
(b) `N`.

Exercise 17.2: **(a)** `DWWDFNDWGDZQ`; **(b)** `UFEUOVUFOUIY`.

Exercise 17.4: **(a)** `ITHINKTHEREFOREIAM`;
(b) $A(x) \equiv 25 - x \mod 26$.

Exercise 17.11: **(c)** It's a really bad idea, because this function $c(x)$ is not invertible. As a result, different plaintexts may encrypt to the same ciphertext, and there is no well-defined method for decryption.

Exercise 17.12: 85

Exercise 17.13: $d = 141$.

Exercise 17.14: $p = 3537197$ and $q = 4888861$ (or vice versa).

❧

Exercise 18.1: 1 mod 4; 1 mod 5; 1, 2, and 4 mod 7.

Exercise 18.2: 1 and 6 mod 7; 1, 5, 8, and 12 mod 13.

Exercise 18.3: Yes, there are a total of eight distinct solutions.

Exercise 18.5: Look for roots!

Exercise 18.8: **(b)** and **(d)** are irreducible. The others can be factored.

❧

Exercise 19.1: **(c)** There is no solution mod 2, and therefore no solution mod 250; **(d)** Use the Chinese remainder theorem to combine solutions mod 3 with your solutions from part (b).

Exercise 19.2: **(b)** There is no solution mod 27.

Exercise 19.3: **(a)** Notice that $x^4 + x + 23 \equiv x^4 + x$. Factor this.
(b) There are 4 distinct solutions mod 25.
What are they?
(c) There are 8 distinct solutions mod 575.
What are they?

Exercise 19.4: Combine the Chinese remainder theorem with the factorizations $999 = 27 \cdot 37$ and $1001 = 7 \cdot 11 \cdot 13$ to simplify this problem.

Exercise 19.6: 1, -1, $2^{n-1} + 1$, $2^{n-1} - 1$.

❧

Exercise 20.1: Since $\phi(23) = 22$, the only possible values for an order mod 23 are 1, 2, 11, and 22. Since 23 is prime, only -1 can have order 2, so every value $k \not\equiv \pm 1$ must have order 11 or 22. These are the only exponents that need to be checked. It turns out that $\mathrm{ord}_{23}(2) = \mathrm{ord}_{23}(3) = 11$, and $\mathrm{ord}_{23}(5) = 22$.

Exercise 20.2: Note that $13 \equiv 3 \mod 10$.

Exercise 20.3: **(a)** 3 and 7 are primitive roots mod 10.
(b) No.

Exercise 20.6: $\mathrm{ord}_8(1) = 1$ and $\mathrm{ord}_8(3) = \mathrm{ord}_8(5) = \mathrm{ord}_8(7) = 2$. There are no primitive roots mod 8.

Exercise 20.11: **(a)** Let $u = r^{\frac{p-1}{2}}$. Show that $u^2 \equiv 1$, but that $u \neq 1$. What other possibilities remain?
(b) *An idea:* Let $\beta = \mathrm{ord}_p(-r)$. What happens if β is even? What if β is odd?
(b) *A different idea:* Use part (a) to show that $-r \equiv r^{\frac{p+1}{2}}$. Combine this with Proposition 20.3 and that fact that $p \equiv 1 \mod 4$.
(c) No! (Why not?)

Exercise 20.14: What is $\mathrm{ord}_n(b^2)$? Is Proposition 20.5 helpful?

Exercise 20.16: Write $\phi(m) = \alpha s$. Show that $\mathrm{ord}_m(r^s) = \alpha$. Then show that if $\mathrm{ord}_m(r^t) = \alpha$, then $s|t$. Now apply Proposition 20.3 to $u = r^s$.

Exercise 20.17:
(a) Since $q|(M^{p-1} + M^{p-2} + \cdots + M + 1)$ we have $M^{p-1} + M^{p-2} + \cdots + M + 1 \equiv 0 \mod q$. What happens if $q = p$?
(b) Same hint as part (a). What happens if $q = q_i$ for some i?
(c) Recall that
$M^p - 1 = (M-1)(M^{p-1} + M^{p-2} + \cdots + M + 1)$.
(d) By part (c) and Proposition 20.1, we have $\alpha|p$. What happens if $\alpha = 1$?
(e) Combine part (d), Fermat's theorem, and Proposition 20.1.

❧

Exercise 21.3: 5, 7, 9, 10, 14, 18.

Exercise 21.4: 2, 3, 4, 6.

Exercise 21.5–21.9: These exercises are most easily understood if done together, in the order presented.

❧

Exercise 22.1: (a) $\{1\}$; **(b)** $\{1,4\}$; **(c)** $\{1\}$; **(d)** $\{1,4\}$; **(e)** $\{1,2,4,8,9,13,15,16\}$; **(f)** $\{1,4,6,9,11,14,16,19,21,24\}$.

Exercise 22.2: (a) -1; **(b)** -1; **(c)** -1; **(d)** -1; **(e)** 1; **(f)** 1; **(g)** -1; **(h)** 1; **(i)** 1.

Exercise 22.7: It helps to recall that $\phi(m)$ is even.

Exercise 22.6: (a) Yes. **(b)** No.

Exercise 22.13: Implication in one direction is still true, but the other direction may fail to hold.

Exercise 22.14: Since each modulus is in this exercise is odd, the quadratic formula applies (why?). Therefore, each polynomial in this exercise has a root in its given modulus iff its discriminant has a square root in that modulus.

Exercise 22.19: (a) True. Prove it! **(b)** False. Find a counterexample.

Exercise 22.20: Yes, you can find one. Keep looking!

❧

Exercise 23.3: If $p \equiv 1 \bmod 12$, what is $p \bmod 4$? Write $p = 12k+1$, and apply the law of quadratic reciprocity to the Legendre symbol $\left(\frac{3}{p}\right)$.

Exercise 23.4: (a) $\left(\frac{5}{p}\right) = 1$ iff $p \equiv \pm 1 \bmod 5$.

Exercise 23.5: Find a condition mod 28.

Exercise 23.6: (a), (d), (i) -1; the rest are 1.

Exercise 23.7: (a) 23; **(b)** 71; **(c)** 7.

Exercise 23.8: First show that $3m \equiv -1 \bmod p$.

Exercise 23.11: -1.

❧

Exercise 25.1: 6 is a quadratic residue mod 25 and mod 95. It is a non-residue mod 35. It is a perfect square mod 75, but not a quadratic residue, since 6 is not a unit mod 75.

Exercise 25.2: All yes except mod 65.

Exercise 25.5: Every odd prime p dividing m must satisfy $p \equiv 1 \bmod 4$. If m is even, then we must have $m \equiv 2 \bmod 4$ as well.

Exercise 25.6: Every prime p dividing m must satisfy $p \equiv \pm 1 \bmod 8$.

Exercise 25.13: $\{0,1\} \bmod 3$; $\{0,1,4,7\} \bmod 9$; $\{0,1,4,6,9,10\} \bmod 15$; and $\{0,1,4,7,9,10,13,16,19,22,25\} \bmod 27$.

❧

Exercise 26.1: (a) 0; **(b)** -1; **(c)** 1; **(d)** 1; **(e)** -1; **(f)** -1; **(g)** 0; **(h)** 1; **(i)** -1.

Exercise 26.2: (a) 1; **(b)** 1; **(c)** 0; **(d)** -1; **(e)** 1; **(f)** -1.

Exercise 26.6: Quadratic reciprocity allows us to invert this Jacobi symbol. Simplifying in the smaller modulus then yields

$$\left(\frac{17292864462617}{17292864462677}\right) = \left(\frac{17292864462677}{17292864462617}\right) = \left(\frac{60}{17292864462617}\right)$$

Now finish the computation. (The final answer is 1.)

Exercise 26.9: (a) Show that, if $a^2 b$ is a quadratic residue in every modulus, then so is b. Then apply the minimality assumption.
(b) The Chinese remainder theorem is helpful at this step.
(c) Similar to (b).
(d) See part (vi) of Proposition 26.1.

❧

Exercise 27.2: (a) $(t^2-a)\alpha$; **(b)** $(t^2-a)^{-1}\alpha$; **(c)** $-a-\alpha$; **(d)** $2t^2 - a + 2t\alpha$.

Exercise 27.4: (a) n must be even. **(b)** If n is even then $\alpha^n \equiv (t^2-a)^{\frac{n}{2}}$. If n is odd then $\alpha^n \equiv (t^2-a)^{\frac{n-1}{2}}\alpha$.

Exercise 27.8: Suppose that $x = r + s\alpha$ is a solution, where $r, s \in \mathbb{Z}_p$. Derive a contradiction. Keep in mind that $p \equiv 1 \bmod 4$.

Exercises 27.11 and 27.12: $\pm 11 \bmod 37$.

Exercise 27.13: (a) $\pm 41 \bmod 151$; **(b)** $\pm 12 \bmod 71$; **(c)** $\pm 29 \bmod 103$; **(d)** $\pm 39 \bmod 101$; **(e)** $\pm 26 \bmod 61$; **(f)** $\pm 16 \bmod 127$.

❧

Exercise 28.3: To prove the assertion without appealing to Euclid's formula, consider what happens mod 8.

Exercise 28.4: What happens mod 4?

Exercise 28.7: Use Euclid's formula (28.2). What happens when $Y = X + 1$?

Exercise 28.9: Use Theorem 28.3. **(b)**, **(d)**, **(e)**, and **(h)** are sums of two integer squares; the rest are not.

Exercise 28.10: Use Lemma 28.4 and the fact that $2 = 1^2 + 1^2$.

Exercise 28.14: Use Lemma 28.4 in two different ways.

❧

Exercise 29.4: Do Exercise 29.3 before this one.

Exercise 29.5: (a) $x_5 = 2$; **(b)** 25.

Exercise 29.6: (a) 8; **(b)** 01010101... (repeating); **(c)** 11010010... (repeating).

Exercise 29.7: (a) 4; **(b)** 01010101... (repeating); **(c)** 1100... (repeating).

Exercise 29.8: (a) 18; **(b)** 101000100101111101... (repeating); **(c)** 111111100011101110... (repeating).

Exercise 29.9: Prove that $x_{n+1} \equiv ax_n + b \bmod 4$, and use this observation to prove parts (a) and (b).

Exercise 29.10: Prove that $x_{n+1} \equiv ax_n + b \bmod p$, and use this observation to complete the proof.

Exercise 29.12:
(a) Prove that $x_n \equiv x_0 \bmod m$ for all $n \geq 0$, resulting in a constant sequence.
(b) Prove that, for some prime $p|m$, we have $x_n \equiv x_0 \bmod p$ for all $n \geq 0$, resulting in a sequence that is constant mod p (and therefore of relatively short period mod m).

❧

Exercise 30.4: Since $2^{9017} \equiv 1399 \bmod 9017$, the integer 9017 cannot be prime.
In case you're curious, $9017 = 71 \cdot 127$, but the point of this exercise is to show that 9017 is composite *without* knowing or finding the actual factorization.

Exercise 30.6: Note that $1105 = 5 \cdot 13 \cdot 17$ and $2465 = 5 \cdot 17 \cdot 29$.

Exercise 30.7: Use Theorem 30.1. Note that the condition $(p-1)|(n-1)$ is equivalent to checking that $n \equiv 1 \bmod p-1$.

❧

Exercise 31.1: Only ± 1 (that is, 1 or 14 mod 15) give false positives. Assuming the algorithm chooses a uniformly random test value $a \in \{1, 2, \ldots, 14\}$, the probability of error in this case is therefore $1/7$ or about 14.3%.
This problem can be solved by brute force (checking all 14 cases), but it can also be reasoned out: Since $\phi(15) = 8$, we have $a^8 \equiv 1 \bmod 15$ for all units. Meanwhile, $(15-1)/2 = 7$, so that $a^7 \equiv \pm 1$ only when $1 \equiv a^8 \equiv \pm a$. Equivalently, this means that $a \equiv \pm 1$ are the only units mod 15 can give values that agree with the Jacobi symbol.

Exercise 31.2: Only ± 1 (that is, 1 or 14 mod 15) give false positives. Assuming the algorithm chooses a uniformly random test value $a \in \{1, 2, \ldots, 14\}$, the probability of error in this case is therefore $1/7$ or about 14.3%.

Exercise 31.3: The integer 90109 is composite. The other integers in the list are prime.

Exercises 31.4 and 31.5: The integers 4115181073 and 1111111111111111111 are prime. The other integers in the list are composite.

Exercises 31.6: The smallest prime integer p greater than 10^{20} is 100000000000000000039.

❧

Exercise 32.1 : $\dfrac{225}{157}$.

Exercise 32.2: $[3, 5, 1, 1, 6]$.

Exercise 32.5: $x = \dfrac{3+\sqrt{21}}{6}$.

❧

Exercise 33.2: (a) We are given

$$\alpha = m + \cfrac{1}{n + \cfrac{1}{\alpha}}$$

so that

$$n\alpha^2 - mn\alpha - m = 0$$

and, since $\alpha > 0$,

$$\alpha = \frac{mn + \sqrt{m^2n^2 + 4mn}}{2n}$$

(b) If $n = 1$, and if we make m even, say $m = 2k$ (in order to clear the denominator), we have

$$\alpha = \frac{2k + \sqrt{4k^2 + 8k)}}{2} = k + \sqrt{k^2 + 2k}$$

If $k = 3$ (that is, $m = 6$) then $k^2 + 2k = 15$, so that

$$3 + \sqrt{15} = [6, 1, 6, 1, 6, 1, 6, \ldots]$$

and

$$\sqrt{15} = [3, 1, 6, 1, 6, 1, 6, \ldots].$$

Exercise 33.4: $\sqrt{7} = [2, \overline{1, 1, 1, 4}]$

Exercise 33.5: First evaluate the purely periodic continued fraction $[\overline{1, 8}]$. Then substitute this value appropriately into the expression $[1, 6, \overline{1, 8}]$ to determine the value of x.

Exercise 33.6: Your answer should be:

$$\sqrt[3]{2} = [1, 3, 1, 5, 1, 1, 4, 1, \ldots].$$

The expansion continues as:

$$[1, 3, 1, 5, 1, 1, 4, 1, 1, 8, 1, 14, 1, 10, 2, 1, 4, 12, 2, 3, \ldots],$$

with no discernible pattern.

❧ ❧ ❧ ❧

References

[1] W. R. Alford, A. Granville, and C. Pomerance, There are infinitely many Carmichael numbers, *Ann. Math.*, **139** (1994), 703-722.

[2] N. Ankeny, Sums of three squares, *Proc. Amer. Math. Soc.*, **8** (1957), 316-319.

[3] T. Apostol, *Introduction to Analytic Number Theory,* Springer, New York, 1976.

[4] M. Artin, *Algebra* (2nd edition), Pearson, New York, 2010.

[5] L. Blum, M. Blum, and M. Shub, A simple unpredictable pseudo-random number generator, *SIAM J. Comput.*, **15** (1986), 364-383.

[6] H. Cohen, *A Course in Computational Algebraic Number Theory*, Springer, New York, 1996.

[7] H. Cohn, A short proof of the simple continued fraction expansion of e, *Amer. Math. Monthly*, **113** (2006), 57-62.

[8] H. Davenport, *The Higher Arithmetic* (8th edition), Cambridge University Press, New York, 2008.

[9] W. Dunham, *Euler: The Master of Us All*, MAA, New York, 1999.

[10] P. Garrett, *Making, Breaking Codes: An Introduction to Cryptology*, Prentice Hall, Upper Saddle River, NJ, 2001.

[11] C. F. Gauss, *Disquisitiones Arithmeticae* (translated by W. C. Waterhouse, A. A. Clarke, et al.), Springer, New York, 1986.

[12] W. Gosper, *Continued fraction arithmetic*, `http://www.tweedledum.com/rwg/cfup.htm`

[13] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers* (6th edition), Oxford University Press, New York, 2006.

[14] A. Khinchin, *Continued Fractions,* Dover, New York, 1979.

[15] D. Knuth, Mathematics and computer science: Coping with finiteness, *Science*, **194** (1976), 1235-1242.

[16] D. Knuth, *The Art of Computer Programming, Volume II* (3rd edition), Addison-Wesley, New York, 1998.

[17] F. Lemmermeyer, *Reciprocity Laws: From Euler to Eisenstein*, Springer, New York, 2000.

[18] P. Mihăilescu, Primary cyclotomic units and a proof of Catalan's conjecture, *J. Reine Angew. Math.*, **572** (2004), 167-195.

[19] G. Miller, Riemann's hypothesis and tests for primality, *Journal of Computer and System Sciences*, **13** (1976), 300-317.

[20] I. Niven, H. Zuckerman, and H. Montgomery, *An Introduction to the Theory of Numbers* (5th edition), John Wiley & Sons, New York, 1991.

[21] C. S. Ogilvy and J. T. Anderson, *Excursions in Number Theory*, Dover, New York, 1988.

[22] J. M. Pollard, A Monte Carlo method for factorization, *Nordisk Tidskr. Informationsbehandling (BIT)*, **15** (1975), 331-334.

[23] M. Rabin, Probabilistic algorithm for testing primality, *J. Number Theory*, **12** (1980), 128-138.

[24] E. Rapaport (translator), *Hungarian Problem Book I*, MAA, New York, 1963.

[25] E. Rapaport (translator), *Hungarian Problem Book II*, MAA, New York, 1963.

[26] A. Rockett and P. Szűsz, *Continued Fractions*, World Scientific, River Edge, NJ, 1992.

[27] K. Rosen, *Elementary Number Theory and its Applications* (6th edition), Addison Wesley, Boston, 2010.

[28] S. Ross, *A First Course in Probability Theory* (8th edition), Prentice Hall, Upper Saddle River, NJ, 2010.

[29] P. Shor, Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer, *SIAM Review*, **41** (1999), 303Ű332.

[30] V. Shoup, *A Computational Introduction to Number Theory and Algebra* (2nd edition), Cambridge University Press, New York, 2009.

[31] R. Solovay and V. Strassen, A fast Monte-Carlo test for primality, *SIAM J. Comput.*, **6** (1977), 84-85.

[32] H. Stark, *An Introduction to Number Theory*, MIT Press, Cambridge, 1987.

[33] J. Stillwell, *Mathematics and Its History* (3rd edition), Springer, New York, 2010.

[34] D. Stinson, *Cryptography Theory and Practice* (3rd edition), Chapman and Hall/CRC Press, Boca Raton, FL, 2006.

[35] G. Suetonius Tranquillus, *The Twelve Caesars*, (translated by R. Graves and J. Rives), Penguin, New York, 2007.

[36] A. Wiles, Modular elliptic curves and Fermat's Last Theorem, *Ann. Math.*, **142** (1995), 443-551.

[37] `http://wims.unice.fr/wims/wims.cgi?module=tool/number/contfrac.en`

Index