Preface

At the University of Waterloo we have been teaching a course in elementary number theory for decades. The idea of offering such a course online gained momentum with the creation of our Master of Mathematics for Teachers. We have adapted our on-campus course to become MATH 640, offered online to mathematical educators and like-minded individuals. In order for such a course to be effective it became clear that a set of coherent notes would be necessary. Thus I wrote A Taste of Number Theory. This online book comprises the contents of MATH 640.

There are many good books on number theory. I have tried to write Taste in a way that is friendly to read by professionals with an already busy schedule, and to keep the mathematics as transparent as possible. Keeping in mind that not everyone comes to the MMT program with the same background knowledge, I have written Taste in a leisurely and informal style, included numerous worked examples, and often repeated the main points.

At one point or another, everyone hits the wall when it comes to any meaningful endeavour, be it running a marathon, mastering the piano, or learning mathematics. Some hit the mathematics wall in high school, others in more advanced studies, and some in their lifetime of research. The latter individuals are called mathematicians. It is my hope that this book will help you push your particular wall a bit farther back.

I should caution that the intricacies of number theory cannot be revealed without doing proofs, and Taste abounds with them. It is in the nature of the subject that explaining how something is so is the same as proving it. Most proofs are quite approachable. By mastering them, you will develop a sense of what the subject is about. Some of the proofs, most notably the primitive root and quadratic reciprocity theorems, demand patience. I have included them here to illustrate that the wall can be real, and in keeping with the attitude that a statement without proof is not really worth much.
At the end of each chapter some exercises have been included. Some are easy, some less so. There is a great overlap between these exercises and the work to be submitted in MATH 640. Should you wish to discuss how to approach any of the exercises or to discuss any topic in Taste please do not hesitate to contact me.

This is the fifth time in which Taste is being used for MATH 640. I have proofread it repeatedly, and in the previous rounds of MATH 640 many students were kind enough to point out a number of corrections. Should any of my dear readers find additional errors or have suggestions for improvements, I would be grateful to hear from you so that I can make the book better.

I hope that my book helps to develop your own taste for number theory.

Frank Zorzitto
University of Waterloo
October 6, 2016
Contents

Preface iii

1 The greatest common divisor 1
  1.1 Divisibility ................................................. 2
  1.2 Greatest common divisors ............................. 7
  1.3 The Euclidean Algorithm ............................... 10
  1.4 Coprimeness ................................................. 20
  1.5 The equation $ax + by = c$ ............................. 21
  1.6 Exercises ................................................ 25

2 Primes 29
  2.1 Factoring into primes ................................. 29
  2.2 Unique factorization .................................... 31
  2.3 A strange ring ........................................... 35
  2.4 Exercises ................................................ 39

3 Modular arithmetic 43
  3.1 Congruences .............................................. 43
  3.2 The replacement principle ............................ 46
  3.3 Residues .................................................. 49
  3.4 Linear congruences ..................................... 53
  3.5 The group of units $\mathbb{Z}_n^*$ ........................ 59
  3.6 The Euler-Fermat theorems ........................... 62
CONTENTS

3.7 The square and multiply algorithm ........................................ 67
3.8 Using Fermat to test for non-primality ................................. 72
3.9 Using Euler to find $k$’th roots modulo $n$ ......................... 75
3.10 Exercises .............................................................. 79

4 The Chinese Remainder Theorem ................................. 85
4.1 Statement, proof and uses of CRT ..................................... 86
4.2 CRT from the point of view of residues .............................. 90
4.3 Calculating the Euler function ........................................... 91
4.4 An application to communications security ....................... 95
4.5 Exercises .............................................................. 102

5 Primitive roots ............................................................ 107
5.1 Primitive roots in $\mathbb{Z}_n$ ............................................... 107
5.2 The order of a residue .................................................. 110
5.3 Polynomials over $\mathbb{Z}_p$ ............................................... 114
5.4 The order of a product and the main theorem .................... 118
5.5 Carmichael numbers ..................................................... 121
5.6 Primitive roots and cryptography .................................... 125
5.7 Exercises .............................................................. 131

6 Quadratic Residues ...................................................... 135
6.1 Quadratic congruences .................................................. 136
6.2 Quadratic residues and primitive roots ............................ 138
6.3 Quadratic reciprocity .................................................... 148
6.4 Quadratic reciprocity-the proof ..................................... 152
6.5 Exercises .............................................................. 164

7 Sums of squares and other powers .................................... 167
7.1 Pythagorean triples ..................................................... 167
7.2 Sums of squares in general ............................................. 174
7.3 Some remarks on Fermat’s Last Theorem ......................... 188
CONTENTS

7.4 A peek at an elliptic curve .............................................. 191
7.5 Exercises ................................................................. 197

8 Multiplicative functions .......................... 199
8.1 Examples and properties ................................. 199
8.2 The $\sigma$-function and perfect numbers .............. 207
8.3 Möbius inversion ...................................................... 211
8.4 Exercises ................................................................. 218
Chapter 1

The greatest common divisor

Ramanujan was enthralled by numbers. On his deathbed as a young man he received a visit from the great Oxford mathematician G.H. Hardy. In making small talk, Hardy remarked that taxi #1729 had brought him to see Ramanujan. Upon hearing that, the feeble young genius perked up and observed that 1729 is a very interesting number, because

“it is the smallest number that can be expressed as the sum of two perfect cubes in two different ways.”

So you see, we can find beauty in unexpected places. Rightly so, it might be in a work of art, a natural setting, a piece of music, or another individual, but, nerdy as it may seem, Ramanujan found it in numbers. To find beauty in such an abstract place, we have to look hard and with patience. I hope that, with a bit of effort on your part, this course will convey a taste of their magic.

Hardy held applied mathematics with a certain degree of scorn. He surely now would be turning over in his grave, knowing that his beloved theory of numbers has risen to the forefront with crucial applications in communications security.

The object of our affection will be the set of integers:

\[
\mathbb{Z} = \{0, \pm 1, \pm 2, \pm 3, \ldots \}.
\]

The fancy letter “zed” comes from the German word for number: “Zahl”. While anybody can scoff that “two and two make four”, most would be hard pressed to explain the following:
CHAPTER 1. THE GREATEST COMMON DIVISOR

- how the greatest common divisor affects the solution of $ax + by = c$ using only integers,
- the pattern for solving $x^2 + y^2 = z^2$ with integers,
- how it is that $\frac{a^p - a}{p}$ is always an integer when $a$ in an integer and $p$ is prime,
- which integers $n$ take the form $n = x^2 + y^2$,
- the proof that $x^4 + y^4 = z^4$ has no integer solutions,
- which primes $p$ and which integers $b$ will permit $x^2 = b + yp$ to have integer solutions $x, y$,
- how it is that infinitely many primes take the form $4n + 1$,
- how to count the number of integers less than $n$ having no proper common factors with $n$,
- how the decimal expansion of $\pi$ is obtained,
- how many consecutive integers are possible without having a prime among them,
- how to estimate the number of primes up to a given $n$,
- whether there are infinitely many prime pairs separated by 2,
- whether every even integer is the sum of two primes,

Some of these matters are not so hard, some are very hard, and some remain unsolved. To approach the answers, we have to proceed with care. This means that we have to speak in the language of precise definitions, theorems, and proofs! This can, at times, require patience and dedication.

1.1 Divisibility

Since $\mathbb{Z}$ contains a “0” and a “1” and since the sum, difference and product of two integers is an integer, $\mathbb{Z}$ is called a ring. There are other rings as well. In particular, the following rings are important.
1.1. DIVISIBILITY

- \( \mathbb{Q} = \{a/b : a, b \in \mathbb{Z} \text{ and } b \neq 0\} \) is the ring of \textit{rational} numbers, or fractions. They are called rational because they comprise the ratio of two integers.

- \( \mathbb{R} = \) the ring of \textit{real} numbers, typically represented as the \( x \)-axis that is fundamental to calculus.

- \( \mathbb{C} = \{a + ib : a, b \in \mathbb{R} \text{ and } i = \sqrt{-1}\} \), the ring of \textit{complex} numbers.

We can see the obvious inclusions: \( \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C} \).

All of these inclusions are proper. Indeed, \( 1/2 \) is in \( \mathbb{Q} \), but not in \( \mathbb{Z} \). Also \( i = \sqrt{-1} \) is in \( \mathbb{C} \), but not in \( \mathbb{R} \). It’s a little more challenging to come up with a number that is in \( \mathbb{R} \) but not in \( \mathbb{Q} \). With some advanced calculus one can show that \( \pi \) is not rational, although it is moderately close to \( 22/7 \). Probably the simplest irrational number is \( \sqrt{2} \).

Let’s warm up with a proof that \( \sqrt{2} \notin \mathbb{Q} \). In search of a contradiction, suppose \( \sqrt{2} = a/b \) where \( a, b \) are integers. By cancelling common factors we can arrange it so that one of \( a \) or \( b \) is odd. From \( a/b = \sqrt{2} \) we get

\[ a^2 = 2b^2. \]

So \( a^2 \), and thereby \( a \), is even. Say \( a = 2c \) where \( c \) is an integer too. Put this into \( a^2 = 2b^2 \) and get

\[ 4c^2 = 2b^2 \text{ and then } 2c^2 = b^2. \]

Now \( b \) is seen to be even, as well. But at least one of \( a \) or \( b \) was odd! The only escape from this contradiction is that \( \sqrt{2} \) is irrational.

The rings \( \mathbb{Q}, \mathbb{R}, \mathbb{C} \) have the additional property that we can carry out division within them. That is, for any \( a, b \) in \( \mathbb{Q}, \mathbb{R} \) or \( \mathbb{C} \), the number \( a/b \) remains in \( \mathbb{Q}, \mathbb{R} \) or \( \mathbb{C} \), respectively, as long as \( b \neq 0 \). A ring with this property is called a \textit{field}.

Now, \( \mathbb{Z} \) is not a field, which makes the business of division inside \( \mathbb{Z} \) \textit{very interesting}.

**Definition 1.1.** If \( a, b \) are integers, we say that \( a \) \textit{divides} \( b \), or that \( a \) is a \textit{factor} of \( b \), when \( b = ak \) for some integer \( k \). We also say at times that \( a \) is a \textit{divisor} of \( b \). When this happens, we write \( a \mid b \), and when this does not happen, we write \( a \nmid b \).

For example, \( -3 \mid 12 \), but \( 6 \nmid 9 \). Every \( a \mid 0 \), since \( 0 = a \cdot 0 \). But \( 0 \nmid a \), when \( a \neq 0 \). For otherwise, we would have some \( k \) such that \( 0 \neq a = 0 \cdot k = 0 \).
The integers \( \pm 1 \) divide every integer \( b \). Indeed, \( b = 1 \cdot b \) and \( b = (-1)(-b) \). Objects with such a property are often called *units*.

Here are the properties of divisibility to be used automatically and often. The proofs are not hard, but they do illustrate the importance of explaining things in terms of basic definitions.

**Proposition 1.2.** Let \( a, b, c, x, y \in \mathbb{Z} \).

1. If \( a \mid b \) and \( b \mid c \), then \( a \mid c \).
2. If \( a \mid b \) and \( a \mid c \), then \( a \mid bx + cy \).
3. If \( a \mid b \) and \( b \neq 0 \), then \( |a| \leq |b| \).
4. If \( a \mid b \) and \( b \mid a \), then \( a = \pm b \).
5. If \( a \mid b \), then \( \pm a \mid \pm b \).

**Proof.**

1. We have \( b = ak \) and \( c = b\ell \) for some integers \( k, \ell \). Then \( c = (ak)\ell = a(k\ell) \), and so \( a \mid c \), since \( k\ell \) is an integer.

2. We have \( b = ak, c = a\ell \) for some integers \( k, \ell \). Then

\[
bx + cy = akx + a\ell y = a(kx + \ell y)
\]

And so \( a \mid bx + cy \), since \( kx + \ell y \) is an integer.

3. We have \( b = ak \) for some integer \( k \). Take absolute values to get \( |b| = |a||k| \). Since \( b \neq 0 \), we get \( |k| > 0 \) and, being an integer, \( |k| \geq 1 \). So \( |a| \leq |b| \).

4. We have \( b = ak \) and \( a = b\ell \) for some integers \( k, \ell \). So \( b = (b\ell)k = b(\ell k) \). If \( b = 0 \), then \( a = 0 \) too, whereby \( a = \pm b \). If \( b \neq 0 \), cancel \( b \) to get \( 1 = \ell k \). Thus \( \ell = \pm 1 \), and then \( a = \pm b \).

5. We have \( b = ak \) for some integer \( k \). Then \( -b = a(-k) \) and so \( a \mid (-b) \). Also \( b = (-a)(-k) \) and so \( -a \mid b \ldots \) and so on for the other two cases.
1.1. DIVISIBILITY

Item 5 in Proposition 1.2 reveals that, as far as divisibility is concerned, we only need to deal with positive integers. The set of positive integers \( \{1, 2, 3, \ldots \} \) is denoted by \( \mathbb{N} \). Sometimes the positive integers are called natural numbers, but what makes them more natural than other numbers, such as 0 or \( \pi \), might be a matter of opinion.

We will use Proposition 1.2 repeatedly, without fuss, and often without even being conscious of it.

Quotients and remainders

When divisibility fails, we do the next best thing and look for remainders.

**Proposition 1.3** (Remainder Theorem). If \( a, b \) are integers and \( a > 0 \), then there exist unique integers \( q, r \) such that

\[
b = aq + r \quad \text{and} \quad 0 \leq r < a.
\]

**Proof.** We offer a proof based on our geometric understanding of the real line. Every real number \( x \) sits between two adjacent integers. More precisely there is a unique integer \( q \) such that

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

This is true in particular for the fraction \( b/a \). Hence there is a unique integer \( q \) such that

\[
q \leq \frac{b}{a} < q + 1.
\]

This is the same as saying that there is a unique integer \( q \) such that

\[
aq \leq b < aq + a,
\]

which is the same as saying there is a unique integer \( q \) such that

\[
0 \leq b - aq < a.
\]

Put \( r = b - aq \). So we have a unique integer \( q \) and a unique integer \( r \) such that

\[
b = aq + r \quad \text{and} \quad 0 \leq r < a,
\]

as desired. \( \square \)
CHAPTER 1. THE GREATEST COMMON DIVISOR

For example,

\[ 159000 = 11000 \cdot 14 + 5000 \text{ and } 0 \leq 5000 < 11000, \]

and also

\[ -159000 = 11000 \cdot (-15) + 6000 \text{ and } 0 \leq 6000 < 11000. \]

The common terminology for the numbers \( a, b, q, r \) in the Remainder Theorem is as follows:

- \( a \) is the **modulus** (from Latin for the ‘measure’ used to do the dividing)
- \( b \) is the **dividend** (from Latin for ‘that which is to be divided’)
- \( q \) is the **quotient** (from Latin for ‘number of times’)
- \( r \) is the **remainder** (from Latin for ‘that which stays’)

In the world of finance, a company might decide to pay a dividend of $159000 to be split among 11000 shares. Each share gets $14, while $5000 are left over. In the world of finance, the quotient of $14 is usually called the ‘dividend’, but when it comes to money, unlike mathematics, such abuse of terminology seems to be well tolerated. In any case the term ‘dividend’ is rarely used in number theory.

Some refer to the ‘modulus’ as the ‘divisor’, but we will restrict the use of the term divisor only to the case where the remainder is zero. We shall have quite a lot more to say about the ‘modulus’ is subsequent chapters.

Occasionally, as in the examples above, the product of two integers \( a, b \) will be denoted by \( a \cdot b \), with a dot between them. This makes it easy on the eye (and less clunky than brackets) when explicit integers are to be multiplied.

Having written \( b = aq + r \) with unique remainder \( r \) such that \( 0 \leq r < a \), we pick up something that is both easy to see and useful to know:

\[ a \mid b \text{ if and only if the remainder } r = 0. \]

**Finding remainders with everyday technology**

The Remainder Theorem can be rewritten as

\[ \frac{b}{a} = q + \frac{r}{a} \text{ where } 0 \leq \frac{r}{a} < 1. \]
This shows us how to find the quotient $q$ and remainder $r$ with any calculator (or smart phone).

- Punch in $\frac{b}{a}$.
- The integer part of your answer is $q$, and the decimal part is $\frac{r}{a}$.
- With $q$ sitting before you, calculate $b - aq$ to also get $r$.

The software Microsoft Excel can also readily find remainders for large numbers. To get the remainder of $b$ when divided by $a$ simply type $= \text{MOD}(b,a)$ into a cell of your spreadsheet. For example, the command $= \text{MOD}(1789675, 34567)$ will produce the output 26758, which is the remainder when 34567 is divided into 1789675. We will need to calculate remainders often, and use of some technology such as Excel, or whatever may feel comfortable, is strongly recommended.

### 1.2 Greatest common divisors

If $a, b \in \mathbb{Z}$, the integers $ax + by$ formed by letting $x, y$ run through $\mathbb{Z}$ are called integer combinations of $a$ and $b$, or more briefly, combinations of $a$ and $b$. For example, 14 is an integer recombination of 6 and 22, just because

$$6 \cdot 28 + 22(-7) = 14.$$ 

But 13 is not an integer combination of 6 and 22. Indeed, any integer combination of 6 and 22 must be even, while 13 is odd. Whatever $a, b$ might be, they are themselves integer combinations of $a$ and $b$, simply because

$$a = a \cdot 1 + b \cdot 0$$

and

$$b = a \cdot 0 + b \cdot 1.$$ 

Obviously, if $a = b = 0$, then $ax + by = 0$ regardless of the chosen $x, y$. But if at least one of $a$ or $b$ is not 0, then, we can always make a combination that is strictly positive. For example, take the integer combination

$$a \cdot a + b \cdot b = a^2 + b^2 > 0.$$ 

Of all the positive integer combinations that can be created, the smallest positive combination is by far the most interesting. Here is a fact which be surprising.
CHAPTER 1. THE GREATEST COMMON DIVISOR

Proposition 1.4. Let $a, b \in \mathbb{Z}$ with at least one of them not 0. If $d$ is the least positive integer combination of $a$ and $b$, then $d$ divides every combination of $a$ and $b$. In particular, $d \mid a$ and $d \mid b$.

Before going to the proof, take for example $a = 6, b = -15$. Obviously,

$$3 = 6 \cdot 3 + (-15) \cdot 1 > 0.$$ 

The lesser integers 1 and 2 cannot be expressed as combinations $6x + 15y$, because any such combination must be divisible by 3, while 1 and 2 are not divisible by 3. So 3 is the least positive combination of 6 and -15. And evidently $3 \mid 6$ and $3 \mid -15$.

Proof. We know that $0 < d = ax + by$ for some integers $x$, $y$, and $d$ is the least such positive integer combination of $a$ and $b$. Now take any combination

$$c = as + bt \text{ where } s, t \text{ are integers.}$$

We want to prove that $d \mid c$. According to the Remainder Theorem 1.3, write

$$c = dq + r \text{ where } 0 \leq r < d.$$ 

Thus,

$$0 \leq r = c - dq = as + bt - (ax + by)q = a(s - xq) + b(t - yq) < d.$$ 

We see that $r$ is an integer combination of $a$ and $b$, which is less than $d$, and non-negative. Because $d$ is the least positive combination of $a$ and $b$, the only option for $r$ is that $r = 0$. Hence $d \mid c$. In particular, $d \mid a$ and $d \mid b$ because $a, b$ are integer combinations of $a$ and $b$. 

Let us continue with $a, b$ in $\mathbb{Z}$, where at least one of them is not 0. The smallest positive combination of $a$ and $b$ is an extremely interesting integer for $a$ and $b$, as the next significant result shows. This more subtle result deserves close attention.

Theorem 1.5 (Greatest Common Divisor). Let $a, b \in \mathbb{Z}$, and suppose at least one of them is not 0. If an integer $d$ satisfies any one of the following properties, then $d$ satisfies them all.
1.2. GREATEST COMMON DIVISORS

1. $d$ is the least positive combination of $a$ and $b$.

2. $d$ is a positive combination of $a$ and $b$, and $d$ divides both $a$ and $b$.

3. $d$ is a positive common divisor of $a$ and $b$, and every common divisor of $a$ and $b$ is a divisor of $d$.

4. $d$ is the greatest common divisor of $a$ and $b$. That is, $d \mid a$, $d \mid b$, and whenever $c \mid a$, $c \mid b$ we have $c \leq d$.

Proof. We will prove that (1) $\implies$ (2) $\implies$ (3) $\implies$ (4) $\implies$ (1).

1 $\implies$ 2. This is part of what Proposition 1.4 says.

2 $\implies$ 3. We have that $1 \leq d = ax + by$ for some integers $x, y$, and $d$ is a common divisor of $a$ and $b$. If $c$ is another common divisor for $a$ and $b$, then $c \mid ax + by$ too. So $c \mid d$.

3 $\implies$ 4. Since every common divisor of $a$ and $b$ is a divisor of $d$ and $d \geq 1$, then $d$ must be bigger than these other common divisors.

4 $\implies$ 1. Now $d$ is the biggest integer that divides both $a$ and $b$. So for sure $1 \leq d$. We want $d$ to be the least positive combination of $a$ and $b$. Well, let $e$ be the least positive combination of $a$ and $b$. Since $e$ satisfies property 1, $e$ also satisfies properties 2, 3 and 4, as was just shown. In particular $e \mid a$, $e \mid b$ and $c \leq e$ whenever $c \mid a$, $c \mid b$. But $d$ has these latter properties too. So, $e \leq d$ and $d \leq e$. Thus $d = e$, which now satisfies property 1.

Naturally, the number $d$ that satisfies any one (and thereby all) of the conditions of Theorem 1.5 will be called the greatest common divisor of $a$ and $b$. We write

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

In order to confirm that an integer $d$ equals $\gcd(a, b)$, we can check that $d$ satisfies any one of the properties 1 to 4 of Theorem 1.5.

On the other hand if we know that $d = \gcd(a, b)$, then we can we use all of properties 1 to 4.

In case $a = b = 0$, it might make sense to say there is no greatest common divisor. Some say that $\gcd(0, 0) = 0$, but in any case the issue for us will not arise.

When $a, b$ are small integers, we can find the $\gcd(a, b)$ by inspection. For instance,

$$\gcd(24, 18) = 6.$$
10

CHAPTER 1. THE GREATEST COMMON DIVISOR

It should be obvious from property 4 of Theorem 1.5, that a change in sign of \( a \) or \( b \) has no effect on \( \gcd(a, b) \). That is,

\[
\gcd(a, b) = \gcd(\pm a, \pm b).
\]

It’s also obvious from property 4 that for any \( a > 0 \), we have \( \gcd(a, 0) = a \).

So, we just need a method for finding \( \gcd(a, b) \) when both \( a > 0 \) and \( b > 0 \), and when these integers are not so small. That famous method is known as the **Euclidean Algorithm** to which we now turn.

### 1.3 The Euclidean Algorithm

If \( u, v \) are integers and \( u = vw + z \) for some integers \( w, z \), the common factors of \( u \) and \( v \) coincide with the common factors of \( v \) and \( z \). Try to see why this is clear. Thus

\[
\gcd(u, v) = \gcd(v, z).
\]

This little but important observation together with the Remainder Theorem 1.3 provide an efficient way to find \( \gcd(a, b) \) for any pair of integers \( a, b \).

Say \( 0 < a < b \).

Apply the Remainder Theorem repeatedly as follows:

\[
\begin{align*}
  b &= aq_1 + r_1 & 0 &< r_1 < a \\
  a &= r_1q_2 + r_2 & 0 &< r_2 < r_1 \\
  r_1 &= r_2q_3 + r_3 & 0 &< r_3 < r_2 \\
  r_2 &= r_3q_4 + r_4 & 0 &< r_4 < r_3
\end{align*}
\]

We have the decreasing remainders \( r_1 > r_2 > r_3 > \cdots \geq 0 \). Sooner or later some remainder becomes 0. In other words, we get to some index \( n \) where

\[
\begin{align*}
  r_{n-3} &= r_{n-2}q_{n-1} + r_{n-1} & 0 &< r_{n-1} < r_{n-2} \\
  r_{n-2} &= r_{n-1}q_n + r_n & 0 &< r_n < r_{n-1} \\
  r_{n-1} &= r_nq_{n+1} + 0 & 0 &= r_{n+1}
\end{align*}
\]

From our little observation at the start, applied repeatedly, we now obtain
1.3. THE EUCLIDEAN ALGORITHM

$$\begin{align*}
r_n &= \gcd(0, r_n) \\
     &= \gcd(r_n, r_{n-1}) \\
     &= \gcd(r_{n-1}, r_{n-2}) \\
     &\vdots \\
     &= \gcd(r_4, r_3) \\
     &= \gcd(r_3, r_2) \\
     &= \gcd(r_2, r_1) \\
     &= \gcd(r_1, a) \\
     &= \gcd(a, b)
\end{align*}$$

The last positive remainder \( r_n \) equals \( \gcd(a, b) \). This famous process for obtaining greatest common divisors is called the **Euclidean Algorithm**.

For a baby example, let’s get \( \gcd(35, 294) \). Well,

\[
\begin{align*}
294 &= 35 \cdot 8 + 14 \\
35 &= 14 \cdot 2 + 7 \\
14 &= 7 \cdot 2 + 0
\end{align*}
\]

Thus \( \gcd(294, 35) = 7 \).

**Using technology to implement the Euclidean Algorithm**

If \( a, b \) are large integers, it makes sense to carry out the Euclidean Algorithm with the aid of a computer. Here is one way, among many, to design the algorithm with an **Excel** spreadsheet. We will carefully build the first two rows of the spreadsheet, corresponding to the first two lines of the Euclidean Algorithm, and then let **Excel** get the rest of the rows automatically, by recycling what we did to the first two rows.

- Start with given integers \( a, b \) whose greatest common divisor is required, and say \( 1 \leq a < b \).
- In cell A1 type in the given value of \( b \).
- In cell B1 type in the given value of \( a \).
CHAPTER 1. THE GREATEST COMMON DIVISOR

• In cell C1 type in: =INT(A1/B1). (Omit the period.)
  This calculates in C1 the integer part of $\frac{b}{a}$, which is the quotient $q_1$ when $a$
is divided into $b$.

• In cell D1 type in: =A1-B1*C1
  This gives the remainder $r_1 = b -aq_1$.
  This can also be achieved by typing in: MOD(A1,B1).

• In cell A2 type in: =B1.
  This copies the value of $a$ into cell A2.

• In cell B2 type in: =D1.
  This copies the remainder $r_1$ into cell B2.

• In cell C2 type in: =INT(A2/B2).
  What comes out in C2 is the integer part of $\frac{a}{r_1}$, which is the quotient $q_2$ when
  $r_1$ is divided into $a$.
  This can also be achieved by dragging the contents of cell C1 into cell C2.

• In cell D2 type in: =A2-B2*C2.
  This gives the remainder $r_2 = a -r_1q_2$. This can also be achieved by
  dragging the contents of cell D1 into cell D2.

• Finish the algorithm by dragging the contents of Row 2 down into subse-
  quent rows far enough that Excel gives division by zero warnings. Just
  ignore these warnings.

• The contents of columns C and D are the appropriate quotients and remain-
  ders. And the last non-zero remainder in Column D is the greatest common
  divisor of $a$ and $b$.

There are other software packages, such as MAPLE, that will carry out the Eu-
clidean Algorithm, but the above home-made approach is fairly straight-forward,
effective and readily available.

For example, let’s find gcd(24684, −41800). We might as well use 24684 and
41800. Using some technology (such as the instructions with Excel given above)
Thus \( \gcd(24684, 41800) = 44 \), the last positive remainder.

If using Excel, save your spread sheet! One can arbitrarily change the positive values of \( b \) and \( a \) in cells A1 and B1, respectively, to get an instant display of the Euclidean Algorithm for new inputs. (It might be necessary to drag the rows a bit farther down until the division by zeroes warning appears, which certifies that the Euclidean Algorithm is complete.) This spread sheet will also be useful for the calculation to follow. It might also be beneficial to add a row of labels indicating what’s what, for future reference.

Writing the GCD as an integer combination

From Theorem 1.5 we know that \( \gcd(a, b) = ax + by \) for some integers \( x, y \). It would be good to find such \( x, y \). Well, not only does the Euclidean Algorithm find \( \gcd(a, b) \), but by working our way back through the algorithm it also yields such \( x, y \).

Indeed, the penultimate (meaning second-last) line of the algorithm shows that \( r_n \) is a combination of \( r_{n-1} \) and \( r_{n-2} \). The line before that shows that \( r_{n-1} \) is a combination of \( r_{n-2} \) and \( r_{n-3} \). Thus, \( r_n \) is a combination of \( r_{n-2} \) and \( r_{n-3} \). And if we repeat such observations we get that \( r_n \) is a combination of \( a \) and \( b \).

For example, let’s retrace our steps in the prior example to write 44 as an
integer combination of 24684 and 41800. Well,

\[
44 = 132 - 88 \cdot 1 \\
= 132 - (220 - 132 \cdot 1) \cdot 1 \\
= 132 \cdot 2 - 220 \cdot 1 \\
= (352 - 220 \cdot 1) \cdot 2 - 220 \cdot 1 \\
= 352 \cdot 2 - 220 \cdot 3 \\
= 352 \cdot 2 - (1628 - 352 \cdot 4) \cdot 3 \\
= 352 \cdot 14 - 1628 \cdot 3 \\
= (1980 - 1628 \cdot 1) \cdot 14 - 1628 \cdot 3 \\
= 1980 \cdot 14 - (7568 - 1980 \cdot 3) \cdot 17 \\
= 1980 \cdot 65 - 7568 \cdot 17 \\
= (17116 - 7568 \cdot 2) \cdot 65 - 7568 \cdot 17 \\
= 17116 \cdot 65 - 7568 \cdot 147 \\
= 17116 \cdot 65 - (24684 - 17116 \cdot 1) \cdot 147 \\
= 17116 \cdot 212 - 24684 \cdot 147 \\
= (41800 - 24684 \cdot 1) \cdot 212 - 24684 \cdot 147 \\
= 41800 \cdot 212 - 24684 \cdot 359
\]

So \( \text{gcd}(24684, 41800) = 24684x + 41800y \) where \( x = -359 \) and \( y = 212 \).

**An extension of the Euclidean Algorithm**

The above procedure is painful to carry out by hand, or even with a basic calculator. Let’s explore a method of calculation, i.e. an algorithm, for solving the equation

\[ ax + by = \text{gcd}(a, b) \]

for \( x, y \), something we can get an application such as Excel to do automatically.

Here’s a method for writing each of the remainders \( r_1, r_2, r_3, \ldots, r_n \) from the Euclidean algorithm as integer combinations of \( a \) and \( b \), successively. We have to build up a pattern with care.

For \( r_1 \) we have

\[ r_1 = b - aq_1. \]
1.3. THE EUCLIDEAN ALGORITHM

So

\[ r_1 = as_1 + bt_1, \text{ where } s_1 = -q_1, t_1 = 1. \]

For \( r_2 \) we have

\[ r_2 = a - r_1 q_2 = a - (b - aq_1)q_2 = a(1 + q_1 q_2) - bq_2. \]

So

\[ r_2 = as_2 + bt_2, \text{ where } s_2 = 1 + q_1 q_2, t_2 = -q_2. \]

For \( r_3 \) an interesting pattern begins to emerge. We will be able to write \( r_3 \) as an integer combination of \( a, b \) using the combinations that we already have for \( r_1 \) and \( r_2 \). Indeed,

\[ r_3 = r_1 - r_2 q_3 = (as_1 + bt_1) - (as_2 + bt_2)q_3 = a(s_1 - s_2 q_3) + b(t_1 - t_2 q_3). \]

So

\[ r_3 = as_3 + bt_3, \text{ where } s_3 = s_1 - s_2 q_3, t_3 = t_1 - t_2 q_3. \]

The formula for \( r_4 \) is a copycat of the one for \( r_3 \). Just kick all subscripts up one notch. Indeed,

\[ r_4 = r_2 - r_3 q_4 = (as_2 + bt_2) - (as_3 + bt_3)q_4 = a(s_2 - s_3 q_4) + b(t_2 - t_3 q_4). \]

So

\[ r_4 = as_4 + bt_4, \text{ where } s_4 = s_2 - s_3 q_4, t_4 = t_2 - t_3 q_4. \]

Quickly the pattern becomes monotonous, and without further explanation the formula for \( r_5 \) is

\[ r_5 = as_5 + bt_5, \text{ where } s_5 = s_3 - s_4 q_5, t_5 = t_3 - t_4 q_5. \]

Proceed in this way until the expression of \( r_n = \gcd(a, b) \) as an integer combination of \( a \) and \( b \) emerges. The final \( s_n, t_n \) are suitable \( x, y \).

The Extended Euclidean Algorithm implemented with technology

Here is a suggestion on how to get these \( s_j, t_j \) using Excel.

Go back to the saved Excel spreadsheet where \( \gcd(a, b) \) was calculated. In column C the quotients \( q_1, q_2, \ldots, q_n \) appear. In column D the remainders \( r_1, r_2, \ldots, r_n \) appear and \( r_n = \gcd(a, b) \). Now all we have to do is generate the \( s_j, t_j \) based on the preceding discussion. We will put these in columns E and F exactly in accordance with the Extended Euclidean Algorithm.
In cell E1 type in: =-C1. (Note the minus sign, and omit the period.) This gives the integer \( s_1 \).

In cell F1 type in: 1. This gives the integer \( t_1 \).

In cell E2 type in: =1+C1*C2. This gives the integer \( s_2 \).

In cell F2 type in: =-C2. This gives the integer \( t_2 \).

In cell E3 type in: =E1-E2*C3. Noting that C3 contains the quotient \( q_3 \), the above instruction will put \( s_3 \) into cell E3.

In cell F3 type in: =F1-F2*C3. Noting that C3 contains the quotient \( q_3 \), the above instruction will put \( t_3 \) into cell F3.

Now drag the contents of the entire row 3 down into the subsequent rows until a zero remainder appears in column D. This will complete the pattern of the Extended Euclidean Algorithm automatically. (In the rows below the zero of column D, division by zero warnings will appear, which can be ignored.) In columns E and F, at the row where \( \gcd(a, b) \) appears just above the zero remainder, suitable \( x, y \) such that \( \gcd(a, b) = ax + by \) will have been found.

The algorithm just described is called the Extended Euclidean Algorithm.

When the Extended Euclidean Algorithm is implemented using the above instructions in Excel to \( a = 24684 \) and \( b = 41800 \), the following output appears.
1.3. THE EUCLIDEAN ALGORITHM

A B C D E F

41800 24684 1 17116 −1 1
24684 17116 1 7568 2 −1
17116 7568 2 1980 −5 3
7568 1980 3 1628 17 −10
1980 1628 1 352 −22 13
1628 352 4 220 105 −62
352 220 1 132 −127 75
220 132 1 88 232 −137
132 88 1 44 −359 212
88 44 2 0 950 −561

We see that \( \gcd(24684, 41800) = 44 \) appears in column D just above the zero. We also learn from the entries in columns E and F, in the row where the 44 appears, that

\[
44 = 24684 \cdot (-359) + 41800 \cdot 212,
\]

which can easily be verified with a calculator.

We should save this spreadsheet. If we need to find \( \gcd(a, b) \), as well as suitable \( x, y \) that give \( \gcd(a, b) = ax + by \), we can type the given \( a, b \) into cells B1, A1, respectively, of our saved spreadsheet. Then drag row 3 far enough down as needed. Those with advanced computer skills may feel inclined to improve the design of the program suggested here.

Why the Euclidean Algorithm is fast

The previous worked example may tempt us to wonder why the Euclidean Algorithm can find the greatest common divisor of two quite large numbers in merely 10 lines. In the Euclidean Algorithm back at the start of this section, we are given \( 1 \leq a < b \). The number of lines (i.e. divisions) in the algorithm needed to get to a remainder \( r_{n+1} = 0 \) is \( n + 1 \). This confirms that \( \gcd(a, b) \) is to be found in the preceding line as \( r_n \). We will now compare the number of lines, \( n + 1 \), to the size of \( b \).

This comparison hinges on one simple observation emerging from the Remainder Theorem 1.3. Namely, if \( 1 \leq a < b \) and if

\[
b = aq + r \text{ where } 0 \leq r < a,
\]
then
\[ r < \frac{b}{2}. \]

To see this, note that
\[ 1 \leq q = \frac{b - r}{a}. \]

Multiply through by \( a \) to get
\[ a \leq b - r \text{ and then } a + r \leq b. \]

And since \( r < a \) we come to
\[ 2r = r + r < a + r \leq b, \]

which is what we wanted.

Here is the comparison of the number of lines in the Euclidean Algorithm to the size of \( b \).

**Proposition 1.6.** If \( 1 \leq a \leq b \), then the number of lines in the Euclidean Algorithm is less than \( 1 + 2 \log_2 b \).

Before going to the proof it might be worthwhile to note why this result is interesting. It’s interesting because the logarithm of a big number is enormously smaller than the number.

For example, take \( b = 987654321234567 \), a rather intimidating quantity. With a calculator we obtain \( \log_2 b \approx 49.8 \), and then \( 1 + 2 \log_2 b \approx 100.6 \). We can thus be sure that the Euclidean Algorithm will find the greatest common divisor of this \( b \) with any \( a \) less than this \( b \) in no more than 100 lines. In comparison with the size of \( b \), that’s not a lot of lines.

For another example, suppose \( 1 \leq a \leq b \leq 2^{500} \). The possibilities for such integers \( a, b \) are huge beyond our imagination. The Euclidean Algorithm will find \( \gcd(a, b) \) in less than \( 1 + 2 \log_2 b \) lines. Clearly
\[ 1 + 2 \log_2 b \leq 1 + 2 \log_2 (2^{500}) = 1 + 2 \cdot 500 = 1001. \]

So, in a mere 1000 lines or fewer we can get the greatest common divisor of pretty much any \( a, b \) we might ever have to confront.

We should also note that even better bounds for the number of lines are known to exist. But our bound has the merits that the proof is relatively straightforward.
Proof. By the observation at the outset of our discussion, along with the first two lines of the algorithm we see that

\[ r_2 < r_1 < \frac{b}{2}. \]

By the observation once more, together with the third and fourth lines of the algorithm, we get

\[ r_4 < r_3 < \frac{r_1}{2} < \frac{b}{2^2}. \]

Similarly, from the fifth and sixth lines:

\[ r_6 < r_5 < \frac{r_3}{2} < \frac{b}{2^3}. \]

Going to the next two lines we get

\[ r_8 < r_7 < \frac{r_5}{2} < \frac{b}{2^4}. \]

The pattern should be clear after that.

Now suppose that the Euclidean algorithm has \( n+1 \) lines, whereby the greatest common divisor \( r_n \) appears on the \( n \)'th line.

If \( n \) is even, it follows from the above pattern of inequalities that

\[ 1 \leq r_n < \frac{b}{2^{n/2}}. \]

From this we get \( 2^{n/2} < b \), and then \( \frac{n}{2} < \log_2 b \), and then \( n < 2 \log_2 b. \)

If \( n \) is odd, the above pattern of inequalities gives

\[ 1 \leq r_n \leq \frac{b}{2^{(n+1)/2}}. \]

Then \( 2^{(n+1)/2} < b \), and then \( \frac{n+1}{2} < \log_2 b \), and then \( n < n + 1 < 2 \log_2 b. \)

Regardless of whether \( n \) is even or odd, we see that the number of lines in the Euclidean Algorithm needed to get to \( r_n = \gcd(a, b) \) is less than \( 2 \log_2 b. \) Since we need the last line to confirm that we are done, the number of lines in the algorithm is less than \( 1 + 2 \log_2 b. \) \( \square \)
1.4 Coprimeness

The situation where $\gcd(a, b) = 1$ deserves to be singled out. When this happens we say that $a, b$ are coprime. Some also say that $a, b$ are relatively prime. Coprime integers only have $\pm 1$ as common divisors. According to Property 1 of Theorem 1.5, two integers $a, b$ are coprime if and only if

$$ax + by = 1 \text{ for some integers } x, y.$$  

This is a fundamental concept in number theory, and we need to become very familiar with it.

For now, here are two key properties of coprimeness.

**Proposition 1.7.** Suppose $a, b, c$ are integers with $a, b$ coprime. If $a \mid bc$, then $a \mid c$.

**Proof.** Due to the coprimeness of $a$ and $b$ we have $ax + by = 1$ for some integers $x, y$. Thus

$$cax + cby = c.$$  

But $a \mid cb$, and for sure $a \mid a$. Hence $a$ divides the integer combination $cax + cby$ of $a$ and $bc$. That is, $a \mid c$. \qed

**Proposition 1.8.** If $a, b$ are coprime integers and $a \mid c$ and $b \mid c$, then $ab \mid c$.

**Proof.** We have integers $x, y$ such that $ax + by = 1$. Thus $c = cax + cby$.

Now $c = ak$ and $c = b\ell$ for some integers $k, \ell$.

Therefore,

$$c = b\ell ax + akby = ab(\ell x + ky).$$  

Evidently, $ab \mid c$. \qed

Proposition 1.7 feels about right. Roughly, it says that if $a$ fits inside $bc$ but no part of $a$ fits inside $b$, then all of $a$ already fits inside $c$. Proposition 1.8 also feels about right. Very roughly, it says that if $a$ and $b$ both fit inside $c$, and $a, b$ do not share a proper factor, then they fit into different parts of $c$, so that $c$ has room for their product. However, these vague remarks should not supersede the proper statements and proofs of Propositions 1.7 and 1.8.

Note that Proposition 1.7 fails when $a, b$ are not coprime. For instance, $4$ and $6$ are not coprime and $4 \mid 12 = 6 \cdot 2$. And yet $4 \nmid 2$. Proposition 1.8 also fails when
1.5. **THE EQUATION** $AX + BY = C$

$a, b$ are not coprime. Indeed, 4 and 6 are not coprime. And clearly $4 \mid 12$ as well as $6 \mid 12$. However, $4 \cdot 6 = 24 \nmid 12$.

And here’s another fact to remember.

If $d = \gcd(a, b)$, then the integers $\frac{a}{d}, \frac{b}{d}$ are coprime. Indeed, we know from Theorem 1.5 that $d = ax + by$ for some integers $x, y$. Hence

$$\frac{a}{d}x + \frac{b}{d}y = 1.$$  

Since 1 is truly the smallest positive combination of $\frac{a}{d}$ and $\frac{b}{d}$, it follows that $1 = \gcd \left( \frac{a}{d}, \frac{b}{d} \right)$, which makes them coprime.

### 1.5 The equation $ax + by = c$

A **Diophantine equation** is a polynomial equation whose parameters and unknowns are integers. For example,

$$x^2 - 3y^2 = 5, \; x^2 + y^2 = z^2, \; x^5 + y^5 = z^5, \; 300x + 482y = 70.$$  

Diophantus of Alexandria, born around 200 BC, is famous for his treatise called *Arithmetica* wherein he raised the matter of solving the equations now named in his honour.

To solve Diophantine equations is a signature endeavour of number theory. They can be truly daunting. For instance, the proof that

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

has no positive integer solutions $x, y, z$ when $n > 2$, sat unknown for hundreds of years. Only about twenty years ago did a fellow called Andrew Wiles prove this result, supported by an assortment of discoveries by mathematicians from around the world. It is famously known as **Fermat’s Last Theorem**.

For starters, let’s work on **linear** Diophantine equations in two unknowns. These take the form:

$$ax + by = c, \; \text{where } a, b, c \text{ are given integers}.$$  

In the trivial situation where $a = b = 0$, the equation $0x + 0y = c$ has a solution $x, y$ if and only if $c = 0$, and in that case any pair of integers $x, y$ forms
a solution. Having disposed of that, we assume that at least one of \(a\) or \(b\) is not zero. This puts us in a position to obtain greatest common divisors.

Obviously, if there is an integer solution \(x, y\), then \(\gcd(a, b)\) must divide \(c\). As it turns out, that’s all it takes for such an equation to have a solution. And after that, all possible solutions are readily discovered.

**Proposition 1.9.** Let \(a, b, c\) be integers, where \(a, b\) are not both 0. The Diophantine equation \(ax + by = c\) has an integer solution if and only if \(\gcd(a, b) \mid c\). In that case, for any initial solution \(x_0, y_0\), the general solution is given by

\[
x = x_0 + \frac{b}{d} \cdot n \quad \text{and} \quad y = y_0 - \frac{a}{d} \cdot n,
\]

where \(d = \gcd(a, b)\) and \(n \in \mathbb{Z}\).

**Proof.** If \(x, y\) form a solution, then \(\gcd(a, b) \mid ax + by\), because \(\gcd(a, b) \mid a\) and \(\gcd(a, b) \mid b\). So \(\gcd(a, b) \mid c\).

For the converse, let \(d = \gcd(a, b)\), and suppose \(d \mid c\). Since \(a, b\) are not both 0, let’s say that \(b \neq 0\). The key to this much more subtle part of the proof lies in Theorem 1.5, which tells us there exist integers \(s, t\) such that

\[
as + bt = d.
\]

We also have \(c = dk\) for some integer \(k\). So

\[
ask + btk = dk = c,
\]

which reveals that \(x = sk, y = tk\) is a solution of \(ax + by = c\).

Now for the general solution, given a specific solution \(x_0, y_0\). If \(x, y\) is any solution, we have

\[
ax + by = c = ax_0 + by_0,
\]

and then

\[
a(x - x_0) = b(y_0 - y).
\]

Divide by \(d = \gcd(a, b)\) to get

\[
\frac{a}{d}(x - x_0) = \frac{b}{d}(y_0 - y).
\]
1.5. THE EQUATION $AX + BY = C$

Here, $\frac{a}{d}, \frac{b}{d}$ may look like fractions, but they are integers. We should also remember that $d > 0$ since $b$ at least is not 0. Since $\frac{a}{d}, \frac{b}{d}$ are coprime integers, Proposition 1.7 tells us that $\frac{b}{d}$ divides $(x - x_0)$. So

$$x - x_0 = \frac{b}{d} \cdot n \text{ for some integer } n.$$ 

By substituting this into the former equation we obtain:

$$\frac{a}{d} \left( \frac{b}{d} \cdot n \right) = \frac{b}{d}(y_0 - y).$$

Cancel the non-zero $\frac{b}{d}$ to obtain

$$\frac{a}{d} \cdot n = y_0 - y.$$ 

Now we can see that

$$x = x_0 + \frac{b}{d} \cdot n \text{ and } y = y_0 - \frac{a}{d} \cdot n, \text{ for some integer } n.$$ 

Having seen that every solution of $ax + by = c$ takes the desired format, we also need to check that numbers in the desired format are, in fact, solutions. But that’s easy, because

$$a(x_0 + \frac{b}{d}n) + b(y_0 - \frac{a}{d}n) = ax_0 + by_0 = c.$$

Proposition 1.9 teaches us how to completely solve equations $ax + by = c$ for given integers $a, b, c$ and unknown integers $x, y$. Here is what to do.

- Find gcd$(a, b)$ using the Euclidean Algorithm, using Excel if one wishes, or by inspection if $a, b$ are small.
- If gcd$(a, b) \nmid c$, there is no solution.
- If $c = \text{gcd}(a, b)k$ for some $k$, then find $s, t$ in $\mathbb{Z}$ such that $as + bt = \text{gcd}(a, b)$ by backtracking along your Euclidean Algorithm, or using something like Excel as discussed already.
• Multiply through by $k$ to get $ask + btk = \gcd(a, b)k = c$.

• With your specific solution $x_0 = sk, y_0 = tk$, write the general solution in accordance with Proposition 1.9.

To illustrate the method, let’s solve

$$858x + 126y = 54.$$  

• First get $\gcd(858, 126)$ by the Euclidean Algorithm. So,

\[
\begin{align*}
858 &= 126 \cdot 6 + 102 \\
126 &= 102 \cdot 1 + 24 \\
102 &= 24 \cdot 4 + 6 \\
24 &= 6 \cdot 4 + 0
\end{align*}
\]

Thus $\gcd(858, 126) = 6$.

• Notice that $6|54$, in fact $54 = 6 \cdot 9$. So our Diophantine equation does have a solution.

• To find a particular solution, first backtrack along the Euclidean Algorithm in order to write 6 as a combination of 858 and 126. Thus,

\[
\begin{align*}
6 &= 102 - 24 \cdot 4 \\
    &= 102 - (126 - 102 \cdot 1) \cdot 4 \\
    &= 102 \cdot 5 - 126 \cdot 4 \\
    &= (858 - 126 \cdot 6) \cdot 5 - 126 \cdot 4 \\
    &= 858 \cdot 5 - 126 \cdot 34 \\
    &= 858 \cdot 5 - 126 \cdot 34.
\end{align*}
\]

Of course, by saving the Excel spreadsheet, which implemented the Extended Euclidean Algorithm, the steps to this point could be done in a blink.

• Now multiply through by 9 to get,

$$858 \cdot 45 + 126 \cdot (-306) = 54.$$  

A particular solution to our equation is $x_0 = 45, y_0 = -306$. 


The general solution, according to Proposition 1.9 is
\[ x = 45 + \frac{126}{6} n = 45 + 21n \]
\[ y = -306 - \frac{858}{6} n = -306 - 143n, \]
where \( n \) is any integer.

1.6 Exercises

1. (a) Find \( \gcd(5388, 618) \).
   
   (b) Find the complete solution of the linear Diophantine equation
   \[ 5388x + 618y = 42. \]
   
   Use technology or time, whichever is preferable.

2. Let \( d = \gcd(1896427914, 48280464440) \).
   
   By means of the Extended Euclidean Algorithm find \( d \), and find suitable integers \( x, y \) that solve
   \[ d = 1896427914x + 48280464440y. \]
   
   Repeat for \( d = \gcd(2452548, 2943234) \).
   
   Use technology, for example, the instructions provided if Excel is to be used.

3. If \( k, a, b \) are positive integers prove that
   
   (a) \( \gcd(ka, kb) = k \gcd(a, b) \)
   
   (b) \( \gcd(a, b + ka) = \gcd(a, b) \).

4. How many integers strictly between \( n^2 \) and \( (n + 1)^2 \) are divisible by \( n \)?

5. Find all integers \( n \) such that \( \frac{n^3 + n + 12}{n - 2} \) is an integer.
   
   Hint. Write the numerator as a polynomial in \( n - 2 \).
6. A rational number $\frac{a}{b}$ is said to be in reduced form when the numerator $a$ is coprime with the denominator $b$. Every rational number $\frac{u}{v}$ can be put in reduced form. Simply, let $z = \gcd(u, v)$ and notice that $$\frac{u}{v} = \frac{u/z}{v/z},$$ where $u/z, v/z$ are now coprime.

This exercise is about explaining the uniqueness of the reduced form, except for a possible alteration in the signs of the numerator and denominator.

If $\frac{a}{b} = \frac{c}{d}$, with all of $a, b, c, d$ positive, and both fractions are in reduced form, prove that $a = c$ and $b = d$.

7. Peter needs to spend all of $\$9.42$ on pickled and jalapeno peppers. Pickled peppers cost 21 cents each, while jalapenos cost 33 cents each.

(a) Find all possible combinations of peppers that Peter can pick to buy.

(b) What is the maximum total number of peppers he can buy?

8. Show how to measure out exactly 2 litres of water using a 27 litre jug and a 16 litre jug. You are allowed to waste water.

9. Find all integers $c$ such that the Diophantine equation $8x + 5y = c$ has exactly one solution $x, y$ where $x > 0$ and $y > 0$.

Hint. Write the general solution $x, y$ to this equation. It will depend on $c$ as well as an arbitrary integer $n$. You need to find those $c$ such that $x > 0$ and $y > 0$ happens for just one $n$.

10. Let $a, b, c, d \in \mathbb{Z}$ and suppose $ad - bc = 1$. If an integer $e$ divides the combinations $ax + by$ and $cx + dy$, show that $e \mid x$ and $e \mid y$.

11. If $a, b$ are coprime integers and $c \mid at$ and $c \mid bt$, show that $c \mid t$.

12. If $a \mid bc$, show that $a \mid b \gcd(a, c)$.

13. If $m, n$ are integers such that $n^3 + n = m^4$, show that $m$ is even. Then show that $n$ is even. After that show that $16 \mid n$.

14. If $a, b, c$ are integers with $a, c$ coprime, prove that $\gcd(ab, c) = \gcd(b, c)$.

15. Mathematical induction is indispensable for proving some results in number theory. If we want to prove that all statements in a sequence of statements $P_n$ hold, we can do this:
1.6. EXERCISES

- check that $P_1$ holds, and
- assuming the statements $P_k$ hold for $k < n$, verify that $P_n$ holds

Sometimes this is called the strong principle of induction. Now, let’s put it into practice with a problem.

Let $m, n$ be positive integers, with $m \leq n$, and let $d = \text{gcd}(n, m)$. If $t$ is any integer prove that

$$\text{gcd}(t^n - 1, t^m - 1) = t^d - 1$$

Hint.
Do an induction on $n$.
If $b = aq + r$, remember that $\text{gcd}(b, a) = \text{gcd}(a, r)$.
This was the main idea that justified the Euclidean Algorithm.
Now for $b = t^n - 1$ and $a = t^m - 1$ find suitable $q, r$.

16. An enormous number of patterns have been discovered about the Fibonacci numbers $f_n$, which are defined recursively by:

$$f_0 = 0, f_1 = 1, \text{ and } f_{n+2} = f_{n+1} + f_n \text{ for } n \geq 0.$$ 

The first few Fibonacci numbers are

$$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144 \ldots$$

Note. Some people start with $f_0 = 1, f_1 = 1$, but this little accounting variation makes no essential difference. We’ll stick with $f_0 = 0, f_1 = 1$.

Prove the following few things, mostly by induction.

(a) $f_n$ is even if and only if $3 | n$.

Hint. You are asked to explain why the $f_n$ take the pattern:

even, odd, odd, even, odd, odd, even, odd, odd, even, odd, odd, even, odd, odd, even, odd, odd, ... 

(b) $f_n, f_{n+1}$ are coprime.

(c) $f_{n+m} = f_m f_{n+1} + f_{m-1} f_n$ for all $m \geq 1$ and $n \geq 0$. 
(d) If \( j, \ell \geq 0 \) and \( j \mid \ell \), then \( f_j \mid f_\ell \).

For example, \( 5 \mid 10 \) and, sure enough, \( f_5 = 5 \mid 55 = f_{10} \). Also, \( 6 \mid 12 \), and we see that \( f_6 = 8 \mid 144 = f_{12} \).

Hint. Rephrase the statement as follows:

\[
\text{if } m, n \geq 0, \text{ then } f_n \mid f_{nm}.
\]

Then do induction on \( n \). You will need part (c).

17. One of the more interesting things about the Fibonacci numbers is their relationship to the so-called golden ratio. The legendary golden ratio is the number \( \tau = \frac{1 + \sqrt{5}}{2} \approx 1.618 \). Here you are asked to prove the crazy formula

\[
f_n = \frac{\tau^n - (1 - \tau)^n}{\sqrt{5}}.
\]

Hint. First verify the following properties of \( \tau \):

\[
\tau^2 = 1 + \tau, \quad \tau - 1 = 1/\tau, \quad 2\tau - 1 = \sqrt{5}, \quad (1 - \tau)^2 = 2 - \tau.
\]

These might come in handy.

In your proof by induction on \( n \), first verify the formula for \( f_0 \) and \( f_1 \). Next, for each \( n \geq 2 \), assume the formula for all \( f_k \) where \( k < n \). Now play around with the definition of \( f_n \) for \( n \geq 2 \) and those identities involving \( \tau \) to deduce the formula for \( f_n \). By playing around just right, it won’t get too messy.
Chapter 2

Primes

Some of the most stubborn problems in all of mathematics concern the distribution of prime numbers. Here’s a few such problems.

- Is every even number beyond four the sum of two odd primes?
- Are there infinitely many pairs \( p, q \) where \( p, q \) are prime and \( q = p + 2 \). This is the twin primes problem.
- Are there infinitely many primes of the form \( x^2 + 1 \)?
- Are there infinitely primes of the from \( 2^q - 1 \)?
- If \( x \) is any real number, and \( \pi(x) \) is the number of primes less than \( x \), how fast does \( \pi(x) \) grow as a function of \( x \)?

These are hard problems. The last one was answered towards the end of the nineteenth century with a memorable result known as the Prime Number Theorem. The others remain unsolved.

Primes are important because every positive integer can be factored into primes in only one way. This needs to be proven, since it is by no means obvious.

We have to begin with the basics.

2.1 Factoring into primes

Definition 2.1. An integer \( p \) is called prime when \( p \neq 0, p \neq \pm 1 \) and the only factors that \( p \) has are \( \pm 1, \pm p \).
Clearly, \( p \) is prime if and only if \(-p\) is prime. To avoid this double counting of primes we shall work only with positive primes, and to be brief we shall usually omit the word “positive”. Here are the primes up to 131:

\[
2 \quad 3 \quad 5 \quad 7 \quad 11 \quad 13 \quad 17 \quad 19 \quad 23 \quad 29 \quad 31 \quad 37 \quad 41 \quad 43 \quad 47 \\
53 \quad 59 \quad 61 \quad 67 \quad 71 \quad 73 \quad 79 \quad 83 \quad 89 \quad 97 \quad 101 \quad 103 \quad 107 \quad 109 \quad 113 \quad 127 \quad 131
\]

As the numbers get bigger, it’s no longer that easy to tell whether a given integer is prime. For example, most people might guess at first that 91 is prime, but 91 = 13 \( \cdot \) 7. The development of tests for primality is still a hot topic of research. Leaving 0 and \( \pm 1 \) out of the picture, the integers that are not prime are called composite. For example, 91 = 13 \( \cdot \) 7 is composite.

Although we might sense that there are infinitely many primes, this is not at all obvious. So we had better get that done right away. But first we make note of a result that we have known since elementary school, and which points to the importance of primes.

**Proposition 2.2.** If \( n \) is an integer and \( n \geq 2 \), then \( n \) can be factored into primes.

**Proof.** If an integer is already prime, we consider it as the product of one prime and thus factored as itself.

Let’s use induction. If \( n = 2 \), then \( n \) is surely a product of one prime. Suppose 2, 3, 4, \ldots, \( n - 1 \) can each be factored into primes. Now look at \( n \). If \( n \) is prime, then \( n \) is a product of primes, namely itself. If \( n \) is composite, write \( n = k \ell \), where \( 1 < k < n \) and \( 1 < \ell < n \). Since \( k, \ell \) are among the integers 2, 3, 4, \ldots, \( n - 1 \), each of them factors into primes. That is,

\[
k = p_1 \cdot p_2 \cdots p_r \quad \text{and} \quad \ell = q_1 \cdot q_2 \cdots q_s, \quad \text{where the } p_j, q_j \text{ are primes.}
\]

Then, of course,

\[
n = k \ell = p_1 \cdots p_r \cdot q_1 \cdots q_s
\]

is a factorization of \( n \) into primes. \( \Box \)

In a degenerate sense the integer 1 is also a product of primes, namely, the product of zero many primes. But there is no need to fret over this.

And now comes a classic discovery with the proof already in Euclid’s books.

**Proposition 2.3.** There are infinitely many primes.
2.2. Unique Factorization

Proof. Given any finite list of primes \( p_1, p_2, \ldots, p_n \), here’s how to come up with one more prime not on such a list. That will show there are infinitely many primes.

Let

\[ n = p_1 \cdot p_2 \cdots p_n + 1. \]

According to Proposition 2.2, \( n \) has a prime factor \( q \). This \( q \) cannot be equal to any \( p_1, \ldots, p_n \). Indeed, if \( q \) equalled some \( p_j \), then we would have that

\[ q \mid n - p_1 \cdot p_2 \cdots p_n, \]

which says that \( q \mid 1 \). Since no prime is a factor of 1, we conclude that \( q \) is a fresh prime, not equal to any of \( p_1, p_2, \ldots, p_n \).

\[ \Box \]

2.2 Unique factorization

The special thing about primes is that there is only one way to factor an integer into primes. Ambiguous factorings such as

\[ 24 = 6 \cdot 4 = 8 \cdot 3 \]

do not occur when only primes are involved in the factors. To prove unique factorization, we need what can only be called the signature property of primes.

Proposition 2.4. If \( p \) is prime and \( p \mid ab \) for some integers \( a, b \), then \( p \mid a \) or \( p \mid b \).

Proof. Say \( p \nmid a \). Let \( d = \gcd(p, a) \). Since \( d \mid p \), the definition of primes forces \( d = 1 \) or \( d = p \), and since \( p \nmid a \), it must be that \( d = 1 \). So \( p \) and \( a \) are coprime. By Proposition 1.7, we get that \( p \mid b \).

\[ \Box \]

Note that Proposition 2.4 fails when \( p \) is not a prime. For instance

\[ 6 \mid 12 = 3 \cdot 4 \text{ while } 6 \nmid 3 \text{ and } 6 \nmid 4. \]

We should also observe that Proposition 2.4 readily extends to the product of several integers. Namely, if a prime \( p \) divides the product \( a_1 a_2 a_3 \cdots a_k \), then \( p \) already divides one of the \( a_j \). The proof to follow, that factorization of integers into primes is unique, rests on the shoulders of Proposition 2.4.
Theorem 2.5 (Unique Factorization). If \( p_1, p_2, \ldots, p_n \) and \( q_1, q_2, \ldots, q_m \) are two lists of primes (positive) with repetitions allowed, and if

\[
p_1 \cdot p_2 \cdots p_n = q_1 \cdot q_2 \cdots q_m,
\]

then \( n = m \), and after a possible rearrangement of the \( q_j \), we have

\[
p_1 = q_1, \ p_2 = q_2, \ldots, \ p_n = q_n.
\]

Proof. Clearly \( p_1 | q_1 \cdot q_2 \cdots q_m \). By Proposition 2.4, \( p_1 \) divides some \( q_j \). Rearrange the \( q_j \), and say \( p_1 | q_1 \). Since \( q_1 \) is prime, we get \( p_1 = 1 \) or \( p_1 = q_1 \). But the first option cannot be true. Thus \( p_1 = q_1 \). Hence,

\[
p_1 \cdot p_2 \cdots p_n = p_1 \cdot q_2 \cdots q_m.
\]

Cancel \( p_1 \) to get

\[
p_2 \cdot p_3 \cdots p_n = q_2 \cdot q_3 \cdots q_m.
\]

Repeat the argument, with the necessary rearrangement of the \( q_j \), to get \( p_2 = q_2 \), and then

\[
p_3 \cdot p_4 \cdots p_n = q_3 \cdot q_4 \cdots q_m.
\]

Continuing in this fashion, after suitable rearrangement of the \( q_j \), we end up with one of the following possibilities.

- \( n < m \), and \( p_1 = q_1, \ldots, p_n = q_n, \ 1 = q_{m-n} \cdots q_m \)
- \( m < n \), and \( p_1 = q_1, \ldots, p_m = q_m, \ p_{n-m} \cdots p_n = 1 \)
- \( m = n \), and \( p_1 = q_1, \ldots, p_n = q_n \)

The first two options cannot happen, since only \( \pm 1 \) can be factors of \( 1 \) and the \( p_j, \ q_j \) are not \( \pm 1 \). So indeed, \( m = n \), and all \( p_j = q_j \) after some relabelling of the \( q_j \). \( \Box \)

The multiplicity of a prime inside an integer

It is customary to collect repeated primes in the unique factorization of a positive integer \( x \) as follows:

\[
x = p_1^{e_1} \cdot p_2^{e_2} \cdots p_n^{e_n},
\]
where $p_j$ are now distinct primes and the exponent integers $e_j > 0$. The $e_j$ is called the multiplicity of $p_j$ inside $x$. Furthermore, it is very convenient at times to allow $e_j = 0$. So, a prime $p$ has multiplicity 0 in $x$ when $p$ is not needed to factor $x$. For example, with

$$x = 2^3 \cdot 3^0 \cdot 5^1 \cdot 11^8 \cdot 29^4$$

we see that 3 has multiplicity 0 in $x$, while the multiplicity of 11 in $x$ is 8.

The Unique Factorization Theorem 2.5 is telling us that for every $x \geq 2$ and every prime $p$, there is a unique multiplicity of $p$ in $x$. This is the number of times that $p$ appears in the unique factorization of $x$. For any $x$, only a finite number of primes have strictly positive multiplicity in $x$. For example, the multiplicity of 2 in 320 is 6 because $320 = 2^6 \cdot 5^1$ while the multiplicity of 7 in 320 is 0. Since no prime is a factor of 1, the multiplicity of every prime $p$ in 1 is 0.

The Unique Factorization Theorem is so deeply embedded into our psyche that we tend to use it without being aware that we are using it.

**Divisibility and unique factorization**

We can now describe the notion of divisibility, in terms of unique factorization and multiplicities. The proof is a bit tedious, but not particularly hard. We can view it as an exercise in organization.

**Proposition 2.6.** A positive integer $a$ divides another positive integer $b$ if and only if, for every prime $p$, the multiplicity of $p$ in $a$ is less than or equal to the multiplicity of $p$ in $b$.

**Proof.** Suppose $a \mid b$. So, $b = ac$ for some $c$. Let $p_1, \ldots, p_n$ be the distinct primes that are used to factor both $a$ and $c$. Write

$$a = p_1^{d_1} \cdot p_2^{d_2} \cdots p_n^{d_n} \text{ and } c = p_1^{e_1} \cdot p_2^{e_2} \cdots p_n^{e_n},$$

where $d_j, e_j$ are the multiplicities of $p_j$ in $a, c$ respectively, and $d_j \geq 0, e_j \geq 0$. Then

$$b = ac = p_1^{d_1+e_1} \cdot p_2^{d_2+e_2} \cdots p_n^{d_n+e_n}.$$

Now the multiplicity of each $p_j$ in $b$ is $d_j + e_j$ and $d_j + e_j \geq d_j$.

Also, for every prime $p$, other than $p_1, \ldots, p_n$, the multiplicity of $p$ in each of $a, c, b$ is 0 and since $0 \leq 0$, our claim holds for these $p$ as well.
Conversely, suppose that for every prime \( p \) the multiplicity of \( p \) in \( a \) is less than or equal to the multiplicity of \( p \) in \( b \). Let \( p_1, \ldots, p_n \) be the primes that are needed to factor \( a \) and \( b \). Write

\[
b = p_1^{e_1} \cdot p_2^{e_2} \cdots p_n^{e_n} \quad \text{and} \quad a = p_1^{d_1} \cdot p_2^{d_2} \cdots p_n^{d_n}
\]

where the \( e_j \), \( d_j \) are the respective multiplicities and \( e_j \geq 0, d_j \geq 0 \). We have assumed \( d_j \leq e_j \). Thus

\[
b = (p_1^{e_1 - d_1} \cdot p_2^{e_2 - d_2} \cdots p_n^{e_n - d_n})(p_1^{d_1} \cdot p_2^{d_2} \cdots p_n^{d_n}) = (p_1^{e_1 - d_1} \cdot p_2^{e_2 - d_2} \cdots p_n^{e_n - d_n})a
\]

Since every \( e_j - d_j \geq 0 \), we see that \( p_1^{e_1 - d_1} \cdot p_2^{e_2 - d_2} \cdots p_n^{e_n - d_n} \) is an integer, say \( c \). Thus \( b = ca \), and so \( a \mid b \). \( \square \)

The greatest common divisor of two positive integers can now be written in terms of the unique factorization of the integers.

**Proposition 2.7.** Suppose that \( a, b, c \) are positive integers with unique factorizations

\[
a = p_1^{d_1} \cdot p_2^{d_2} \cdots p_n^{d_n}, \quad b = p_1^{e_1} \cdot p_2^{e_2} \cdots p_n^{e_n}, \quad c = p_1^{\ell_1} \cdot p_2^{\ell_2} \cdots p_n^{\ell_n},
\]

where the \( p_j \) are distinct primes and the \( e_j, d_j, \ell_j \) are their respective multiplicities in \( a, b, c \). The number \( c \) equals \( \gcd(a, b) \) if and only if every \( \ell_j = \min(d_j, e_j) \).

**Proof.** To say that \( c = \gcd(a, b) \) means that \( c \) is the biggest integer that divides both \( a \) and \( b \). According to Proposition 2.6, \( c \) can only come from taking \( \ell_j \) as big as possible subject to the requirement that \( \ell_j \leq d_j \) and \( \ell_j \leq e_j \). The biggest such \( \ell_j \) are given by \( \ell_j = \min(d_j, e_j) \). \( \square \)

To illustrate Proposition 2.7,

\[
\gcd(2^3 \cdot 3^0 \cdot 7^5 \cdot 13^4 \cdot 29^8, 2^5 \cdot 3^7 \cdot 7^3 \cdot 13^4 \cdot 29^0) = 2^3 \cdot 3^0 \cdot 7^3 \cdot 13^4 \cdot 29^0.
\]

**Coprimeness in terms of primes**

Two positive integers \( a, b \) are coprime if and only if 1 is their greatest common divisor. By referring to Proposition 2.7, we see that this happens if and only if \( \min(e_j, d_j) = 0 \) for all \( j \). In other words, \( a, b \) are coprime if and only if there is no prime with positive multiplicity in both \( a \) and \( b \). To repeat, \( a, b \) are coprime if and
only if there is no prime \( p \) such that \( p | a \) and \( p | b \). Another way to say it is that \( a, b \) are coprime when the unique factorizations of \( a \) and \( b \) can be presented as

\[
a = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k} \quad \text{and} \quad b = q_1^{d_1} q_2^{d_2} \cdots q_ℓ^{d_ℓ},
\]

where the distinct primes \( p_j \) never overlap with the distinct primes \( q_j \). We shall have frequent occasion to use these little observations.

The unique factorization of every positive integer into a product of primes is the keystone that holds number theory together.

**Why is 1 not a prime?**

Some might argue that the integer 1 deserves to be called a prime. After all, it cannot be factored down any further. However, if we allow 1 to be a prime, then unique factorization goes out the window. Indeed, we can factor the integer 1 from any integer \( a \) as much as we like:

\[
a = 1 \cdot 1 \cdot 1 \cdot 1 \cdots 1 \cdot a.
\]

True primes don’t do that. The number of times they appear in the unique factorization of \( a \) is unique. That’s what allows the factorization to be called “unique”. Better to leave 1 out of the basket of integers known as primes.

### 2.3 A strange ring

We may wonder if there was any point to proving the Unique Factorization Theorem. After all, is it not intuitively clear that anyone who factors an integer into primes will get the same primes with the same multiplicities as anyone else? Here we discuss an exotic example to make the case that unique factorization is not to be taken as self evident.

The negative integer \(-5\) has two square roots, which are themselves complex numbers. The notation \( \sqrt{-5} \) is thereby ambiguous. Nevertheless, we shall pick one of the two square roots, and denote our choice by \( \sqrt{-5} \).

Now let

\[
A = \{ a + b\sqrt{-5} : a, b \in \ZZ \}.
\]

This is a set of complex numbers. For example,

\[
0 = 0+0\sqrt{-5} \in A, \quad 1 = 1+0\sqrt{-5} \in A, \quad \text{and every integer } n = n + 0\sqrt{-5} \in A.
\]
Also
\[ -7 + 12\sqrt{-5} \in A, \text{ but } \frac{1}{2} + \frac{2}{3}\sqrt{-5} \notin A. \]

It is very easy to see that if \( x, y \in A \), then \( x \pm y \in A \), and the product \( xy \in A \). In other words, \( A \) is closed under the usual arithmetic operations. So, \( A \) is what is known as a ring.

Just to be sure, let’s check that the product \( xy \in A \), whenever \( x, y \in A \). Well, let \( x = a + b\sqrt{-5}, y = c + d\sqrt{-5}, \) where \( a, b, c, d \in \mathbb{Z} \). Then
\[
xy = (a + b\sqrt{-5})(c + d\sqrt{-5}) = (ac - 5bd) + (bc + ad)\sqrt{-5} \in A,
\]
since \( ac - 5bd \in \mathbb{Z} \) and \( bc + ad \in \mathbb{Z} \). The ultra-routine verification that the sum \( x + y \in A \) is omitted.

The norm function

A useful tool for the study of \( A \) is the norm function defined as follows. For \( x = a + b\sqrt{-5} \) in \( A \), let
\[
N(x) = a^2 + 5b^2.
\]
Note that \( N(x) \) is a non-negative integer, and that \( N(x) = 0 \) if and only if \( x = 0 \).

Also, we can write \( x = a + \sqrt{5}bi \) where \( i = \sqrt{-1} \). After noting that
\[
|x| = \sqrt{a^2 + (\sqrt{5}b)^2} = \sqrt{a^2 + 5b^2},
\]
we see that \( N(x) = |x|^2 \), the square of the absolute value of \( x \). It is well known that \( |uv| = |u||v| \) for every \( u, v \) in \( \mathbb{C} \), and thus
\[
|uv|^2 = (|u||v|)^2 = |u|^2|v|^2.
\]

When this is applied to the norm on \( A \), we obtain the following useful multiplicative property of norms:
\[
N(xy) = N(x)N(y) \text{ for every } x, y \in A.
\]

Divisibility and units

If \( x, y \in A \), we say \( x \) divides \( y \) in \( A \), and write \( x | y \), provided \( y = xz \) for some \( z \) in \( A \). For example, \( 1 + 2\sqrt{-5} \) divides \( 21 + 0\sqrt{-5} \) in \( A \) because
\[
21 = (1 + 2\sqrt{-5})(1 - 2\sqrt{-5}).
\]
2.3. A STRANGE RING

Clearly, ±1 divides every \( y \) in \( A \). Are there other elements of \( A \) that divide every number in \( A \)?

We say that \( x \) in \( A \) is a **unit** provided \( x \) divides every \( y \) in \( A \). That is, for every \( y \) there is a \( z \) in \( A \) such that \( y = xz \).

Let us now verify that the only units of \( A \) are ±1. Suppose \( x \) is a unit of \( A \). Thus \( x \mid 1 \), and so \( xz = 1 \) for some \( z \) in \( A \). Write \( x = a + b\sqrt{-5}, z = c + d\sqrt{-5} \) where \( a, b, c, d \in \mathbb{Z} \). Taking norms we get

\[
1 = N(1) = N(xz) = N(x)N(z) = (a^2 + 5b^2)(c^2 + 5d^2).
\]

Since \( a^2 + 5b^2 \) is a positive integer factor of 1, it follows that \( a^2 + 5b^2 = 1 \). By inspection, we deduce that \( b = 0 \), and from \( a^2 = 1 \), we get \( a = ±1 \). So

\[
x = ±1 + 0\sqrt{-5} = ±1.
\]

Up to this point, \( A \) is just like \( \mathbb{Z} \) in that its only units are ±1.

**Primes in the strange ring**

As we did in \( \mathbb{Z} \), we can define the concept of a **prime** in \( A \). We say that \( x \) is an \( A \)-prime provided \( x \) is not 0 and not ±1, and the only divisors of \( x \) are ±1 and \( ±x \).

For example, let us verify that \( 3 = 3 + 0\sqrt{-5} \) is an \( A \)-prime. Accordingly, suppose that

\[
x = a + b\sqrt{-5} \in A \text{ and that } x \mid 3.
\]

So

\[
3 = xy \text{ for some } y = c + d\sqrt{-5} \text{ in } A.
\]

Apply the norm to get

\[
9 = N(3) = N(xy) = N(x)N(y) = (a^2 + 5b^2)(c^2 + 5d^2).
\]

Hence \( a^2 + 5b^2 \) equals one of 1, 3, 9.

- If \( a^2 + 5b^2 = 1 \), we can see as before that \( b = 0, a = ±1 \) and so \( x = ±1 \).

- The possibility \( a^2 + 5b^2 = 3 \) does not occur. Indeed, if \( b \neq 0 \), then \( a^2 + 5b^2 \geq 5 \), and thus \( a^2 + 5b^2 \neq 3 \). While if \( b = 0 \), then \( a^2 + 5b^2 = a^2 \), and this perfect square is not equal to 3.
• If $a^2 + 5b^2 = 9$, then $c^2 + 5d^2 = 1$, and as seen already this forces $y = \pm 1$, which forces $x = \pm 3$.

Thus 3 is (in addition to being a prime in $\mathbb{Z}$) an $A$-prime.

By imitating the preceding argument, we would see that 2 is also an $A$-prime.

For yet another $A$-prime look at $1 + \sqrt{−5}$. To see that this number is an $A$-prime, suppose $x = a + b\sqrt{−5}$ and $y = c + d\sqrt{−5}$ give

$$1 + \sqrt{−5} = xy.$$ 

Apply the norm to get

$$6 = N(1 + \sqrt{−5}) = N(xy) = N(x)N(y) = (a^2 + 5b^2)(c^2 + 5d^2).$$

Thus the integer $a^2 + 5b^2$ is one of 1, 2, 3, 6.

• If $a^2 + 5b^2 = 1$, we get as already seen that $x = \pm 1$.

• The option $a^2 + 5b^2 = 3$ does not happen, as we already saw. By a similar argument, $a^2 + 5b^2 = 2$ does not happen.

• Finally, if $a^2 + 5b^2 = 6$, then $c^2 + 5d^2 = 1$ which gives $y = \pm 1$, and then $x = \pm (1 + \sqrt{−5})$.

Hence $1 + \sqrt{−5}$ is an $A$-prime.

Similarly, we can prove that $1 − \sqrt{−5}$ is an $A$-prime.

**Something unexpected**

Look at the following factorizations of 6 inside $A$.

$$2 \cdot 3 = 6 = (1 + \sqrt{−5})(1 − \sqrt{−5}).$$

As we have seen, all of 2, 3, $1 + \sqrt{−5}, 1 − \sqrt{−5}$ are $A$-primes. They do not factor down any further in $A$.

We have just factored $6 + 0\sqrt{−5}$ in *two fundamentally different ways* into $A$-primes. There is also no way to change the sign of 2 or 3 by multiplying them by $\pm 1$ and get $1 + \sqrt{−5}$ or $1 − \sqrt{−5}$.

In $A$ there is factorization into primes, but it is *not* unique. The history of Fermat’s Last Theorem that $x^n + y^n = z^n$ has no solution in positive integers when $n > 2$, is riddled with erroneous proofs. One of the more common fallacies was the presumption that rings such as $A$ possessed unique factorization.
2.4 Exercises

1. If $a$ is an integer such that $40|a^2$, prove that $20|a$.

2. If $n$ is a positive integer and $r$ is an integer between 0 and $n$, the binomial coefficient $\binom{n}{r}$ is the integer $\frac{n!}{r!(n-r)!}$. If $n$ is prime, show that $n$ divides $\binom{n}{r}$. Does this result hold when $n$ is not prime?

3. How many zeroes are at the end of the decimal expansion of $100!$?

4. If an integer $n \geq 2$ and if $n|(n - 1)! + 1$, prove that $n$ is prime.
   
   Hint. What happens if $n$ is not prime?

5. Let $a, b$ be positive integers. The least common multiple of $a$ and $b$ is the smallest integer that they both divide. For example, 6 and 4 both divide their product, which is 24, but that is not their least common multiple, which is 12. We denote this least common multiple by $\text{lcm}(a, b)$.

   (a) Let $a = p_1^{e_1} \cdot p_2^{e_2} \cdots p_n^{e_n}$ and $b = p_1^{d_1} \cdot p_2^{d_2} \cdots p_n^{d_n}$ be the unique factorizations of $a$ and $b$ into distinct primes with multiplicities as indicated.

   Write a formula for $\text{lcm}(a, b)$ as a product of distinct primes with suitable multiplicities.

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

   (c) If $a|c$ and $b|c$, prove that $\text{lcm}(a, b)|c$.

6. If $a, b$ are integers and $a^2|b^2$, prove that $a|b$.

7. If $a, b$ are coprime integers and $ab = c^2$ for some integer $c$, show that $a = t^2$ and $b = s^2$ for some integers $t$ and $s$.

   Does this result still hold if we replace the 2 by any positive exponent $n$?

   Hint. The easiest thing is to use the Unique Factorization Theorem.
8. Show that the Diophantine equation \( x^3 + x = y^2 \) has no integer solution \( x, y \), where \( x > 0 \) and \( y > 0 \).

Hint. If \( x, y \) are positive solutions, explain why \( x, x^2 + 1 \) are coprime, and then note that \( x(x^2+1) \) is a perfect square. Then use the preceding problem.

9. Let \( \mathbb{Z}[i] \) denote the set of complex numbers of the form \( a + bi \) where \( a, b \) are integers and \( i = \sqrt{-1} \). Such complex numbers are called Gaussian integers. We shall study them more closely in Chapter 7.

For example, \( 1 = 0 + 1i, 0 = 0 + 0i, 2 - 7i, 43 - 8i, i = 0 + i \) are Gaussian integers, but \( 1/2 + \sqrt{2}i \) is definitely not a Gaussian integer. By a quick inspection, we see that the sum, difference and product of any two Gaussian integers is another Gaussian integer. So we can we speak of the ring \( \mathbb{Z}[i] \) of Gaussian integers.

As with the usual integers, we have divisibility. If \( x, y \in \mathbb{Z}[i] \) we say that \( x \mid y \) when \( y = xz \) for some other \( z \in \mathbb{Z}[i] \).

Then we say that \( x \) in \( \mathbb{Z}[i] \) is a unit provided \( x \) divides every element of \( \mathbb{Z}[i] \). We will discover in the exercises that follow that the only units in \( \mathbb{Z}[i] \) are \( \pm 1 \) and \( \pm i \).

We can speak about primes in \( \mathbb{Z}[i] \). A Gaussian integer \( p \) is called a Gaussian prime provided the only factors of \( x \) are \( \pm 1, \pm i, \pm p, \pm ip \). In other words, if you want to factor \( p \), one of the factors must be a unit. It’s impossible to stop units from being factors of \( p \), because units are factors of everything in the ring.

A useful tool for understanding the Gaussian integers is the norm.
For \( z = a + ib \) in \( \mathbb{Z}[i] \) let

\[
N(z) = a^2 + b^2.
\]

Obviously \( N(z) \in \mathbb{Z} \) and its easy to see that \( N(xy) = N(x)N(y) \) for every \( x, y \) in \( \mathbb{Z}[i] \).

(a) Show that the four elements \( \pm 1, \pm i \) are units of \( \mathbb{Z}[i] \).
(b) Show that the only units of \( \mathbb{Z}[i] \) are \( \pm 1, \pm i \).
(c) Show that 2 is not prime in \( \mathbb{Z}[i] \).
(d) Show that 3 is prime in \( \mathbb{Z}[i] \), i.e. 3 is a Gaussian prime.
2.4. **EXERCISES**

(e) Show that $1 + i$ is a prime in $\mathbb{Z}[i]$.

Hint. The norm comes in handy for (b), (d) and (e).

10. An integer $n$ is a sum of two squares when there are integers $a, b$ such that $n = a^2 + b^2$. If $m, n$ are each sums of two squares, prove that their product $mn$ is also a sum of two squares.

Hint. For a quick proof, think about factoring into Gaussian integers.

11. If $n \geq 1$ and $2^n - 1$ is a prime, prove that $n$ is prime.

Such primes are called **Mersenne primes**. Some exceptionally large Mersenne primes have been discovered.

Hint. Start with supposing $n$ is not prime. Then use the following factorization, which holds for any $x$ and any positive $m$.

$$x^m - 1 = (x - 1)(x^{m-1} + x^{m-2} + \cdots + x^2 + x + 1).$$

12. A positive integer is called **perfect** (for reasons that seem obscure) provided the sum of its proper positive divisors is $x$. For example, 6 is perfect because

$$6 = 1 + 2 + 3.$$ 

Another perfect number is 496. Indeed, $496 = 31 \cdot 16$, and we see that the positive proper divisors of 496 are $1, 2, 4, 8, 16, 31, 31 \cdot 2 = 62, 31 \cdot 4 = 124, 31 \cdot 8 = 248$. Then we add them up and get

$$1 + 2 + 4 + 8 + 16 + 31 + 62 + 124 + 248 = 496.$$ 

If we include $x$ as divisor of itself, we see that $x$ is perfect if and only if the sum of all of its positive divisors (including $x$) is $2x$.

A number of the type $2^n - 1$ is sometimes a prime, then called a Mersenne prime.

If $2^n - 1$ is a prime, show that $x = 2^{n-1}(2^n - 1)$ is a perfect number.

Hint. You should be able to see a pattern in the factors of $x$ and also to add them up. Don’t forget that $2^n - 1$ is given to be prime.

13. This exercise is about the strange ring $A$ made up of numbers of the form $a + b\sqrt{-5}$ where $a, b$ are integers.
(a) Find a prime \( p \) in \( A \) such that \( p \nmid x \), \( p \nmid y \), and yet \( p \mid xy \) in \( A \).

(b) Show that 41 is not an \( A \)-prime.

   Hint. \( 41 = 36 + 5 \).

(c) If \( x \in A \) and the integer \( N(x) \) is a usual prime in \( \mathbb{Z} \), show that \( x \) is a prime in \( A \).

(d) Show that \( 3 + 4\sqrt{-5} \) is a prime in \( A \).

(e) Find an element \( v \) in \( A \) with the signature property of primes. Namely, that whenever \( v \mid xy \) in \( A \), then \( v \mid x \) or \( v \mid y \) in \( A \).

14. If \( n > 1 \), prove that sum

\[
1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{n}
\]

is not an integer.

This one is tricky.
Chapter 3

Modular arithmetic

3.1 Congruences

We fix an integer $n \geq 1$, and call it a “modulus”. This comes from the Latin word for a “measure”, or as we might say, a “yardstick”. This modulus is used to compare two integers.

If $a$ is any integer, the Remainder Theorem gives us $q$ and $r$ where

\[ a = nq + r \quad \text{and} \quad 0 \leq r < n. \]

The remainder $r$ between 0 and $n-1$ is uniquely determined by $a$.

Definition 3.1. Two integers $a, b$ are said to be \emph{congruent modulo} $n$, and we write

\[ a \equiv b \mod n, \]

provided $a, b$ have equal remainders between 0 and $n - 1$, when they are each divided by $n$.

When $n = 1$, we see that the only possible remainder upon division by 1 is 0. In this trivial and uninteresting case, every $a$ is congruent to every $b$. So, to keep the story worthwhile, \textit{we typically assume that} $n \geq 2$.

For instance, when the modulus is $n = 6$, here are all of the integers congruent to each other with a remainder of 2:

\[ \ldots, -10, -4, 2, 8, 14, 20, 26, \ldots \]
These are the integers that take the form $6q + 2$, where $q$ is any integer.

It should be plain to see with a modulus of 6, that the infinite set of integers $\mathbb{Z}$ gets carved into 6 disjoint pieces as shown.

- The integers with a remainder of 0:
  $$\{6q : q \in \mathbb{Z}\} = \{0, \pm 6, \pm 12, \ldots \}$$

  This is just the set of integers divisible by 6.

- Those with a remainder of 1:
  $$\{6q + 1 : q \in \mathbb{Z}\} = \{\ldots, -11, -5, 1, 7, 13, \ldots \}$$

- Those with a remainder of 2:
  $$\{6q + 2, q \in \mathbb{Z}\} = \{\ldots, -10, -4, 2, 8, 14, \ldots \}$$

- Those with a remainder of 3: $\{6q + 3 : q \in \mathbb{Z}\}$

- Those with a remainder of 4: $\{6q + 4 : q \in \mathbb{Z}\}$

and finally...

- Those with a remainder of 5: $\{6q + 5 : q \in \mathbb{Z}\}$.

The same observation applies to any modulus $n$. Given a modulus $n \geq 2$, the set of integers $\mathbb{Z}$ gets partitioned into $n$ disjoint pieces according to the $n$ possible remainders $0, 1, 2, \ldots, n - 1$. We will return to this very important idea shortly.

It also does not take much to see the following.

- If $a$ is any integer, then $a \equiv a \mod n$.

- If $a \equiv b \mod n$, then $b \equiv a \mod n$

- If $a \equiv b$ and $b \equiv c \mod n$, then $a \equiv c \mod n$.

Thus congruences can be treated pretty much like equations.
3.1. CONGRUENCES

A test for congruence without using remainders

Here’s a quick, alternate way to tell if $a \equiv b \mod n$.

**Proposition 3.2.** Two integers $a, b$ are congruent modulo $n$ if and only if $n \mid b - a$.

**Proof.** If $a \equiv b \mod n$, then

$$b = nq + r \text{ and } a = nt + r,$$

for some integers $q, t$ and some integer $r$ where $0 \leq r < n$. Then

$$b - a = n(q - t),$$

which clearly reveals that $n \mid b - a$.

Conversely, suppose $n \mid b - a$. According to the Remainder Theorem we have integers $q, t, r, s$ such that

$$b = nq + r, \ 0 \leq r < n \text{ and } a = nt + s, \ 0 \leq s < n.$$

Then

$$r - s = (b - nq) - (a - nt) = b - a + n(t - q).$$

Since $n \mid b - a$, we see that $n \mid r - s$.

Now, if $r - s \neq 0$, item 3 of Proposition 1.2 shows us that $n \leq |r - s|$. But since $0 \leq r < n$ and $0 \leq s < n$, we also have that $|r - s| < n$. The resulting contradiction, that $n < n$, forces us to conclude $r - s = 0$, which means $r = s$. This proves that $a \equiv b \mod n$, in accordance with Definition 3.1.

Since $a \equiv b \mod n$ if and only if

$$b - a = nq \text{ for some integer } q,$$

we learn from Proposition 3.2 why the word “modulus” for $n$ makes sense as a synonym for “yardstick”. Two integers are congruent when the distance between them can be measured off in integer multiples of the “yardstick” $n$. 

3.2 The replacement principle

We can readily see that $3 \equiv -6 \mod 9$ and that $11 \equiv 2 \mod 9$. Now $3 + 11 = 14$ while $-6 + 2 = -4$, and we can’t help but notice that $14 \equiv -4 \mod 9$. Furthermore, $3 \cdot 11 = 33$ while $-6 \cdot 2 = -12$. Again it hits us that $33 \equiv -12 \mod 9$. These are not coincidences, as we now can prove.

**Proposition 3.3.** Let $n$ be a modulus, and suppose

$$a \equiv a_1 \mod n \quad \text{and} \quad b \equiv b_1 \mod n.$$  

Then

$$a \pm b \equiv a_1 \pm b_1 \quad \text{and} \quad ab \equiv a_1b_1 \mod n.$$  

**Proof.** First, let’s check that

$$a + b \equiv a_1 + b_1 \mod n.$$  

This comes down to showing that $n \mid ((a + b) - (a_1 + b_1))$, in other words that $n \mid (a - a_1) - (b - b_1)$. But this is obviously true, since $n \mid a - a_1$ and $n \mid b - b_1$.

The proof that $a - b \equiv a_1 - b_1 \mod n$ imitates the one for addition so much that we can skip it.

Now, we check

$$ab \equiv a_1b_1 \mod n.$$  

We need to show that $n \mid (ab - a_1b_1)$. Notice

$$ab - a_1b_1 = ab - a_1b + a_1b - a_1b_1 = (a - a_1)b + a_1(b - b_1),$$  

which is an integer combination of $a - a_1$ and $b - b_1$. From this it becomes obvious that $n \mid (ab - a_1b_1)$, because $n \mid a - a_1$ and $n \mid b - b_1$. \hfill \Box

Proposition 3.3 is telling us that after any sum, difference or product of integers is replaced by a sum, difference or product of respectively congruent integers, the resulting sum, difference or product remains congruent to the original. For instance, if

$$a \equiv a_1, \ b \equiv b_1, \ \text{and} \ c \equiv c_1 \mod n,$$

then

$$a^{100} \equiv a_1^{100} \mod n,$$

$$(a^2 - 5b)c \equiv (a_1^2 - 5b_1)c_1 \mod n,$$  

and

$$(17 + abc)ac + bc^4 \equiv ((17 + 2n) + a_1b_1c_1)a_1c_1 + b_1c_1^4 \mod n.$$
All such formulas involving only the addition, subtraction and multiplication of integers are called *polynomial expressions*.

In general, the replacement principle says the following.

**Proposition 3.4** (Replacement Principle). If \( f(a, b, c, d, \ldots) \) is any polynomial expression in integers \( a, b, c, d, \ldots \), and

\[
a \equiv a_1, \quad b \equiv b_1, \quad c \equiv c_1, \quad d \equiv d_1, \ldots \quad \text{mod } n,
\]

then

\[
f(a, b, c, d, \ldots) \equiv f(a_1, b_1, c_1, d_1, \ldots) \quad \text{mod } n.
\]

The replacement principle holds because we can apply Proposition 3.3 repeatedly to all of the additions, subtractions and multiplications involved in building the polynomial \( f(a, b, c, d, \ldots) \).

As we shall come to appreciate, replacement is key to solving a wide array of problems about integers.

**Reduction modulo \( n \)**

The process of finding the remainder of a given integer \( a \) upon division by \( n \) will be called reduction modulo \( n \). Another way to think about it is that the *reduction of a modulo \( n \)* is the unique integer \( r \) between 0 and \( n - 1 \) such that \( a \equiv r \mod n \).

Reductions of not-too-large integers, such as \( 306 \equiv 5 \mod 7 \) and \( 625 \equiv 2 \mod 7 \) can easily be found by using a calculator to implement the Remainder Theorem. But replacement can, at times, be a handier way to get the reduction.

**Example 3.5.** Let us find the remainder when \( 306^{100} \) is divided by 7. In other words, let us reduce this large integer modulo 7.
CHAPTER 3. MODULAR ARITHMETIC

Well, modulo 7 we have:

\[ 306^{100} \equiv 5^{100} \text{ since } 306 \equiv 5 \mod 7, \text{ so replace } \]
\[ = (5^4)^{25} \]
\[ = 625^{25} \equiv 2^{25} \text{ since } 625 \equiv 2 \mod 7, \text{ so replace } \]
\[ = (2^6)^4 \cdot 2 \]
\[ = 64^4 \cdot 2 \equiv 1^4 \cdot 2 \text{ since } 64 \equiv 1 \mod 7, \text{ so replace } \]
\[ = 2. \]

The reduction modulo 7 is 2.

More illustrations of replacement

Example 3.6. Let us show that the Diophantine equation

\[ x^2 + y^2 = 4z + 3 \]

has no integer solutions \( x, y, z \).

Since there are infinitely many possibilities for \( x, y, z \), it seems a bit daunting to show that none of them work. But a little trick with congruences and replacement makes this problem quite straight-forward.

Indeed, we are really asked to show that the congruence

\[ x^2 + y^2 \equiv 3 \mod 4 \]

has no integer solution \( x, y \).

If such a solution \( x, y \) did exist, then any other \( x_1, y_1 \) such that \( x \equiv x_1 \mod 4 \) and \( y \equiv y_1 \mod 4 \) would also be a solution, by replacement. But \( x, y \) have to each be congruent \( \mod 4 \) to one of their possible remainders 0, 1, 2, 3. Thus, if there is a solution \( x, y \), then there will also be a solution \( x, y \) among the four possibilities 0, 1, 2, 3. So, all we need to do is check that these possibilities never solve the congruence.

Notice that

\[ 0^2 \equiv 0, \ 1^2 \equiv 1, \ 2^2 \equiv 0, \ 3^2 \equiv 1 \mod 4. \]
Thus with $x, y$ congruent $\text{mod } 4$ to one of $0, 1, 2, 3$, we see that
\[ x^2 + y^2 \equiv \text{one of } 0 + 0 = 0, 0 + 1 = 1, 1 + 0 = 1, 1 + 1 = 2 \text{ mod } 4. \]
It’s now clear that $x^2 + y^2 \not\equiv 3 \text{ mod } 4$, ever.

Here’s another example on the use of replacement.

**Example 3.7.** Let’s show that $x^5 \equiv x \text{ mod } 5$ for all integers $x$.

Well, every integer $x$ is congruent $\text{mod } 5$ to one of its possible remainders $0, 1, 2, 3, 4$. If the desired congruence holds for these remainders, then, by replacement, the congruence holds for any integer $x$. By routine calculation we see that
\[ 0^5 = 0 \equiv 0, \quad 1^5 = 1 \equiv 1, \quad 2^5 = 32 \equiv 2, \quad 3^5 = 243 \equiv 3, \quad 4^5 = 1024 \equiv 4 \text{ mod } 5. \]
Having verified the result on the five possible remainders, replacement gives the result for all integers.

### 3.3 Residues

As we saw in the preceding examples, the replacement principle allows us to reduce congruence problems about the infinite set of integers to a matter of finitely many cases. This important idea needs to be considered carefully.

For any integer $a$, the “box notation” $[a]$ will stand for the set of integers that have the same remainder as $a$ when divided by $n$. In other words,
\[ [a] = \{ b \in \mathbb{Z} : b \equiv a \text{ mod } n \}. \]

We have to get used to thinking of $[a]$ as one thing even though that one thing consists of an infinite set of numbers. (By way of analogy, we can think of the one thing we call “red” as the collection of all red objects. For another analogy, the different looking fractions $2/3, 6/9, (-6)/(-9), 20/30, 10/15, \ldots$ are ”one thing”.)

For example, with $n = 9$ we have that $[15]$ is the set of integers congruent to $15 \text{ mod } 9$. In other words, $[15]$ consists of the set of integers which leave the same remainder as $15$ does when divided by $9$. That remainder is $6$. So $[15]$ equals the set of integers that have a remainder of $6$ when divided by $9$. Thus
\[ [15] = \{ \ldots, -21, 21, -6, 15, 24, 33, 42, \ldots \}. \]
A careful look at the definition of the box notation shows that

\[ [a] = [b] \text{ if and only if } a \equiv b \mod n. \]

We can (and will) pass automatically back and forth between the box notation and congruences.

**Definition 3.8.** The set \([a]\) is called the *residue class*, or *residue* of the integer \(a\) using some modulus \(n\). Some mathematicians also call this the *congruence class* of \(a\) modulo \(n\), but we will stick with the term “residue”. The integer \(a\) is called a *representative* for the set of integers congruent to \(a\) modulo \(n\).

As noted above, two integers represent the same residue class if and only if they are congruent \(\mod n\).

Every integer is congruent to its remainder when divided by \(n\). The possible remainders are

\[ 0, 1, 2, \ldots, n - 1. \]

None of these remainders are congruent to each other, since \(n\) never divides the difference of any two of them. Thus there are only a *finitely many* residue classes, in fact exactly \(n\) of them.

This finite set of residues modulo \(n\) will henceforth be denoted by \(\mathbb{Z}_n\).

For example, here are the five residues of \(\mathbb{Z}_5\).

- \([0] = \{0 + 5q : q \in \mathbb{Z}\}\)
- \([1] = \{1 + 5q : q \in \mathbb{Z}\}\)
- \([2] = \{2 + 5q : q \in \mathbb{Z}\}\)
- \([3] = \{3 + 5q : q \in \mathbb{Z}\}\)
- \([4] = \{4 + 5q : q \in \mathbb{Z}\}\)

The representatives 0, 1, 2, 3, 4 chosen above are often preferred because they are the remainders that can result upon division by 5, but we should not lose sight of the possibility of using other representatives. For example, the five residues of \(\mathbb{Z}_5\) can just as well be represented as

- \([0] = \{0 + 5q : q \in \mathbb{Z}\} = [10]\)
3.3. RESIDUES

- $[1] = \{1 + 5q : q \in \mathbb{Z}\} = \{0\}$
- $[-1] = \{-1 + 5q : q \in \mathbb{Z}\} = \{4\}
- \{2\} = \{2 + 5q : q \in \mathbb{Z}\} = \{0\}
- \{2\} = \{-2 + 5q : q \in \mathbb{Z}\} = \{3\}

The ring of residues

We want $\mathbb{Z}_n$ to become a ring too. In other words, we would like to add and multiply residues modulo $n$, and get other residues modulo $n$. The replacement principle of Proposition 3.3 makes it all possible.

Let us translate Proposition 3.3 into the language of residues. It tells us the following.

If $[a] = [a_1]$ and $[b] = [b_1]$ in $\mathbb{Z}_n$,

then $[a + b] = [a_1 + b_1]$, $[a - b] = [a_1 - b_1]$ and $[ab] = [a_1b_1]$.

This gives us the right to define the following operations of addition, subtraction, and multiplication in $\mathbb{Z}_n$. For $[a], [b]$ in $\mathbb{Z}_n$, let

$[a] + [b] = [a + b]$, $[a] - [b] = [a - b]$, $[a][b] = [ab]$.

To add, subtract or multiply two residues we simply add, subtract or multiply their representatives, respectively. The beauty of the replacement principle, as rephrased just above, is that

this procedure gives the same result no matter which representatives are used for our residues.

When we carry out such operations in $\mathbb{Z}_n$, we are doing modular arithmetic. The beauty of $\mathbb{Z}_n$ is that it is a finite ring, through which we can approach some major problems concerning the infinite ring of integers.

Like the ring $\mathbb{Z}$ of integers, the ring $\mathbb{Z}_n$ has a “zero”, namely $[0]$, and a “one”, namely $[1]$. And we see that for any residue $[a]$: $[0] + [a] = [0 + a] = [a]$ and $[1][a] = [1 \cdot a] = [a]$. 
just like “zeroes” and “ones” are supposed to behave.

For example, in $\mathbb{Z}_{11}$ we have calculations such as

$$[8] + [21] = [29] = [7] = [-3] + [10].$$

and

$$[8][21] = [168] = [3] = [-30] = [-3][10].$$

To do modular arithmetic:

- carry out ordinary arithmetic and freely replace the integers involved
  by congruent integers modulo $n$.

However, we must caution against a habit that we take for granted within the ring $\mathbb{Z}$. We know for integers $a, b, c$ that if $ab = ac$ and $a \neq 0$, then $b = c$. In other words, we can cancel non-zero integers. There is no luxury of cancellation in $\mathbb{Z}_n$.

For example, take $\mathbb{Z}_6$. Here

$$[3] \neq [0] \text{ and } [3][2] = [6] = [0] = [3][0].$$

Even though $[3]$ is not $[0]$ in $\mathbb{Z}_6$ we dare not cancel the $[3]$, for that would lead to the false conclusion $[2] = [0]$.

For another example take $\mathbb{Z}_{20}$. Notice that in $\mathbb{Z}_{20}$:

$$[8][3] = [24] = [4] \text{ and } [8][8] = [64] = [4].$$

So $[8][3] = [8][8]$, but cancelling the $[8]$ would lead to the falsehood $[3] = [8]$.

What we have so far

At the risk of being unduly repetitive, here is what to keep in mind for a given modulus $n \geq 2$.

- For any integer $a$, its residue is the set
  $$[a] = \{ b : b \equiv a \mod n \}$$
  $$= \{ b : a \text{ and } b \text{ have equal remainders upon division by } n \}.$$

- The set of residues is denoted by $\mathbb{Z}_n$. This comprises exactly $n$ things.
3.4. LINEAR CONGRUENCES

- \([a] = [b]\) in \(\mathbb{Z}_n\) if and only if \(a \equiv b \mod n\).
  Thus equations in \(\mathbb{Z}_n\) can be converted into congruences in \(\mathbb{Z}\) modulo \(n\).

- \(\mathbb{Z}_n\) is finite with exactly \(n\) residues in it. Here is a full listing of the possible residues modulo \(n\):

  \([0], [1], [2], [3], \ldots, [n - 1]\).

  This listing is complete because every integer is congruent modulo \(n\) to exactly one remainder from 0 to \(n - 1\). The residues in the list do not repeat themselves because no two integers from 0 to \(n - 1\) are congruent \(\mod n\).

- \(\mathbb{Z}_n\) is a ring with the operations defined as follows:

  \([a] + [b] = [a + b], [a] - [b] = [a - b]\) and \([a][b] = [ab]\).

  The replacement principle tells us that if \(a \equiv c \mod n\) and \(b \equiv d \mod n\), then \(a \pm b \equiv c \pm d \mod n\) and \(ab \equiv cd \mod n\).

  This is the same as saying the following about \(\mathbb{Z}_n\). If \([a] = [c]\) and \([b] = [d]\) then \([a \pm b] = [c \pm d]\) and \([ab] = [cd]\).

  So, the operations of \(\mathbb{Z}_n\) do not depend on the chosen representative \(x\) for a residue class \([x]\). They are well-defined.

- For those who have studied abstract algebra, \(\mathbb{Z}_n\) is nothing but the quotient ring of \(\mathbb{Z}\) modulo the ideal \(n\mathbb{Z}\) integers divisible by \(n\). But we won’t need to know this.

- We cannot automatically cancel non-zero residues.

3.4 Linear congruences

We can begin to discuss equations in \(\mathbb{Z}_n\). A good place to start is with the linear equation

\([a][x] = [b]\) where \([a], [b]\) are given and \([x]\) is unknown.

This problem is the same as that of solving the linear congruence

\(ax \equiv b \mod n\).
If the integers involved are not too big, the solution can be found by a finite inspection. For instance, let’s solve

\[ [3][x] = [8] \text{ in } \mathbb{Z}_{11}, \]

which is the same as solving the congruence

\[ 3x \equiv 8 \pmod{11}. \]

We can plug in all possible residues \([0],[1],[2],\ldots,[10]\) for \([x]\), and see which ones work. The only residue that works is \([x] = [10]\). This was discovered by doing a bit of mental modular arithmetic.

For another example, consider

\[ [2][x] = [6] \text{ in } \mathbb{Z}_{8}, \text{ or equivalently } 2x \equiv 6 \pmod{8}. \]

Here we see that \([x] = [3]\), as well as \([x] = [7]\), are solutions. We should notice immediately the unusual situation of a linear equation having more than one solution.

For a third example take,

\[ [2][x] = [3] \text{ in } \mathbb{Z}_{8}. \]

Here, a quick inspection of the possibilities \([0],[1],[2],\ldots,[7]\) reveals that no solution \([x]\) exists.

**The method for solving a linear congruence**

To solve a linear congruence, a systematic method is required when the integers in question become large. The next result shows us that the problem of linear equations in \(\mathbb{Z}_n\) comes down to solving a linear Diophantine equation in \(\mathbb{Z}\). This brings it all back to the Euclidean Algorithm.

**Proposition 3.9.** A residue \([x]\) in \(\mathbb{Z}_n\) is a solution of the equation

\[ [a][x] = [b] \]

if and only if its representative \(x\) is a solution of the linear Diophantine equation

\[ ax - ny = b, \]

along with some other integer \(y\). Thus, \([a][x] = [b]\) has a solution if and only if \(\gcd(a,n)\mid b\).
3.4. LINEAR CONGRUENCES

Proof. Since \([a][x] = [ax]\), the problem is to solve \([ax] = [b]\), which is the same as solving the congruence \(ax \equiv b \mod n\). An integer \(x\) solves this congruence if and only if \(n \mid ax - b\). This comes down to saying that \(ax - b = ny\) for some integer \(y\), or equivalently that \(ax - ny = b\). Thus, the first part of our result is shown. The second part follows from Proposition 1.9.

To solve \([a][x] = [b]\), we have just seen that we need to solve \(ax - ny = b\), and this we can do readily by the Euclidean Algorithm discussed after Proposition 1.9.

Example 3.10. Let’s solve \([102][x] = [54]\) in \(\mathbb{Z}_{126}\).

According to Proposition 3.9, we need to solve the Diophantine equation

\[
102x - 126y = 54.
\]

By the Euclidean Algorithm we first get \(\gcd(126, 102)\).

\[
\begin{align*}
126 &= 102 \cdot 1 + 24 \\
102 &= 24 \cdot 4 + 6 \\
24 &= 6 \cdot 4 + 0.
\end{align*}
\]

Since \(\gcd(126, 102) = 6\) and \(6 \mid 54\), our equation in \(\mathbb{Z}_{126}\) has a solution. To find the residue \([x]\) obtain the general solution of \(102x - 126y = 54\). Backtrack along the Euclidean Algorithm to see that

\[
6 = 102 - 24 \cdot 4 = 102 - (126 - 102 \cdot 1) \cdot 4 = 102 \cdot 5 - 126 \cdot 4.
\]

Of course, the above procedure can be expedited by means of software such as Excel.

Since \(54 = 9 \cdot 6\), we get

\[
102 \cdot 45 - 126 \cdot 36 = 54.
\]

The general solution for \(x\) to our Diophantine equation is

\[
x = 45 + 21n, \text{ where } n \in \mathbb{Z}.
\]

We won’t need the \(y\)-solution.
So, all $x \equiv 45 \mod 21$ will give $[102][x] = 54$ in $\mathbb{Z}_{126}$. By replacement this means that all $x \equiv 3 \mod 21$ will yield $[102][x] = [54]$ in $\mathbb{Z}_{126}$.

Now $\mathbb{Z}_{126}$ consists of $[0], [1], [2], [3], [4], [5], [6], \ldots, [123], [124], [125]$. We pick off those $[x]$ from this list for which $x \equiv 3 \mod 21$. That would be $[x] = \text{any one of } [3], [24], [45], [66], [87], [108]$.

It may be somewhat unexpected that in $\mathbb{Z}_{126}$, the linear equation $[102][x] = [54]$ has 6 solutions, and that $\gcd(126, 102) = 6$. We might wonder whether it is always true that if $\gcd(a, n) | b$, then the number of solutions to $[a][x] = [b]$ in $\mathbb{Z}_n$ equals $\gcd(a, n)$. Yes, it’s true, but we’ll let that matter go for now.

Units and cancellation modulo $n$

As we have seen, linear equations $[a][x] = [b]$ in $\mathbb{Z}_n$ might have one, several, or no solutions. The special situation where they always have just one solution is very important. The result that follows addresses this matter. Many of the statements in this result are minor variations of each other. Yet all facets of this result deserve to be well understood.

**Proposition 3.11.** Let $n$ be a modulus, and $a$ be any integer. The following statements about $a$ are equivalent to each other.

1. The integer $a$ is coprime with the modulus $n$.
2. The congruence $ax \equiv 1 \mod n$ has a solution.
3. For every integer $b$, the congruence $ax \equiv b \mod n$ has a solution.
4. For every integer $b$, the equation $[a][x] = [b]$ in $\mathbb{Z}_n$ has a solution.
5. The equation $[a][x] = [1]$ in $\mathbb{Z}_n$ has a solution.
6. Whenever $[a][b] = [a][c]$ in $\mathbb{Z}_n$, the $[a]$ can be cancelled to give $[b] = [c]$.
7. Whenever the congruence $ab \equiv ac \mod n$ holds, the $a$ can be cancelled to give $b \equiv c \mod n$. 
If a satisfies any (and thereby all) of the above conditions, the equation \([a][x] = [b]\) will have exactly one solution \([x]\) in \(\mathbb{Z}_n\).

Proof. We will prove that \(1 \rightarrow 2 \rightarrow 3 \rightarrow 5 \rightarrow 6 \rightarrow 7 \rightarrow 1\), which will show that every item implies every other item.

Assume item 1, that \(a, n\) are coprime. This tells us that there exist integers \(x, y\) such that \(ax + ny = 1\). Thus \(n | 1 - ax\) for some \(x\). This says that \(ax \equiv 1 \mod n\) for some \(x\), which is item 2.

Assume item 2 and take any integer \(b\). The congruence \(ax \equiv 1 \mod n\) has a solution, say \(x\). Multiply through by \(b\) to see that \(axb \equiv 1 \cdot b = b \mod n\), giving the solution \(xb\) to the congruence in item 3.

Assume item 3. Now item 4 is just a restatement of item 3 in the language of residues.

Assume item 4. In that case item 5 is just the special case of item 4 where \([b] = [1]\).

Assume item 5 and that \(b, c\) are integers such that \([a][b] = [a][c]\) in \(\mathbb{Z}_n\). Item 5 gives an integer \(x\) such that \([a][x] = [1]\) in \(\mathbb{Z}_n\). From \([a][b] = [a][c]\) we get \([x][a][b] = [x][a][c]\). Since \([x][a] = [a][x] = [1]\), we conclude that \([1][b] = [1][c]\), and then \([b] = [c]\), which is item 6.

Assume item 6 that \([a]\) can be cancelled in \(\mathbb{Z}_n\). Now, item 7 says the same thing all over again using the language of congruencies.

Finally assume item 7. To prove that \(a\) is coprime with \(n\), we suppose they are not coprime and find a congruence in which \(a\) cannot be cancelled.

With that supposition, there is an integer \(p > 1\) that divides both \(a\) and \(n\). Thus \(a = pq\) and \(n = pr\) for some integers \(q, r\). Furthermore, \(1 \leq r < n\), since \(p > 1\). This implies that

\[
r \not\equiv 0 \mod n.
\]

On the other hand, we also have

\[
ar = pqr = prq = nq \equiv 0q \equiv 0 = a0 \mod n.
\]

Evidently now, the \(a\) cannot be cancelled, for that would lead to \(r \equiv 0 \mod n\).

That proves the equivalence of all the items.

For the final statement of the result, we know that \([a][x] = [b]\) does have a solution \([x]\). But suppose there were another solution \([y]\). In that case we would have

\[
[a][x] = [b] = [a][y] \text{ in } \mathbb{Z}_n.
\]
Cancel the $[a]$ to get $[x] = [y]$. When $a$ is coprime with $n$, the linear equation $[a][x] = [b]$ has just one solution.

Just a small caution. The uniqueness of the solution $[x]$ in the equation $[a][x] = [b]$ does not mean that the integer $x$ is unique. After all the unique residue $[x]$ can be represented by any other integer congruent to $x$ modulo $n$.

**Definition 3.12.** With $n \geq 2$ as our modulus, an element $[a]$ in $\mathbb{Z}_n$ is a *unit* provided the equation $[a][x] = [1]$ has a solution $x$. The integer $a$ representing $[a]$ is then called a *unit modulo* $n$.

Units are the elements $[a]$ that fulfill any and thereby all of the seven conditions of Proposition 3.11.

The origins of the curious term ‘unit’ seem hard to track down. Possibly it’s connected to the way the term is used in Science. In Science and also in everyday matters, a unit is a standard amount of something such that every other amount is a multiple of the fixed amount. Looking at item 4 of Proposition 3.11 we see that in $\mathbb{Z}_n$ a unit is a residue $[a]$ in $\mathbb{Z}_n$ such that every other residue in $\mathbb{Z}_n$ is a multiple of $[a]$. Of course in Science, any fixed amount can play the role of a unit (except for the amount 0). But in number theory only some residues qualify for such a role. Proposition 3.11 tells the varied ways to think about units. In particular, $[a]$ is a unit if and only if $a$ is coprime with $n$.

**Units modulo a prime**

A truly special case of Proposition 3.11 commands our attention. Namely, if the modulus is a prime. Call it $p$. In that case, an integer $a$ is coprime with $p$ if and only if $p \nmid a$, for then, the only positive common factor of $p$ and $a$ is 1. The statement $p \nmid a$ is equivalent to the statement $[a] \neq [0]$ in $\mathbb{Z}_p$. Thus, Proposition 3.11 includes within it the following important fact.

**Proposition 3.13.** If $p$ is a prime and $[a] \neq [0]$ in $\mathbb{Z}_p$, then $[a]$ is a unit.

The importance of Proposition 3.13 lies in the fact that $\mathbb{Z}_p$ behaves just like the well understood sets of numbers $\mathbb{Q}, \mathbb{R}, \mathbb{C}$. Namely, $\mathbb{Z}_p$ admits addition, subtraction, multiplication AND division by everything except zero. Indeed, if $[a] \neq [0]$ in $\mathbb{Z}_p$, then $[a][x] = [b]$ always has a solution. So we can ‘divide’ $[a]$ into $[b]$. Systems that admit all four of the arithmetic operations are usually called *fields*. 


3.5 The group of units $\mathbb{Z}_n^*$

According to Proposition 3.11, $[a]$ is a unit in $\mathbb{Z}_n$ if and only if $a$ is coprime with $n$. These coincide with the residues that can be cancelled. If $[a]$ is a unit, then the equation $[a][x] = [b]$ always has a unique solution $[x]$. In particular, $[a][x] = [1]$ has only one solution, according to Proposition 3.11.

**Definition 3.14.** If $[a]$ is a unit, the unique $[x]$ in $\mathbb{Z}_n$ such that $[a][x] = [1]$ is called the inverse of $[a]$ in $\mathbb{Z}_n$.

The set of units in $\mathbb{Z}_n$ is denoted by $\mathbb{Z}_n^*$.

The set of units $\mathbb{Z}_n^*$ is only a part of the full set of residues $\mathbb{Z}_n$. We can tell whether $[a] \in \mathbb{Z}_n^*$ by simply checking if $a$ is coprime with $n$.

**Examples of units**

The ring $\mathbb{Z}_{15}$ consists of the 15 residues $[0], [1], [2], [3], \ldots, [13], [14]$. The units among them are those $[a]$ for which $a$ is coprime with 15. By a simple inspection, the set of units of $\mathbb{Z}_{15}$ is

$$\mathbb{Z}_{15}^* = \{ [1], [2], [4], [7], [8], [11], [13], [14] \}.$$

There are 8 units in $\mathbb{Z}_{15}$.

What are the units of $\mathbb{Z}_{101}$? Well, these are the residues $[a]$ among $[0], [1], [2], \ldots, [100]$ for which $a$ is coprime with 101. Since 101 is itself a prime, every integer from 1 to 100 is coprime with 101. So $\mathbb{Z}_{101}^*$ contains 100 units, namely $[1], [2], [3], \ldots, [99], [100]$. Every residue in $\mathbb{Z}_{101}$ is a unit, except for the zero residue $[0]$.

To repeat Proposition 3.13, for any prime $p$, the set $\mathbb{Z}_p^*$ of units consists of $[1], [2], [3], \ldots, [p - 1]$, omitting only the residue $[0]$. This is because every representative from 1 to $p - 1$ is coprime with the prime number $p$.

Let us now highlight the five core properties of the set of units $\mathbb{Z}_n^*$, regardless of whether or not $n$ is prime.

**Proposition 3.15.** Let $n$ be a modulus, and let $\mathbb{Z}_n^*$ be the set of units in $\mathbb{Z}_n$.

1. The residue $[1]$ is always a unit, i.e. $[1] \in \mathbb{Z}_n^*$. 
2. **The product of units is commutative.**
   That is, \([a][b] = [b][a]\) for all \([a], [b]\) in \(\mathbb{Z}_n^*\).

3. **The product of units is associative.**
   That is, \(((a)\,[b])\,[c] = [a] \,( [b] \,[c] )\) for all \([a], [b], [c]\) in \(\mathbb{Z}_n^*\).

4. **The product of units is another unit.**
   That is, if \([a], [b] \in \mathbb{Z}_n^*\), then \([a][b] \in \mathbb{Z}_n^*\).

5. **The inverse of a unit is a also a unit.**
   That is, if \([a] \in \mathbb{Z}_n^*\) and \([x] \in \mathbb{Z}_n\) is the unique residue that gives \([a][x] = [1]\), then \([x] \in \mathbb{Z}_n^*\).

**Proof.** The proofs of these items are mundane, but still they need to be put down.

1. This is obvious because \([1][1] = [1]\), making \([1]\) its own inverse.

2. This commutative law is inherited from the same law in \(\mathbb{Z}\) and holds for all residues, be they units or not. Indeed, here is the very dull explanation:

   \([a][b] = [ab] = [ba] = [b][a]\).

3. This associative law is also inherited from the same law which holds in \(\mathbb{Z}\).
   Indeed,

   \(((a)[b])[c] = [abc] = [a][bc] = [a]([b][c])\).

4. There are various ways to explain this item.
   One way is to note that since \([a], [b]\) are units, then \(a, b\) are each coprime with \(n\). This means that no prime divisor of \(n\) can appear in the unique factorizations of \(a\) and \(b\). Thus no prime divisor of \(n\) can appear in the unique factorization of \(ab\). This shows that \(ab\) remains coprime with \(n\) and thereby that \([a][b] = [ab]\) is a unit in accordance with Proposition 3.11.

   Another way is more abstract. Let \([x], [y]\) be the inverses of \([a], [b]\), respectively. Then we check in a truly pedestrian way that the inverse of \([a][b]\) is
3.5. THE GROUP OF UNITS $\mathbb{Z}_n^*$

$[xy]$. Well,

\[
([a][b]) [xy] = [abxy], \text{ by the definition of residue multiplication in } \mathbb{Z}_n,
\]

\[
= [a.xby], \text{ by the definition of residue multiplication in } \mathbb{Z}_n
\]

\[
= ([a][x])([b][y]), \text{ by the definition of residue multiplication in } \mathbb{Z}_n
\]

\[
= [1][1], \text{ by the nature of inverses}
\]

\[
= [1].
\]

5. Since $[a]$ has $[x]$ as its inverse, obviously then, $[x]$ has $[a]$ as its inverse.

The core properties of units are now verified.

Any set that enjoys the five properties of Proposition 3.15 is called an abelian group. But we’ll call it a group for brevity.

The associative law, item 3, for multiplication is very convenient because it tells us that when we multiply three or more residues, we need not worry about brackets, in which case we don’t even need to write such brackets. So, $[a][b][c]$ can be interpreted as $([a][b]) [c]$ or as $[a] ([b][c])$, and it still means the same thing.

The word “commutative” in item 2 of Proposition 3.15 has interesting origins and connections. It comes for the Latin “mutare”, meaning “to change”. Thus we get words such as “mutation” to describe a genetic change, and “immutable” to mean “unchanging”. When we have “exchangeable” feelings, or protection agreements, we call them “mutual”. When we “commute to work”, we are “interchanging” our place between home and work. And the “commuted” value of our pension is the present value taken as an “interchange” with its value as a sequence of monthly payments. For mathematics, the “commutative law” is the ability to “interchange” the order of multiplication of two elements.

Negative exponents in the group $\mathbb{Z}_n^*$

A residue $[a]$ in $\mathbb{Z}_n$ can be multiplied to itself any number of times. Thus we get the notation

\[
[a]^k = [a][a][a] \cdots [a], \text{ where the number of } [a]'s \text{ being multiplied is } k.
\]
That is,
\[ [a]^1 = [a], [a]^2 = [a][a], [a]^3 = [a][a][a], \ldots \]

By the definition of multiplication of residues, it’s also clear that \( [a]^k = [a^k] \) for all positive integers \( k \). With this understanding, the laws of exponents are just the way the are with ordinary numbers. Namely,
\[ [a]^k [a]^\ell = [a]^{k+\ell} \quad \text{and} \quad [a]^k [b]^k = ([a][b])^k. \]

In order to retain the laws of exponents, we also put \( [a]^0 = [1] \). This ensures that
\[ [a]^0 [a]^k = [1][a]^k = [a]^k = [a]^{0+k}. \]
as desired.

But what are we to make of negative exponents? For instance, what should \( [a]^{-1} \) stand for? Well, if the laws of exponents are to remain true, we had better have
\[ [a]^1 [a]^{-1} = [a]^{1-1} = [a]^0 = [1]. \]
The only way to make sense of \( [a]^{-1} \) in a way that the exponent laws prevail, is to have \( [a]^{-1} \) be the inverse of \( [a] \). But only units have inverses. This tells us that if we are to use negative exponents, then we had better stick to just the residues in the group \( \mathbb{Z}_n^* \). So, if \( [a] \in \mathbb{Z}_n^* \) and \( k \) is a positive integer, we define the negative power \( [a]^{-k} \) by the rule
\[ [a]^{-k} = [a]^{-1} [a]^{-1} [a]^{-1} \ldots [a]^{-1}, \]
where the inverse of \( [a] \) is multiplied by itself \( k \) times. In this way, as long as \( [a] \) is a unit (i.e., \( [a] \in \mathbb{Z}_n^* \)), the notation \( [a]^k \) makes sense for all integers, both positive and negative. And with this notation, the laws of exponents prevail for all integer exponents.

Although we can often avoid using negative exponents, they remain a useful instrument to keep in mind.

### 3.6 The Euler-Fermat theorems

The time has come to launch into the first surprising, and not-so-simple, result about numbers. We will obtain it by exploiting the group properties of \( \mathbb{Z}_n^* \).
For every positive integer \( n \), let \( \varphi(n) \) denote the number of units in \( \mathbb{Z}_n \). In other words, \( \varphi(n) \) equals the number of residues in \( \mathbb{Z}_n^\ast \). In practical terms, \( \varphi(n) \) is the number of integers in the list \( 0, 1, 2, 3, 4, \ldots, n - 1 \) that are coprime with \( n \). The resulting function \( \varphi \), from the set of positive integers to the set of positive integers, is known as Euler’s \( \varphi \)-function.

Here are some sample calculations of \( \varphi \).

- To get \( \varphi(18) \), list the numbers from 0 to 17 that are coprime with 18, and then count them. These are 1, 5, 7, 11, 13, 17. So \( \varphi(18) = 6 \).
- For \( \varphi(32) \), we see that 1, 3, 5, 7, 9, 11, \ldots, 31 are coprime with 32. Thus \( \varphi(32) = 16 \).
- For \( \varphi(101) \), we notice that 101 is prime and the full list 1, 2, 3, \ldots, 100 (leaving 0 out) is coprime with 101. So \( \varphi(101) = 100 \).
- For any prime \( p \), all integers in the full list 1, 2, 3, \ldots, \( p - 1 \) are coprime with \( p \).
  
  Thus \( \varphi(p) = p - 1 \) for every prime \( p \).

Here comes a major result, known as Euler’s Theorem.

To avoid notational clutter in the proof, let’s denote the residues \([a]\) of \( \mathbb{Z}_n^\ast \) by simple letters like \( x, y, u, v, \ldots \), without the box notation. We also denote the special residue \([1]\) by the symbol 1.

**Theorem 3.16 (Euler).** If \( v \in \mathbb{Z}_n^\ast \), then \( v^{\varphi(n)} = 1 \).

**Proof.** Just a reminder that the above 1 is a residue, and not an integer.

Let \( k = \varphi(n) \), just to write less. List the residues of \( \mathbb{Z}_n^\ast \) as

\[ u_1, u_2, u_3, \ldots, u_k, \]

without repetitions in the list. This list contains the residue \( v \).

Form the new list

\[ v \cdot u_1, v \cdot u_2, v \cdot u_3, \ldots, v \cdot u_k. \]

Since \( \mathbb{Z}_n^\ast \) is a group, this list of residues is also in \( \mathbb{Z}_n^\ast \). The latter list never repeats itself either. Indeed, if we had \( v \cdot u_j = v \cdot u_i \) for some \( u_j, u_i \) in the original list,
we could cancel the unit \( v \) to get \( u_j = u_i \). Hence, the second list is a permutation of the original list.

It follows that the product of the residues in the first list equals the product of the residues in the second list. After all, the two lists contain the same residues, only written in a different order. Thus we obtain

\[
 u_1 \cdot u_2 \cdots u_k = (v \cdot u_1) \cdot (v \cdot u_2) \cdots (v \cdot u_k).
\]

By rearranging the order of multiplication and throwing in a harmless 1, this equation becomes

\[
 1 \cdot u_1 \cdot u_2 \cdots u_k = v^k \cdot u_1 \cdot u_2 \cdots u_k.
\]

For this we used the commutative law in the group \( \mathbb{Z}_n^\ast \), and, by not having any brackets, the associative law too.

Now cancel the unit \( u_1 \cdot u_2 \cdots u_k \), to conclude that \( 1 = v^k \), as desired. \( \square \)

Getting back to box notation for residues, we have just seen that

\[
 [a]^{\varphi(n)} = [1].
\]

for every residue \([a] \in \mathbb{Z}_n^\ast\). Keeping in mind how we multiply residues along with the fact \([a] \in \mathbb{Z}_n^\ast \) whenever \( a \) is coprime with \( n \), the above says that

\[
 [a^{\varphi(n)}] = [1],
\]

whenever \( a \) is coprime with \( n \).

Turning this into the language of congruences we come to the following, fairly stunning result.

**Theorem 3.17** (Euler, Alternate Version). If \( a \in \mathbb{Z} \) and \( a \) is coprime with \( n \), then

\[
 a^{\varphi(n)} \equiv 1 \ mod \ n.
\]

In other words, if \( a \) is coprime with a modulus \( n \), then

\[
 \frac{a^{\varphi(n)} - 1}{n} \]

is an integer.

For example, 2011 is a prime and so \( \varphi(2011) = 2010 \). Also 793 is coprime with 2011. Thus we know that the monstrously large calculation

\[
 \frac{793^{2010} - 1}{2011}
\]

will yield an integer. Our calculator will never tell us that by a brute force calculation.
**Euler’s Theorem when the modulus is prime**

When the modulus is a prime, say \( p \), we know that \( \varphi(p) = p - 1 \). Thus Euler’s theorem specializes to another famous result attributed to Fermat.

**Theorem 3.18** (Fermat’s Little Theorem). *If \( p \) is prime and \( p \nmid a \), then*

\[
a^{p-1} \equiv 1 \mod p.
\]

We shall return to this famous result quite regularly.

A small variant of Fermat’s theorem often comes in handy.

**Corollary 3.19** (Fermat variant). *If \( p \) is prime and \( a \) is any integer, then*

\[
a^p \equiv a \mod p.
\]

**Proof.** If \( p \nmid a \), then \( a^{p-1} \equiv 1 \mod p \) by Fermat’s theorem. Multiply through by \( a \) to get

\[
a^p = a \cdot a^{p-1} \equiv a \cdot 1 = a \mod p.
\]

On the other hand, if \( p \mid a \), then \( a \equiv 0 \mod p \), in which case it’s obvious that

\[
a^p \equiv 0^p = 0 \equiv a \mod p.
\]

So the result holds for every integer \( a \). \( \square \)

In anticipation of something to come, let’s note that the variant of Fermat’s theorem affords a neat way to prove that an integer \( n \) is NOT prime. Just fish for an integer \( a \) such that \( a^n \not\equiv a \mod n \). But, we’ll come back to this later.

**A useful shortcut for working with high powers**

Next comes a corollary of Euler’s Theorem, which helps us decide when two high powers of an integer are congruent.

**Proposition 3.20.** *If \( n \) is a modulus and \( a \) is coprime with \( n \), and \( k, \ell \) are non-negative integers such that*

\[
k \equiv \ell \mod \varphi(n),
\]

*then*

\[
a^k \equiv a^\ell \mod n,
\]

*for every \( a \) that is coprime with \( n \).*
CHAPTER 3. MODULAR ARITHMETIC

Proof. Say $k \leq \ell$. We are given that $\ell = q\varphi(n) + k$ for some $q \geq 0$. Then

$$a^\ell = a^{q\varphi(n)+k} = (a^{\varphi(n)})^q a^k \equiv 1^q a^k = a^k \mod n.$$  

The congruence in the line above is true by Euler’s Theorem along with the replacement principle.

We caution that Proposition 3.20 does not hold when $a$ is not coprime with $n$. For example, take $n = 9 = 3^2$ and $a = 3$. Here, $\varphi(9) = 6$. Clearly $7 \equiv 1 \mod 6$, and yet $3^7 \equiv 0 \not\equiv 3^1 \mod 9$.

Illustrations of Euler-Fermat

Let’s work out a couple of small examples that exploit the Euler-Fermat theorems. Replacement will be used persistently.

Example 3.21. Let us find all $n \geq 0$ such that $5$ divides $3^n + 1$.

We first notice that $\varphi(5) = 4$. After that, we see from Proposition 3.20 that,

- if $n \equiv 0 \mod 4$, then $3^n + 1 \equiv 3^0 + 1 = 2 \mod 5$,
- if $n \equiv 1 \mod 4$, then $3^n + 1 \equiv 3^1 + 1 = 4 \mod 5$,
- if $n \equiv 2 \mod 4$, then $3^n + 1 \equiv 3^2 + 1 \equiv 0 \mod 5$, and
- if $n \equiv 3 \mod 4$, then $3^n + 1 \equiv 3^3 + 1 = 28 \equiv 3 \mod 5$.

From the above chart it is clear that those $n \equiv 2 \mod 4$ satisfy $5|3^n + 1$.

Recall that the expression “reducing a number modulo $n$ ” means finding the remainder of the number upon division by $n$.

Example 3.22. Let’s reduce $8^{2^{100}}$ modulo 11. Since $2^{100}$ is already very big, $8^{2^{100}}$ is truly huge, and a brute force reduction is out of the question.

Here we notice that $\varphi(11) = 10$. So it makes sense to first reduce $2^{100}$ modulo 10 and then use Proposition 3.20, which applies because 8 and 11 are coprime. Well, modulo 10:

$$2^{100} = (2^5)^{20} = 32^{20} \equiv 2^{20} = (2^5)^4 = 32^4 \equiv 2^4 \equiv 6.$$  

Notice how we replaced 32 by 2 modulo 10. Then by Proposition 3.20 we see that $8^{2^{100}} \equiv 8^6 \mod 11$. And thus, modulo 11:

$$8^{2^{100}} \equiv 8^6 \equiv (-2)^3 = -8 \equiv 3.$$  

Since 3 is indeed between 0 and 10, our reduction of $8^{2^{100}}$ modulo 11 is done.
3.7 The square and multiply algorithm

Let us remind ourselves that when we find the remainder of a number $b$ upon division by some positive integer $n$ we are reducing $b$ modulo $n$.

We have seen from Proposition 3.20 that, as long as $a$ is coprime with $n$, the reduction of $a^k$ modulo some $n$ can be expedited by reducing $k$ modulo $\varphi(n)$ first. Namely, if $k$ reduces to $\ell$ modulo $\varphi(n)$, then $a^k \equiv a^\ell \mod n$. But we are left with the problem of reducing $a^\ell$. When $n$ is rather large, this still leaves us with a bit of work to do. An efficient method would be desirable.

The so-called square and multiply algorithm is widely recognized as being the best for the job. It works reasonably fast with a calculator when $n$ is of moderate size, and exceptionally fast with a computer even when $n$ is of enormous size.

Suppose that $n$ is a modulus, $a$ is any integer and $k$ is some exponent. We want to reduce $a^k$ modulo $n$. That is, we want the remainder when $n$ is divided into $a^k$.

Here is the square and multiply algorithm that does it efficiently.

- First, write $k$ in its binary expansion.

  $$k = k_0 + k_1 \cdot 2 + k_2 \cdot 2^2 + \cdots + k_\ell \cdot 2^\ell,$$

  where $k_j = 0$ or $1$ and $k_\ell = 1$

  This can be done often by inspection, but if necessary by repeated division by 2, as follows.

  $$
  \begin{align*}
  k &= 2q_0 + k_0 & k_0 &= 0, 1 \text{ and } q_0 < k \\
  q_0 &= 2q_1 + k_1 & k_1 &= 0, 1 \text{ and } q_1 < q_0 \\
  q_1 &= 2q_2 + k_2 & k_2 &= 0, 1 \text{ and } q_2 < q_1 \\
  q_2 &= 2q_3 + k_3 & k_3 &= 0, 1 \text{ and } q_3 < q_2 \\
  &\vdots
  \end{align*}
  $$

  Keep going until some $q_\ell = 0$ to end up with

  $$q_{\ell-1} = 2q_\ell + k_\ell \quad k_\ell = 1, \text{ and } 0 = q_\ell < q_{\ell-1} \text{ while } 0 \neq q_{\ell-1} = k_\ell.$$
Then, by repeated substitution, we get
\[ k = 2q_0 + k_0 \]
\[ = 2(2q_1 + k_1) + k_0 \]
\[ = 2^2q_1 + 2k_1 + k_0 \]
\[ = 2^2(2q_2 + k_2) + 2k_1 + k_0 \]
\[ = 2^3q_2 + 2^2k_2 + 2k_1 + k_0 \]
\[ \vdots \]
\[ = 2^\ell k_\ell + \cdots + 2^2k_2 + 2k_1 + k_0 \]

- Then observe that
\[ a^k = a^{k_0 + 2k_1 + 2^2k_2 + \cdots + 2^\ell k_\ell} = a^{k_0} \cdot (a^2)^{k_1} \cdot (a^2)^{k_2} \cdots (a^2)^{k_\ell}, \text{ where } k_j = 0, 1. \]

- Next obtain the powers \( a^{2^j} \) by repeated squaring and reduction modulo \( n \).

\[
\begin{align*}
a & \equiv b_0 & & \text{where } 0 \leq b_0 < n \\
a^2 & \equiv b_0^2 \equiv b_1 & & \text{where } 0 \leq b_1 < n \\
a^{2^2} & \equiv b_1^2 \equiv b_2 & & \text{where } 0 \leq b_2 < n \\
a^{2^3} & \equiv b_2^2 \equiv b_3 & & \text{where } 0 \leq b_3 < n \\
\vdots & & & \\
a^{2^\ell} & \equiv b_\ell^2 \equiv b_\ell & & \text{where } 0 \leq b_\ell < n
\end{align*}
\]

Notice that in the above squarings and reductions we never encounter a number bigger than \( n^2 \). Here the number of squarings modulo \( n \) was exactly \( \ell \).

- Finally multiply the \( b_j^{k_j} \), which are congruent to \( \left(a^{2^j}\right)^{k_j} \), and reduce the products modulo \( n \) as indicated next.

\[
\begin{align*}
b_0^{k_0}b_1^{k_1} & \equiv c_1 & & \text{where } 0 \leq c_1 < n \\
b_0^{k_0}b_1^{k_1}b_2^{k_2} & \equiv c_1b_2^{k_2} \equiv c_2 & & \text{where } 0 \leq c_2 < n \\
b_0^{k_0}b_1^{k_1}b_2^{k_2}b_3^{k_3} & \equiv c_2b_3^{k_3} \equiv c_3 & & \text{where } 0 \leq c_3 < n \\
\vdots & & & 
\end{align*}
\]

This final process also called for \( \ell \) multiplications and reductions modulo \( n \).
3.7. THE SQUARE AND MULTIPLY ALGORITHM

We managed to reduce $a^k \mod n$ using $\ell$ squarings and $\ell$ multiplications modulo $n$. This gave us $2\ell$ fairly simple steps in all. Each step involves only a multiplication of two integers from 0 to $n - 1$ plus a reduction modulo $n$ of a number no bigger than $n^2$.

For instance, if $k$ has 1000 binary digits, then $k \geq 2^{1000}$, which is huge beyond any possible imagination. Yet a mere 2000 simple multiplications and reductions modulo $n$ will yield the remainder of $a^k \mod n$. On a properly programmed laptop computer the reduction

$$7^{10^{200000}} \equiv 787 \mod 853$$

can be done in under one second.

**Example 3.23.** Rather than remember the messy notations of the square and multiply algorithm, it is probably better to just work out an example with some more modest numbers. Let’s reduce

$$7^{327} \mod 853$$

with a pencil and a cheap calculator by the square and multiply method.

First get the binary expansion of 327:

$$327 = 1 + 2 + 4 + 64 + 256 = 1 + 2^2 + 2^6 + 2^8.$$  

Thus

$$7^{327} \equiv 7 \cdot 7^2 \cdot 7^4 \cdot 7^{64} \cdot 7^{256}. $$

Now, start with 7 and keep squaring and reducing modulo 853 to get:

$$7 \equiv 7$$
$$7^2 \equiv 49$$
$$7^4 \equiv 49^2 \equiv 695$$
$$7^8 \equiv 695^2 \equiv 227$$
$$7^{16} \equiv 227^2 \equiv 349$$
$$7^{32} \equiv 349^2 \equiv 675$$
$$7^{64} \equiv 675^2 \equiv 123$$
$$7^{128} \equiv 123^2 \equiv 628$$
$$7^{256} \equiv 628^2 \equiv 298$$
Then, multiply the reductions of $7, 7^2, 7^4, 7^{64}, 7^{256}$ modulo 853 to obtain:

$$7^{327} \equiv 7 \cdot 49 \cdot 695 \cdot 123 \cdot 298 \equiv 343 \cdot 695 \cdot 123 \cdot 298 \equiv 398 \cdot 123 \cdot 298 \equiv 333 \cdot 298 \equiv 286.$$ 

The above took a bit of effort, but it would have been overwhelming to actually calculate $7^{327}$ and then divide by 853 to come up with the remainder.

**Squaring and multiplying with the help of technology**

Here is a suggestion for implementing the square and multiply algorithm using Microsoft Excel.

The main functions to use are:

- INT($b/a$), which gives the quotient when $a$ is divided into $b$, and
- MOD($b,n$), which gives the remainder when $n$ is divided into $b$, in other words, the reduction of $b$ modulo $n$.

To warm up, try reducing $735^{9824}$ mod 487 in any cell by using the command =MOD(735^9824,487). A warning comes up, indicating that Excel cannot handle such a monster as $735^{9824}$. But with a proper implementation of the square and multiply algorithm, the reduction is very efficient.

Here is a possible way (among many, and likely not even the best) to program the algorithm in Excel. There are a number of steps, but to someone familiar with Excel they should all be easy and natural.

- We are given positive integers $k, n, a$, and we need to reduce $a^k$ modulo $n$.
- First work on getting the binary expansion of $k$.
  In cell A1 type in the value of the exponent $k$.
- In cell A2 type in: =INT(A1/2).
  This gives the quotient $q_0$ when 2 is divided into $k$, and it prepares us to get all of the other quotients down column A.
3.7. THE SQUARE AND MULTIPLY ALGORITHM

- In cell B2 type in: =MOD(A1,2).
  This puts the remainder $k_0$ into cell B2, when 2 is divided into $k$. It will either be a 0 or a 1.

- Select cells A2 and B2 and drag them down until the first zero appears in column A. Starting with cell A2, the strictly decreasing quotients $q_0, q_1, q_2, \ldots, q_\ell = 0$ appear in column A. In column B, starting with cell B2, the remainders $k_0, k_1, k_2, \ldots, k_\ell$ appear. These will either 0 or 1, but the last remainder next to the first zero quotient $q_\ell$ will be a 1.

- In cell C1 type in the given value of $n$.

- In cell C2 type in: =C1. This just copies $n$ inside cell C2.

- Select cell C2, which just repeats the modulus $n$ in cell C1, and drag cell C2 down along column C until the row where the first zero quotient appears in column A. This has the effect of copying $n$ all the way down column C.

- In cell D1 type in the value of $a$.

- In cell E1 type in: =MOD(D1,C1).
  This reduces $a$ modulo $n$, in case $a$ was bigger than $n$, which ensures that the size of the numbers does not get out of hand.

- In cell E2 type in: =E1.
  This just copies the reduction $b_0$ of $a$ modulo $n$ into row 2 so it can be nicely lined up with the remainder $k_0$ in cell B2. This is just a bit of housekeeping.

- In cell E3 type in: =MOD(E2*E2, C2).
  This squares the reduction of $a$ and reduces that modulo $n$. In other words, the integer $b_1$ appears in cell E3, which is in the same row as the remainder $k_1$ over in cell B3.

- Select cell E3 and drag it down until the last row where the zero quotient $q_\ell$ and remainder $k_\ell = 1$ appears.
  In column E, starting with E2, the repeated squares
  $$a^1, a^2, a^4, a^8, a^{16}, \ldots,$$
  reduced modulo $n$, will appear. These are the numbers $b_0, b_1, b_2, \ldots, b_\ell$ in the general discussion of the algorithm.
• Now it’s time to pick out which of these $b_j$ actually need to be multiplied and reduced modulo $n$.

• In cell F2 type in: =E2** B2.
  This will keep $b_0$ if a 1 appears in B2, or turn it into a 1 if a 0 appears in B2.

• Select cell F2 and drag it down to the last row with that zero quotient in column A. The numbers in column F, starting with F2, will either repeat those in column E or be equal to 1. These numbers in column F need to be multiplied and reduced modulo $n$.

• In cell G2 type in: =F2.
  This copies the number $b_{k_0}$ into cell G2.

• In cell G3 type in: =MOD(F3*G2,C2).
  This multiplies and reduces the product $b_{k_0} \cdot b_{k_1}$.

• Select cell G3 and drag it down as far as the last row with the zero quotient in column A. This has the effect of multiplying all of the $b_{k_j}$ and reducing that product modulo $n$.

• In the bottom corner of the spreadsheet in column G, the reduction of $a^k \mod n$ will appear.

Save the spreadsheet! By simply changing the inputs for $k$, $n$, $a$ in cells A1, C1 and D1, respectively, the spreadsheet just designed will automatically calculate the reduction of $a^k \mod n$ in that bottom right corner. It might be necessary to drag the last row a bit farther down until the appropriate zero quotient appears in column A. This spreadsheet will certainly come in handy for the non-primality testing which is coming up next.

3.8 Using Fermat to test for non-primality

The variant of Fermat’s Theorem 3.19 offers a lovely application. If we are given a large integer $n$ and if we can find some integer $a$ such that

$$a^n \not\equiv a \mod n,$$

then we can be sure $n$ is not a prime. Thus we have a practical, sufficient condition for an integer $n$ to be composite.
3.8. **USING FERMAT TO TEST FOR NON-PRIMALITY**

Any integer \( a \) such that \( a^n \not\equiv a \mod n \), is known as a *witness* for the non-primality of \( n \). A good question to be considered in due course is whether every non-prime has a witness. It seems that most non-primes do have witnesses.

**Example 3.24.** Take \( n = 8633 \). It’s a bit of a chore to try factoring \( n \). If \( n \) were much bigger, it becomes a truly severe problem. However, using \( a = 2 \) as a tester, let’s reduce \( 2^{8633} \mod 8633 \). If the reduction is not equal to 2, we may conclude that 8633 is not prime. We can do the reduction by our reliable (if somewhat tedious) square and multiply algorithm. The binary expansion of 8633 is

\[
8633 = 8192 + 256 + 128 + 32 + 16 + 8 + 1 = 2^{13} + 2^8 + 2^7 + 2^5 + 2^4 + 2^3 + 1.
\]

Thus

\[
2^{8633} = 2^{8192} \cdot 2^{256} \cdot 2^{128} \cdot 2^{32} \cdot 2^{16} \cdot 2^8 \cdot 2.
\]

Now square and reduce modulo 8633, with the help of a calculator, to obtain:

\[
\begin{align*}
2 &\equiv 2 \\
2^2 &\equiv 4 \\
2^4 &\equiv 4^2 \equiv 16 \\
2^8 &\equiv 16^2 \equiv 256 \\
2^{16} &\equiv 256^2 \equiv 5105 \\
2^{32} &\equiv 5105^2 \equiv 6631 \\
2^{64} &\equiv 6631^2 \equiv 2292 \\
2^{128} &\equiv 2292^2 \equiv 4400 \\
2^{256} &\equiv 4400^2 \equiv 4814 \\
2^{512} &\equiv 4814^2 \equiv 3624 \\
2^{1024} &\equiv 3624^2 \equiv 2583 \\
2^{2048} &\equiv 2583^2 \equiv 7213 \\
2^{4096} &\equiv 7213^2 \equiv 4911 \\
2^{8192} &\equiv 4911^2 \equiv 5952.
\end{align*}
\]

Then, modulo 8633,

\[
2^{8633} \equiv 5952 \cdot 4814 \cdot 4400 \cdot 6631 \cdot 5105 \cdot 256 \cdot 2 \equiv 5407 \neq 2.
\]
Since $2^{8633} \not\equiv 2 \mod 8633$, the integer 2 is a witness to the fact 8633 is not prime. Admittedly, we have no clue as to its factors.

Somebody might complain that it is probably easier just to try and factor 8633 directly, and thereby avoid the tedious squaring operations, multiplications and reductions. For this particular number 8633, it’s debatable which way is faster. But when the integer $n$ gets significantly bigger, the square and multiply method, assisted by a decent program in our computer, wins hands down.

Testing for non-primality with technology (the better way)

Using the saved Excel spreadsheet for the square and multiply algorithm, the reduction of $a^n$ modulo $n$ can be done in a flash, making our test for non-primality effortless. Given a large test integer $n$, pick a convenient small base $a$ such as one of 2, 3, 5, 6 or 7. Then reduce $a^n$ modulo $n$ by typing in:

- the value of $n$ into cells A1 and C1
- the value of the chosen base $a$ into cell D1

If the reduction of $a^n$ modulo $n$ comes back different than the chosen $a$, the test integer $n$ is surely not prime. If the reduction of $a^n$ modulo $n$ comes back equal to $a$, try a different base. Should $a^n$ keep reducing to $a$ for several different choices of $a$, the integer $n$ is very probably prime, although there is no guarantee.

The possible failure of the non-primality test

The test for non-primality based on Fermat can on occasion disappoint. There do exist composite numbers $n$ such that $2^n \equiv 2 \mod n$. For example, the integer $341 = 11 \cdot 31$ is certainly not prime. However, let’s calculate the reduction of $2^{341}$ modulo 341. We have

$$2^{341} = 2^{256} \cdot 2^{64} \cdot 2^{16} \cdot 2^{4} \cdot 2^1.$$
3.9 Using Euler to find \( k \)'th roots modulo \( n \)

By repeated squaring and reduction modulo 341 we obtain

\[
\begin{align*}
2 & \equiv 2 \\
2^2 & \equiv 4 \\
2^4 & \equiv 16 \\
2^8 & \equiv 256 \\
2^{16} & \equiv 256^2 \equiv 64 \\
2^{32} & \equiv 64^2 \equiv 4 \\
2^{64} & \equiv 4^2 = 16 \\
2^{128} & \equiv 16^2 = 256 \\
2^{256} & \equiv 256^2 \equiv 64.
\end{align*}
\]

Then

\[
2^{341} \equiv 64 \cdot 16 \cdot 64 \cdot 16 \cdot 2 \equiv 2 \mod 341.
\]

Of course, the suggested square and multiply program using Excel, could carry out the above calculation in an instant.

Integers \( n \) that are not prime, but act as if the were by giving \( 2^n \equiv 2 \mod n \), are sometimes known as pseudo-primes for the base 2. There could be pseudo-primes for other bases as well. Fortunately the chances of hitting a pseudo-prime are low, but we would need to go very far afield to pursue such matters. In Chapter 5 we will examine this test for non-primality more deeply.

3.9 Using Euler to find \( k \)'th roots modulo \( n \)

Having looked at ways to reduce the \( k \)'th power of an integer \( a \) modulo \( n \), it might be worthwhile to turn this around and discuss the extraction of \( k \)'th roots modulo \( n \). We now discuss a mechanism by which at least some congruences of the form

\[
x^k \equiv b \mod n
\]

can be solved for \( x \), given \( n, k \) and \( b \).

The result that follows tells what to do at least for some choices of \( b \) and \( k \). It hinges on Euler’s \( \varphi \)-function.
Proposition 3.25. Let $n$ be a modulus, and let $b, k$ be integers such that $b$ is coprime with $n$ while $k$ is coprime with $\varphi(n)$. If $\ell$ is any integer solution to

$$k\ell \equiv 1 \mod \varphi(n),$$

then the one and only solution to

$$x^k \equiv b \mod n$$

is given by

$$x \equiv b^\ell \mod n.$$

Proof. By way of analogy, we know, for positive real numbers, that the solution of $x^k = b$ is $x = b^{1/k}$. With our congruence problem, the $\ell$ plays the role of the $1/k$. But this is merely an analogy, so let’s do a proper proof.

We should notice, by the way, that since $k$ and $\varphi(n)$ are coprime, the congruence $k\ell \equiv 1 \mod \varphi(n)$ does have a unique solution $\ell$ due to the key Proposition 3.11. In other words, the residue $[k]$ is a unit in $\mathbb{Z}_{\varphi(n)}$, and $[\ell]$ is its inverse.

Since $b$ is coprime with $n$ and $k\ell \equiv 1 \mod \varphi(n)$, Proposition 3.20 tells us that $b^{k\ell} \equiv b^1 = b \mod n$. Therefore,

$$\left(b^\ell\right)^k = b^{k\ell} \equiv b \mod n.$$

So, $b^\ell$ is indeed a solution of $x^k \equiv b \mod n$.

To see that $b^\ell$ is the only solution, suppose $x$ is any solution of $x^k \equiv b \mod n$. Since $b$ is coprime with $n$, so is $x$ coprime with $n$. Indeed, if some integer $p > 1$ divided $x$ and $n$, then $p$ would also divide $b$. This is true because

$$p \mid n, n \mid b - x^k \text{ and } b = (b - x^k) + x^k.$$

Thus, Proposition 3.20 kicks in again to tell us that

$$x = x^1 \equiv x^{k\ell} = \left(x^k\right)^\ell \equiv b^\ell \mod n.$$

Our result is shown. \qed

We should caution that Proposition 3.25 is only valid if $b$ is coprime with $n$, and if $k$ is coprime with $\varphi(n)$. Otherwise, it’s possible for some $b$ to not have a $k$’th root modulo $n$ or maybe several $k$’th roots that are not congruent modulo $n$. 
For a simple example, take
\[ x^2 \equiv 3 \mod 4. \]
There is no \( x \) that works here, since the possible remainders 0, 1, 2, 3 for \( x \), when squared, give
\[ 0^2 = 0, 1^2 = 1, 2^2 \equiv 0, 3^2 \equiv 1 \mod 4, \]
but never give 3. What went wrong is that the exponent 2 is not coprime with \( \varphi(4) \), which also happens to be 2.

For another baby example, take
\[ x^3 \equiv 0 \mod 4. \]
Here, by inspection we find solutions \( x = 0 \) and \( x = 2 \) are solutions, but \( 0 \not\equiv 2 \mod 4 \). Here 0 is not coprime with 4.

The general problem of finding \( k \)’th roots modulo \( n \) is more subtle than it might first appear. For now, here is a worked example to illustrate Proposition 3.25.

**Example 3.26.** Let’s solve
\[ x^{113} \equiv 347 \mod 463. \]
We could test all possibilities for \( x \) from 0 to 462, and see which ones work, but that would be slightly insane. Instead, we notice after a bit of checking (for instance, from a table) that 463 is prime, and thus 463 is coprime with 347. Also \( \varphi(463) = 462 \), and with a bit of work using the Euclidean Algorithm (or maybe a saved Excel spreadsheet), we see that \( \gcd(113, 462) = 1 \). The conditions of Proposition 3.25 apply.

According to Proposition 3.25, the solution of this congruence is \( x \equiv 347^\ell \), where \( \ell \) is the solution of the linear congruence \( 113\ell \equiv 1 \mod 462 \). To solve this linear congruence, we need to solve the Diophantine equation
\[ 113\ell - 462j = 1. \]
This is solved using the Euclidean Algorithm as discussed in Chapter 1, (either by hand or with some software). With the calculations omitted, the solution is \( \ell \equiv 323 \). We can check that this \( \ell \) is correct by dividing \( 113 \cdot 323 = 36499 \) by 462, and discovering that the remainder is 1.
The solution to the original congruence is \( x \equiv 347^{323} \mod 463 \). Now, it’s only polite to reduce this answer modulo 463, which we can do by the square and multiply algorithm. Let’s carry out that calculation, for the sake of practice (or, even better, with an appropriately saved software program).

First notice that

\[
323 = 256 + 64 + 2 + 1 = 2^8 + 2^6 + 2^1 + 2^0.
\]

Thus

\[
347^{323} = 347^{256} \cdot 347^{64} \cdot 347^2 \cdot 347.
\]

Next start with 347, repeatedly square it, and reduce modulo 463:

\[
\begin{align*}
347 & \equiv 347 \\
347^2 & \equiv 120409 \equiv 29 \\
347^4 & \equiv 29^2 \equiv 378 \\
347^8 & \equiv 378^2 \equiv 280 \\
347^{16} & \equiv 280^2 \equiv 153 \\
347^{32} & \equiv 153^2 \equiv 259 \\
347^{64} & \equiv 259^2 \equiv 409 \\
347^{128} & \equiv 409^2 \equiv 138 \\
347^{256} & \equiv 138^2 \equiv 61.
\end{align*}
\]

Then reduce the product \( 347^{256} \cdot 347^{64} \cdot 347^2 \cdot 347 \mod 463 \) to get:

\[
347^{256} \cdot 347^{64} \cdot 347^2 \cdot 347 \equiv 61 \cdot 409 \cdot 29 \cdot 347 \equiv 37.
\]

The solution to our congruence is

\[
x \equiv 37 \mod 463.
\]

To make sure no mistake happened, one could square and multiply to confirm that \( 37^{113} \mod 463 \) does reduce to 347. Rest assured that this verification has been done.

As we can see from the above example, some computing is needed to get answers in modular arithmetic once the numbers start to get big. The only sane thing is to use software such as that saved Excel spreadsheet (or some other suitably designed program) for reducing high powers \( a^k \mod n \).

We should keep in mind that, before using Proposition 3.25 to solve \( x^k \equiv b \mod n \), we must first confirm the coprimeness of \( b \) with \( n \) and of \( k \) with \( \varphi(n) \).
3.10. Exercises

The non-issue of negative exponents and \( k^{th} \) roots

Let’s look back at Example 3.26.

Having seen that \( \gcd(113, 462) = 1 \), we solved the congruence \( 113\ell \equiv 1 \mod 462 \) and obtained a solution \( \ell = 323 \). Then we reduced \( 347^{323} \mod 463 \) to obtain our answer \( x \equiv 37 \mod 463 \), in accordance with Proposition 3.25.

However, any \( \ell \equiv 323 \mod 462 \) is a solution of the preceding congruence. For instance, the negative number \( \ell = 323 - 462 = -139 \) is a perfectly good solution to our congruence \( 113\ell \equiv 1 \mod 462 \). According to Proposition 3.25, the unique solution to \( x^{113} \equiv 347 \mod 463 \) is \( x \equiv 347^{-139} \mod 463 \). This raises the issue of what to do with that negative exponent. One way to make it go away is to use positive \( \ell \)'s such as our originally chosen 323. While this avoids the issue, it does not explain it.

The explanation comes from thinking inside the group \( \mathbb{Z}_{463}^* \) of units modulo 463. The original Example 3.26 written in the language of residues becomes

\[
[x]^{113} = [347] \text{ in } \mathbb{Z}_{463}.
\]

Proposition 3.25, properly interpreted in terms of residues, says that \( [x] = [347]^{-139} \) calculated in \( \mathbb{Z}_{463} \) gives the solution. But now we can interpret what \( [347]^{-139} \) is supposed to mean. Namely, this is the same as \( ([347]^{-1})^{139} \) where \( [347]^{-1} \) is the inverse of \( [347] \) inside the group \( \mathbb{Z}_{463}^* \).

Using the Euclidean algorithm to solve the Diophantine equation \( 347t + 463u = 1 \), we discover the solution \( t = 459 \), and thereby that \( [t] = [459] \) is the inverse of \( [347] \) in \( \mathbb{Z}_{463}^* \). So the solution of our original problem in terms of residues is

\[
[459]^{139} \text{ calculated in } \mathbb{Z}_{463}^*.
\]

When \( 459^{139} \) is reduced modulo 463 (using square and multiply), we get the answer \( x \equiv 37 \mod 463 \), exactly as we got in the original solution to Example 3.26.

3.10 Exercises

1. Show that a number of the from \( x^2 + 1 \) is never divisible by 19.

2. If \( a, b, c \) are integers and \( a^2 + b^2 = c^2 \), show that \( a, b \) cannot both be odd.
   
   Hint. Use arithmetic modulo 4.
3. If $p$ is an odd prime and $x^2 \equiv 1 \pmod{p}$, show that $x \equiv \pm 1 \pmod{p}$. Then prove this holds if we replace $p$ by $p^2$. Then show it if $p$ is replaced by $p^3$.

4. List the units of $\mathbb{Z}_{13}$, and the units of $\mathbb{Z}_{40}$.

5. Use Fermat’s Theorem to reduce $2^{17} \pmod{13}$ modulo 13.

6. Reduce $25^{782} \pmod{901}$, by using the square and multiply method.

7. Use the square and multiply algorithm to find the last three decimal digits of $107^{107}$.

8. Find the last three digits of $1033^{2009}$. In other words, reduce this integer mod 1000.

9. Show that $10981$ is not prime, by using Fermat’s theorem. You do not need to factor this number, but you will need your calculator or other computing device to square and multiply modulo $10981$.

10. Solve the congruence $x^{53} \equiv 17 \pmod{101}$.

11. If $p, q$ are distinct primes and $a \equiv b \pmod{p}$ and $a \equiv b \pmod{q}$, show that $a \equiv b \pmod{pq}$.

12. Prove that every integer is congruent modulo 9 to the sum of its digits. Then deduce that an integer is divisible by 9, if and only if the sum of its digits is divisible by 9.

For your amusement ...

You are now a magician.

· Tell your audience to pick out an integer $a$ with a good number of digits.
· Tell them to figure out the sum of the digits $b$.
· Tell them to calculate $c = a - b$.
· Tell them to secretly delete one of the non-zero digits in $c$.
· Tell them to reveal the remaining digits of $c$.
· If you are good magician, you instantly tell them the digit they deleted.
· Are you a good magician?
13. Solve
\[ x^{600} + 29x^{543} - 19x^{482} + 199x^{301} + 82x^{182} - 75x^{121} + 34x^{63} - 60 \equiv 0 \mod 61. \]

Hint. The modulus 61 is prime, and Fermat’s Theorem can help lower those high powers. It should come down to a cubic, and after that do some trial and error searching. The MOD command in an Excel spreadsheet might help speed things up.

14. Let’s find out a bit about integers of the form \( 2^n + 1 \), where \( n = 1, 2, 3, \ldots \).

Here is a short list of such numbers.

<table>
<thead>
<tr>
<th>( n )</th>
<th>( 2^n + 1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>17</td>
</tr>
<tr>
<td>5</td>
<td>33</td>
</tr>
<tr>
<td>6</td>
<td>65</td>
</tr>
<tr>
<td>7</td>
<td>129</td>
</tr>
<tr>
<td>8</td>
<td>257</td>
</tr>
<tr>
<td>9</td>
<td>513</td>
</tr>
<tr>
<td>10</td>
<td>1025</td>
</tr>
<tr>
<td>11</td>
<td>2049</td>
</tr>
<tr>
<td>12</td>
<td>4097</td>
</tr>
<tr>
<td>13</td>
<td>8193</td>
</tr>
<tr>
<td>14</td>
<td>16385</td>
</tr>
<tr>
<td>15</td>
<td>32769</td>
</tr>
<tr>
<td>16</td>
<td>65537</td>
</tr>
<tr>
<td>17</td>
<td>131073</td>
</tr>
<tr>
<td>18</td>
<td>262145</td>
</tr>
</tbody>
</table>

(a) After looking at them closely, you might notice that when the exponent \( n \) is not a power of 2, then \( 2^n + 1 \) is not a prime.

For example, \( 2^{17} + 1 = 131073 \) is divisible by 3, and \( 2^{14} + 1 = 16385 \) is divisible by 5.

• Prove that this observation holds in general.

Hint. Here is a little factoring trick that works for odd powers \( m \):

\[ x^m + 1 = (x + 1)(x^{m-1} - x^{m-2} + x^{m-3} - \cdots + x^2 - x + 1) \]

When \( m \) is even this does not work.

Now if \( n \) is not a power of 2, write \( n = 2^k m \), where \( m \) is an odd number and \( m > 1 \) and \( k \) is the multiplicity of 2 inside \( n \).

(b) Somewhat more surprising in the above chart is that when the exponent \( n \) is a power of 2, then \( 2^n + 1 \) seems to be a prime, namely 3, 5, 17, 257 and 65537. Fermat thought he was onto something and suggested that every integer of the form \( 2^{2^k} + 1 \) would be a prime. Then we would have a machine for making really big primes in a hurry. It came as a heartbreak when Euler discovered that

\[ 2^{32} + 1 = 4294967297 = 641 \cdot 6700417, \]
and, in doing so, he dashed our hopes of finding a machine for building
giant primes. The numbers

\[ F_k = 2^{2^k} + 1 \text{ for } k = 0, 1, 2, 3, \ldots \]

are called Fermat numbers and if they happen to be prime, such as
\( F_0, F_1, F_2, F_3, F_4 \) happen to be, they are called Fermat primes. Not a
single Fermat prime has been discovered beyond \( F_4 \). So, one way to
become famous is to find the next Fermat prime, or else show there are
not any more.

But the Fermat numbers are still good for something.

• If \( 1 \leq k < \ell \), prove that \( F_\ell \) and \( F_k \) are coprime.

  Hint. Suppose \( p \) is a prime common factor of \( F_k \) and \( F_\ell \). Explain why
  \( p \neq 2 \) and use congruences modulo \( p \) to get a contradiction.

(c) Use part (b) to give an alternate proof that there are infinitely many
primes.

15. Prove there are infinitely many primes of the form \( 4k + 3 \). This is the same
as showing there are infinitely many primes congruent to 3 modulo 4.

  Hint. Start with a finite list \( p_1, p_2, \ldots, p_n \) of such primes, but leave 3 out of
the list. Show how to get one more by looking at
\[ n = 4p_1 \cdot p_2 \cdots p_n + 3. \]
Is it possible for all of the prime factors of \( n \) to be congruent to 1 modulo
4? Also explain why \( n \) does not have prime factors congruent to 0 or 2
mod 4? Some organized thinking is needed.

16. For this problem let \( p \) be a prime and let \( p \) be odd. So \( p \neq 2 \). In \( \mathbb{Z}_p^* \) there are
\( \varphi(p) = p - 1 \) residues, and this is an even number.

  We have seen that \( \mathbb{Z}_p^* \) can be listed as the residues

\[ [1], [2], [3], \ldots, [p - 1] \]

represented by the non-zero remainders \( 1, 2, 3, \ldots, p - 1 \). Another good
choice of representatives of \( \mathbb{Z}_p^* \) is the set of integers

\[ \pm 1, \pm 2, \pm 3, \ldots, \pm \frac{p - 1}{2}. \]
Indeed, there are \( p - 1 \) integers here and no two of them are congruent mod \( p \), because the distance between any two of them is less than \( p \). Hence

\[
\mathbb{Z}_p^* = \left\{ \pm[1], \pm[2], \ldots, \pm \left[ \frac{p-1}{2} \right] \right\}.
\]

A quadratic residue modulo \( p \) is any unit \([a]\) of \( \mathbb{Z}_p^* \) that takes the form \([a] = [x]^2\) for some other \([x]\) in \( \mathbb{Z}_p^* \). In other words, the quadratic residues are the perfect squares of \( \mathbb{Z}_p^* \). In the language of congruences, the integer \( a \) is said to have a quadratic residue modulo \( p \) when \( p \nmid a \) and \( a \equiv x^2 \mod p \) has a solution \( x \).

The exercises that follow are each quite short, once the situation is understood.

(a) By squaring all elements of \( \mathbb{Z}_7^* \), list the quadratic residues of \( \mathbb{Z}_7^* \). Repeat for \( \mathbb{Z}_{17}^* \).

(b) If \([x] \in \mathbb{Z}_p^*\), explain briefly why \([x] \neq -[x]\).

(c) Let \([x], [y] \in \mathbb{Z}_p^*\). Prove that \([x]^2 = [y]^2\) if and only if \([y] = \pm [x]\).

(d) Explain why \( \mathbb{Z}_p \) has exactly \( \frac{p-1}{2} \) quadratic residues.

Hint. Show that the list \([1]^2, [2]^2, [3]^2, \ldots, \left[ \frac{p-1}{2} \right]^2\) never repeats and picks up all possible quadratic residues.

(e) If \([a]\) is a quadratic residue, and \([b]\) is a quadratic residue, show that \([a][b] = [ab]\) is also a quadratic residue.

(f) Show that the inverse in \( \mathbb{Z}_p \) of every quadratic residue \([a]\) is also a quadratic residue.

Items (e) and (f) demonstrate that the set of quadratic residues is a subgroup of \( \mathbb{Z}_p^* \).

(g) If \([a]\) is a quadratic residue, while \([b]\) is not a quadratic residue, show that \([a][b]\) is not a quadratic residue.

Hint. Suppose it is, and find a contradiction.

(e) If \([a], [b]\) are both not quadratic residues, show that \([a][b]\) is a quadratic residue.

This one is more subtle than the previous ones.

17. Let \( p \) be an odd prime, i.e. \( p \neq 2 \).
(a) If \([x] \in \mathbb{Z}_p^*, \) and \([x]\) equals its own inverse, show that \([x] = \pm[1].\)

(b) You have just seen in part (a) that, except for the elements \(\pm[1]\) in \(\mathbb{Z}_p^*,\)
each of the other elements has an inverse different than itself.

Simplify the product


This should only take a couple of lines to obtain and explain the answer.

(c) Prove that \((p - 1)! \equiv -1 \mod p\) for every odd prime \(p.\)

This result is known as **Wilson’s theorem**. The proof is essentially captured by part (b).

(d) In the Exercises of Chapter 2, we had seen the converse of Wilson’s theorem. Namely, if \((p - 1)! \equiv -1 \mod p,\) then \(p\) is prime. Thus an odd number \(p\) is prime if and only if \((p - 1)! \equiv -1 \mod p.\)

Very briefly discuss why such a test as this might not be that great of a primality test, when \(p\) gets very big.
Chapter 4

The Chinese Remainder Theorem

Around the year 300 a solution to the following mathematical problem appeared in the mathematical manual of Chinese Master, Sun Tzu Suan Ching.

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

The master was asking us to solve the three simultaneous congruences:

\[ x \equiv 2 \mod 3 \]
\[ x \equiv 3 \mod 5 \]
\[ x \equiv 2 \mod 7 \]

The Chinese Remainder Theorem is a simple exploitation of Proposition 1.8. Here is a reminder of that result.

If \( m, n \) are coprime integers, their product \( mn \) will divide an integer \( a \) if and only if \( m \) and \( n \) individually divide \( a \). This does not hold when \( m, n \) are not coprime. For instance, 4 and 6 are not coprime. They individually divide 12, but their product 24 does not divide 12. In terms of congruences, Proposition 1.8 said:

\[ a \equiv 0 \mod mn \text{ if and only if } a \equiv 0 \mod m \text{ and } a \equiv 0 \mod n. \]

If we replace \( a \) by a difference \( a - b \), the above tells us a very useful piece of information.
Proposition 4.1. If $a, b$ are integers and $m, n$ are coprime moduli, the congruence
\[ a \equiv b \mod mn \]
is true if and only if the pair of congruences
\[ a \equiv b \mod m \text{ and } a \equiv b \mod n. \]
are true.

Thus any congruence problem involving a modulus that is the product of co-prime integers can be reduced to a pair of congruence problems involving the factors that made up the product. We will be using this fact persistently.

4.1 Statement, proof and uses of CRT

Here is the renowned Chinese Remainder Theorem.

Theorem 4.2 (Chinese Remainder). If $m, n$ are coprime moduli, and $a, b$ are any integers, then the congruences
\[ x \equiv a \mod m \text{ and } x \equiv b \mod n \]
have a common solution $x$. Furthermore, any two solutions $x, y$ to this pair of congruences must be such that $x \equiv y \mod mn$.

Proof. Since $m, n$ are coprime, the Diophantine equation
\[ mt - ns = b - a \]
has a solution $t, s$. With such solutions $s, t$ we obtain $mt + a = ns + b$. Now, let $x$ be this common number. Clearly
\[ x \equiv a \mod m \text{ and } x \equiv b \mod n, \]
which makes $x$ a solution to the simultaneous congruences.

If $y$ is another solution to the simultaneous congruences, then
\[ x \equiv y \mod m \text{ and } x \equiv y \mod n. \]
According to Proposition 4.1, we conclude $x \equiv y \mod mn$. \qed
This Chinese Remainder Theorem can be jacked up to cover more than two simultaneous congruences. Here is the more general statement of the theorem, but we shall skip the proof, which goes along the same lines as that Theorem 4.2.

**Theorem 4.3.** Suppose $n_1, n_2, \ldots, n_k$ are moduli that are pairwise coprime. That is, $n_i$ and $n_j$ are coprime when $i \neq j$. If $a_1, \ldots, a_k \in \mathbb{Z}$, then there is an integer $x$ such that

\[
\begin{align*}
x &\equiv a_1 \mod n_1 \\
x &\equiv a_2 \mod n_2 \\
& \vdots \\
x &\equiv a_k \mod n_k.
\end{align*}
\]

If $x_0$ is such a solution of these congruences, then the complete solution is given by all

\[x \equiv x_0 \mod n_1 n_2 \cdots n_k.
\]

Regarding the name of this theorem, we can see that the Chinese Master was asking for a number whose “remainders” with respect to different moduli were specified.

Once we find a particular solution to our simultaneous congruences, the general solution is readily available as specified in Theorem 4.3. To actually find a particular solution $x_0$, we do some work with linear Diophantine equations. Here is an example.

**Example 4.4.** Let’s solve the simultaneous congruences

\[
\begin{align*}
x &\equiv 2 \mod 3 \\
x &\equiv 3 \mod 10 \\
x &\equiv 5 \mod 11.
\end{align*}
\]

Since 3, 10, 11 are pairwise coprime, we know that once we find one solution $x_0$, then the complete solution is given by all $x \equiv x_0 \mod 330$.

In order to satisfy the first congruence we need an $x = 2 + 3t$ for some $t$ in $\mathbb{Z}$. For such $x$ to satisfy the second congruence we need a $t$ such that

\[2 + 3t \equiv 3 \mod 10.
\]
This leads us to find a $t$ such that

$$3t \equiv 1 \mod 10$$

By inspection we see that

$$t \equiv 7 \mod 10.$$ 

For the last step, we could have solved $3t - 10r = 1$, but instead we just multiplied both sides by 7 and used $7 \cdot 3 \equiv 1 \mod 10$. In other words, we multiplied by the inverse of $[3]$ in $\mathbb{Z}_{10}$. So $t = 7 + 10s$ for some $s$.

We have learned that in order to solve the first two congruences we need

$$x = 2 + 3(7 + 10s) = 23 + 30s \text{ for some integer } s.$$ 

For the third congruence to hold, we require

$$23 + 30s \equiv 5 \mod 11.$$ 

This comes down to

$$8s \equiv 4 \mod 11,$$

and then

$$s \equiv 6 \mod 11.$$ 

So we need $s = 6 + 11u$ for some $u$.

Thus,

$$x = 23 + 30(6 + 11u) = 203 + 330u$$

where $u \in \mathbb{Z}$, will solve all three congruences. We readily see that the general solution to our congruences is all $x \equiv 203 \mod 330$.

Notice at one point that we succeeded in getting from $3t \equiv 1 \mod 10$ to $t \equiv 7 \mod 10$. This was possible because 3 and 10 were coprime.

At another point, we succeeded in getting from $8s \equiv 4 \mod 11$ to $s \equiv 6 \mod 11$. That was because $8 \equiv 30 = 3 \cdot 10 \mod 11$, and 3 and 10 were coprime with 11, making 30 coprime with 11. It was no accident that 8 had an inverse $\mod 11$.

The preceding laborious method will succeed in solving any number of simultaneous congruences with pairwise coprime moduli.

Here's comes an example that combines the Chinese Remainder Theorem with Euler's Theorem.
4.1. STATEMENT, PROOF AND USES OF CRT

Example 4.5. Let's solve \( x^9 \equiv 4 \mod 55 \).

According to Proposition 4.1, an integer \( x \) solves this congruence if and only if,

\[
x^9 \equiv 4 \mod 5 \quad \text{and} \quad x^9 \equiv 4 \mod 11.
\]

Before solving the original congruence we need to solve these simpler congruences first.

For \( x^9 \equiv 4 \mod 5 \), we see that the solution \( x \equiv 0 \mod 5 \) does not work. So any solution \( x \) satisfies \( x^4 \equiv 1 \mod 5 \), by Fermat’s Little Theorem. Thus

\[
x^9 = x^4 \cdot x^4 \cdot x \equiv x \mod 5.
\]

So the solution to \( x^9 \equiv 4 \mod 5 \) is \( x \equiv 4 \mod 5 \).

For \( x^9 \equiv 4 \mod 11 \), we see that \( x \equiv 0 \mod 11 \) never works. So any solution \( x \) satisfies \( x^{10} \equiv 1 \mod 11 \), by Fermat. Thus, \( x^9 \equiv 4 \mod 11 \) if and only if \( x^{10} \equiv 4x \mod 11 \), which happens if and only if \( 4x \equiv 1 \mod 11 \). Since \( 4 \cdot 3 \equiv 1 \mod 11 \) we see that \( x \equiv 3 \mod 11 \).

Thus \( x^9 \equiv 4 \mod 55 \) if and only if

\[
x \equiv 4 \mod 5 \quad \text{and} \quad x \equiv 3 \mod 11.
\]

Since 5 and 11 are coprime, the Chinese Remainder Theorem ensures that a solution \( x \) exists. From the first equation we need \( x = 4 + 5t \) for some \( t \). The second equation forces \( 4 + 5t = x \equiv 3 \mod 11 \). Then we get \( 5t \equiv 10 \mod 11 \) and then, by inspection, \( t \equiv 2 \mod 11 \). So \( t = 2 + 11s \) for some \( s \), and the solution to our simultaneous congruence problem, and thereby our original congruence, is \( x = 4 + 5(2 + 11s) = 14 + 55s \). Thus

\[
x \equiv 14 \mod 55 \text{ solves } x^9 \equiv 4 \mod 55.
\]

The idea of using the Chinese Remainder Theorem, as in this example, to break down a congruence problem into simpler congruence problems is valuable. But we may also wonder if this example could have been solved by using Proposition 3.25. In our example we have \( \varphi(55) = 40 \). This can be seen by counting all numbers from 0 to 54 that are coprime with 55, or by a theorem to be revealed soon. Since 4 and 55 are coprime and 9 and 40 are coprime, Proposition 3.25 is indeed applicable.
To solve our congruence by the method of Proposition 3.25 we first need to solve
\[ 9\ell \equiv 1 \mod 40, \]
and then write \( x \equiv 4\ell \mod 55 \). Here we have one bigger problem instead of a number of smaller problems. Which one is better might depend on the situation. By a lucky inspection we see that \( \ell \equiv 9 \mod 40 \) solves the linear congruence. Thus the solution to our original problem is
\[ x \equiv 4^9 \mod 55. \]
A quick reduction reveals that \( x \equiv 14 \mod 55 \). In this example, Proposition 3.25 with the help of a calculator was probably a touch faster.

### 4.2 CRT from the point of view of residues

The rings of residues \( \mathbb{Z}_n \) provide an elegant way to restate the Chinese Remainder Theorem. This, in turn, will enable us to establish an important formula for the Euler \( \varphi \)-function.

Let \( m, n \) be two moduli. Form the residue rings \( \mathbb{Z}_m \) and \( \mathbb{Z}_n \), each containing exactly \( m \) and \( n \) residue classes, respectively. The residues of \( \mathbb{Z}_m \) will be denoted by \([x]_m\), while the residues of \( \mathbb{Z}_n \) will be denoted by \([x]_n\), in order to not get them mixed up. Then form the so called Cartesian product:

\[
\mathbb{Z}_m \times \mathbb{Z}_n = \{ ([a]_m, [b]_n) : [a]_m \in \mathbb{Z}_m \text{ and } [b]_n \in \mathbb{Z}_n \}.
\]

In other words, \( \mathbb{Z}_m \times \mathbb{Z}_n \) is the set of all possible pairs of residues, where first residue comes from \( \mathbb{Z}_m \) and the second residue comes from \( \mathbb{Z}_n \). (By way of analogy, the Cartesian product of the real line \( \mathbb{R} \) with itself, is the usual Cartesian plane \( \mathbb{R} \times \mathbb{R} = \mathbb{R}^2 \), also known as the \( xy \)-plane.)

Since there are exactly \( m \) possible residues for the first entry and \( n \) possible residues for the second entry, the set \( \mathbb{Z}_m \times \mathbb{Z}_n \) has exactly \( mn \) pairs of residues.

We remember as well that the residue ring \( \mathbb{Z}_{mn} \) also has exactly \( mn \) elements. There ought to be some decent one-to-one correspondence between \( \mathbb{Z}_m \times \mathbb{Z}_n \) and \( \mathbb{Z}_{mn} \).

Let’s take a close look at the function
\[
f : \mathbb{Z}_{mn} \rightarrow \mathbb{Z}_m \times \mathbb{Z}_n \text{ where } f([x]_{mn}) = ([x]_m, [x]_n).
\]
This $f$ is well-defined. That is, $f$ does not depend on the representative $x$ used to get $[x]_{mn}$. Indeed, suppose $[x]_{mn} = [y]_{mn}$. This means that $x \equiv y \mod mn$. By the meaning of congruences we see that

$$x \equiv y \mod m \text{ and } x \equiv y \mod n.$$ 

Therefore,

$$[x]_m = [y]_m \text{ and } [x]_n = [y]_n,$$

meaning that $([x]_m, [x]_n) = ([y]_m, [y]_n)$.

For example, let $m = 5$ and $n = 3$. Our mapping $f : \mathbb{Z}_{15} \to \mathbb{Z}_5 \times \mathbb{Z}_3$ will give

$$f([11]_{15}) = ([11]_5, [11]_3) = ([1]_5, [2]_3) \text{ and } f([9]_{15}) = ([9]_5, [9]_3) = ([4]_5, [0]_3).$$

**Theorem 4.6.** If $m, n$ are coprime, then the mapping

$$f : \mathbb{Z}_{mn} \to \mathbb{Z}_m \times \mathbb{Z}_n, \text{ given by } [x]_{mn} \mapsto ([x]_m, [x]_n),$$

is both one-to-one and onto.

**Proof.** To say that $f$ is one-to-one and onto means that for every pair of residues $([a]_m, [b]_n)$ in $\mathbb{Z}_m \times \mathbb{Z}_n$ there is exactly one residue $[x]_{mn}$ in $\mathbb{Z}_{mn}$ such that

$$([x]_m, [x]_n) = ([a]_m, [b]_n).$$

In terms of representatives this says that for any $a, b$ in $\mathbb{Z}$, there is an $x$ in $\mathbb{Z}$ such that

$$x \equiv a \mod m \text{ and } x \equiv b \mod n,$$

and that any $y$ that solves these is just another representative of the residue $[x]_{mn}$. Well, Theorem 4.2 says exactly that. \qed

Theorem 4.6 is useful to develop an instrument for computing the values of $\varphi(n)$.

### 4.3 Calculating the Euler function

Let’s go back to a pair of coprime moduli $m, n$. We have our one-to-one and onto mapping

$$f : \mathbb{Z}_{mn} \to \mathbb{Z}_m \times \mathbb{Z}_n, \text{ where } f([x]_{mn}) = ([x]_m, [x]_n).$$
As well, recall that $\mathbb{Z}_{mn}^*$ is the group of units of $\mathbb{Z}_{mn}$, and that by the very definition of Euler’s function, $\varphi(mn)$ equals the number of elements of $\mathbb{Z}_{mn}^*$. In more concrete terms, $\varphi(mn)$ equals the number of integers from 1 to $mn - 1$ that are coprime with $mn$.

The next result will give us a breakthrough on how to calculate $\varphi(mn)$ in terms of $\varphi(m)$ and $\varphi(n)$.

**Proposition 4.7.** If $m, n$ are coprime integers and

$$f : \mathbb{Z}_{mn} \rightarrow \mathbb{Z}_m \times \mathbb{Z}_n \text{ is given by } [x]_{mn} \mapsto ([x]_m, [x]_n),$$

then

$$f(\mathbb{Z}_{mn}^*) = \mathbb{Z}_m^* \times \mathbb{Z}_n^*.$$

**Proof.** First we check that

$$f(\mathbb{Z}_{mn}^*) \subseteq \mathbb{Z}_m^* \times \mathbb{Z}_n^*.$$

Well, if $[x]_{mn} \in \mathbb{Z}_{mn}^*$, this means that $x$ is coprime with $mn$. Hence, $x$ is coprime with $m$, and $x$ is coprime with $n$. And so,

$$[x]_m \in \mathbb{Z}_m^* \text{ and } [x]_n \in \mathbb{Z}_n^*,$$

which is saying that

$$([x]_m, [x]_n) \in \mathbb{Z}_m^* \times \mathbb{Z}_n^*.$$

Next, we check that for every $([a]_m, [b]_n)$ in $\mathbb{Z}_m^* \times \mathbb{Z}_n^*$, there is an $[x]_{mn}$ in $\mathbb{Z}_{mn}^*$ such that $f([x]_{mn}) = ([a]_m, [b]_n)$. By Theorem 4.6 we do have some $[x]_{mn}$ in $\mathbb{Z}_{mn}$ such that

$$f([x]_{mn}) = ([x]_m, [x]_n) = ([a]_m, [b]_n).$$

This means that

$$[x]_m = [a]_m \in \mathbb{Z}_m^* \text{ and } [x]_n = [a]_n \in \mathbb{Z}_n^*.$$

Thus $x$ is coprime with $m$ and $x$ is coprime with $n$. Hence $x$ is coprime with $mn$. Indeed, if $x$ and $mn$ shared a prime factor, then $x$ would share that prime factor with one of $m$ or $n$, by the Unique Factorization Theorem. Thereby we learn that $[x]_{mn} \in \mathbb{Z}_{mn}^*$.
4.3. CALCULATING THE EULER FUNCTION

Here is an important consequence of Proposition 4.7. Once more, recall that \( \varphi(n) \) the number of elements in \( \mathbb{Z}_n^* \), which is the same as the number of units in \( \mathbb{Z}_n \), which in turn is the same as the number of integers between 1 and \( n - 1 \) that are coprime with \( n \).

First a bit of notation. The number of elements in a finite set \( A \) will be denoted by \( \#A \). This is sometimes called the cardinality of \( A \).

**Proposition 4.8.** If \( m, n \) are coprime positive integers, then

\[
\varphi(mn) = \varphi(m)\varphi(n).
\]

**Proof.** Proposition 4.7 tells us that there is a one-to-one and onto correspondence between \( \mathbb{Z}_{mn}^* \) and \( \mathbb{Z}_m^* \times \mathbb{Z}_n^* \). Therefore,

\[
\#\mathbb{Z}_{mn}^* = \#(\mathbb{Z}_m^* \times \mathbb{Z}_n^*).
\]

But

\[
\#(\mathbb{Z}_m^* \times \mathbb{Z}_n^*) = \#\mathbb{Z}_m^* \cdot \#\mathbb{Z}_n^*.
\]

This is because the cardinality of a Cartesian product equals the product of the cardinalities of the individual sets. Therefore

\[
\#\mathbb{Z}_{mn}^* = \#\mathbb{Z}_m^* \cdot \#\mathbb{Z}_n^*.
\]

According to the definition of Euler’s function this says that

\[
\varphi(mn) = \varphi(m)\varphi(n),
\]

and we are done. \( \square \)

Before Proposition 4.8 can be used, we must make sure \( m, n \) are coprime.

By way of example, we see by simple counting that \( \varphi(25) = 20 \) and \( \varphi(4) = 2 \). Since 25 and 4 are coprime, we conclude:

\[
\varphi(100) = \varphi(25 \cdot 4) = \varphi(25) \cdot \varphi(4) = 20 \cdot 2 = 40.
\]

For another example, since 101 and 7 are prime, we know that \( \varphi(101) = 100 \) and \( \varphi(7) = 6 \), and since 101 and 7 are coprime we conclude:

\[
\varphi(707) = \varphi(101 \cdot 7) = \varphi(101) \cdot \varphi(7) = 100 \cdot 6 = 600.
\]
CHAPTER 4. THE CHINESE REMAINDER THEOREM

More generally, if \( p, q \) are distinct primes, then, as above, we obtain:
\[
\varphi(pq) = \varphi(p)\varphi(q) = (p - 1)(q - 1).
\]

The fact that \( \varphi(mn) = \varphi(m)\varphi(n) \) whenever \( m, n \) are coprime is known as the multiplicative property of \( \varphi \).

It should be clear as well that \( \varphi \) is multiplicative on the product of any number of pairwise coprime integers. That is
\[
\varphi(m_1m_2\ldots m_k) = \varphi(m_1)\varphi(m_2)\ldots\varphi(m_k).
\]
whenever all pairs \( m_i, m_j \) are coprime. This tells us what to do with an arbitrary integer \( x \). Suppose that
\[
x = p_1^{e_1} \cdot p_2^{e_2} \cdots p_k^{e_k}
\]
is the unique factorization of \( x \) into distinct primes \( p_j \) with respective multiplicities \( e_j \). Proposition 4.8 tells us that
\[
\varphi(x) = \varphi(p_1^{e_1}) \cdot \varphi(p_2^{e_2}) \cdots \varphi(p_k^{e_k}).
\]

We can now see that a formula for \( \varphi(x) \) would be known if we had a formula for \( \varphi(p^e) \) when \( p \) is prime and \( e \geq 0 \). To figure out that formula, let’s think about the integers from 1 to \( p^e \) and count the number that are not coprime with \( p^e \). Such numbers are the ones with \( p \) as a factor. Well, how many multiples of \( p \) are there from 1 to \( p^e \)? Since \( p^e = p^{e-1} \cdot p \), we see that there are \( p^{e-1} \) such multiples of \( p \) from 1 to \( p^e \). For instance, from 1 to 8, there are 8 multiples of \( p \), namely
\[1p, 2p, 3p, 4p, 5p, 6p, 7p, 8p.\]
The remaining numbers from 1 to \( p^e \) are coprime with \( p^e \). The number of those remaining numbers is precisely
\[
\varphi(p^e) = p^e - p^{e-1} = p^{e-1}(p - 1).
\]
By putting together the preceding remarks we have an interesting result.

**Proposition 4.9.** If
\[
x = p_1^{e_1} \cdot p_2^{e_2} \cdots p_k^{e_k}
\]
is the factorization of \( x \) into distinct primes \( p_j \) with positive multiplicities \( e_j \), then
\[
\varphi(x) = (p_1^{e_1} - p_1^{e_1-1})(p_2^{e_2} - p_2^{e_2-1}) \cdots (p_k^{e_k} - p_k^{e_k-1})
\]
\[
= p_1^{e_1-1}(p_1 - 1) \cdot p_2^{e_2-1}(p_2 - 1) \cdots p_k^{e_k-1}(p_k - 1).
\]
There are other ways of presenting the formula for $\varphi(x)$, such as:

$$
\varphi(x) = p_1^{e_1} \left(1 - \frac{1}{p_1}\right) \cdot p_2^{e_2} \left(1 - \frac{1}{p_2}\right) \cdots p_k^{e_k} \left(1 - \frac{1}{p_k}\right) \\
= p_1^{e_1} \cdot p_2^{e_2} \cdots p_k^{e_k} \left(1 - \frac{1}{p_1}\right) \cdot \left(1 - \frac{1}{p_2}\right) \cdots \left(1 - \frac{1}{p_k}\right) \\
= x \left(1 - \frac{1}{p_1}\right) \cdot \left(1 - \frac{1}{p_2}\right) \cdots \left(1 - \frac{1}{p_k}\right).
$$

Here are some illustrations of the formula for $\varphi$.

$$
\varphi(408) = \varphi(2^3 \cdot 3 \cdot 17) = 408 \left(1 - \frac{1}{2}\right) \cdot \left(1 - \frac{1}{3}\right) \cdot \left(1 - \frac{1}{17}\right) = 128.
$$

$$
\varphi(5282739) = \varphi(3^4 \cdot 7^2 \cdot 11^3) = (3^4 - 3^3)(7^2 - 7)(11^3 - 11^2) = 2744280.
$$

As long as we can factor $x$, we can calculate $\varphi(x)$ with ease.

### 4.4 An application to communications security

Mathematics can be peculiar at times. We are, by and large, motivated to learn the truth about numbers and their patterns out of sheer curiosity. Such motivation led Euler to his discovery that $a^{\varphi(n)} \equiv 1 \pmod{n}$. No one could have expected, 300 years ago, that such an insight could be useful in today’s society, but a little over 30 years ago his theorem hit the jackpot. It became the central fact about integers that supports a commonly used scheme for data encryption, a word from ancient Greek meaning “to hide”. Here we will describe this so called RSA-scheme, named after its discoverers Ronal Rivest, Adi Shamir, and Leonard Adelman.

The RSA-scheme relies on the fact that, even with a powerful computer, no one to this day has found a practical way to factor huge integers into primes. It’s relatively easy (with computers and advanced number theory) to find huge primes, and very easy to multiply them. But, given the product of two giant primes, say each made up of about 100 digits, it is technically impossible, at this point in our history, to factor the resulting number having roughly 200 digits.

**Alice, Bob and Eve**

The drama of encryption is played out by hypothetical individuals who wish to exchange a secret message. Let us call them Alice and Bob. A message needs
to be sent between them, encrypted in such a way that Eve, the eavesdropper or spy, cannot break the code. All three characters have powerful computers, and the channel that carries messages between Alice and Bob, for instance the internet, is not secure. Eve can see everything that is going through the communication channel. The RSA-scheme is one method (among several) by which Alice and Bob can exchange secret messages in such an open environment.

This field of cryptography has blossomed into a huge endeavour of applied research, with number theory as a core piece of mathematics within it.

**A message is a number**

Any message of any sort can be encoded as an integer. For instance, to encode a passage of English text we can encode letters, spaces and punctuation marks via a simple scheme as shown.

“space” \( \to 00, A \to 01, B \to 02, C \to 03, \ldots W \to 25, Z \to 26, “,” \to 27 \ldots \)

and so on for other punctuation marks and lower case letters. So, for example, the message “HI THERE” becomes the integer 0809002008051805.

Since any message can be encoded as an integer, we can simply think of a message as actually being an integer.

**The theory behind RSA**

The RSA scheme for sending secret messages rests on the following result, based on the Euler-Fermat Theorems.

**Proposition 4.10.** Suppose that \( p, q \) are distinct primes and let \( n = pq \).

Also suppose \( e, d \) are positive integers such that

\[
ed d \equiv 1 \mod (p - 1)(q - 1).
\]

Let \( m \) be an integer between 0 and \( n - 1. \)

If \( m^e \) reduces to \( c \) modulo \( n \), then \( c^d \) reduces to \( m \) modulo \( n \).

**Proof.** We have to show that \( c^d \equiv m \mod n. \)
Since \( n = pq \), we have that \( \varphi(n) = (p - 1)(q - 1) \). In case \( m \) is coprime with \( n \), Proposition 3.20 (itself an immediate spin-off from Euler’s Theorem) quickly reveals what we want:

\[
c^d \equiv (m^c)^d = m^{ed} \equiv m^1 = m \mod n.
\]

So, we need to worry about the case where \( m \) is not coprime with \( pq \). In this case, one of \( p \) or \( q \) divides \( m \), and since \( 0 \leq m < pq \), they cannot both divide \( m \). Say \( p \mid m \) but \( q \nmid m \).

Since \( ed \equiv 1 \mod (p - 1)(q - 1) \) it follows that \( ed \equiv 1 \mod q - 1 \). Also since \( c \equiv m^e \mod pq \), it follows that \( c \equiv m^e \mod q \). Furthermore \( m \) and \( q \) are coprime, due to the fact \( q \nmid m \). Because of these observations, Proposition 3.20 once more gives:

\[
c^d \equiv (m^c)^d = m^{ed} \equiv m^1 = m \mod q.
\]

In addition, \( m \equiv 0 \mod p \), and since \( c \equiv m^e \mod pq \), it follows that \( c \equiv m^e \mod p \). Thus \( c \equiv 0 \mod p \), and thereby

\[
c^d \equiv 0^d = 0 \equiv m \mod p.
\]

Having just seen that \( c^d \equiv m \mod p \) and \( \mod q \), it follows from Proposition 4.1 that

\[
c^d \equiv m \mod pq.
\]

Thus, the result is explained for all possible \( m \).

**The RSA-scheme**

We need to keep Proposition 4.10 in mind. To assist with mnemonics we may think of the symbols in Proposition 4.10 as playing the following roles.

- \( n \) is the product of two primes, that a spy cannot factor.
- \( m \) is the message, taken as a number between 0 and \( n - 1 \).
- \( e \) is the encoding parameter used to create the code \( c \equiv m^e \mod n \) where \( c \) is between 0 and \( n - 1 \).
- \( d \) is the decoding parameter that gives back the original message \( m \) as the reduction of \( c^d \).
Alice wants to receive an encrypted message from Bob. Here is what she does using her computer.

- Alice picks two very large primes \( p, q \) and multiplies them to get \( n = pq \).
- She knows \( \varphi(n) = (p - 1)(q - 1) \), and picks a number \( e \) that is coprime with \( \varphi(n) \).
- She solves the congruence \( ed \equiv 1 \mod \varphi(n) \) for \( d \).
  She chooses \( d \) positive and less than \( \varphi(n) \), and keeps it to herself.
- Alice publishes the encoding parameters \( n, e \).
  Everyone can see these, Bob, Eve, the government, you name it. But currently, only Alice knows the factors \( p, q \) that give \( n = pq \). So, only she knows \( \varphi(n) \) and \( d \).

Bob has a message to send to Alice. Here is what he does.

- Knowing \( n \), he converts his message to a number \( m \) between 0 and \( n - 1 \).
  If his message is too long he can always break it up into a few blocks of numbers between 0 and \( n - 1 \), and separately send each block.
  Furthermore, Bob is careful not to use the numbers 0, 1 and \( n - 1 \) as his message \( m \) because these don’t scramble well in the subsequent reduction.
- With his computer (and probably the square and multiply algorithm) he finds the reduction \( c \) of \( m^e \) modulo \( n \). Don’t forget that he has the \( n \) and \( e \) published by Alice.
- He publishes \( c \) for the world to see, including Alice and Eve.

To unscramble the message that Bob sent, here is what Alice does.

- Alice takes \( c \), which is now public, and reduces \( c^d \mod n \) to get the original message \( m \).

Why is this secure?

- To get \( m \) from \( c \), Eve needs to know \( d \), the solution of \( ed \equiv 1 \mod \varphi(n) \).
4.4. AN APPLICATION TO COMMUNICATIONS SECURITY

- For that, Eve needs to know $\varphi(n)$.

- Thus, Eve needs to factor $n$ in order to see that $\varphi(n) = (p - 1)(q - 1)$. But with her high powered computer Eve cannot, to this day, effectively factor $n$.

A toy example of RSA using small primes

One day Bob, a teacher, gets the following e-mail from his student Alice.

Dear Professor Bob,

I am currently out of town, and I need to know my number theory mark, but I would prefer it if my mother Eve did not see it. If you post my grade on the course web page encrypted using the RSA-scheme, she will never find out what I got.

My encoding parameters are $n = 119$ and $e = 35$.

Eve is very savvy with computers and hacking, but she never learned how to factor anything, including 119. This makes it safe to use RSA.

Sincerely,

Alice

A few moments later Alice sees that her RSA-encrypted grade, posted on the web page by Bob, is 53.

How does Alice find out her true mark?

- To start with, she had picked the two primes $p = 17$ and $q = 7$ to obtain

$$n = pq = 17 \cdot 7 = 119.$$ 

- Then she figured out $\varphi(119) = 16 \cdot 6 = 96$.

- After that she picked a number coprime with 96. In her case she decided to use $e = 35$.

- Then she solved

$$35d \equiv 1 \mod 96$$

for $d$. 
For that, she worked on the Diophantine equation
\[ 35d - 96t = 1. \]

With the Euclidean algorithm she got \( d \equiv 11 \mod 96 \).

- Then she published \( n = 119 \) and \( e = 35 \) for Bob to see, but she kept the \( p = 17, q = 7 \) and \( d = 11 \) to herself.

- She knew that Bob would encrypt her true mark \( m \) by reducing \( m^{35} \mod 119 \) to get the 53. In other words, she knew that
  \[ m^{35} \equiv 53 \mod 119. \]

- To get \( m \), all she had to do was pull out her secret \( d = 11 \) and reduce \( 53^{11} \mod 119 \). After doing the necessary calculation (just look below), she determined that her mark was \( m = 93 \).

Just to be sure Alice got it right, let’s now check with a calculator (or better, a program such as the one in Excel suggested in Chapter 3) that \( 53^{11} \) reduces down to 93 modulo 119, using the square and multiply algorithm.

First we put 11 into binary form by noticing that
\[ 11 = 8 + 2 + 1 = 2^3 + 2^1 + 2^0. \]

Thus
\[ 53^{11} = 53^8 \cdot 53^2 \cdot 53^1. \]

Now keep squaring and reducing 53 modulo 119 to get:
\[
\begin{align*}
53^1 &= 53 \\
53^2 &= 2809 \equiv 72 \\
53^4 &= 72^2 = 5184 \equiv 67 \\
53^8 &= 67^2 = 4489 \equiv 86.
\end{align*}
\]

And so we get
\[ 53^{11} = 53^8 \cdot 53^2 \cdot 53^1 \equiv 86 \cdot 72 \cdot 53 = 328176 \equiv 93. \]

Alice calculated correctly.
4.4. AN APPLICATION TO COMMUNICATIONS SECURITY

Just to be sure Bob did not make a mistake, let’s check that $93^{35}$ does reduce to 53 modulo 119.

First we put 35 into binary form by noticing that

$$35 = 32 + 2 + 1 = 2^5 + 2^1 + 2^0.$$ 

Thus

$$93^{35} = 93^{32} \cdot 93^2 \cdot 93^1.$$ 

Next keep squaring and reducing 93 modulo 119 to get:

- $93^1 \equiv 93$
- $93^2 \equiv 8649 \equiv 81$
- $93^4 \equiv 81^2 \equiv 6561 \equiv 16$
- $93^8 \equiv 16^2 \equiv 256 \equiv 18$
- $93^{16} \equiv 18^2 \equiv 324 \equiv 86$
- $93^{32} \equiv 86^2 \equiv 7396 \equiv 18$

And so we get

$$93^{35} = 93^{32} \cdot 93^2 \cdot 93^1 \equiv 18 \cdot 81 \cdot 93 = 135594 \equiv 53.$$ 

Bob did his calculation correctly too.

Messages to avoid

If Alice publishes $n = pq$, where $p, q$ are distinct primes, Bob had better avoid some messages $m$ between 0 and $n - 1$ because the exponents $m^e$, after reduction modulo $n$, do not scramble effectively.

For instance, $m = 0$ and $m = 1$ are poor choices for a message, because $0^e = 0$ and $1^e = 1$. Also $m = n - 1$ is poor, because $(n - 1)^e \equiv (-1)^e = \pm 1$ mod $n$.

Bob should also avoid $m$ such that either $p | m$ or $q | m$. To see why, suppose that Bob has a message $m$ where $p | m$. He reduces $m^e$ modulo $n$ down to $c$, and sends out $c$. Since $m^e \equiv c \mod pq$, it’s clear that $m^e \equiv c \mod p$. Since $p | m$, we have $c \equiv m^e \equiv 0^e \equiv 0 \mod p$, and thus $p | c$.

Now, with her computer, Eve readily finds $\gcd(c, n)$ by using the Euclidean Algorithm. The Euclidean Algorithm works effectively even when the integers
involved are huge. Since \( p | c \) and \( p | n \) and \( n = pq \), we get that \( \gcd(c, n) \) has to be \( p \). The only other possibility would be \( \gcd(c, n) = n \), but \( c \) is too small for that to happen. Thus Eve has captured \( p \), one of the factors of \( n \). Then she gets \( q = n/p \), and then \( (p - 1)(q - 1) \), and then the solution \( d \) to \( ed \equiv 1 \mod (p - 1)(q - 1) \). Since she now knows everything that Alice knows, Eve can decode \( c \) to get \( m \).

The trouble for Bob, however, is that he does not know \( p \) or \( q \). So how can he avoid messages that are not divisible by \( p \) or \( q \)? Fortunately, the chances of a message \( m \) being divisible by \( p \) or \( q \) are extremely small. To see how small, notice that the number of integers between \( 0 \) and \( n - 1 \) that are coprime with \( n \) is \( \varphi(n) = (p - 1)(q - 1) \). So, the number of integers between \( 0 \) and \( n - 1 \) that are NOT coprime with \( n \) is

\[
n - (p - 1)(q - 1) = n - pq + p + q - 1 = p + q - 1.
\]

The number of possible messages \( m \) between \( 0 \) and \( n - 1 \) is \( n = pq \). Hence, the probability that one of these messages is not coprime with \( n \) is

\[
\frac{p + q - 1}{pq} = \frac{1}{q} + \frac{1}{p} - \frac{1}{pq}.
\]

This is the probability that a message is divisible by either \( p \) or \( q \). When \( p, q \) are huge primes this probability is extremely small. For instance, when \( p, q \) have roughly 100 digits, this probability is roughly \( 1/10^{100} \). That is an outrageously small probability.

As one can imagine from the above discussion, the practical implementation of RSA involves an assortment of engineering details, conventions and protocols. All we can do is marvel that a result of number theory sits at the core of this enterprise.

### 4.5 Exercises

1. Use the Chinese remainder theorem to solve simultaneously:

\[
x \equiv 3 \mod 17 \quad \text{and} \quad x \equiv 5 \mod 23
\]

2. Solve the problem of Sun Tzu.

3. Reduce \( 7^{4002} \mod 100 \). Here you could use the square and multiply algorithm, but you can also do it by factoring the modulus, then making a reduction modulo the factors, and then using the Chinese Remainder Theorem.
4. Solve \( x^2 \equiv 1 \mod 85 \).

5. Solve \( x^7 \equiv x \mod 77 \).

6. Solve \( x^{39} + x^{25} + x^{14} + 1 \equiv 0 \mod 91 \).
   
   Hint. Since 91 = 13 \cdot 7, solve the congruence modulo the primes first. Use Fermat to simplify the exponents. Then use the Chinese Remainder Theorem to construct the solution to the original problem.

7. Suppose that \( m, n \) are coprime integers.
   
   If \( x \equiv m \mod n \) and \( x \equiv n \mod m \), show that \( x \equiv m + n \mod mn \).

8. (a) Determine \( \varphi(18900) \).
   
   (b) Determine \( \varphi(14!) \). Give your answer as a prime factorization.

9. Find all \( n \) such that \( \varphi(n) = 6 \).

10. Find all positive integers \( n \) such that \( \varphi(n) \) is a prime.
    
    Hint. To stay organized, first show that any prime \( p > 3 \) is not a factor of \( n \). That lets you write \( n = 2^d \cdot 3^e \) where \( d, e \geq 0 \). Use the formula for \( \varphi(n) \), along with the fact \( \varphi(n) \) is prime, to rule out all but a few possibilities for \( d \) and \( e \).

11. (a) Suppose \( n = p_1p_2 \cdots p_k \) and the \( p_j \) are all distinct. If \( m \equiv 1 \mod \varphi(n) \), show that \( a^m \equiv a \mod n \) for all integers \( a \).
    
    (b) Does this result hold for \( n = 49 = 7^2 \)?

12. (a) Suppose \( n = pq \) where \( p, q \) are primes. Let \( m = \varphi(n) \). Find explicit formulas for \( p \) and \( q \) as functions of \( n \) and \( m \).
    
    Hint. It should not be too hard to get \( p + q \) in terms of \( n \) and \( m \). Then verify and use the identity \( (p - q)^2 = (p + q)^2 - 4pq \), to get \( p - q \), and from that \( p \) and \( q \), in terms of \( n \) and \( m \).
    
    (b) You are given that \( m = 14398199 \) is a product of two primes, and that \( \varphi(m) = 14390400 \). Use your calculator to find the primes.

13. Note that \( 143 = 11 \cdot 13 \).
    
    (a) If \( a \) is coprime with 143 explain, very quickly, why \( a^{120} \equiv 1 \mod 143 \).
(b) If \(a\) is coprime with 143, show that, in fact, \(a^{60} \equiv 1 \mod 143\).
Hint. Use Proposition 4.1.

14. Alice picks two secret primes \(p = 17, q = 23\). Then she publishes the encoding parameters \(n = 391, e = 51\). Bob sends out the message \(c = 100\) encoded by the RSA-scheme using Alice’s published parameters.

Only Alice knows that \(391 = 17 \cdot 23\). Using this secret knowledge, belonging to Alice, find the original message \(m\) that Bob intended.

15. This problem is about counting the number of solutions to the congruence

\[x^2 \equiv 1 \mod n.\]

(a) If \(p\) is an odd prime, and \(x^2 \equiv 1 \mod p\), show that \(x \equiv \pm 1 \mod p\). This should be very easy by now.

(b) If \(p\) is an odd prime, and \(e \geq 1\) and \(x^2 \equiv 1 \mod p^e\), show that \(x \equiv \pm 1 \mod p^e\). This is a small enhancement of part (a).

(c) Solve \(x^2 \equiv 1 \mod 561\). Hint. \(561 = 17 \cdot 11 \cdot 3\).

(d) If \(n = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}\) where the \(p_j\) are odd, non-repeated primes. Show that \(x^2 \equiv 1 \mod n\) has exactly \(2^k\) solutions modulo \(n\).

Hint. Put the Chinese Remainder Theorem to use.

16. This problem is a continuation of the preceding problem by taking into account the possibility of \(n\) being even. We need to consider first the tricky case where \(n\) is a power of 2.

The congruence \(x^2 \equiv 1 \mod 2\) has one solution, namely \(x \equiv 1 \mod 2\). And the congruence \(x^2 \equiv 1 \mod 4\) has two solutions \(x \equiv 1, 3 \mod 4\). The solutions to \(x^2 \equiv 1 \mod 8\) are \(x \equiv 1, 3, 5, 7 \mod 8\), which gives four solutions. At this time we are tempted to guess that \(x^2 \equiv 1 \mod 16\) has 8 solutions. But it does not. The solutions to \(x^2 \equiv 1 \mod 16\) are \(x \equiv 1, 7, 9, 15 \mod 16\), providing only four solutions. As a matter of fact once \(e \geq 3\), the number of solutions to \(x^2 \equiv 1 \mod 2^e\) remains stuck at 4.

(a) If \(e \geq 3\), show that \(x^2 \equiv 1 \mod 2^e\) has the following solutions \(x \equiv 1, 2^e - 1, 2^{e-1} - 1, 2^{e-1} + 1 \mod 2^e\), by just squaring these numbers and reducing them modulo \(2^e\).
(b) Suppose $e \geq 3$ and that some $x$ between 0 and $2^e - 1$ is a solution to the congruence $x^2 \equiv 1 \mod 2^e$. Show $x$ is one of the four numbers from part (a).

Hint. Deduce that $2^e | (x - 1)(x + 1)$. By unique factorization, there are exponents $a, b$ such that

$$0 \leq a \leq b \leq e, \ a + b = e, \ 2^a | x - 1, \ 2^b | x + 1 \text{ or } 2^a | x + 1, \ 2^b | x - 1.$$ 

Deduce that $a = 0$ or $a = 1$. Each of these two options determines $b$, and from that two of the desired four solutions.

(c) Suppose that $n$ is an even modulus. Write $n = 2^e m$ where $e \geq 1$ and $m$ is odd with $k$ distinct odd primes in its factorization. If $e = 1$, how many solutions does $x^2 \equiv 1 \mod n$ have? If $e = 2$, how many solutions does this congruence have? If $e \geq 3$, how many solutions does the congruence have?
Chapter 5

Primitive roots

We might recall from DeMoivre’s Theorem that the complex number $\omega = e^{2\pi i/n}$ is a root of $X^n - 1$. What’s interesting is that all the roots of this polynomial are given by

$$\omega, \omega^2, \omega^3, \ldots, \omega^{n-1}, \omega^n = 1.$$

Since all $n$’th roots of 1 emerge as powers of a single root, this single root is called a primitive root of unity.

Not all roots of $X^n - 1$ are primitive. For example, $X^4 - 1$ has the four roots

$$i, i^2 = -1, i^3 = -i, i^4 = 1.$$

All of the roots are expressible as powers of one root $i$. So $i$ is a primitive root. Also, $-i$ is primitive, but the roots $\pm 1$ are not primitive.

In this chapter we intend to explore how the above phenomenon plays out in the group of units $\mathbb{Z}_n^*$. The concepts become deeper, and clear thinking will be required.

5.1 Primitive roots in $\mathbb{Z}_n$

We shall prefer to denote residues in the finite rings $\mathbb{Z}_n$ by Greek letters $\alpha, \beta, \gamma, \ldots$, in favour of the more cumbersome box notations $[a], [b], [c], \ldots$. We shall also take the liberty of denoting the special residue $[1]$ by the unadorned symbol “1”.

Recall that $\varphi(n)$ is the number of elements in the group of units $\mathbb{Z}_n^*$, and Euler’s
CHAPTER 5. PRIMITIVE ROOTS

Theorem told us that

\[ \alpha^{\varphi(n)} = 1, \]

for every \( \alpha \) in \( \mathbb{Z}_n^* \). Hence the group of units \( \mathbb{Z}_n^* \) consists of roots of \( X^{\varphi(n)} - 1 \). Conversely, every \( \alpha \) in \( \mathbb{Z}_n \) that is a root of \( X^{\varphi(n)} - 1 \) is automatically in \( \mathbb{Z}_n^* \). Indeed, \( \alpha^{\varphi(n)} = 1 \), and from that we see that \( \alpha \) is a unit with inverse \( \alpha^{\varphi(n)-1} \). So \( X^{\varphi(n)} - 1 \) has precisely \( \varphi(n) \) roots in \( \mathbb{Z}_n \), and these roots make up the group \( \mathbb{Z}_n^* \).

Could it be, as it is in \( \mathbb{C} \), that, for some special \( \alpha \) in \( \mathbb{Z}_n^* \), the list

\[ \alpha, \alpha^2, \alpha^3, \ldots, \alpha^{\varphi(n)-1}, \alpha^{\varphi(n)} = 1 \]

captures all of \( \mathbb{Z}_n^* \) and thereby all roots of \( X^{\varphi(n)} - 1 \)?

Let’s look at some examples.

**Example 5.1.** Take \( \mathbb{Z}_8^* \) whose units are \([1],[3],[5],[7]\). Here \( \varphi(8) = 4 \). These four units are the roots of \( X^4 - 1 \). Let’s calculate the powers of these residues inside \( \mathbb{Z}_8^* \).

- For \( \alpha = [1] \), we see that

\[ [1] = \alpha = \alpha^2 = \alpha^3 = \alpha^4 = \ldots, \]

and obviously the powers of this \( \alpha \) do not produce the other three units.

- For \( \alpha = [3] \) we get

\[ \alpha = [3], \alpha^2 = [1], \alpha^3 = \alpha \alpha^2 = \alpha, \alpha^4 = \alpha \alpha^3 = \alpha^2, \alpha^5 = \alpha, \ldots \]

Here the only roots of \( X^4 - 1 \) produced by the powers of this \( \alpha \) are \([3]\) and \([1]\).

- For \( \alpha = [5] \) we calculate its powers, and learn that the only roots produced by the powers of \( \alpha \) are \([5]\) and \([1]\).

- Finally, for \( \alpha = [7] \), its powers only produce \([7]\) and \([1]\).

We just discovered that in \( \mathbb{Z}_8^* \) there is no single residue whose powers produce all of the four residues in \( \mathbb{Z}_8^* \).
5.1. PRIMITIVE ROOTS IN $\mathbb{Z}_N$

Example 5.2. Let’s look at $\mathbb{Z}_7$. Here $\varphi(7) = 6$ and

$$\mathbb{Z}_7^\ast = \{[1], [2], [3], [4], [5], [6]\}.$$  

The powers of $\alpha = [2]$ give:

$$\alpha = [2], \alpha^2 = [4], \alpha^3 = [1], \alpha^4 = \alpha, \alpha^5 = \alpha^2, \alpha^6 = [1], \ldots$$

The powers of $[2]$ seem to be picking up only $[2], [4]$ and $[1]$, which is definitely not all of $\mathbb{Z}_7^\ast$.

Maybe $\alpha = [3]$ is lucky. Its powers come out to be:

$$\alpha = [3], \alpha^2 = [2], \alpha^3 = [6], \alpha^4 = [4], \alpha^5 = [5], \alpha^6 = [1].$$

We hit a winner, since using $\alpha = [3]$ we have:

$$\mathbb{Z}_7^\ast = \{\alpha, \alpha^2, \alpha^3, \alpha^4, \alpha^5, \alpha^6 = 1\}.$$  

Definition 5.3. An element $\alpha$ in $\mathbb{Z}_n^\ast$ is called a primitive root provided the list of powers

$$\alpha, \alpha^2, \alpha^3, \ldots, \alpha^{\varphi(n)-1}, \alpha^{\varphi(n)} = 1$$

picks up every element of $\mathbb{Z}_n^\ast$.

In the language of congruences, we say that an integer $a$ is a primitive root modulo $n$, provided $a$ is coprime with $n$ and the powers of its residue $[a]$ pick up every element of $\mathbb{Z}_n^\ast$. This is telling us that the list of powers

$$a, a^2, a^3, \ldots, a^{p-1};$$

after reduction modulo $n$, picks up every remainder from 0 to $n-1$ that is coprime with $n$.

For much of the theoretical development it is convenient to work inside $\mathbb{Z}_n^\ast$ rather than with integers. However, translation between integers and residues should not be far from our minds.

As we saw in the preceding examples, for $n = 8$, there are no primitive roots in $\mathbb{Z}_8^\ast$, whereas with $n = 7$, the residue $\alpha = [3]$ is a primitive root in $\mathbb{Z}_7^\ast$.

Our main goal is to prove that if $p$ is prime, then $\mathbb{Z}_p^\ast$ has a primitive root. After that we’ll discuss the discrete logarithm and its application to another cryptographic scheme. To obtain the rather difficult primitive root theorem, we need a run of propositions each requiring close attention.
5.2 The order of a residue

If \( \alpha \in \mathbb{Z}_n^\times \), Euler’s theorem gives \( \alpha^{\varphi(n)} = 1 \). In particular, there is some exponent \( k \geq 1 \) such that \( \alpha^k = 1 \). The following concept is indispensable.

**Definition 5.4.** If \( \alpha \in \mathbb{Z}_n^\times \), the **order** of \( \alpha \) is the smallest exponent \( k \geq 1 \) such that \( \alpha^k = 1 \). The order is denoted by

\[
\text{ord}(\alpha) = k.
\]

According to the examples in the preceding section, the order of \( \alpha = [5] \) in \( \mathbb{Z}_8^\times \) is 2. The order of \( \alpha = [3] \) in \( \mathbb{Z}_7^\times \) is 6.

The concept of order can be reframed in terms of integers and congruences. If \( a \) is an integer coprime with \( n \), the order of \( a \) modulo \( n \) is the order of the residue \( \alpha = [a] \) in \( \mathbb{Z}_n^\times \). In other words, the order of \( a \) is the smallest \( k \geq 1 \) such that

\[
a^k \equiv 1 \pmod{n}.
\]

Even though the order concept is really about integers, for doing proofs, it is convenient to work with the finite groups \( \mathbb{Z}_n^\times \).

Because of Euler’s theorem, it’s clear that \( \text{ord}(\alpha) \leq \varphi(n) \), for every \( \alpha \in \mathbb{Z}_n^\times \). But much more needs to be said about the order. The upcoming theory requires patient reading.

**Proposition 5.5.** Let \( \alpha \in \mathbb{Z}_n^\times \). A positive integer \( m \) satisfies \( \alpha^m = 1 \) if and only if \( \text{ord}(\alpha) | m \). Consequently,

\[
\text{ord}(\alpha) | \varphi(n).
\]

**Proof.** Let \( k = \text{ord}(\alpha) \), just to be brief. According to the Remainder Theorem write

\[
m = kq + r \text{ where } 0 \leq r < k.
\]

Then, since \( \alpha^k = 1 \), we obtain

\[
1 = \alpha^m = \alpha^{kq+r} = (\alpha^k)^q \alpha^r = 1^q \alpha^r = \alpha^r.
\]

By the minimality of \( k \), it must be that \( r = 0 \), and so \( k | m \).

The converse is easier. If \( m = kq \) for some integer \( q \), then

\[
\alpha^m = \alpha^{kq} = (\alpha^k)^q = 1^q = 1.
\]

Finally, since Euler guarantees that \( \alpha^{\varphi(n)} = 1 \), the above result applies to \( m = \varphi(n) \) and yields \( \text{ord}(\alpha) | \varphi(n) \).

\qed
Proposition 5.5 restricts the possible orders of a residue in \( \mathbb{Z}_n^* \) to the divisors of \( \varphi(n) \). For example, take \( n = 9 \). Here \( \varphi(9) = 6 \), and \( \mathbb{Z}_9^* \) consists of the six units 

\[ [1], [2], [4], [5], [7], [8]. \]

The possible orders of any one of these \( \alpha \) in \( \mathbb{Z}_9^* \) are the divisors of 6, which are 1, 2, 3, 6. For instance, take \( \alpha = [2] \). We have

\[ \alpha^1 = [2], \alpha^2 = [4], \alpha^3 = [8], \]

and already we see that \( \text{ord}(\alpha) = 6 \), because \( \alpha \) does not have order 1, 2, or 3.

**Proposition 5.6.** If \( \alpha \in \mathbb{Z}_n^* \) and \( k = \text{ord}(\alpha) \), then the list

\[ \alpha, \alpha^2, \alpha^3, \ldots, \alpha^{k-1}, \alpha^k = 1 \]

does not repeat itself.

**Proof.** Say we had a repetition \( \alpha^i = \alpha^j \) where \( 1 \leq i < j \leq k \). Thus

\[ \alpha^i \cdot 1 = \alpha^i \alpha^{j-i}. \]

Cancel the unit \( \alpha^i \) and get \( 1 = \alpha^{j-i} \). Since \( 1 \leq j - i < k \), this contradicts the minimality of \( k \) as the order of \( \alpha \).

Proposition 5.6 is useful in detecting primitive roots. For if we bump into an \( \alpha \) in \( \mathbb{Z}_n^* \) of order \( \varphi(n) \), then the list

\[ \alpha, \alpha^2, \alpha^3, \ldots, \alpha^{\varphi(n)-1}, \alpha^{\varphi(n)} = 1 \]

will have no repetitions. And since this list has \( \varphi(n) \) entries, it will pick up all of \( \mathbb{Z}_n^* \). This is worth recording.

**Proposition 5.7.** A residue \( \alpha \) in \( \mathbb{Z}_n^* \) is a primitive root if and only if \( \text{ord}(\alpha) = \varphi(n) \). In particular, for a prime \( p \), an \( \alpha \) in \( \mathbb{Z}_p^* \) is a primitive root if and only if \( \text{ord}(\alpha) = p - 1 \).

Next comes a more subtle result about the order of a residue and its powers.

**Proposition 5.8.** If \( \alpha \in \mathbb{Z}_n^* \) and \( k = \text{ord}(\alpha) \), then

\[ \text{ord}(\alpha^j) = k / \gcd(k, j). \]
Before going to the proof, let’s digest what Proposition 5.8 is saying by looking at examples.

Suppose \( n = 19 \). Here \( \varphi(n) = 18 \), and say we have an \( \alpha \) such that \( \text{ord}(\alpha) = 18 \). (In fact \( \alpha = [2] \) will do, but we can just call it \( \alpha \).)

What is \( \text{ord}(\alpha^3) \)? Well, for \( \beta = \alpha^3 \), the order of \( \beta \) is one of the divisors 1, 2, 3, 6, 9, 18 of 18, according to Proposition 5.5. So calculate the powers of \( \beta \) for these divisors and see what comes out. The powers of \( \beta \) for these divisors are:

\[
\beta = \alpha^3, \quad \beta^2 = \alpha^6, \quad \beta^3 = \alpha^9, \quad \beta^6 = \alpha^{18} = 1.
\]

We stopped at \( \beta^6 \) since that gave us our 1. Evidently

\[
\text{ord}(\beta) = 6 = 18/3 = 18/\gcd(18, 3).
\]

What about \( \text{ord}(\alpha^4) \)? Well, let \( \beta = \alpha^4 \). To find \( \text{ord}(\beta) \) we need the powers of \( \beta \) for the divisors of 18. Note that \( \alpha^{18} = 1 \), and so higher powers of \( \alpha \) will simplify. For instance, \( \alpha^{24} = \alpha^{18} \alpha^6 = \alpha^6 \). Thus,

\[
\beta = \alpha^4, \quad \beta^2 = \alpha^8, \quad \beta^3 = \alpha^{12}, \quad \beta^6 = \alpha^{24} = \alpha^6, \quad \beta^9 = \alpha^{18} = 1.
\]

Evidently

\[
\text{ord}(\alpha^4) = 9 = 18/2 = 18/\gcd(18, 4).
\]

What might \( \text{ord}(\alpha^5) \) be? For \( \beta = \alpha^5 \) we have

\[
\beta = \alpha^5, \quad \beta^2 = \alpha^{10}, \quad \beta^3 = \alpha^{15}, \quad \beta^6 = \alpha^{12}, \quad \beta^9 = \alpha^9.
\]

This partial table shows that \( \text{ord}(\beta) = 18 \). Why? Well, \( \text{ord}(\beta) \) has to be a divisor of 18, and the divisors 1, 2, 3, 6, 9 did not bring \( \beta \) down to 1. So

\[
\text{ord}(\alpha^5) = 18/1 = 18/\gcd(18, 5).
\]

Now we have to prove what the above examples illustrate. The proof demands some concentration.

**Proof.** It all hinges on Proposition 5.5. Let the order of \( \alpha^j \) be \( \ell \). What we want is \( \ell = k/\gcd(k, j) \). To get that, we will show that the integer on each side of this desired equation is a divisor of the integer on the other side.

Since \( \ell \) is the order of \( \alpha^j \) we get that

\[
\alpha^{j\ell} = (\alpha^j)^\ell = 1.
\]
By Proposition 5.5, applied to the order of \( \alpha \), we obtain that \( k \mid j\ell \). That is, \( j\ell = ku \) for some integer \( u \). And then,
\[
\frac{j}{\gcd(k, j)} \ell = \frac{k}{\gcd(k, j)} u.
\]
Since \( j/ \gcd(k, j) \) and \( k/ \gcd(k, j) \) are coprime, Proposition 1.7 yields that \( k/ \gcd(k, j) \) divides \( \ell \), which is one of the things we were looking for.

On the other hand, since \( k \) is the order of \( \alpha \), we also see that
\[
(\alpha^j)^{k/ \gcd(k, j)} = (\alpha)^{jk/ \gcd(k, j)} = (\alpha^k)^{j/ \gcd(k, j)} = 1/ \gcd(k, j) = 1.
\]
By Proposition 5.5, applied to the order of \( \alpha^j \), we obtain that \( \ell \mid k/ \gcd(k, j) \), which is the other thing we were looking for. \( \square \)

A special case of Proposition 5.8 needs to be singled out.

**Proposition 5.9.** For \( \alpha \) in \( \mathbb{Z}_n^\ast \) and \( j \) a positive integer,
\[
\text{ord}(\alpha^j) = \text{ord}(\alpha) \text{ if and only if } j \text{ is coprime with } \text{ord}(\alpha).
\]

**Proof.** Proposition 5.8 tells us that
\[
\text{ord}(\alpha^j) = \text{ord}(\alpha)/ \gcd(\text{ord}(\alpha), j).
\]
Clearly now
\[
\text{ord}(\alpha^j) = \text{ord}(\alpha) \text{ if and only if } \gcd(\text{ord}(\alpha), j) = 1,
\]
which is what we wanted. \( \square \)

And now a cute formula pops out.

**Proposition 5.10.** If \( \mathbb{Z}_n^\ast \) has a primitive root, then the total number of primitive roots in \( \mathbb{Z}_n^\ast \) is \( \varphi(\varphi(n)) \).

**Proof.** According to Proposition 5.7 the primitive roots of \( \mathbb{Z}_n^\ast \) are those \( \alpha \) for which \( \text{ord}(\alpha) = \varphi(n) \). If \( \alpha \) is one such primitive root, the powers
\[
\alpha, \alpha^2, \alpha^3, \ldots, \alpha^{\varphi(n)-1}, \alpha^\varphi(n) = 1
\]
cover all of \( \mathbb{Z}_n^\ast \) without repetition. The other primitive roots are those powers \( \alpha^j \) in the list for which
\[
\text{ord}(\alpha^j) = \varphi(n) = \text{ord}(\alpha).
\]
According to Proposition 5.9, these are the powers \( \alpha^j \) where \( j \) from 1 to \( \varphi(n) \) is coprime with \( \varphi(n) \). We know there are \( \varphi(\varphi(n)) \) such \( j \). \( \square \)
We caution that in order to apply Proposition 5.10, we have to first make sure \( \mathbb{Z}_n^* \) has a primitive root. For instance, we may recall that \( \mathbb{Z}_8^* \) had no primitive roots at all. Thus it would be a mistake to say that the number of primitive roots of \( \mathbb{Z}_8^* \) is \( \phi(\phi(8)) = \phi(4) = 2 \).

**Example 5.11.** Take \( n = 17 \), and then \( \phi(n) = 16 \). Let’s first verify that \( \alpha = [3] \) is primitive. The possible orders of \( \alpha \) are 1, 2, 4, 8, 16. Now
\[
\alpha = [3], \alpha^2 = [9], \alpha^4 = [81] = [13] \text{ and } \alpha^8 = [169] = [16] = [−1]
\]
in \( \mathbb{Z}_{17}^* \). Since none of 1, 2, 4, 8 equals \( \text{ord}(\alpha) \), then \( 16 = \text{ord}(\alpha) \) and thus \( \alpha = [3] \) is primitive. According to Proposition 5.10, \( \mathbb{Z}_{17}^* \) has \( \phi(16) = 8 \) primitive roots in all.

**Example 5.12.** Does \( \mathbb{Z}_{50}^* \) have a primitive root and if so how many are there?

First obtain
\[
\phi(50) = \phi(5^2 \cdot 2) = \phi(5^2)\phi(2) = 20.
\]
We need to find in \( \mathbb{Z}_{50}^* \) an element of order 20. Our best bet is trial and error. Note that \( [2] \notin \mathbb{Z}_{50}^* \) since 2 and 50 are not coprime. Hoping for good luck, let’s see if \( \alpha = [3] \) is primitive. The possible orders of \( \alpha \) are 1, 2, 4, 5, 10, 20, the divisors of 20. Let’s calculate these powers of \( \alpha \) in \( \mathbb{Z}_{50}^* \).
\[
\begin{align*}
\alpha^1 &= [3] & \alpha^2 &= [9] & \alpha^4 &= [81] = [31] \\
\alpha^5 &= [93] = [43] & \alpha^{10} &= [1849] = [49] = [−1] & \alpha^{20} &= [1]
\end{align*}
\]
Thus \( \text{ord}(\alpha) = 20 \), and \( \alpha = [3] \) primitive. According to Proposition 5.10, the number of primitive roots in \( \mathbb{Z}_{50}^* \) is
\[
\phi(20) = \phi(5 \cdot 4) = \phi(5)\phi(4) = 4 \cdot 2 = 8.
\]
In \( \mathbb{Z}_{50}^* \) there are 8 primitive roots. They consist of those \( \alpha^j \) where \( j \) is coprime with 20. Namely,
\[
\begin{align*}
\alpha &= [3], & \alpha^3 &= [27], & \alpha^7 &= [37], & \alpha^9 &= [33], \\
\alpha^{11} &= [47], & \alpha^{13} &= [23], & \alpha^{17} &= [13], & \alpha^{19} &= [17].
\end{align*}
\]

### 5.3 Polynomials over \( \mathbb{Z}_p \)

It’s time go deeper in our effort to prove that if \( p \) is prime, then \( \mathbb{Z}_p^* \) has a primitive root. Some of the upcoming material is quite demanding.
5.3. POLYNOMIALS OVER \(\mathbb{Z}_p\)

We need to digress into a discussion of polynomials over \(\mathbb{Z}_p\) when \(p\) is prime. The importance of \(p\) being a prime is that every \(\alpha\) in \(\mathbb{Z}_p\), except for 0, is a unit. In the parlance of algebra we would say that \(\mathbb{Z}_p\) is a field. Thus,

\[
\mathbb{Z}_p^* = \{[1], [2], \ldots, [p-1]\}.
\]

If \(\alpha, \beta \in \mathbb{Z}_p^*\), then so is \(\alpha\beta\) in \(\mathbb{Z}_p^*\). That is, if \(\alpha \neq 0\) and \(\beta \neq 0\), then \(\alpha\beta \neq 0\). (Here 0 means \([0]\).) In still other words, \(\mathbb{Z}_p^*\) is group.

**Definition 5.13.** Any expression of the form

\[
\alpha_n X^n + \alpha_{n-1} X^{n-1} + \cdots + \alpha_1 X + \alpha_0
\]

where \(n \geq 0\), \(\alpha_j \in \mathbb{Z}_p\) and \(\alpha_n \neq 0\) (i.e. \(\alpha_n \in \mathbb{Z}_p^*\)) is called a **non-zero polynomial in \(X\) of degree \(n\) with coefficients in \(\mathbb{Z}_p\).** We also say this is a polynomial over \(\mathbb{Z}_p\).

The degree is written as \(n = \deg f(X)\).

We can add, subtract, multiply and factor polynomials with coefficients in \(\mathbb{Z}_p\) pretty much the same as we do with polynomials having coefficients in \(\mathbb{Q}\). For instance, over \(\mathbb{Z}_5\) we have

\[
(3X^2 + 2X + 4)(4X^3 + X + 2) = 12X^5 + 3X^3 + 6X^2 + 8X^4 + 2X^2 + 4X + 16X^3 + 4X + 8
\]
\[
= 12X^5 + 8X^4 + 19X^3 + 8X^2 + 8X + 8
\]
\[
= 2X^5 + 3X^4 + 4X^3 + 3X^2 + 3X + 3
\]

Note that, for tidiness of appearance, we omitted the box notation in the coefficients above. These coefficients are actually in \(\mathbb{Z}_5\), even though they look like integers. Consequently, we invoked the replacement principle liberally.

We should notice right away that the degree of the product of two polynomials equals the sum of the degrees of the original polynomials. That is, if \(f(X), g(X)\) are polynomials over \(\mathbb{Z}_p\), then

\[
\deg f(X)g(X) = \deg f(X) + \deg g(X).
\]

To see this, let

\[
f(X) = \alpha_n X^n + \cdots + \alpha_1 X + \alpha_0 \quad \text{and} \quad g(X) = \beta_m X^m + \cdots + \beta_1 X + \beta_0,
\]

where \(\alpha_j, \beta_j \in \mathbb{Z}_p\), and \(\alpha_n, \beta_m\) are not 0. Then

\[
f(X)g(X) = \alpha_n \beta_m X^{n+m} + \text{terms of lower degree}.
\]
Also $\alpha_n\beta_m \neq 0$ because $\alpha_n, \beta_m$ are in the group $\mathbb{Z}_p^*$. By looking, we see that the degree of the product equals the sum of the degrees.

A residue $\beta$ in $\mathbb{Z}_p$ is called a root of the polynomial

$$f(X) = \alpha_nX^n + \cdots + \alpha_1X + \alpha_0$$

when

$$f(\beta) = \alpha_n\beta^n + \cdots + \alpha_1\beta + \alpha_0 = 0 \text{ in } \mathbb{Z}_p.$$

For example, every $\beta$ in $\mathbb{Z}_p^*$ is a root of $X^{p-1} - 1$, due to Fermat’s Little Theorem.

In order to contrast with what’s coming up, notice that the polynomial


has two roots. Namely, $[2]$ and $[5]$. But its degree is only 1. This goes against our experience with polynomials over $\mathbb{Q}$ where the number of roots never goes above the degree. The problem here is that 6 is not a prime. When $p$ is a prime, things get back to normal, and polynomials over $\mathbb{Z}_p$ never have more roots than their degree allows.

**Proposition 5.14.** If $p$ is prime, and $f(X)$ is a polynomial over $\mathbb{Z}_p$ of degree $n$, then $f(X)$ has at most $n$ roots in $\mathbb{Z}_p$.

**Proof.** Let’s do a proof by induction on the degree of $f(X)$.

If $\deg f(X) = 0$, we have $f(X) = \alpha_0$ and $\alpha_0 \neq 0$ in $\mathbb{Z}_p$. It’s clear that such a constant polynomial has no root at all, since the only possible value such a polynomial can take is the constant, non-zero value $\alpha_0$. So the number of roots, which is 0, is no more than the degree, which is also 0.

Supposing all polynomials over $\mathbb{Z}_p$ of degree less than $n$ have no more roots than their degree allows, consider a polynomial

$$f(X) = \alpha_nX^n + \alpha_{n-1}X^{n-1} + \cdots + \alpha_1X + \alpha_0$$

of degree $n$. If $f(X)$ has no roots, then surely $0 \leq n$. Otherwise, $f(X)$ has a root, say $\beta$. Then

$$f(X) = f(X) - 0$$

$$= f(X) - f(\beta)$$

$$= \alpha_n(X^n - \beta^n) + \alpha_{n-1}(X^{n-1} - \beta^{n-1}) + \cdots + \alpha_1(X - \beta).$$
Notice the factorization
\[ X^k - \beta^k = (X - \beta)(X^{k-1} + X^{k-2}\beta + X^{k-3}\beta^2 + \cdots + X\beta^{k-2} + \beta^{k-1}). \]

Thus, by inspection, the polynomial \( f(X) \) has \( X - \beta \) as a factor. Indeed, \( X - \beta \) is a factor of each of the summands \( \alpha_k(X^k - \beta^k) \). Let’s write
\[ f(X) = (X - \beta)g(X) \]
for some other polynomial \( g(X) \) over \( \mathbb{Z}_p \).

Since the degree of a product is the sum of the degrees of the factors, we see that \( \deg g(X) = n - 1 \). By the inductive hypothesis \( g(X) \) has at most \( n - 1 \) roots.

Now any root \( \gamma \) of \( f(X) \) other than \( \beta \) is such that
\[ 0 = f(\gamma) = (\gamma - \beta)g(\gamma). \]

Since \( \gamma - \beta \neq 0 \), it follows that \( g(\gamma) = 0 \). Indeed, the alternative would be that \( g(\gamma) \neq 0 \), and then we would have two non-zero elements of \( \mathbb{Z}_p \) multiplying out to 0, which never happens in \( \mathbb{Z}_p \) when \( p \) is prime.

Having seen that a root of \( f(X) \) is either \( \beta \) or a root of \( g(X) \), and that \( g(X) \) has at most \( n - 1 \) roots, it becomes clear that \( f(X) \) has at most \( n \) roots.
5.4 The order of a product and the main theorem

The proof of the next result is very subtle. It hinges on the fact that if \( \gamma \in \mathbb{Z}_p^* \) and if there are coprime exponents \( m, n \) such that \( \gamma^n = 1 \) and \( \gamma^m = 1 \), then \( \gamma \) is already the residue 1. (To keep notations clean we write 1 instead of \([1]\) to denote this residue.) To see this claim just invoke Proposition 5.5. It tells us that the order of \( \gamma \) divides both \( m \) and \( n \), which are coprime. Thereby the order of \( \gamma \) is 1, meaning that \( \gamma \) is the residue 1. For another way to see this fact, just notice that

\[
ms + nt = 1 
\]

for some integers \( s, t \).

Then

\[
\gamma = \gamma^1 = \gamma^{ms + nt} = (\gamma^m)^s(\gamma^n)^t = 1^s1^t = 1.
\]

We used the usual laws of exponents, which do apply in \( \mathbb{Z}_p^* \).

**Proposition 5.16.** Let \( p \) be a prime and let \( \alpha, \beta \in \mathbb{Z}_p^* \) have orders \( k, \ell \) respectively. If \( k, \ell \) are coprime, then

\[
\text{ord}(\alpha\beta) = k\ell.
\]

**Proof.** Let \( m = \text{ord}(\alpha\beta) \). Since

\[
(\alpha\beta)^{k\ell} = \alpha^{k\ell}\beta^{k\ell} = (\alpha^k)^\ell(\beta^\ell)^k = 1^\ell1^k = 1,
\]

we see from Proposition 5.5 that \( m \mid k\ell \).

To get \( m = k\ell \), it suffices to prove \( k\ell \mid m \), and for that it suffices to prove \( k \mid m \) and \( \ell \mid m \), due to the fact \( k, \ell \) are coprime.

Clearly,

\[
(\alpha^m)^k = \alpha^{mk} = (\alpha^k)^m = 1^m = 1.
\]

And likewise,

\[
(\beta^m)^\ell = \beta^{m\ell} = (\beta^\ell)^m = 1^m = 1.
\]

But also notice that

\[
(\alpha^m)^\ell = (\alpha^m)^\ell \cdot 1 = (\alpha^m)^\ell(\beta^m)^\ell = (\alpha^m\beta^m)^\ell \text{ by the usual exponent laws} = (\alpha\beta)^{\ell m} \text{ by exponent laws} = 1^\ell \text{ since } m \text{ is the order of } \alpha\beta = 1.
\]
Since $k, \ell$ are coprime it follows, as noted just before this result, that $\alpha^m = 1$. By Proposition 5.5, $k \mid m$. For identical reasons, $\ell \mid m$ too, and we are done. □

The primitive root theorem

Here is a major result about $\mathbb{Z}_p^*$ when $p$ is a prime. As is typical of important discoveries, the proof is not so plain to see.

**Theorem 5.17.** If $p$ is a prime, then $\mathbb{Z}_p^*$ has a primitive root.

**Proof.** Start with any $\alpha$ in $\mathbb{Z}_p^*$.

- If $\text{ord} \alpha = p - 1$, then our $\alpha$ is primitive, and we are done.
- If $\text{ord} \alpha = k < p - 1$, we present a way to find some $\beta$ whose order is bigger than the order of $\alpha$.

According to Proposition 5.15, the list $\alpha, \alpha^2, \ldots, \alpha^k = 1$ picks up all roots of $X^k - 1$ in $\mathbb{Z}_p^*$. Since $k < p - 1$, there is some $\gamma$ in $\mathbb{Z}_p^*$, but not on the list. Thus $\gamma^k \neq 1$.

Let $\ell = \text{ord}(\gamma)$. Notice that if $\ell \mid k$, then $\gamma^k = 1$. So $\ell \nmid k$. Thus, in the unique factorizations of $k$ and $\ell$, there is a prime $q$ that appears more often in $\ell$ than it does in $k$. Thus we may write

$$k = q^d k_1 \text{ and } \ell = q^e \ell_1 \text{ where } 0 \leq d < e \text{ and } q \nmid k_1, q \nmid \ell_1.$$

Let $\beta = \alpha^{q^d} \gamma^{\ell_1}$. Proposition 5.8 tells us that

$$\text{ord}(\alpha^{q^d}) = k / \text{gcd}(k, q^d) = k / q^d = k_1,$$

and

$$\text{ord}(\gamma^{\ell_1}) = \ell / \text{gcd}(\ell, \ell_1) = \ell / \ell_1 = q^e.$$

But $k_1$ and $q^e$ are coprime. By Proposition 5.16 we get

$$\text{ord} \beta = q^e k_1 > q^d k_1 = k = \text{ord} \alpha.$$

In this way, new elements of strictly increasing order can be found in $\mathbb{Z}_p^*$, until we hit an element of order $p - 1$. And that’s when we get our primitive root. □

It might be worthwhile to interpret the primitive root theorem in terms of congruences. Here is that interpretation, hoping that the translation from residues in $\mathbb{Z}_p^*$ to integers is clear.
Theorem 5.18. If $p$ is a prime, then there is an integer $g$ coprime with $p$ and such that every integer from 1 to $p - 1$ is congruent modulo $p$ to one of the powers $g, g^2, g^3, \ldots, g^{p-1}$.

Obviously $g$ gives the primitive root $[g]$ in $\mathbb{Z}_p^*$. The integer $g$ itself is also known as a primitive root modulo $p$.

Actually finding a primitive root

In practice, for primes that are not too big, primitive roots are found by trial and error. We can start start with [2], then they try [3], then [5], and so on hoping to hit a winner quickly. (By the way, there was no point in trying [4] because this is already a power of 2.)

How might we, in practice, test if a candidate $\alpha$ is primitive in $\mathbb{Z}_p^*$ when $p$ starts to get big? We have to determine the order of $\alpha$. The square and multiply algorithm lets us calculate a given power $\alpha^k$ in $\mathbb{Z}_p^*$ pretty efficiently. But if $p$ is huge, there are too many exponents $k$ to test in order to find the order of $\alpha$. One way to expedite matters a bit, is to factor $p - 1$ into primes:

$$p - 1 = p_1 \cdot p_2 \cdot p_3 \cdots p_n.$$ 

This in itself could be a practical challenge, but suppose we succeeded. Then we simply calculate the powers $\alpha^{(p-1)/p_j}$ for each $j = 1, 2, \ldots, n$.

If $k = \text{ord}(\alpha)$, we know that $k \mid p - 1$. If $k < p - 1$, then $k$ will divide one of the exponents $(p - 1)/p_j$. This is because a proper divisor of $p - 1$ has to have at least one prime appearing in its unique factorization less often than it appears in $p - 1$. Since $k$ divides some $\frac{p-1}{p_j}$, one of $\alpha^{(p-1)/p_j}$ will come out to equal 1. Thus, if none of $\alpha^{(p-1)/p_j}$ equals 1, we know that $\alpha$ is a primitive root. This method is pretty fast on a computer, but it hinges on factoring the very large number $p - 1$.

A famous conjecture of Emil Artin back in the 1920’s is that if $g$ is a prime, then there are infinitely many primes $p$ for which the residue residue $[g]$ is a primitive root in $\mathbb{Z}_p^*$. Two Canadian mathematicians, together with the work of an Englishman, showed that this holds for all but maybe two primes $g$. The exceptional two, if they even exist, are not known. One way to become a famous mathematician would be to actually prove Artin’s conjecture.
5.5. CARMICHAEL NUMBERS

We shall have at least three occasions to use the Primitive Root Theorem, in this and subsequent chapters.

5.5 Carmichael numbers

It’s not easy to tell, in practice, if a large number $n$ is prime. It’s fairly clear that if $n$ is not a prime, then one of its prime factors has to be at most $\sqrt{n}$. For if all prime factors were more than $\sqrt{n}$, then the presence of at least two such prime factors would make their product bigger than $n$. So we could plod along and consider every integer $k$ from 1 to $\sqrt{n}$ and see if $k \mid n$. And if that never happened, we could safely say that $n$ is prime. However, when $n$ is huge, this is far too much work, even for a fast computer.

The variant of Fermat’s Theorem 3.19 can help in part. It told us that if $n$ is prime, then

$$a^n \equiv a \mod n,$$

for every integer $a$.

In other words, if

$$a^n \not\equiv a, \text{ for some integer } a,$$

then $n$ is not a prime. This is an interesting and definitive way to decide that an integer $n$ is not prime. Just find an integer $a$ such that $a^n \not\equiv a \mod n$. The reduction of $a^n$ is completely feasible by the square and multiply algorithm.

Definition 5.19. An integer $a$ is called a witness for the non-primality of $n$ when $a^n \not\equiv a \mod n$.

If we can find such a witness for a given $n$, then we have an outstanding and efficient test for non-primality. If we sample quite a few composite integers $n$, we often succeed in producing a witness. For instance, take $n = 91$ and see if $a = 2$ is a witness. We need to reduce $2^{91} \mod 91$. In accordance with the square and multiply method first observe that

$$91 = 64 + 16 + 8 + 2 + 1,$$

and so

$$2^{91} = 2^{64} \cdot 2^{16} \cdot 2^{8} \cdot 2^{2} \cdot 2^{1}.$$
With the help of a calculator or some software package, start with 2 and keep squaring modulo 91:

\[
\begin{align*}
2 & \equiv 2 \\
2^2 & \equiv 4 \\
2^8 & \equiv 256 \equiv 74 \\
2^{16} & \equiv 74^2 \equiv 16 \\
2^{32} & \equiv 16^2 \equiv 74 \\
2^{64} & \equiv 74^2 \equiv 16
\end{align*}
\]

Thus

\[2^{91} \equiv 16 \cdot 16 \cdot 74 \cdot 4 \cdot 2 = 15155 \equiv 37 \mod 91.\]

Since \(2^{91} \not\equiv 2 \mod 91\), we discover that 91 is not a prime. The above could all be done in an instant as well by using a computer algorithm, such as the one we saved in Excel.

Of course, we knew all along that \(91 = 13 \cdot 7\), but when the integer \(n\) gets big, this is something that won’t be so obvious, and finding a witness for the non-primality of \(n\) might be lot easier than finding a factor of \(n\). It’s also maddening that finding a witness for \(n\) not to be prime provides us with no clue as to what the factors of \(n\) could be. It’s a lot easier to discover that a number is composite, than it is to actually factor it.

The first Carmichael number

If every non-prime had a witness, things would be dandy. Until we come to the strange case of

\[n = 561 = 3 \cdot 11 \cdot 17.\]

It turns out that this non-prime has no witnesses. Indeed, consider any integer \(a\).

If \(3 \nmid a\), Fermat’s theorem for the prime 3 gives:

\[a^{561} = a^{280 \cdot 2 + 1} = (a^{280})^2 \cdot a \equiv 1 \cdot a = a \mod 3.\]

And if \(3 \mid a\), then

\[a^{561} \equiv 0 \equiv a \mod 3.\]

So \(a^{561} \equiv a \mod 3\) regardless of \(a\).

If \(11 \nmid a\), Fermat for the prime 11 gives:

\[a^{561} = a^{56 \cdot 11 + 1} = (a^{56})^{10} \cdot a \equiv 1 \cdot a.\]

And if \(11 \mid a\), then

\[a^{561} \equiv 0 \equiv a \mod 11.\]
5.5. CARMICHAEL NUMBERS

So \( a^{561} \equiv a \mod 11 \) regardless of \( a \).

If \( 17 \nmid a \), Fermat for the prime 17 gives:

\[
a^{561} = a^{35 \cdot 16 + 1} = (a^{35})^{16} \cdot a \equiv 1 \cdot a = a \mod 17.
\]

And if \( 17 \mid a \), then

\[
a^{561} \equiv 0 \equiv a \mod 17.
\]

So \( a^{561} \equiv a \mod 17 \) regardless of \( a \).

We have just seen that for all \( a \),

\[
a^{561} \equiv a \mod 3, \text{ and } \mod 11, \text{ and } \mod 17.
\]

Thus

\[
a^{561} \equiv a \mod 3 \cdot 11 \cdot 17,
\]

since 3, 11, 17 are pairwise coprime and Proposition 4.1 applies. Thus we have shown

\[
a^{561} \equiv a \mod 561 \text{ for all } a.
\]

In other words, 561 is pretending to be a prime and there are no witnesses to prove otherwise.

**Definition 5.20.** An integer \( n \) is called a **Carmichael number** whenever \( n \geq 2 \) and \( n \) is not a prime, but still

\[
a^n \equiv a \mod n \text{ for all } a.
\]

Robert Carmichael, an American mathematician, was first to discover in 1910 that 561 had no witnesses. Let’s now see how we might go about discovering more of these strange beasts.

**A test for Carmichael numbers**

The next result tells us how to spot a Carmichael number by looking at its unique factorization. This result is known as Korselt’s criterion. The German Alwin Korselt found this test in 1899, well before any Carmichael numbers had even been discovered.

**Theorem 5.21.** An integer \( n \) is a **Carmichael number** if and only if
• \( n = p_1 \cdot p_2 \cdots p_k \) where the \( p_j \) are primes without repetitions.

• every \( p_j - 1 \) divides \( n - 1 \), and

• \( k > 1 \),

**Proof.** First suppose that \( n \) satisfies the three conditions.

The \( k > 1 \) guarantees that \( n \) is not prime. In order to ensure that for every integer \( a \) we get \( a^n \equiv a \mod n \), it suffices, by Proposition 4.1, to see that

\[
a^n \equiv a \mod p_j \text{ for every prime factor } p_j \text{ of } n.
\]

Note how the non-repetition of the \( p_j \) is crucial in the preceding observation.

We are given that \( n - 1 = (p_j - 1)\ell_j \) for some \( \ell_j \). If \( p_j \nmid a \), Fermat’s Theorem 3.18 gives

\[
a^n = a^{p_j - 1}a = (a^1)^{p_j - 1}a \equiv 1 \cdot a = a \mod p_j.
\]

And if \( p_j | a \) we easily get

\[
a^n \equiv 0 \equiv a \mod p_j.
\]

Since \( a^n \equiv a \mod p_j \) for all \( a \), our \( n \) is a Carmichael number.

Next suppose that \( n \) is a Carmichael number.

To show the first of the three desired conditions, assume to the contrary that \( p \) is a prime factor of \( n \) with multiplicity more than 1. Thus \( p^2 \mid n \). We know that \( p^n \equiv p \mod n \) since \( n \) is Carmichael. Because

\[
p^2 \mid n \text{ and } n \mid p^n - p = p(p^{n-1} - 1),
\]

it follows that \( p^2 \mid p(p^{n-1} - 1) \), which forces \( p \mid p^{n-1} - 1 \). From this we get the contradiction that \( p \mid 1 \). Thus the unique factorization of \( n \) has no repeated primes.

For the second condition, we need to show that every prime factor \( p \) of the Carmichael number \( n \) satisfies \( p - 1 \mid n - 1 \). This is where we invoke the Primitive Root Theorem 5.17 within a delicate argument. Let \( a \) be such that \( [a] \) is a primitive root of \( \mathbb{Z}_p^* \). From \( a^n \equiv a \mod n \) we get \( a^n \equiv a \mod p \). So \( [a]^n = [a] \) in \( \mathbb{Z}_p^* \). Then \( [a][a]^{n-1} = [a] \) in \( \mathbb{Z}_p^* \). Cancel the unit \( [a] \) to get \( [a]^{n-1} = [1] \) in \( \mathbb{Z}_p^* \). By Proposition 5.5, \( \text{ord}([a]) \) divides \( n - 1 \). Since \( [a] \) is a primitive root in \( \mathbb{Z}_p^* \), we have \( \text{ord}([a]) = p - 1 \). Hence \( p - 1 \mid n - 1 \).

For the last item, note that the Carmichael number \( n \) is not a prime, and so \( k > 1 \). \( \square \)
For instance, \( n = 1105 = 5 \cdot 13 \cdot 17 \) is Carmichael, because
\[
4 \mid 1104, \ 12 \mid 1104 \ \text{and} \ \ 16 \mid 1104.
\]
Also \( n = 2821 = 7 \cdot 13 \cdot 31 \) is Carmichael since
\[
6 \mid 2820, \ 12 \mid 2820 \ \text{and} \ \ 30 \mid 2820.
\]
It was only verified in 1994 that there exist infinitely many Carmichael numbers. Also, it was proven that for all large \( n \) there are at least \( n^{2/7} \) Carmichael numbers between 1 and \( n \). For instance with \( n = 10^6 \), we have \( n^{2/7} \approx 52 \). There are at least 52 Carmichael numbers from 1 to one-million.

In the above discussion of Carmichael numbers we needed to get our hands on a primitive root in order to prove Theorem 5.21. In the next section we demonstrate that primitive roots are not only needed to develop the theory of numbers, but they also have practical value.

## 5.6 Primitive roots and cryptography

Let \( p \) be a huge prime. For example, \( p \) could have 200 decimal digits. There is a primitive root \( \alpha \) for \( \mathbb{Z}_p^* \). Letting \( \alpha = [g] \) for some representative integer \( g \), this means that the powers
\[
g, g^2, g^3, \ldots, g^{p-1} \equiv 1,
\]
after reduction modulo \( p \), never duplicate, and will pick up every possible remainder
\[
1, 2, \ldots, p - 1.
\]
Apparently it takes some work to obtain such \( p \) and \( g \), but tables of huge primes along with their primitive roots are out there. The first list, reduced modulo \( p \), is a horrible scrambling of the second list of possible remainders. By the square and multiply method we can readily reduce \( g^\ell \) modulo \( p \) for any exponent \( \ell \) from 1 to \( p - 1 \). But, to this date, no one has yet found an effective way to find \( \ell \), given the reduced form \( f \) (i.e. remainder modulo \( p \)) of \( g^\ell \).

**Definition 5.22.** Let \( p \) be a prime and \( g \) a primitive root for \( p \). That is, \([g]\) is a primitive root in \( \mathbb{Z}_p^* \). If \( f \) is a remainder modulo \( p \) between 1 and \( p - 1 \), the **discrete logarithm** of \( f \) relative to the primitive root \( g \) is the unique exponent \( \ell \) between 1 and \( p - 1 \) such that
\[
g^\ell \equiv f \mod p.
\]
This $\ell$ is also known as the index of $f$ relative to $g$.

**The difficulty of computing discrete logarithms**

As noted just above, for a huge prime $p$ with primitive root $g$, there is no known effective algorithm that can compute the discrete logarithm of an integer $f$ relative to $g$.

To get an idea of how hard it is to find the discrete logarithm $\ell$, suppose $f$ is given between 1 and $p - 1$. With a powerful computer we decide to reduce the powers $g, g^1, g^2, g^3, g^4, \ldots$ in succession modulo $p$. Then we compare each power as we go along to see if we get $f$. How long would that take? Say $p$ has approximately 200 digits, and so $p - 1 \approx 10^{200}$. Also suppose that in one second we can reduce one trillion ($10^{12}$) of the powers $g^j$ and compare to see if we hit upon $f$. There are much fewer than 100 million ($10^8$) seconds in a year. So in one year we would have reduced and compared much fewer than

$$10^{12} \cdot 10^8 = 10^{20}$$

of the powers $g^j$.

There are approximately $10^{200}$ non-congruent powers to check. So after one year, the proportion of the powers we can check, out of all possible powers, is less than

$$\frac{10^{20}}{10^{200}} = \frac{1}{10^{180}}.$$  

In one year we have barely started to scratch the surface in our search. We could get lucky and hit upon $f$ in one year. But the probability of hitting $f$ in one year is less than $1/10^{180}$. The chances of that happening are much less than the chances of winning the top prize in LOTTO-649, 11 times in a row. We do not have good odds.

**The Diffie-Hellman key exchange**

This difficulty of finding the discrete logarithm makes it possible to design some pretty effective cryptographic schemes. We present one famous such protocol, named after the American cryptographers Whitfield Diffie and Martin Hellman, who discovered it back in the 1970’s.
5.6. PRIMITIVE ROOTS AND CRYPTOGRAPHY

As usual, we have our protagonists, Alice and Bob, who need to exchange a secret, without Eve, the eavesdropper, being able to obtain it. Recall that a message, after a simple and openly agreed upon encoding, is a just a number.

The first thing Alice and Bob do is exchange a key right under Eve’s nose. Here is how they do it.

• They publicly agree on a huge prime $p$ with primitive root $g$. Thus $g$ is an integer whose reduced powers

$$g, g^2, \ldots, g^{p-1}$$

pick up all the remainders

$$1, 2, \ldots, p - 1.$$ 

• Then Alice picks an integer $\ell$ at random from 1 to $p - 1$. Bob also picks a random integer $k$ from 1 to $p - 1$. (Preferably not 1 or $p - 1$.) There is an enormous number of integers from which they can pick. Only Alice knows $\ell$, and only Bob knows $k$.

• Alice reduces $g^\ell$ modulo $p$ down to $f$ by squaring and multiplying, and publishes $f$. Thus

$$f \equiv g^\ell \mod p$$

where $1 \leq f \leq p - 1$.

Bob, Eve and anybody else can see $f$.

• Bob does the same thing with his secret $k$. He reduces $g^k$ modulo $p$ down to $h$ and publishes $h$. Thus

$$h \equiv g^k \mod p$$

where $1 \leq h \leq p - 1$.

Everybody can see $h$.

• Eve’s frustration is that she is unable to obtain the discrete logarithms $\ell, k$ of $f, h$, respectively, because an effective way to carry out such a calculation for huge primes is not currently available.

• Now Bob takes the $f$, which Alice has published, and, using his secret logarithm $k$, he reduces $f^k$ down to some $w$ modulo $p$. Thus,

$$w \equiv f^k \equiv (g^\ell)^k = g^{\ell k} \mod p,$$

and $1 \leq w \leq p - 1$. 

Alice does the same thing. She takes the $h$, which Bob has published, and, using her secret logarithm $\ell$, she reduces $h^\ell$ down to some $z$ modulo $p$. Thus,

$$z \equiv h^\ell \equiv (g^k)^\ell = g^{k\ell} \mod p, \quad \text{and} \quad 1 \leq z \leq p - 1.$$ 

Since $g^{k\ell} = g^{\ell k}$ and since both $w$ and $z$ are congruent to this power of $g$ we have that $w \equiv z \mod p$. And because both $w$ and $z$ are between 1 and $p - 1$ we conclude that

$$w = z.$$ 

Thus Alice and Bob are in possession of a common number $z$ between 1 and $p - 1$, while Eve is unable to find this number.

Now that Alice and Bob have exchanged the key $z$, they separately calculate the inverse $u$ of $z$ modulo $p$. That is, they solve

$$uz \equiv 1 \mod p \text{ for } u.$$ 

This is readily doable by means of the Euclidean Algorithm.

They can now use the key to exchange messages. Since all messages can be encoded openly as numbers, and since long messages can be blocked off into pieces, we can take a message to be a number $x$ between 1 and $p - 1$. Suppose Bob wants to send such a message $x$ to Alice in encrypted form. He simply reduces $zx$ modulo $p$ down to $y$ and sends out $y$, which is an encryption of $x$. Thus

$$y \equiv zx \mod p \text{ and } 1 \leq y \leq p - 1.$$ 

Now Alice takes the $y$ and reduces $uy$ modulo $p$ to get

$$uy \equiv uzx \equiv 1 \cdot x = x.$$ 

Alice has decrypted $y$ to recover the message $x$.

Eve, in the meantime, cannot get $x$ from $y$ because she does not possess the key $z$, nor its inverse $u$.

Here is an illustration of the Diffie-Hellman scheme using a rather small prime.
Example 5.23. Alice wants Professor Bob to send her MATH grade by e-mail. Since her mother Eve checks the e-mail, Alice wants her grade to be encrypted by Bob. Eve is known to be unable to figure out discrete logarithms, and so a Diffie-Hellman key exchange will work.

- Alice chooses the prime $p = 101$ with primitive root $g = 2$.
  
  To be sure, let’s check $2$ is primitive modulo $101$. Its possible orders modulo $101$ are the divisors of $100$. These are $1, 2, 4, 5, 10, 20, 25, 50, 100$. Modulo $101$ we get:
  
  $2 \equiv 2$ \quad $2^2 \equiv 4$ \quad $2^4 \equiv 16$

  $2^5 \equiv 32$ \quad $2^{10} \equiv 14$ \quad $2^{20} \equiv 95$

  $2^{25} \equiv 10$ \quad $2^{50} \equiv 100 \equiv -1$ \quad $2^{100} \equiv 1$.

  So the order of $2$ is indeed $100$.

  Alice publishes $p = 101$ and $g = 2$ for Bob and Eve and anybody to see. Bob, and even Eve, know that these will be used for a Diffie-Hellman key exchange.

- Next, Bob and Alice each pick an integer at random from $2$ to $99$. Alice picks $\ell = 41$, and reveals it to no one. Bob picks $k = 17$, and reveals it to no one.

- Alice reduces $2^{41}$ modulo $101$ to get:
  
  $2^{41} \equiv (2^{20})^2 \cdot 2 \equiv 95^2 \cdot 2 \equiv 72 \mod 101$.

  She publishes $f = 72$ for Bob and all to see.

- Bob reduces $2^{17}$ modulo $101$ to get:
  
  $2^{17} \equiv 2^{10} \cdot 2^5 \cdot 2^2 \equiv 14 \cdot 32 \cdot 4 \equiv 75 \mod 101$.

  He publishes $h = 75$ for Alice and all to see.

- Alice reduces $75^{41}$ modulo $101$, and gets $z = 38$.

  Let’s do the work for her to make sure she’s right.
Note $41 = 32 + 8 + 1$, and so $75^{41} \equiv 75^{32} \cdot 75^8 \cdot 75^1$. Squaring and multiplying modulo 101 we obtain:

\begin{align*}
75^2 & \equiv 70 \\
75^4 & \equiv 70^2 \equiv 52 \\
75^8 & \equiv 52^2 \equiv 78 \\
75^{16} & \equiv 78^2 \equiv 24 \\
75^{32} & \equiv 24^2 \equiv 71
\end{align*}

And thus

$75^{41} \equiv 71 \cdot 78 \cdot 75 \equiv 38 \mod 101$.

- Bob reduces $72^{17}$ modulo 101, and also gets $w = 38$, as expected. Let’s do the work for Bob to make sure he’s right.

Note $72^{17} \cdot 85 = 72^{16} \cdot 72 \cdot 85$, Squaring and multiplying modulo 101 we obtain:

\begin{align*}
72^2 & \equiv 33 \\
72^4 & \equiv 33^2 \equiv 79 \\
72^8 & \equiv 79^2 \equiv 80 \\
72^{16} & \equiv 80^2 \equiv 37
\end{align*}

Then

$72^{17} \equiv 37 \cdot 72 \equiv 38 \mod 101$.

- Eve does not know $z = 38$, since she could not compute the logarithms 41, 17 of 72, 75, respectively.

Aside. Of course, with such a small prime as 101, Eve could easily get these discrete logarithms. But let’s pretend 101 is a huge prime, and that Eve’s computer cannot find the discrete logarithms. That’s the way it is with truly huge primes.

- Alice next solves

$$u \cdot 38 \equiv 1 \mod 101.$$  

With a quick use of the Euclidean Algorithm, she discovers that $u = 8$ is a solution.
Bob has the grade $x = 85$ to encrypt and send to Alice. For that he reduces $zx = 38 \cdot 85$ modulo 101 and gets $y = 99$. Indeed, we can easily check that

$$38 \cdot 85 \equiv 99 \mod 101.$$ 

So Bob sends out $y = 99$ to Alice as her encrypted grade.

Eve can see this 99. She knows it’s encrypted, but cannot recover the original 85 because she does not possess the key $z = 38$ nor its inverse $u = 8$ modulo 101.

But Alice can. She reduces $uy = 8 \cdot 99$ modulo 101 to get the original grade $x = 85$. Indeed, we can easily check that

$$8 \cdot 99 = 792 \equiv 85 \mod 101.$$ 

And that’s how Alice got her mark $x = 85$.

### 5.7 Exercises

1. If $n$ is a positive integer and there exists a residue in $\mathbb{Z}_n^*$ of order $n - 1$, show that $n$ is prime.

2. (a) Suppose that $p, q$ are distinct primes and that $a$ is an integer such that $a \not\equiv 1 \mod q$. If $a^p \equiv 1 \mod q$, show that $q \equiv 1 \mod p$.

   **Hint.** First explain why $q \nmid a$. Then explain why the residue $[a]$ taken in $\mathbb{Z}_q^*$ has order $p$. Then deduce that $p \mid q - 1$. Proposition 5.5 comes into play.

   (b) The result in part (a) can come in handy to test for the primality of some integers. Decide if $2^{13} - 1$ is prime.

   **Hint.** In looking for a potential prime factor $q$ of $2^{13} - 1$, you only need to test for primes up to the square root of this number, which is less than $2^7 = 128$. You also only need to test for such primes $q$ that are congruent to 1 modulo 13. Find those primes, and then reduce $2^{13} - 1$ modulo those $q$.

   (c) Decide if $2^{37} - 1$ is prime.

3. Find a primitive root for the prime 51.
4. Show that 19 is a primitive root for the prime $p = 191$.
   Hint. It might be more efficient to first factor 118.

5. For a Diffie-Hellman key exchange, Alice and Bob agree to use the prime $p = 43$ with primitive root $g = 3$. Alice picks the discrete logarithm $\ell = 13$. Bob picks the discrete logarithm $k = 29$.
   (a) Verify that $3$ is a primitive root for the prime 43.
   (b) Find the key $z$ that gets exchanged between Alice and Bob.
   (c) Bob wants to send the message $x$ to Alice in encrypted form. The encrypted message that Bob sends is $y = 15$. How does Alice decrypt $y$ to get $x$? Find the message $x$.
   (d) If Bob wants to send the message $x = 12$, what is the encrypted message $y$ that he transmits?

6. How many primitive roots are there in $\mathbb{Z}_{101}^*$?

7. Find all primitive roots in $\mathbb{Z}_{19}^*$.

8. Recall that the $n$'th Fermat number is $F_n = 2^{2^n} + 1$. Only the first few are known to be primes, but there might be more. Suppose that $n$ is such that $F_n$ is prime.
   Show that $\alpha = [2]$ is not a primitive root in $\mathbb{Z}_{F_n}$.
   Hint. First factor $\alpha^{2^n+1} - 1$ to show that $\alpha$ has order at most $2^{n+1}$ in $\mathbb{Z}_{F_n}$.
   Then notice that $\varphi(F_n)$ is bigger than $2^{n+1}$.

9. If $p, q$ are distinct, odd primes (e.g. 7 and 19), show that $\mathbb{Z}_{pq}^*$ has no primitive root.
   Hint. Observe that $\varphi(pq) = (p - 1)(q - 1)$.
   Then explain why, for any $\alpha$ in $\mathbb{Z}_{pq}^*$, we have $\alpha^{\frac{(p-1)(q-1)}{2}} = [1]$.
   You need to recall an important general principle. Namely, for $m, n$ coprime we always have
   
   $a \equiv b \mod mn$ if and only if $a \equiv b \mod m$ and $a \equiv b \mod n$.

   This was an essential aspect of the Chinese remainder theorem.
10. If \( n \) is a Carmichael number, prove that \( n \) is an odd number.
   Hint. There are no witnesses for \( n \), not even \( n - 1 \).

11. If \( n = pq \), where \( p \) and \( q \) are distinct odd primes, prove that \( n \) is not a
    Carmichael number.
   Hint. You can use Korselt’s criterion.
   The previous two exercises, taken together, show that it takes three or more
   distinct odd primes to build a Carmichael number.

12. Find a witness for the non-primality of 8633.
   Hint. Often 2 works, and if not, try 3.

13. If \( k \) is a positive integer and \( p = 6k + 1 \), \( q = 12k + 1 \), \( r = 18k + 1 \) are all
    primes, show that \( n = pqr \) is a Carmichael number.

14. If \( \varphi(n) \mid n - 1 \), show that \( n \) is either prime or Carmichael.
   A famous unsolved problem is to rule out the possibility that \( n \) is Carmichael.

15. This exercise offers a possible way to confirm that a given integer is prime.
   Let \( n \) be that given integer, possibly quite large. If \( a^n \equiv a \mod n \) for all
   integers \( a \), our \( n \) is Carmichael, but more likely than not it is prime. Here’s
   how we might be able to confirm that \( n \) is prime, in the lucky event that we
   have a factorization of \( n - 1 \) into primes:

   \[
   n - 1 = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}, \quad \text{where} \ e_j \geq 1.
   \]

   The numbers \( r_j = \frac{n-1}{p_j} \) are all integers.

   Suppose \( a \) is such that \( a^{n-1} \equiv 1 \mod n \). For instance, \( a \) could be something
   as simple as \( a = 2 \). Thus \( n \) is acting like it wants to be prime. Such \( n \) are
   called \textit{pseudo-primes}, because they have passed the first screening test for
   primality.

   (a) If every \( a^{r_j} \neq 1 \mod n \), prove that \( n \) is prime, and that \( a \) is a primitive
   root for \( p \).

   Hint. If \( \ell = \text{ord}(a) \mod n \), explain why \( k \mid n - 1 \). If you can show
   that \( \text{ord}(a) = n - 1 \), then you can finish the proof by using Exercise
   1 of this chapter. Well, show that if \( k < n - 1 \), then some \( a^{r_j} \equiv 1 \mod n \). It helps to read the discussion on finding primitive roots just
   after the proof of the Primitive Root Theorem.
(b) To apply the primality test of part (a), here is what to do.

- Pick an $a$ such that $a^{n-1} \equiv 1 \mod n$. If this is to happen, chances are it will happen after few small $a$'s are tested such as 2, 3, 5, 7, . . . .

- Factor $n - 1 = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}$ into primes. This could be hard to do, but sometimes we get lucky.

- Let $r_j = \frac{n-1}{p_j}$.

- Reduce each $a^{r_j}$ modulo $n$. You can square and multiply modulo $n$ to do this.

- If you never get $a^{r_j} \equiv 1$, then you know for sure that $n$ is prime with $a$ as your primitive root.

Show that 2549 is prime with 2 as its primitive root.

(c) Show that 27943 is prime. Use technology to help with calculations.
Chapter 6

Quadratic Residues

Linear congruences of the type \( ax \equiv b \mod n \) are readily solved by means of linear Diophantine equations and the Euclidean Algorithm. But what do we do with quadratic congruences

\[ ax^2 + bx + c \equiv 0 \mod n, \text{ where } a, b, c \in \mathbb{Z}? \]

Equivalently, we can consider the quadratic equations

\[ ax^2 + \beta x + \gamma = 0, \]

where \( \alpha, \beta, \gamma \in \mathbb{Z}_n \) and \( x \) is an unknown residue in \( \mathbb{Z}_n \).

Before we go any further, we should require that \( n \nmid a \). For otherwise, the quadratic congruence collapses to the linear congruence \( bx + c \equiv 0 \mod n \). When we speak of a quadratic congruence modulo \( n \), we shall always assume that the coefficient \( a \) in front of \( x^2 \) is not divisible by \( n \).

Also if \( n = 2 \), Fermat’s theorem tells us that \( x^2 \equiv x \mod 2 \) regardless of \( x \). Hence

\[ ax^2 + bx + c \equiv ax + bx + c \equiv (a + b)x + c \mod 2. \]

For \( n = 2 \) all quadratic congruences collapse into linear congruences. We might as well assume that \( n \geq 3 \).

It also makes matters much simpler if we take \( n \) to be a prime, which is best remembered by using the letter \( p \). If \( p \) is an odd prime (i.e. \( p \neq 2 \)), the integer \( p - 1 \) is even, and thereby \( \frac{p - 1}{2} \) is an integer, a fact we shall need to keep in mind.
For such \( p \), Proposition 5.14 ensures that the polynomial \( \alpha X^2 + \beta X + \gamma \) over \( \mathbb{Z}_p \) has at most two roots in \( \mathbb{Z}_p \). This is the same as saying that the congruence \( ax^2 + bx + c \equiv 0 \mod p \) has at most two non-congruent integer solutions. But the congruence may have no solutions, so how are we to know?

### 6.1 Quadratic congruences

We might recall from high school that a quadratic equation \( ax^2 + bx + c = 0 \) has a real solution \( x \) if and only if the discriminant \( b^2 - 4ac \geq 0 \). Since, among the real numbers, only the non-negative numbers are squares of other numbers, the condition \( b^2 - 4ac \geq 0 \) translates into saying that the equation \( y^2 = b^2 - 4ac \) has a real number solution. Well, pretty much the same idea holds with quadratic congruences.

**Proposition 6.1.** Let \( p \) be an odd prime, and \( a, b, c \) integers where \( p \nmid a \). The quadratic congruence

\[
ax^2 + bx + c \equiv 0 \mod p
\]

has a solution \( x \) if and only if the congruence

\[
y^2 \equiv b^2 - 4ac \mod p
\]

has a solution \( y \). In that case, \( y \equiv 2ax + b \mod p \).

**Proof.** Suppose the quadratic congruence has a solution \( x \). Multiply through by \( 4a \) to get

\[
4a^2 x^2 + 4abx + 4ac \equiv 0 \mod p.
\]

This can be rewritten as

\[
(2ax + b)^2 - b^2 + 4ac \equiv 0 \mod p,
\]

which is the same as

\[
(2ax + b)^2 \equiv b^2 - 4ac \mod p.
\]

Thus \( y = 2ax + b \) solves the congruence \( y^2 \equiv b^2 - 4ac \mod p \).

Conversely, suppose that \( y \) is a solution of \( y^2 \equiv b^2 - 4ac \mod p \). The linear congruence \( 2ax + b \equiv y \mod p \) can be rewritten as \( 2ax \equiv y - b \mod p \). Since \( p \) is
6.1. QUADRATIC CONGRUENCES

odd and \( p \nmid a \), we see that \( p \nmid 2a \), and therefore this linear congruence has a solution \( x \). Thus we have an \( x \) such that

\[(2ax + b)^2 \equiv b^2 - 4ac \mod p.\]

After expanding out and rearranging we obtain

\[4a^2x^2 + 4abx + 4ac \equiv 0 \mod p.\]

Factor out \( 4a \) to get

\[4a(ax^2 + bx + c) \equiv 0 \mod p.\]

Since \( p \nmid 4a \), we can cancel the \( 4a \) to deduce

\[ax^2 + bx + c \equiv 0 \mod p.\]

Thus \( x \) solves our quadratic congruence.

As in high school, the integer \( d = b^2 - 4ac \) will be called the discriminant of the quadratic \( aX^2 + bX + c \). The significance of Proposition 6.1 is that quadratic congruences come down to solving congruences of the form \( x^2 \equiv d \mod p \). Here is an illustration.

**Example 6.2.** Let us solve

\[2x^2 + 18x + 3 \equiv 0 \mod 23.\]

The discriminant of this quadratic, reduced modulo 23, is

\[18^2 - 4 \cdot 2 \cdot 3 = 300 \equiv 1 \mod 23.\]

According to Proposition 6.1, we should first solve the congruence

\[y^2 \equiv 1 \mod 23.\]

By inspection, we can see that \( y \equiv \pm 1 \) provide solutions. To see that these are the only solutions, let \( y \) be any integer such that \( y^2 \equiv 1 \mod 23 \). Hence

\[23 \mid y^2 - 1 = (y - 1)(y + 1).\]

Since 23 is prime, it follows that \( 23 \mid y - 1 \) or \( 23 \mid y + 1 \). From this we deduce \( y \equiv \pm 1 \mod 23 \). Another way to see that these are the only solutions is by invoking
Proposition 5.14, telling us that since 23 is prime, the quadratic $Y^2 - 1$ has at most two roots modulo 23.

According to Proposition 6.1, we next solve the congruences

$$2 \cdot 2x + 18 \equiv \pm 1 \mod 23.$$  

These come down to

$$4x \equiv 4 \mod 23 \text{ and } 4x \equiv 6 \mod 23.$$  

For the first congruence, cancel the 4 to get $x \equiv 1 \mod 23$. For second congruence, a bit of work with the Euclidean algorithm, alternately with some patient testing, gives $x \equiv 13 \mod 23$.

The solutions to the original quadratic congruence are

$$x \equiv 1 \mod 23 \text{ and } x \equiv 13 \mod 23.$$  

We can be sure there are no more solutions since 23 is prime, which, in accordance with Proposition 5.14, permits our quadratic to have at most two solutions.

### 6.2 Quadratic residues and primitive roots

**Definition 6.3.** A residue $\alpha$ in $\Z_p$ is called a **quadratic residue** when

$$\alpha \in \Z_p^* \text{ and } \alpha = \beta^2 \text{ for some other residue } \beta \text{ in } \Z_p^*.$$  

We also say that an integer $a$ has a **quadratic residue** modulo an odd prime $p$ provided

$$p \nmid a \text{ and } a \equiv x^2 \mod p \text{ for some integer } x.$$  

In other words, an integer $a$ has a quadratic residue modulo $p$ whenever its residue $\alpha = [a]$ in $\Z_p^*$ is quadratic.

We should be comfortable by now in translating back and forth between the language of congruences modulo $p$ and the language of residues in $\Z_p$. The advantage of working sometimes inside $\Z_p^*$ is that it is a finite set with exactly $p - 1$ elements. And we should not forget that $p - 1$ is even.
For example, let’s find the quadratic residues of \( \mathbb{Z}_{11}^* \) by squaring all residues in \( \mathbb{Z}_{11}^* \). Thus

\[
\begin{align*}
\end{align*}
\]

The quadratic residues in \( \mathbb{Z}_{11}^* \) are

\[ [1], [3], [4], [5], [9], \]

a seemingly haphazard list. Alternately, we can say that the integers having quadratic residues modulo 11 are those that are congruent to 1, 3, 4, 5, 9.

**The number of quadratic residues**

In \( \mathbb{Z}_p^* \) exactly half of its residues are quadratic. To see this, write the \( p - 1 \) residues in \( \mathbb{Z}_p^* \) as

\[ \pm[1], \pm[2], \pm[3], \ldots, \pm \left[ \frac{p - 1}{2} \right]. \]

Since \([a]^2 = (-[a])^2\), there can be at most \( \frac{p-1}{2} \) quadratic residues. To see that there are exactly \( \frac{p-1}{2} \) such residues, we can simply note that when \( 1 \leq a < b \leq \frac{p-1}{2} \), the squares \([a]^2\) and \([b]^2\) are not equal. Indeed, say we had \([a]^2\) and \([b]^2\) both equal to some residue \([c]\). Thus both \([a]\) and \([b]\) would be roots of \(X^2 - [c]\). However, \(-[a]\) and \(-[b]\) would also be roots, and that would cause this quadratic to have too many roots in contradiction to Proposition 5.14. Another way to put it is that exactly half of the integers from 1 to \( p - 1 \) have a quadratic residue modulo \( p \).

We shall soon see other explanations as to why the number of quadratic residues in \( \mathbb{Z}_p^* \) is \( \frac{p-1}{2} \).

**The Legendre symbol**

Our problem will be to decide, for a given odd prime \( p \) and a given integer \( a \), if \( a \) has a quadratic residue modulo \( p \) or not.

Back in 1798 the French mathematician Adrien-Marie Legendre introduced a handy symbol to mark this distinction. It has endured since then.
Definition 6.4. For an odd prime $p$ and an integer $a$ coprime with $p$, we let

$$\left( \frac{a}{p} \right) = \begin{cases} +1 & \text{if } a \text{ has a quadratic residue modulo } p \\ -1 & \text{if } a \text{ does not have a quadratic residue modulo } p. \end{cases}$$

The symbol $\left( \frac{a}{p} \right)$ is called the Legendre symbol for $a$ modulo $p$.

Keep in mind that the Legendre symbol is not a fraction, even though it sort of looks like one.

For example,

$$\left( \frac{3}{11} \right) = +1 \text{ while } \left( \frac{7}{11} \right) = -1.$$ 

It should be clear that $1$ has a quadratic residue for any odd prime $p$, i.e. $\left( \frac{1}{p} \right) = 1$.

But $-1$ might or might not have a quadratic residue, depending on the prime. For instance, by inspection $x^2 \equiv -1 \mod 3$ has no solution $x$, while $x^2 \equiv -1 \mod 5$ has the solution $x \equiv 2 \mod 5$. So we have

$$\left( \frac{-1}{3} \right) = -1 \text{ while } \left( \frac{-1}{5} \right) = 1.$$ 

By looking for solutions to the congruences $x^2 \equiv a \mod p$ we can decide on the value of the Legendre symbol, a time consuming process that reveals no patterns. It’s one thing to define the Legendre symbol, but it’s quite another to actually calculate it. We have our work cut out for us.

Detecting quadratic residues with primitive roots

Since $p$ is prime, the group $\mathbb{Z}_p^*$ has a primitive root $\beta$. In fact, the number of such primitive roots is $\varphi(p-1)$. The powers $\beta^k$ exhaust $\mathbb{Z}_p^*$. By looking at the exponents $k$ we can decide if $\beta^k$ is a quadratic residue or not.

**Proposition 6.5.** If $\beta$ is a primitive root in $\mathbb{Z}_p^*$ and $\beta^i = \beta^j$ for some positive exponents $i$, $j$, then these exponents will be both even or both odd.

**Proof.** Say $1 \leq i \leq j$. Then

$$\beta^i = \beta^i \beta^{j-i}.$$ 

Cancel $\beta^i$ to get $1 = \beta^{j-i}$. According to Proposition 5.5, $\text{ord}\, \beta \mid j - i$. But $\text{ord}\, \beta = p - 1$. Thus $p - 1 \mid j - i$, and since $p - 1$ is even, so is $j - i$ is even. Therefore, either both $j$ and $i$ are even, or both $j$ and $i$ are odd. \hfill $\square$

What Proposition 6.5 tells us is that, even though a residue can be expressed using different powers of $\beta$, it cannot be written both as an even power and an odd power of $\beta$.

**Proposition 6.6.** Let $p$ be an odd prime, and let $\beta$ in $\mathbb{Z}_p^*$ be a primitive root. An element $\alpha$ in $\mathbb{Z}_p^*$ is a quadratic residue if and only if $\alpha$ is an even power of $\beta$.

**Proof.** Say $\alpha = \beta^{2k}$ for some integer $k$. Since $\alpha = (\beta^k)^2$, our $\alpha$ is a quadratic residue.

Conversely, suppose $\alpha$ is a quadratic residue. So $\alpha = \gamma^2$ for some $\gamma$ in $\mathbb{Z}_p^*$. Since $\beta$ is a primitive root, we can write $\gamma = \beta^j$ for some exponent $j$. Thus

$$\alpha = \gamma^2 = (\beta^j)^2 = \beta^{2j},$$

which is clearly an even power of $\beta$. \hfill $\square$

Here’s a reminder of the notion of index or discrete logarithm. If $\beta$ is our primitive root in $\mathbb{Z}_p^*$, the finite list

$$\beta, \beta^2, \ldots, \beta^{p-1} = 1$$

covers all of $\mathbb{Z}_p^*$ without any duplication. For $\alpha$ in $\mathbb{Z}_p^*$, the unique $\ell$ from 1 to $p - 1$ that gives $\alpha = \beta^\ell$ is called the index or the discrete logarithm of $\alpha$ with base $\beta$. This gives us another way to state Proposition 6.6.

**Proposition 6.7.** For an odd prime $p$, a residue in $\mathbb{Z}_p^*$ is a quadratic residue if and only if its index with respect to a primitive root is even.

From Propositions 6.5, 6.6 and 6.7 some observations fall into our lap. Here $\beta$ is taken to be a primitive root of $\mathbb{Z}_p^*$:

- A primitive root $\beta$ is never a quadratic residue, since $\beta = \beta^1$, and 1 is odd.
- The number of quadratic residues in $\mathbb{Z}_p^*$ equals the number of non-quadratic residues, since the number of even indices from 1 to $p - 1$ equals the number of odd indices.
• The product of two quadratic residues is a quadratic residue, since an even power times an even power of $\beta$ is another even power of $\beta$.

• The product of a quadratic residue with a non-quadratic residue is a non-quadratic residue, since an even power of $\beta$ times an odd power of $\beta$ is an odd power of $\beta$.

• The product of two non-quadratic residues is a quadratic residue, since an odd power of $\beta$ times an odd power of $\beta$ is an even power of $\beta$.

While Proposition 6.7 is nice in theory, it cannot be very practical. In order to know if a given $\alpha$ in $\mathbb{Z}_p^*$ is a quadratic residue, we need to have a primitive root, and then we face the daunting task of actually computing the index of $\alpha$, something computers have not been good at so far.

**Euler’s calculation of the Legendre symbol**

On the other hand, we don’t need the actual index of a given $\alpha$ in $\mathbb{Z}_p^*$. All we need to know is whether such index is even or odd. Another result of Euler comes to the rescue in this regard. It lets us calculate the Legendre symbol for an integer $b$, and thereby decide on the quadratic character of $b$, by simply reducing one of its powers.

**Proposition 6.8 (Euler’s test).** If $p$ is an odd prime and $a$ is an integer coprime with $p$ (i.e. $p \nmid a$), then

$$a^{\frac{p-1}{2}} \equiv \left(\frac{a}{p}\right) \mod p.$$  

*Keeping in mind the meaning of the Legendre symbol, this says the following. If $a$ has a quadratic residue modulo $p$, then $a^{\frac{p-1}{2}} \equiv +1 \mod p$, and if $a$ does not have a quadratic residue modulo $p$, then $a^{\frac{p-1}{2}} \equiv -1 \mod p$.*

**Proof.** Let the integer $b$ represent a primitive root in $\mathbb{Z}_p^*$. This means that every integer coprime with $p$ is congruent modulo $p$ to an integer in the list

$$b, b^2, \ldots, b^{p-1}$$

In particular $a \equiv b^k \mod p$, where $k$ is the index of $a$ using the base $b$. 
If \( a \) has a quadratic residue, then \( k \) is even according to Proposition 6.6. Say \( k = 2j \) for some \( j \). After that, Fermat’s theorem gives
\[
a^{\frac{p-1}{2}} \equiv (b^{2j})^{\frac{p-1}{2}} = (b^j)^{p-1} \equiv 1 \pmod{p}.
\]

If the residue of \( a \) is not quadratic, then \( k \) is odd. Say \( k = 2j + 1 \) for some \( j \). Then, using Fermat once more:
\[
a^{\frac{p-1}{2}} \equiv (b^{2j+1})^{\frac{p-1}{2}} = (b^j)^{p-1}b^{\frac{p-1}{2}} \equiv 1 \cdot b^{\frac{p-1}{2}} = b^{\frac{p-1}{2}}.
\]

Since \( b \) represents a primitive root, its residue \([b]\) has order \( p - 1 \). Thus \( b^{\frac{p-1}{2}} \not\equiv 1 \pmod{p} \). However since
\[
\left(b^{\frac{p-1}{2}}\right)^2 = b^{p-1} \equiv 1 \pmod{p},
\]

a little exercise shows that \( b^{\frac{p-1}{2}} \equiv \pm 1 \pmod{p} \). Having ruled out 1 as a possibility, we conclude that \( b^{\frac{p-1}{2}} \equiv -1 \pmod{p} \). \(\square\)

Here is an illustration of Euler’s test.

**Example 6.9.** Does 79 have a quadratic residue modulo 31? According to Euler, just reduce 79^{15} modulo 31. First notice 79^{15} \equiv 17^{15} \pmod{31}. Then reduce 17^{15} by the square and multiply algorithm. Note that 15 = 1 + 2 + 4 + 8 and so 17^{15} = 17^1 \cdot 17^2 \cdot 17^4 \cdot 17^8. Then get squaring modulo 31:
\[
17 \equiv 17, \quad 17^2 \equiv 10, \quad 17^4 \equiv 7, \quad 17^8 \equiv 18
\]

Thus,
\[
17^{15} \equiv 17 \cdot 10 \cdot 7 \cdot 18 \equiv 30 \equiv -1 \pmod{31}.
\]

Of course, the above calculation can also be done by machine. According to Euler’s test, 79 does not have a quadratic residue modulo 31. In the Language of the Legendre symbol we have found that
\[
\left(\frac{79}{31}\right) = -1.
\]
The interesting case of $-1$

We might ask for which odd primes $p$ does $-1$ have a quadratic residue modulo $p$? According to Euler’s test

$$\left(\frac{-1}{p}\right) \equiv (-1)^{\frac{p-1}{2}} \mod p.$$ 

Since both sides of this congruence can only equal $\pm 1$, the congruence is actually an equality. But obviously

$$(-1)^{\frac{p-1}{2}} = \begin{cases} 1 & \text{if } \frac{p-1}{2} \text{ is even} \\ -1 & \text{if } \frac{p-1}{2} \text{ is odd}. \end{cases}$$

Now $\frac{p-1}{2}$ is even if and only if $4 | p - 1$.

Thus we encounter a neat little fact.

**Proposition 6.10.** The integer $-1$ has a quadratic residue modulo an odd prime $p$ if and only if $p \equiv 1 \mod 4$.

For instance $-1$ has a quadratic residue modulo 101, but $-1$ does not have a quadratic residue modulo 31. This is because $101 \equiv 1 \mod 4$, while $31 \not\equiv 1 \mod 4$.

**Euler’s test and primes congruent to 1 modulo 4**

A bit of consideration reveals that every odd prime is congruent to 1 or to 3 modulo 4. There are infinitely many primes, but are there infinitely many congruent to 1 modulo 4? What about congruent to 3 modulo 4? The answers to these questions are far from obvious. Let’s show that the number of primes congruent to 1 modulo 4 is infinite.

**Proposition 6.11.** There are infinitely many primes congruent to 1 modulo 4.

**Proof.** Suppose we had a finite list of primes $p_1, p_2, \ldots, p_n$, all congruent to 1 modulo 4. Here is how to get one more prime $q \equiv 1 \mod 4$, and not on this list. Let

$$x = (2 \cdot p_1 \cdot p_2 \cdots p_n)^2 + 1.$$
and let \( q \) be any prime factor of \( x \). If \( q \) equals any one of \( 2, p_1, \ldots, p_n \), then \( q \mid 1 \), which is impossible. So \( q \) is a prime other than \( 2, p_1, p_2, \ldots, p_n \). Since \( q \mid x \), we see that

\[
-1 \equiv (2 \cdot p_1 \cdot p_2 \cdots p_n)^2 \mod q.
\]

Thus \(-1\) has a quadratic residue modulo \( q \). By Proposition 6.10, \( q \equiv 1 \mod 4 \).

Since any finite list of primes congruent to \( 1 \) modulo \( 4 \) can be augmented with an additional such prime, there must be infinitely many such primes. \( \square \)

The case of \( 2 \)

If \( p \) is an odd prime, we can reduce \( 2^{\frac{p-1}{2}} \mod p \) and decide by Euler’s test whether \( 2 \) has a quadratic residue modulo \( p \). Here we will highlight an even better way to decide, due to the inimitable C.F. Gauss.

First a couple of examples to warm up with.

**Example 6.12.** Let’s decide if \( 2 \) has a quadratic residue modulo \( 19 \). Here \( p = 19 \), and \( \frac{p-1}{2} = 9 \). According to Euler’s test we need to reduce \( 2^9 \) modulo \( 19 \) and see if we get \( 1 \) or \(-1 \). Here’s Gauss’ clever way to make the reduction.

Examine the product of the even integers \( 2, 4, 6, 8, \ldots, 18 \) in two different ways. First

\[
2 \cdot 4 \cdot 6 \cdot 8 \cdot 10 \cdot 12 \cdot 14 \cdot 16 \cdot 18 = 2^9 \cdot 9!.
\]

But also

\[
2 \cdot 4 \cdot 6 \cdot 8 \cdot 10 \cdot 12 \cdot 14 \cdot 16 \cdot 18 \equiv 2 \cdot 4 \cdot 6 \cdot 8 \cdot (-9) \cdot (-7) \cdot (-5) \cdot (-3) \cdot (-1) \mod 19.
\]

Thus

\[
2^9 \cdot 9! \equiv (-1)^5 \cdot 9! = -9! \mod 19.
\]

Since \( 19 \nmid 9! \), we can cancel \( 9! \) and get \( 2^9 \equiv -1 \mod 19 \). By Euler’s test, \( 2 \) does not have a quadratic residue modulo \( 19 \).

**Example 6.13.** Does \( 2 \) have a quadratic residue modulo \( 23 \)? According to Euler’s test we need to reduce \( 2^{11} \) modulo \( 23 \). As in the preceding example, look at the product of \( 2, 4, 6, 8, \ldots, 22 \) modulo \( 23 \) in two different ways, to get

\[
2 \cdot 4 \cdot 6 \cdot 8 \cdot 10 \cdot 12 \cdot 14 \cdot 16 \cdot 18 \cdot 20 \cdot 22
\]

\[
\equiv 2 \cdot 4 \cdot 6 \cdot 8 \cdot 10 \cdot (-11) \cdot (-9) \cdot (-7) \cdot (-5) \cdot (-3) \cdot (-1) \mod 23.
\]
Look closely to get
\[ 2^{11} \cdot 11! \equiv (-1)^6 \cdot 11! = 11! \mod 23. \]

Cancel 11! and obtain \( 2^{11} \equiv 1 \mod 23 \). Thus 2 has a quadratic residue modulo 23.

On the basis of the tricks in the above examples Gauss provided a general test for 2 to have a quadratic residue modulo any prime \( p \). Note, in advance, that an odd prime \( p \) has to be congruent modulo 8 to one of 1, 3, 5, 7.

**Proposition 6.14.** If \( p \) is a prime and \( p \equiv 1 \mod 8 \) or \( p \equiv 7 \mod 8 \), then 2 has a quadratic residue modulo \( p \). If \( p \equiv 3 \mod 8 \) or \( p \equiv 5 \mod 8 \), then 2 does not have a quadratic residue modulo \( p \).

**Proof.** The proof involves Gauss’ trick plus some careful (and possibly time consuming) book-keeping. There are four cases to prove, and they take a while. Since they are quite repetitive, we shall be content to prove only two of the cases.

Suppose \( p \equiv 1 \mod 8 \). Thus \( p = 8k + 1 \) and \( p - 1 = 8k \) for some \( k \). There are \( 4k = \frac{p-1}{2} \) even integers from 2 to 8\( k \). Here they are:

\[ 2, 4, 6, \ldots, 4k - 2, 4k, 4k + 2, 4k + 4, \ldots, 8k - 2, 8k. \]

Their product comes out to be (look closely):

\[ x = 2 \cdot 4 \cdot 6 \cdots (4k - 2) \cdot (4k) \cdot (4k + 2) \cdot (4k + 4) \cdots (8k - 2) \cdot (8k) \]
\[ = 2^{4k} \cdot 1 \cdot 2 \cdot 3 \cdots (2k) \cdot (2k + 1) \cdot (2k + 2) \cdots (4k - 1) \cdot (4k) \]
\[ = 2^{4k} \cdot (4k)! \]

Looking at the second half of our list and remembering that \( p = 8k + 1 \), we notice that

\[ 4k + 2 \equiv 1 - 4k \mod p \]
\[ 4k + 4 \equiv 3 - 4k \mod p \]
\[ \vdots \]
\[ 8k - 2 \equiv -2 \mod p \]
\[ 8k \equiv -1 \mod p \]
6.2. QUADRATIC RESIDUES AND PRIMITIVE ROOTS

By replacement modulo \( p \) and patiently keeping track we get

\[
x \equiv 2 \cdot 4 \cdot 6 \cdots (4k - 2) \cdot (4k) \cdot (1 - 4k) \cdot (3 - 4k) \cdot (5 - 4k) \cdots (-2) \cdot (-1)^{2k} \\
= 2 \cdot 4 \cdot 6 \cdots (4k - 2) \cdot (4k) \cdot (4k - 1) \cdot (4k - 3) \cdot (4k - 5) \cdots 3 \cdot 1 \cdot (4k)! \\
= (4k)! 
\]

The \((-1)^{2k}\) appeared by collecting \(-1\) from each factor in the second half of the product.

Having seen that \(2^{4k}(4k)! \equiv (4k)! \mod p\),
cancel \((4k)!\) in this congruence and obtain

\[
2^{\frac{p-1}{2}} = 2^{4k} \equiv 1 \mod p. 
\]

By Euler’s test, 2 has a quadratic residue modulo \( p \).

Moving on to another case, suppose \( p \equiv 3 \mod 8 \). Now there is an integer \( k \) such that

\[
p = 8k + 3, \ p - 1 = 8k + 2 \text{ and } \frac{p - 1}{2} = 4k + 1 \text{ (an odd number)}. 
\]

Again multiply the even integers from 2 to \( p - 1 \). There are \( \frac{p - 1}{2} = 4k + 1 \) such even integers. Here is their product:

\[
x = 2 \cdot 4 \cdot 6 \cdots (4k - 2) \cdot (4k) \cdot (4k + 2) \cdot (4k + 4) \cdots (8k) \cdot (8k + 2) \\
= 2^{4k+1} \cdot (4k + 1)! 
\]

Just to keep track, notice that up to \( 4k \) there are \( 2k \) factors, while from \( 4k + 2 \) to \( 8k + 2 \) there are \( 2k + 1 \) factors in the product that gave \( x \). Keeping in mind \( p = 8k + 3 \), notice that

\[
4k + 2 \equiv -4k - 1 \mod p \\
4k + 4 \equiv -4k + 1 \mod p \\
4k + 6 \equiv -4k + 3 \mod p \\
\vdots \\
8k \equiv -3 \mod p \\
8k + 2 \equiv -1 \mod p. 
\]
By replacement modulo \( p \) we get
\[
x \equiv 2 \cdot 4 \cdot 6 \cdot (4k - 2) \cdot (4k) \cdot (-4k - 1) \cdot (-4k + 1) \cdot (-4k + 3) \cdots (-3) \cdot (1)
\]
\[
= (-1)^{2k+1} (2 \cdot 4 \cdot 6 \cdot (4k - 2) \cdot (4k) \cdot (4k + 1) \cdot (4k - 1) \cdot (4k - 3) \cdots 3 \cdot 1
\]
\[
= (-1)^{2k+1} \cdot (4k + 1)! \quad \text{(Look very closely!)}
\]
\[
= -(4k + 1)!
\]

We collected \(-1\) from each factor from \(-4k - 1\) to \(-1\), and there are \(2k + 1\) such factors. That explains the \((-1)^{2k+1}\).

Thus we arrive at
\[
2^{4k+1} \cdot (4k + 1)! \equiv -(4k + 1)! \mod p,
\]

Cancel \((4k + 1)!\) to get
\[
2^{\frac{p-1}{2}} = 2^{4k+1} \equiv -1 \mod p.
\]

By Euler’s test, \(2\) is not a quadratic residue modulo \( p \).

The proofs of the cases \( p \equiv 5 \) and \( p \equiv 7 \) are omitted because they repeat the lengthy discussions we have already undertaken.

Now, for example, we see that
- \(2\) has a quadratic residue modulo \(41\) because \(41 \equiv 1 \mod 8\)
- \(2\) has no quadratic residue modulo \(43\) because \(43 \equiv 3 \mod 8\)
- \(2\) has a quadratic residue modulo \(31\) because \(31 \equiv 7 \mod 8\)
- \(2\) has no quadratic residue modulo \(109\) because \(109 \equiv 5 \mod 8\).

### 6.3 Quadratic reciprocity

Having seen how to decide with ease when \(-1\) and \(2\) have quadratic residues modulo an odd prime \( p \), we might become bold and see if we can give a clean, definitive answer for any integer \( a \), which goes beyond Euler’s test. In other words, we want a fast algorithm for calculating the Legendre symbol \( \left( \frac{a}{p} \right) \) for any integer \( a \).
and any odd prime \( p \). It was Gauss who succeeded in giving us that method, after the efforts of many, including Legendre.

That method, known as the law of quadratic reciprocity is easy to state and pleasant to use. It is here that the Legendre symbol comes into its own.

The ensuing easy properties of the Legendre symbol help with its calculation.

**Proposition 6.15.** The Legendre symbol is multiplicative. Namely, if \( p \) is an odd prime, and \( a, b \) are integers coprime with \( p \), then

\[
\left( \frac{ab}{p} \right) = \left( \frac{a}{p} \right) \left( \frac{b}{p} \right).
\]

Furthermore, if \( a \equiv b \mod p \), then

\[
\left( \frac{a}{p} \right) = \left( \frac{b}{p} \right).
\]

**Proof.** The second statement of the proposition is obvious because the residue is the same for all congruent integers.

The first part comes easily from Euler’s test 6.8 as follows:

\[
\left( \frac{a}{p} \right) \left( \frac{b}{p} \right) \equiv \left( \frac{a^{p-1}}{p} \right) \left( \frac{b^{p-1}}{p} \right) = \left( \frac{ab}{p} \right)^{p-1} \equiv \left( \frac{ab}{p} \right) \mod p.
\]

Since \( \left( \frac{a}{p} \right) \left( \frac{b}{p} \right) = \pm 1 \) and \( \left( \frac{ab}{p} \right) = \pm 1 \), and these two integers are congruent modulo \( p \), they have to be equal. \( \square \)

Every positive integer \( a \) is a product of primes \( a = q_1 \cdot q_2 \cdots q_k \), with repetitions allowed. By Proposition 6.15 we see that if \( p \nmid a \), then

\[
\left( \frac{a}{p} \right) = \left( \frac{q_1}{p} \right) \left( \frac{q_2}{p} \right) \cdots \left( \frac{q_k}{p} \right).
\]

Also, if \( a \) is a negative integer, write \( a = -1 \cdot b \) where \( b > 0 \), to get:

\[
\left( \frac{a}{p} \right) = \left( \frac{-1}{p} \right) \left( \frac{b}{p} \right).
\]

In light of Proposition 6.10, telling us how to determine \( \left( \frac{-1}{p} \right) \), and the above multiplicative property of the Legendre symbol, we are reduced to finding the
values of $\left( \frac{q}{p} \right)$ where $q$ is a positive prime. Also, since Proposition 6.14 tells us how to determine $\left( \frac{2}{p} \right)$, the tough job remaining is to figure out $\left( \frac{q}{p} \right)$ where $q$ and $p$ are both odd and distinct primes.

It was Gauss, at the age of 19, who told us what to do in a highly celebrated theorem. Here is that theorem. After stating it we shall work with it a little bit, and after that delve into the intricate proof.

**Theorem 6.16** (Gauss’ law of quadratic reciprocity). *Let $p, q$ be odd and distinct primes.*

*If at least one of $p \equiv 1 \mod 4$ or $q \equiv 1 \mod 4$, then*

$$\left( \frac{q}{p} \right) = \left( \frac{p}{q} \right).$$

*If both $p \equiv 3 \mod 4$ and $q \equiv 3 \mod 4$, then*

$$\left( \frac{q}{p} \right) = -\left( \frac{p}{q} \right).$$

The law of quadratic reciprocity is, without a doubt, mysterious.

**Working with the Legendre symbol**

Before getting to the difficult proof of the reciprocity law, let’s work out a few examples.

**Example 6.17.** To decide if 7 has a quadratic residue modulo 109, we can use quadratic reciprocity to determine the Legendre symbol $\left( \frac{7}{109} \right)$. Thus,

$$\left( \frac{7}{109} \right) = \left( \frac{109}{7} \right) \quad \text{since } 109 \equiv 1 \mod 4$$

$$= \left( \frac{4}{7} \right) \quad \text{since } 109 \equiv 4 \mod 7$$

$$\equiv \left( \frac{2}{7} \right) \left( \frac{2}{7} \right) \quad \text{since the Legendre symbol is multiplicative}$$

$$= \left( \frac{2}{7} \right)^2 = 1.$$
We learn that 7 has a quadratic residue modulo 109.

**Example 6.18.** Does 55 have a quadratic residue modulo 179? By applying quadratic reciprocity judiciously we get:

\[
\left( \frac{55}{179} \right) = \left( \frac{5}{179} \right) \left( \frac{11}{179} \right)
\]

because the Legendre symbol is multiplicative

\[
= \left( \frac{179}{5} \right) \left( -\frac{179}{11} \right)
\]

since \( 5 \equiv 1 \mod 4, 179 \equiv 3 \mod 4, 11 \equiv 3 \mod 4 \)

\[
= -\left( \frac{4}{5} \right) \left( \frac{3}{11} \right)
\]

since \( 179 \equiv 4 \mod 5, 179 \equiv 3 \mod 11 \)

\[
= -\left( -\frac{1}{5} \right) \left( -\frac{11}{3} \right)
\]

since \( 4 \equiv -1 \mod 5, 11 \equiv 3 \mod 4 \) and \( 3 \equiv 3 \mod 4 \)

\[
= \left( -\frac{1}{5} \right) \left( \frac{2}{3} \right)
\]

since \( 11 \equiv 2 \mod 3 \)

\[
= 1 \left( \frac{2}{3} \right)
\]

because of Proposition 6.10 and \( 5 \equiv 1 \mod 4 \)

\[
= -1
\]

because of Proposition 6.14 and \( 3 \equiv 3 \mod 8 \)

Thus 55 does not have a quadratic residue modulo 179.

Here is one more illustration of quadratic reciprocity.
Example 6.19. Does $-299$ have a quadratic residue modulo $397$? Well,

\[
\left(\frac{-299}{397}\right) = \left(\frac{-1}{397}\right) \left(\frac{299}{397}\right) \quad \text{since the Legendre symbol is multiplicative}
\]

\[
= \left(\frac{299}{397}\right) \quad \text{since } 397 \equiv 1 \pmod{4}
\]

\[
= \left(\frac{13}{397}\right) \left(\frac{23}{397}\right) \quad \text{since } 13 \cdot 23 = 299
\]

\[
= \left(\frac{397}{13}\right) \left(\frac{397}{23}\right) \quad \text{since } 397 \equiv 1 \pmod{4}
\]

\[
= \left(\frac{7}{13}\right) \left(\frac{6}{23}\right) \quad \text{since } 397 \equiv 7 \pmod{13} \text{ and } 397 \equiv 6 \pmod{23}
\]

\[
= \left(\frac{7}{13}\right) \left(\frac{2}{23}\right) \left(\frac{3}{23}\right) \quad \text{since } 6 = 2 \cdot 3
\]

\[
= \left(\frac{13}{7}\right) \left(\frac{2}{23}\right) \left(\frac{3}{23}\right) \quad \text{since } 13 \equiv 1 \pmod{4}
\]

\[
= \left(\frac{-1}{7}\right) \left(\frac{2}{23}\right) \left(\frac{3}{23}\right) \quad \text{since } 13 \equiv -1 \pmod{7}
\]

\[
= (-1)(+1) \left(\frac{3}{23}\right) \quad \text{since } 7 \equiv 3 \pmod{4} \text{ and } 23 \equiv 7 \pmod{8}
\]

\[
= - \left(-\left(\frac{23}{3}\right)\right) \quad \text{since } 3 \equiv 3 \pmod{4} \text{ and } 23 \equiv 3 \pmod{4}
\]

\[
= \left(\frac{2}{3}\right) \quad \text{since } 23 \equiv 2 \pmod{3}
\]

\[
= -1 \quad \text{since } 3 \equiv 3 \pmod{8}
\]

Thus $-299$ does not have a quadratic residue modulo $397$.

### 6.4 Quadratic reciprocity—the proof

The discovery of quadratic reciprocity is one of Gauss’ crowning achievements. In order to appreciate it, one should make an attempt to understand at least one of its proofs. Scores of proofs have followed Gauss’ original. He himself came up with no less than six proofs. The one offered here is broken into a number of small steps. Notwithstanding, some diligent reading will be required.
Symmetric reduction modulo $p$

We need some terminology to facilitate the discussion. As usual, $p$ is an odd prime. We know that every integer $a$, where $p \nmid a$, is congruent modulo $p$ to its remainder in the list

$$1, 2, 3, \ldots, \frac{p - 1}{2}, \frac{p + 1}{2}, \ldots, p - 3, p - 2, p - 1.$$ 

By subtracting $a$ from the integers

$$\frac{p + 1}{2}, \ldots, p - 3, p - 2, p - 1,$$

in the second half of this list, we see that they are respectively congruent to

$$-\frac{p - 1}{2}, \ldots, -3, -2, -1.$$

Thus every integer $a$, where $p \nmid a$, is congruent modulo $p$ to exactly one of the integers in the symmetric list

$$-\frac{p - 1}{2}, \ldots, -3, -2, -1, 1, 2, 3, \ldots, \frac{p - 1}{2}.$$ 

For example, using $p = 13$ every integer $a$, not divisible by 13, is congruent to exactly one integer in the list

$$-6, -5, -4, -3, -2, -1, 1, 2, 3, 4, 5, 6.$$ 

We shall refer to the above list as the symmetric system of remainders for $p$. Half of the integers in the symmetric system are negative, and half are positive. (The remainder 0 does not become involved since $p \nmid a$.) If $p \mid a$, the unique integer $s$ in the symmetric system congruent to $a$ will be called the symmetric reduction of $a$ modulo $p$.

For some $a$ the symmetric reduction is negative, and for others it is positive. For example, the symmetric reduction of 34 modulo 13 is $-5$, while the symmetric reduction of 29 modulo 13 is 3.

If $p \nmid a$, the symmetric reductions of the integers

$$a, 2a, 3a, \ldots, \frac{p - 1}{2}a,$$
play a major role in the proof of the reciprocity law. Since no two of these integers are congruent modulo \( p \), their symmetric reductions never repeat.

For example, take \( p = 13 \) and \( a = 10 \). Then \( \frac{p-1}{2} = 6 \). As \( k \) runs from 1 to 6, the symmetric reductions of \( k \cdot 10 \) modulo 13 are given as follows:

- \( 1 \cdot 10 \equiv -3 \)
- \( 2 \cdot 10 \equiv -6 \)
- \( 3 \cdot 10 \equiv 4 \)
- \( 4 \cdot 10 \equiv 1 \)
- \( 5 \cdot 10 \equiv -2 \)
- \( 6 \cdot 10 \equiv -5 \).

Something interesting to notice is that, as \( k \) ran from 1 to 6, the symmetric reductions of \( k \cdot 10 \) gave back exactly the integers from 1 to 6, with only a possible corruption of their sign. For instance, we did not get back both +4 and −4. This is a general phenomenon.

**Proposition 6.20.** Let \( p \) be an odd prime, and \( a \) an integer such that \( p \nmid a \). Let

\[
s_1, s_2, s_3, \ldots, s_{\frac{p-1}{2}}
\]

be the symmetric reductions, modulo \( p \), of the integers

\[
a, 2a, 3a, \ldots, \frac{p-1}{2}a.
\]

Then the list of their absolute values,

\[
|s_1|, |s_2|, |s_3|, \ldots, |s_{\frac{p-1}{2}}|
\]

picks up every integer from 1 to \( \frac{p-1}{2} \).

**Proof.** As \( k \) runs from 1 to \( \frac{p-1}{2} \), the absolute values \( |s_k| \) are integers from 1 to \( \frac{p-1}{2} \). To see that the \( |s_k| \) pick up all of these integers from 1 to \( \frac{p-1}{2} \), it suffices to notice that the list

\[
|s_1|, |s_2|, \ldots, |s_{\frac{p-1}{2}}|
\]

never repeats itself.

Well, if \( |s_k| = |s_\ell| \) for some \( k, \ell \) from 1 to \( \frac{p-1}{2} \), we have either \( s_k = s_\ell \) or \( s_k = -s_\ell \). In the first case, we deduce \( k = \ell \), since the symmetric reductions are distinct for differing \( k \)'s. And the second case never happens. For if it happened, we would have:

\[
(k + \ell)a = ka + \ell a \equiv s_k + s_\ell = 0 \mod p.
\]
Since \( p \nmid a \), we could cancel \( a \) to get \( k + \ell \equiv 0 \mod p \). But \( k, \ell \) run from 1 to \( \frac{p-1}{2} \), and therefore \( k + \ell \) is somewhere between 2 and \( p - 1 \), whence \( k + \ell \neq 0 \mod p \). \( \square \)

As \( k \) runs from 1 to \( \frac{p-1}{2} \), some of the symmetric reductions of the \( ka \)'s will be negative and some positive. It is important for us to keep track of the number that are negative. Define

\[
\mu_a = \text{the number of } k \text{ from } 1 \text{ to } \frac{p-1}{2}
\]

for which the symmetric reduction of \( ka \) is negative.

In the example preceding Proposition 6.20, with \( p = 13 \) and \( a = 10 \), notice that \( \mu_{10} = 4 \). Another thing to notice in that example is that 10 has a quadratic residue modulo 13. Indeed, \( 10 \equiv 6^2 \mod 13 \). The fact that 10 has a quadratic residue modulo 13 and that \( \mu_{10} \) is even is no coincidence. For that is the essence of the next general fact, known as Gauss' Lemma.

Gauss' lemma

Many proofs of quadratic reciprocity make use of Gauss' Lemma. The proof of Gauss' Lemma somewhat imitates the proof of Euler's Theorem 3.16, but with considerable added complexity.

**Proposition 6.21 (Gauss' Lemma).** If \( p \) is an odd prime and \( p \) does not divide an integer \( a \), then

\[
\left( \frac{a}{p} \right) = (-1)^{\mu_a}.
\]

In other words, \( a \) has a quadratic residue modulo \( p \) if and only if the list

\[
a, 2a, 3a, \ldots, \frac{p - 1}{2}a
\]

yields an even number of negative symmetric reductions, modulo \( p \).

**Proof.** As \( k \) runs from 1 to \( \frac{p-1}{2} \), let \( s_k \) be the symmetric reduction of \( ka \).

The number of \( s_k \) that are negative is precisely \( \mu_a \).

By replacement modulo \( p \) we have:

\[
\frac{p - 1}{2}! \cdot a^\frac{p-1}{2} = a \cdot 2a \cdot 3a \cdots \frac{p - 1}{2}a \equiv s_1 \cdot s_2 \cdot s_3 \cdots s_{\frac{p-1}{2}} \mod p.
\]
According to Proposition 6.20, the $|s_k|$ pick up all integers from 1 to $\frac{p-1}{2}$. After a bit of thinking, it becomes clear that

$$s_1 \cdot s_2 \cdot s_3 \cdots s_{\frac{p-1}{2}} = (-1)^{\mu_a} \cdot 1 \cdot 2 \cdot 3 \cdots \frac{p-1}{2} = \frac{p-1}{2}! \cdot (-1)^{\mu_a}.$$ 

Hence,

$$\frac{p-1}{2}! \cdot a^{\frac{p-1}{2}} \equiv \frac{p-1}{2}! \cdot (-1)^{\mu_a} \mod p.$$

Since $p \nmid \frac{p-1}{2}!$, cancel the $\frac{p-1}{2}!$ and obtain

$$a^{\frac{p-1}{2}} \equiv (-1)^{\mu_a} \mod p.$$

Since Euler’s test 6.8 gives

$$a^{\frac{p-1}{2}} \equiv \left( \frac{a}{p} \right) \mod p,$$

the desired conclusion follows. \qed

Gauss’ Lemma is the first hurdle under our belt. The next challenge is to have a way of figuring out $\mu_a$, the number of negative symmetric reductions of the $ka$ as $k$ runs from 1 to $\frac{p-1}{2}$. Actually, in order to exploit Gauss’ Lemma, all we need is a way to tell whether $\mu_a$ is even or odd.

Calculating the parity of $\mu_a$ using quotient sums when $a$ is odd

Continue with an odd prime $p$ and an integer $a$ where $p \nmid a$, and now suppose that $a$ is also odd. Any integer congruent to $\mu_a$ modulo 2 will be even when $\mu_a$ is even, and odd when $\mu_a$ is odd. Now we go and find such an integer which we can systematically calculate.

**Proposition 6.22.** Let $p$ be an odd prime and $a$ an odd, positive integer such that $p \nmid a$. For each $k$ from 1 to $\frac{p-1}{2}$, let $q_k, r_k$ be the respective quotient and remainder obtained when $ka$ is divided by $p$. Thus,

$$ka = q_k p + r_k \text{ where } 0 \leq r_k < p.$$ 

The parity of $\mu_a$ is the same as the parity of $\sum_{k=1}^{\frac{p-1}{2}} q_k$. In other words,

$$\mu_a \equiv \sum_{k=1}^{\frac{p-1}{2}} q_k \mod 2.$$
Proof. Clearly,

\[
a \sum_{k=1}^{\frac{p-1}{2}} k = \sum_{k=1}^{\frac{p-1}{2}} ka = \sum_{k=1}^{\frac{p-1}{2}} (q_k p + r_k) = p \sum_{k=1}^{\frac{p-1}{2}} q_k + \sum_{k=1}^{\frac{p-1}{2}} r_k.
\]

For each \( k \) from 1 to \( \frac{p-1}{2} \), let \( s_k \) be the symmetric reduction of \( k \) modulo \( p \). As shown in Proposition 6.20, the absolute values of these reductions give back every integer from 1 to \( \frac{p-1}{2} \). Thus

\[
\sum_{k=1}^{\frac{p-1}{2}} k = \sum_{k=1}^{\frac{p-1}{2}} |s_k|.
\]

By putting together the above observations we come to

\[
a \sum_{k=1}^{\frac{p-1}{2}} |s_k| = p \sum_{k=1}^{\frac{p-1}{2}} q_k + \sum_{k=1}^{\frac{p-1}{2}} r_k.
\]

Next notice that for every \( s_k \) (in fact, for every integer)

\[
|s_k| \equiv s_k \mod 2.
\]

And the fact \( a, p \) are odd means

\[
a \equiv 1 \text{ and } p \equiv 1 \mod 2.
\]

Thus the preceding equation leads to the following congruence modulo \( 2 \):

\[
\sum_{k=1}^{\frac{p-1}{2}} s_k \equiv \sum_{k=1}^{\frac{p-1}{2}} q_k + \sum_{k=1}^{\frac{p-1}{2}} r_k \mod 2.
\]

Since \( q_k \equiv -q_k \mod 2 \) the above congruence becomes

\[
\sum_{k=1}^{\frac{p-1}{2}} q_k \equiv \sum_{k=1}^{\frac{p-1}{2}} (r_k - s_k) \mod 2.
\]

Now comes a key observation about the symmetric reductions \( s_k \). If \( s_k > 0 \), then \( s_k = r_k \), and so \( r_k - s_k = 0 \). But if \( s_k < 0 \), then \( r_k - s_k = p \). (If this is not clear,
just notice that \(-\frac{p-1}{2} \leq s_k \leq -1\). Add \(p\) to get \(\frac{p+1}{2} \leq s_k + p \leq p - 1\), which reveals that \(s_k + p = r_k\).

The congruence modulo 2 comes down to

\[
\sum_{k=1}^{\frac{p-1}{2}} q_k \equiv \sum_{s_k < 0} p,
\]

where the second sum is saying to add \(p\) to itself once for each time \(s_k < 0\). Recalling again that \(p \equiv 1 \mod 2\) we get

\[
\sum_{k=1}^{\frac{p-1}{2}} q_k \equiv \sum_{s_k < 0} 1.
\]

Since the last summation is nothing but \(\mu_a\), the number of times that \(s_k < 0\), our result is shown.

From Propositions 6.22 and 6.21 it becomes apparent that an odd integer \(a\) has a quadratic residue modulo \(p\) if and only if the sum of the quotients for the first \(\frac{p-1}{2}\) multiples of \(a\), upon division by \(p\), is an even number.

For example, let’s decide if 17 has a quadratic residue modulo 31, by means of Proposition 6.22. With a calculator, divide 31 into each of the multiples

\[
1 \cdot 17, 2 \cdot 17, 3 \cdot 17, \ldots, 15 \cdot 17
\]

to get the quotients, and then add them up. Omitting the tedious calculation (which can be done fast on a spreadsheet), we obtain that the quotients add up to 59, an odd number. Thus 17 does not have a quadratic residue modulo 31.

**The quotients as integer parts**

If \(p\) is an odd prime and \(a\) is an odd integer and \(p \nmid a\), we have the quotients and remainders \(q_k, r_k\) for each \(k\) from 1 to \(\frac{p-1}{2}\) given by

\[
ka = q_kp + r_k \text{ where } 0 < r_k < p.
\]

We might note that since \(p \nmid a\) and \(p \nmid k\) then \(p \nmid ka\) either, and so our remainders \(r_k\) are strictly positive. Since

\[
\frac{ka}{p} = q_k + \frac{r_k}{p} \text{ and } 0 < \frac{r_k}{p} < 1,
\]
the quotient $q_k$ is nothing but the last integer prior to $\frac{ka}{p}$. The greatest integer less than or equal to a real number $t$ is called the integer part of $t$. The common notation for the integer part of $t$ is $\lfloor t \rfloor$. Some also refer to this as the floor of $t$. Our quotients $q_k = \lfloor \frac{ka}{p} \rfloor$.

As noted already, the sum $\sum_{k=1}^{\frac{p-1}{2}} \left\lfloor \frac{ka}{p} \right\rfloor$ of the integer parts of the $\frac{ka}{p}$ is important to us, because, in accordance with Proposition 6.22 and Gauss’ Lemma 6.21, the integer $a$ has a quadratic residue modulo $p$ if and only if this sum is even. We can record that fact in terms of the Legendre symbol.

**Proposition 6.23.** If $p$ is an odd prime, and $a$ is an odd integer $a$ for which $p \nmid a$, then

$$\left( \frac{a}{p} \right) = (-1)^{\sum_{k=1}^{\frac{p-1}{2}} \left\lfloor \frac{ka}{p} \right\rfloor}.$$ 

A geometric interpretation of the sum of the integer parts

The next result provides a nice geometrical interpretation of the sum $\sum_{k=1}^{\frac{p-1}{2}} \left\lfloor \frac{ka}{p} \right\rfloor$ in terms of integer lattice points. A point $(x, y)$ in the $xy$-plane will be called an integer lattice point when both $x$ and $y$ are integers. For instance, $(-2, 3), (0, 0), (5, -8)$ and $(1, 1)$ are integer lattice points while $(\sqrt{2}, 7), (3/8, 1/3)$ are not. The integer lattice points in the plane are in a grid formation, as shown.
Proposition 6.24. Let \( p \) be an odd prime that does not divide an odd integer \( a \).

The sum
\[
\sum_{k=1}^{\frac{p-1}{2}} \left\lfloor \frac{ka}{p} \right\rfloor
\]
equals the number of integer lattice points strictly inside the triangle bounded by the \( x \)-axis, the vertical line \( x = \frac{p}{2} \) and the line \( y = \frac{a}{p} x \), as shown.

Proof. For each \( k \) from 1 to \( \frac{p-1}{2} \), the integer lattice points
\[(k, 1), (k, 2), (k, 3), \ldots, (k, j), \ldots\]
along the vertical line \( x = k \), will remain below the line \( y = \frac{a}{p} x \) as long as \( 1 \leq j \leq \left\lfloor \frac{ka}{p} \right\rfloor \). (This is the key point in this proof.) After that they go above this line. Note that since \( 1 \leq k \leq \frac{p-1}{2} \) and \( p \nmid a \), then \( p \nmid ka \), and so there are no integer lattice points on the line \( y = \frac{a}{p} x \) when \( 1 \leq x \leq \frac{p-1}{2} \).

So, the number of integer lattice points along the line \( x = k \) and strictly inside our triangle equals \( \left\lfloor \frac{ka}{p} \right\rfloor \). The number of integer lattice points strictly inside the
given triangle is just the total of the number of integer lattice points inside the
triangle and along the lines $x = k$, as $k$ varies from 1 to $\frac{p-1}{2}$. That total is precisely
our desired quantity $\sum_{k=1}^{\frac{p-1}{2}} \left\lfloor \frac{ka}{p} \right\rfloor$.

Proof of quadratic reciprocity-final steps

To finish proving Theorem 6.16, we count the number of integer lattice points
inside a rectangle in two different ways.

**Proposition 6.25.** If $p, q$ are odd and distinct primes, then

$$\frac{p-1}{2} \cdot \frac{q-1}{2} = \sum_{k=1}^{\frac{p-1}{2}} \left\lfloor \frac{kq}{p} \right\rfloor + \sum_{\ell=1}^{\frac{q-1}{2}} \left\lfloor \frac{\ell p}{q} \right\rfloor.$$

**Proof.** Examine the rectangle in the $xy$-plane whose vertices are

$$(0, 0), \left(\frac{p}{2}, 0\right), \left(\frac{p}{2}, \frac{q}{2}\right), \left(0, \frac{q}{2}\right).$$

The diagonal line from $(0, 0)$ to $\left(\frac{p}{2}, \frac{q}{2}\right)$ is given by $y = \frac{q}{p}x$, as shown.
CHAPTER 6. QUADRATIC RESIDUES

The integer lattice points strictly inside this rectangle make up the rectangular grid of points \((k, \ell)\) where \(k\) runs from 1 to \(p - \frac{1}{2}\) and \(\ell\) runs from 1 to \(q - \frac{1}{2}\). There are \(\frac{p - 1}{2} \cdot \frac{q - 1}{2}\) such integer lattice points.

There are no integer lattice points on the diagonal of this rectangle and strictly inside the rectangle. Indeed, if \((k, \ell)\) were an integer lattice point on this diagonal line, we would have \(p\ell = qk\). Since \(p \nmid q\), we would get \(p \mid k\). But that’s not possible since \(1 \leq k \leq \frac{p - 1}{2}\).

So, the integer lattice points in the interior of this rectangle are either in the interior of the triangle below the line \(y = \frac{q}{p}x\), or in the interior of the triangle above this diagonal line. According to Proposition 6.24 (with \(q = a\)) the number of integer lattice points in the interior of the lower triangle is

\[
\sum_{k=1}^{\frac{p-1}{2}} \left\lfloor \frac{kp}{p} \right\rfloor.
\]
6.4. QUADRATIC RECIPROCITY - THE PROOF

By interchanging the roles of $p$ and $q$, we see that the number of integer lattice points in interior of the upper triangle is

$$\sum_{\ell=1}^{q-1} \left\lfloor \frac{\ell p}{q} \right\rfloor$$

Since all integer lattice points in the interior of the rectangle are in the interior of the lower or the interior of the upper triangle, we take our counts for the integer lattice points in each of the three interiors to get:

$$\frac{p-1}{2} \cdot \frac{q-1}{2} = \sum_{k=1}^{p-1} \left\lfloor \frac{kq}{p} \right\rfloor + \sum_{\ell=1}^{q-1} \left\lfloor \frac{\ell p}{q} \right\rfloor,$$

as claimed. \qed

And now comes Gauss’ famous quadratic reciprocity theorem. The formula may seem at first obscure, but after its proof we will interpret what it means.

**Theorem 6.26.** If $p, q$ are odd, distinct primes, then

$$\left( \frac{p}{q} \right) \left( \frac{q}{p} \right) = (-1)^{\frac{p-1}{2} \cdot \frac{q-1}{2}}.$$

**Proof.** By Propositions 6.23 and 6.25 we come to

$$\left( \frac{p}{q} \right) \left( \frac{q}{p} \right) = (-1)^{\sum_{k=1}^{p-1} \left\lfloor \frac{kq}{p} \right\rfloor} \cdot (-1)^{\sum_{\ell=1}^{q-1} \left\lfloor \frac{\ell p}{q} \right\rfloor}$$

$$= (-1)^{\sum_{k=1}^{p-1} \left\lfloor \frac{kq}{p} \right\rfloor + \sum_{\ell=1}^{q-1} \left\lfloor \frac{\ell p}{q} \right\rfloor}$$

$$= (-1)^{\frac{p-1}{2} \cdot \frac{q-1}{2}}.$$

And that (painful as it may have felt) is it. \qed

To see what Theorem 6.26 means, suppose that one of $p$ or $q$ is congruent to 1 modulo 4. Say $p \equiv 1 \mod 4$. Since $4 \mid p - 1$, the integer $\frac{p-1}{2}$ remains even, and so does $\frac{q-1}{2}$. Then the right hand side of the formula in Theorem 6.26 gives $+1$. Thus $\left( \frac{p}{q} \right)$ and $\left( \frac{q}{p} \right)$ have the same sign, and since each of them takes one of the values $\pm 1$, they are both 1 or both $-1$. 
In other words, if one of \( p \) or \( q \) is congruent to 1 modulo 4, then \( \left( \frac{p}{q} \right) = \left( \frac{q}{p} \right) \).

On the other hand, if both \( p \) and \( q \) are congruent to 3 modulo 4, and \( 4 | q - 3 \). So, \( \frac{p-3}{2} \) and \( \frac{q-3}{2} \) remain even. Since \( \frac{p-1}{2} = \frac{p-3}{2} + 1 \) and \( \frac{q-1}{2} = \frac{q-3}{2} + 1 \), we see that both \( \frac{p-1}{2} \) and \( \frac{q-1}{2} \) are odd. Then their product is odd, whence the right side of the formula in Theorem 6.26 gives \( -1 \). Consequently, \( \left( \frac{p}{q} \right) \) and \( \left( \frac{q}{p} \right) \) have opposite sign. This means that when one of them is 1, the other one is \(-1\).

In other words, if \( p \) and \( q \) are congruent to 3 modulo 4, then \( \left( \frac{p}{q} \right) = - \left( \frac{q}{p} \right) \).

And so we see that Theorem 6.26 really comes down to the famous quadratic reciprocity law stated in Theorem 6.16. If nothing else, the proof of quadratic reciprocity shows us that significant mathematics demands serious effort.

### 6.5 Exercises

1. If an integer \( g \) represents a primitive root modulo an odd prime \( p \), show that
   \[
   \left( \frac{g^k}{p} \right) = (-1)^k.
   \]

2. If \( p \) is a prime and \( p \geq 5 \), show that either all three of 2, 3, 6 have quadratic residues modulo \( p \) or just one of them has a quadratic residue modulo \( p \).

3. Compute the following Legendre symbols.
   - (a) \( \left( \frac{85}{101} \right) \)
   - (b) \( \left( \frac{29}{541} \right) \)
   - (c) \( \left( \frac{101}{1987} \right) \)

4. Determine if the following congruences are solvable. You do not need to actually solve them.
   - (a) \( 3x^2 + 6x + 5 \equiv 0 \pmod{89} \)
   - (b) \( x^2 - 3x - 1 \equiv 0 \pmod{31957} \). You can assume that 31957 is prime.

5. Solve \( 12x^2 + 28x + 1 \equiv 0 \pmod{35} \)
   
   Hint. First break the problem down to two congruences \( \pmod{5} \) and \( \pmod{7} \).
6.5. EXERCISES

6. Prove there exist infinitely many primes congruent to $-1$ modulo $8$.

Hint. If $p_1, p_2, \ldots, p_k$ is a list of such primes, look at $x = (4p_1p_2 \cdots p_k)^2 - 2$. You may need the test for $2$ to be a quadratic residue modulo a prime. First deduce that the prime factors of $x$ are congruent to $1$ or $7 \mod 8$. Then show they all cannot be congruent to $1 \mod 8$.

7. If $p$ is a prime and $p \equiv 3 \mod 4$ and if $x \not\equiv 0 \mod p$, show that $x^3 + x \not\equiv 0 \mod p$.

8. Using quadratic reciprocity determine which primes $p \geq 5$ are such that $3$ has a quadratic residue $\mod p$. In other words devise a rule for finding\[ \left( \frac{3}{p} \right). \]

9. The integer $p = 1 + 8 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 \cdot 17 \cdot 19 \cdot 23 = 892371481$ is prime.

Show that $\left( \frac{q}{p} \right) = 1$ for all primes $q$ from 2 to 23.

Then evaluate $\left( \frac{a}{p} \right)$ for all integers $a$ from 1 to 28.
Chapter 7

Sums of squares and other powers

In this chapter we aim to provide definitive answers to two problems involving the sum of squares. First we shall discover all possible ways that a sum of two squares \(a^2 + b^2\) is itself another perfect square, such as, for example, \(3^2 + 4^2 = 5^2\). In other words, we find all integer solutions \(x, y, z\) to the Diophantine equation
\[
x^2 + y^2 = z^2.
\]
Then we will look for a test to decide if an arbitrary integer \(n\) is a sum of two squares. In other words, we decide when the Diophantine equation
\[
x^2 + y^2 = n
\]
has a solution \(x, y\) by simply looking at \(n\).

Let us start with the first problem, which is probably the easier of the two.

7.1 Pythagorean triples

**Definition 7.1.** A triplet of integers \((a, b, c)\) is a *Pythagorean triple* if
\[
a^2 + b^2 = c^2 \text{ and not all } a, b, c = 0.
\]

In a Pythagorean triple, \(c = 0\) forces \(a = b = 0\). Since a Pythagorean triple requires that \(c \neq 0\), we shall feel free to divide by \(c\).

With a bit of mental arithmetic we see that the triples \((1, 0, 1)\), \((3, 4, 5)\) and \((5, 12, 13)\) are Pythagorean.

Our goal is to find the general solution to \(x^2 + y^2 = z^2\).
Pythagorean triples and the unit circle

If \((a, b, c)\) is a Pythagorean triple, then
\[
\left(\frac{a}{c}\right)^2 + \left(\frac{b}{c}\right)^2 = 1,
\]
and so the point \(\left(\frac{a}{c}, \frac{b}{c}\right)\) lies on the unit circle given by the equation \(x^2 + y^2 = 1\).

The coordinates \(\frac{a}{c}, \frac{b}{c}\) of this point are rational numbers. Such points on the unit circle will be called rational points. For instance, \((\frac{4}{5}, 35)\) is a rational point on the unit circle, while \((\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})\) is on the unit circle, but is not a rational point.

The key to finding all Pythagorean triples is to grab the rational points on the unit circle, as the next result explains.

**Proposition 7.2.** For every rational number \(m\), the rational point
\[
\left(\frac{1 - m^2}{1 + m^2}, \frac{2m}{1 + m^2}\right)
\]
lies on the unit circle. Furthermore, every rational point on the circle, except for \((-1, 0)\), takes the above form.
Proof. Clearly, points of the given form are rational points. Also clearly,

\[
\left( \frac{1 - m^2}{1 + m^2} \right)^2 + \left( \frac{2m}{1 + m^2} \right)^2 = \frac{1 - 2m^2 + m^4 + 4m^2}{(1 + m^2)^2} = \frac{1 + 2m^2 + m^4}{(1 + m^2)^2} = \frac{(1 + m^2)^2}{(1 + m^2)^2} = 1,
\]

which confirms that \( \left( \frac{1 - m^2}{1 + m^2}, \frac{2m}{1 + m^2} \right) \) is a rational point on the circle.

The next, and harder, step is to see that every rational point on the circle, except for \((-1, 0)\), originates in this way. Let \((r, s)\) be a rational point on the circle, other than \((-1, 0)\), as shown.

Here

\[(r, s) \neq (-1, 0), \ r^2 + s^2 = 1 \text{ and } r, s \in \mathbb{Q}.
\]

The slope \(m\) of the line joining \((-1, 0)\) to \((r, s)\) is the rational number

\[m = \frac{s}{r + 1}.
\]

Note, by the way, \(r \neq -1\). The equation of this line is \(y = m(x + 1)\). Thus
\[ s = m(1 + r). \text{ But } (r, s) \text{ is also on the circle whose equation is } x^2 + y^2 = 1. \text{ Thus} \]
\[ r^2 + (m(r + 1))^2 = 1 \]
\[ r^2 + m^2r^2 + 2m^2r + m^2 = 1 \]
\[ (1 + m^2)r^2 + 2m^2r + m^2 - 1 = 0 \]

By the quadratic formula,
\[ r = \frac{-2m^2 \pm \sqrt{4m^4 - 4(1 + m^2)(m^2 - 1)}}{2(1 + m^2)} \]
\[ = \frac{-2m^2 \pm \sqrt{4}}{2(1 + m^2)} \]
\[ = \frac{-m^2 \pm 1}{1 + m^2} \]
\[ = \frac{1 - m^2}{1 + m^2} \text{ or } r = -1. \]

Since \( r \neq -1 \), we conclude
\[ r = \frac{1 - m^2}{1 + m^2}. \]

And then,
\[ s = m(r + 1) = m \left( \frac{1 - m^2}{1 + m^2} + 1 \right) = m \left( \frac{1 - m^2 + 1 + m^2}{1 + m^2} \right) = \frac{2m}{1 + m^2}. \]

Therefore the rational point \((r, s)\) on the unit circle is represented as \(\left( \frac{1 - m^2}{1 + m^2}, \frac{2m}{1 + m^2} \right)\), as long as \( r \neq -1 \).

**Primitive Pythagorean triples are all that matter**

Getting back to Pythagorean triples \((a, b, c)\) where \( c \neq 0 \), notice that for any non-zero integer \(d\), the triple \((da, db, dc)\) remains Pythagorean. That is,
\[ (da)^2 + (db)^2 = d^2(a^2 + b^2) = (dc)^2. \]

Because of this, the only interesting triples \((a, b, c)\) are those for which \(a, b, c\) do not have a common prime factor. When there is no prime \(p\) such that \(p \mid a, p \mid b\) and \(p \mid c\) we say that \((a, b, c)\) is a **primitive Pythagorean triple**.
7.1. PYTHAGOREAN TRIPLES

It’s worth noting that if \((a, b, c)\) is a primitive Pythagorean triple, then the stronger condition holds that any two of \(a, b, c\) are coprime. Indeed, suppose for instance that \(a, b\) are not coprime. From the equation \(a^2 + b^2 = c^2\) it follows that \(p\) also divides \(c\), and so the triple is not primitive.

Here is a useful little property that primitive triples come with.

**Proposition 7.3.** If \((a, b, c)\) is a primitive Pythagorean triple, then \(c\) is odd. Also \(a, b\) cannot be both even or both odd.

**Proof.** If \(c\) is even so is \(c^2 = a^2 + b^2\) and thus both \(a, b\) are even or both \(a, b\) are odd. Now \(a, b\) can’t both be even, for then 2 would divide all of \(a, b, c\), contrary to the primitiveness of \((a, b, c)\). And if both \(a, b\) are odd, then both \(a\) and \(b\) are congruent to \(\pm 1\) modulo 4, and thus

\[
c^2 = a^2 + b^2 \equiv (\pm 1)^2 + (\pm 1)^2 = 2 \mod 4.
\]

However,

\[
0^2 \equiv 0, \ 1^2 \equiv 1, \ 2^2 \equiv 0, \ 3^2 \equiv 1 \mod 4,
\]

which tells us that \(c^2 \equiv 2 \mod 4\) is impossible. The only way out is to have \(c\) be odd. This automatically forces one of \(a\) or \(b\) to be even and the other one to be odd. \(\square\)

**The parametrization of primitive triples**

A couple of baby things need to be sorted out. First is that a change of sign in any of \(a, b, c\) will not alter whether \((a, b, c)\) is Pythagorean or not. So we may as well suppose that all of \(a, b, c \geq 0\). Let’s call such a triple **non-negative**. Second is that if \((a, b, c)\) is Pythagorean, so is \((b, a, c)\). The roles of \(a\) and \(b\) are interchangeable. When \((a, b, c)\) is primitive one of \(a\) of \(b\) is odd while the other one is even. Since the roles of \(a\) and \(b\) can be interchanged, there is no harm done in supposing that

\[
a \text{ is odd, while } b \text{ is even.}
\]

We have enough information collected to decide how a non-negative, primitive Pythagorean triple must arise.
Theorem 7.4. If \((a, b, c)\) is a non-negative, primitive Pythagorean triple, with \(a\) odd and \(b\) even, then there are coprime integers \(r, s\) such that

\[
\begin{align*}
    r, s &\text{ have different parity} \\
    0 &\leq r < s \\
    a &\equiv s^2 - r^2 \\
    b &\equiv 2rs \\
    c &\equiv s^2 + r^2.
\end{align*}
\]

Proof. Since \(a\) is odd, \(a\) is not 0. If \(b = 0\), our primitive triple has to be \((1, 0, 1)\), in which case, use \(r = 0, s = 1\) to get \(a, b, c\) as prescribed. So we are down to the case where \(a, b, c\) are all strictly positive, which is really the only case that’s interesting.

Since \(a > 0, b > 0\) and \(a^2 + b^2 = c^2\), it follows that \(c > a\) and \(c > b\), and thus

\[
0 < \frac{a}{c} < 1 \quad \text{and} \quad 0 < \frac{b}{c} < 1.
\]

Thereby \((\frac{a}{c}, \frac{b}{c})\) is a rational point in the first quadrant of the unit circle.

From Proposition 7.2

\[
\frac{a}{c} = \frac{1 - m^2}{1 + m^2} \quad \text{and} \quad \frac{b}{c} = \frac{2m}{1 + m^2},
\]

where \(m\) is a rational number representing the slope of the line from \((-1, 0)\) to \((\frac{a}{c}, \frac{b}{c})\). Since \((\frac{a}{c}, \frac{b}{c})\) is in the first quadrant and on the circle, we get \(0 < m < 1\). Write \(m = \frac{r}{s}\), where \(r, s\) are coprime integers such that \(0 < r < s\). Put this into the equations above to arrive at

\[
\frac{a}{c} = \frac{1 - \left(\frac{r}{s}\right)^2}{1 + \left(\frac{r}{s}\right)^2} = \frac{s^2 - r^2}{s^2 + r^2}.
\]
7.1. PYTHAGOREAN TRIPLES

and

\[
\frac{b}{c} = \frac{2 \left( \frac{r}{s} \right)}{1 + \left( \frac{r}{s} \right)^2} = \frac{2rs}{s^2 + r^2}.
\]

At this point it would be desirable to equate numerators and denominators in the above equalities of rational numbers, but first a bit care is needed to ensure that the fractions involved are in reduced form.

Remember that \( b \) is even and \( c \) is odd. The second of the equations above gives

\[
(s^2 + r^2) \frac{b}{2} = rsc.
\]

Since \( r, s \) are coprime, they cannot both be even. And if they are both odd, we get that \( (r^2 + s^2) \frac{b}{2} \) is even while \( rsc \) is odd. So \( r, s \) have different parity, which tells us that both \( s^2 + r^2 \) and \( s^2 - r^2 \) are odd.

Now we verify that \( s^2 + r^2 \) and \( s^2 - r^2 \) are coprime. Well, if \( p \mid s^2 + r^2 \) and \( p \mid s^2 - r^2 \) for some prime \( p \), we can add and subtract to see that \( p \mid 2s^2 \) and \( p \mid 2r^2 \), and since \( p \) has to be odd, we get \( p \mid s \) and \( p \mid r \), contrary to the fact \( s, r \) are coprime. Thus \( s^2 + r^2, s^2 - r^2 \) are coprime.

From

\[
\frac{a}{c} = \frac{s^2 - r^2}{s^2 + r^2},
\]

where \( a, c \) are coprime and \( s^2 + r^2, s^2 - r^2 \) are coprime, we deduce

\[
a = s^2 - r^2 \text{ and } c = s^2 + r^2.
\]

And then from \( \frac{b}{c} = \frac{2rs}{s^2 + r^2} \), we get \( b = 2rs \).

We need to finish off this section with the easier converse of Theorem 7.4.

**Proposition 7.5.** If \( r, s \) are coprime integers of different parity with \( 0 \leq r < s \) and if

\[
a = s^2 - r^2, \quad b = 2rs \text{ and } c = s^2 + r^2,
\]

then \((a, b, c)\) is a non-negative, primitive Pythagorean triple, wherein \( a \) is odd.
**CHAPTER 7. SUMS OF SQUARES AND OTHER POWERS**

**Proof.** To see the triple is Pythagorean just calculate

\[
\begin{align*}
    a^2 + b^2 &= (s^2 - r^2)^2 + (2rs)^2 \\
    &= s^4 - 2s^2r^2 + r^4 + 4r^2s^2 \\
    &= s^4 + 2s^2r^2 + r^4 \\
    &= (s^2 + r^2)^2 \\
    &= c^2.
\end{align*}
\]

To see that \((a, b, c)\) is primitive, suppose some prime \(p\) divides all of \(a, b, c\). Since \(s, r\) have different parity, \(a = s^2 - r^2\) is odd. So \(p \neq 2\). Since \(p | 2rs\) deduce that \(p\) divides one of \(r\) or \(s\). Then using the fact \(p | r^2 + s^2\) deduce \(p\) divides both \(r\) and \(s\), which goes against the fact \(r, s\) are coprime. Therefore, the triple \((a, b, c)\) is primitive.

The facts that \(a\) is odd and that the triple \((a, b, c)\) is non-negative are plain to see from the definitions of \(a, b, c\).

To exploit our discovery here’s some Pythagorean triples using various \(r, s\).

With \(r = 2, s = 3\), we get

\[
a = 3^2 - 2^2 = 5, \quad b = 12, \quad c = 3^2 + 2^2 = 13
\]

and hence the primitive Pythagorean triple \((5, 12, 13)\).

With \(r = 3, s = 4\), we get

\[
a = 4^2 - 3^2 = 7, \quad b = 2 \cdot 3 \cdot 4 = 24, \quad c = 25,
\]

and hence the primitive Pythagorean triple \((7, 24, 25)\).

And with \(r = 4, s = 7\), what comes out is

\[
a = 7^2 - 4^2 = 33, \quad b = 2 \cdot 4 \cdot 7 = 56, \quad c = 7^2 + 4^2 = 65,
\]

which give the primitive Pythagorean triple \((33, 56, 65)\).

### 7.2 Sums of squares in general

Which integers are the sum of two squares? This problem is quite a bit harder than the Pythagorean triples problem. To get started, let’s sample the first 36 integers.
7.2. SUMS OF SQUARES IN GENERAL

If there is a pattern, it’s well hidden. However, there sure seem to be a lot of No’s in that third column. Let’s grab that piece of low hanging fruit.

**Proposition 7.6.** If \( n \equiv 3 \mod 4 \), then \( n \) is not a sum of two squares.

**Proof.** Say \( a^2 + b^2 = n \) for some integers \( a, b \). Replace \( a, b \) by any of its possible remainders \( 0, 1, 2, 3 \) to get each of \( a^2, b^2 \) congruent, modulo 4, to one of \( 0, 1 \). Then the sum \( a^2 + b^2 \) is congruent to one of

\[
0^2 + 0^2 = 0, \quad 0^2 + 1^2 = 1 \quad \text{or} \quad 1^2 + 1^2 = 2 \mod 4,
\]

but never congruent to 3. \( \square \)

**Sums of squares and factoring among Gaussian integers**

We notice that

\[ n = a^2 + b^2 \quad \text{if and only if} \quad n = (a + ib)(a - ib) \quad \text{where} \quad i = \sqrt{-1}. \]

The sum of squares problem is a factoring problem in disguise!

Because of the above factorization of \( a^2 + b^2 \), we are forced to look at numbers of the form \( a + ib \) where \( a, b \in \mathbb{Z} \).

**Definition 7.7.** A complex number of type \( a + bi \), where \( a, b \in \mathbb{Z} \) is called a **Gaussian integer**. The set of Gaussian integers is denoted by \( \mathbb{Z}[i] \), to suggest that they are a composite of numbers from the ordinary integers \( \mathbb{Z} \) and \( i \).
Notice right away that the sum, difference and product of two Gaussian integers is again a Gaussian integer. Indeed, suppose \( a + bi, c + di \in \mathbb{Z}[i] \). Then
\[
(a + bi) \pm (c + di) = (a \pm c) + (b \pm d)i \in \mathbb{Z}[i].
\]
Also
\[
(a + bi)(c + di) = (ac - bd) + (ad + bc)i \in \mathbb{Z}[i].
\]
Furthermore, since \( 0 = 0 + 0i \) and \( 1 = 1 + 0i \) belong to \( \mathbb{Z}[i] \), the Gaussian integers form a ring. That is, a system of numbers with a \( 0 \) and a \( 1 \), and closed under addition, subtraction and multiplication. Since every ordinary integer \( a \) can be written as \( a + 0i \), we see that \( \mathbb{Z} \subseteq \mathbb{Z}[i] \).

Here’s a geometrical way to think of \( \mathbb{Z}[i] \), which will come in handy:

The dots constitute \( \mathbb{Z}[i] \). Geometrically, the Gaussian integers are the integer lattice points in the plane.

**Divisibility among Gaussian integers**

We can readily begin to factor in the ring \( \mathbb{Z}[i] \). For instance,
\[
2 = (1 + i)(1 - i), \quad 18 + i = (2 - i)(7 + 4i), \quad 5 = (2 + i)(2 - i).
\]

Just as we did with ordinary integers, we say that a Gaussian integer \( z \) divides, or is a factor of another Gaussian integer \( w \), and we write \( z \mid w \) in \( \mathbb{Z}[i] \), when
\[
w = zu \text{ for some other } u \text{ in } \mathbb{Z}[i].
\]

For instance, \( 1+i \) divides \( 2 \) in \( \mathbb{Z}[i] \) with the other factor being \( 1-i \). It is interesting to note that \( \pm 1, \pm i \) divide every \( a + bi \). Indeed,
\[
a + bi = 1(a + bi) = (-1)(-a - bi) = i(b - ai) = (-i)(-b + ai).
\]

The problem of divisibility in \( \mathbb{Z}[i] \), benefits from an important tool, that relates it back to the more familiar divisibility in \( \mathbb{Z} \). That tool is called the norm.
The norm of a Gaussian integer

The *norm function* is defined to be

\[ N : \mathbb{Z}[i] \to \{0, 1, 2, 3, \ldots\} \text{ where } N(a + bi) = a^2 + b^2. \]

Here is another way to define the norm. If \( z = a + bi \), the *complex conjugate* of \( z \) is \( \bar{z} = a - bi \). The *absolute value* of \( z \) is

\[ |z| = \sqrt{a^2 + b^2} = \sqrt{(a + ib)(a - ib)} = \sqrt{z \bar{z}}. \]

With the conjugate and absolute value notations, we have

\[ N(z) = (a + ib)(a - ib) = z \bar{z} = a^2 + b^2 = |z|^2. \]

It is well known that \( |zw| = |z||w| \), and so

\[ N(zw) = |zw|^2 = (|z||w|)^2 = |z|^2|w|^2 = N(z)N(w). \]

We will often be exploiting this fact that \( N \) is a *multiplicative function* sending Gaussian integers into non-negative integers.

Different Gaussian integers can have equal norm. For instance \( 3 + 4i \) and \( 5 + 0i \) both have norm equal to 25. But only the zero Gaussian integer can have zero norm. Namely,

\[ N(z) = 0 \text{ if and only if } z = 0. \]

Another thing to keep in mind is that an ordinary integer \( n \) is a sum of two squares if and only if \( n = N(z) \) for some \( z \) in \( \mathbb{Z}[i] \).

The classification of Gaussian units

**Definition 7.8.** A Gaussian integer \( u \) is a *unit* of \( \mathbb{Z}[i] \) when \( u | w \) for all \( w \) in \( \mathbb{Z}[i] \).

As already noted, \( \pm 1, \pm i \) are units, but could there be other units? Here is the proof that the ring \( \mathbb{Z}[i] \) only has these four units. The result also gives us a number of ways to think about units.

**Proposition 7.9.** Any one of the following statements about a Gaussian integer \( z \) implies all of the other ones.

1. \( z \) is a unit in \( \mathbb{Z}[i] \)
2. \( N(z) = 1 \)

3. \( z = \pm 1, \pm i \)

4. the inverse complex number \( z^{-1} \) is also a Gaussian integer.

**Proof.** We will prove that \( 1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 4 \Rightarrow 1 \).

Say \( z \) is a unit. Then \( z \mid 1 \), since \( z \) divides every Gaussian integer. So \( 1 = zw \) for some \( w \) in \( \mathbb{Z}[i] \). Then

\[
1 = 1^2 + 0^2 = N(1) = N(zw) = N(z)N(w).
\]

Since \( N(z), N(w) \) are positive integers, we deduce \( N(z) = 1 \) (and \( N(w) = 1 \)).

Say \( N(z) = 1 \). Then write \( z = a + bi \) where \( a, b \in \mathbb{Z} \).

We have \( a^2 + b^2 = 1 \) and so \( a^2 = 1, b = 0 \) or \( a = 0, b^2 = 1 \). Thus

\[
a = \pm 1, b = 0 \text{ which gives } z = \pm 1,
\]

or

\[
a = 0, b = \pm 1 \text{ which gives } z = \pm i.
\]

If \( z \) is one of \( 1, -1, i, -i \), its inverse is \( 1, -1, -i, i \), respectively, and these are again Gaussian integers.

Finally, say \( z \) and \( z^{-1} \) are Gaussian integers. If \( w \) is any other Gaussian integer we see that \( z \mid w \) simply by looking at \( w = z(z^{-1}w) \) and remembering that \( z^{-1}w \) is a Gaussian integer. \( \blacksquare \)

The various ways of thinking about the units of \( \mathbb{Z}[i] \), given in Proposition 7.9, will be fruitful.

**Gaussian primes**

Next we look at primes in \( \mathbb{Z}[i] \). We know that we cannot prevent any of the four units from being a factor of any Gaussian integer. Those Gaussian integers that tolerate only this trivial form of factorization need to be singled out.

**Definition 7.10.** We say that \( z \) in \( \mathbb{Z}[i] \) is a **Gaussian prime** provided

- \( z \) is not a unit, and
any factorization \( z = wu \) in \( \mathbb{Z}[i] \) forces \( w \) or \( u \) to be a unit.

This is just like ordinary integer primes, which can only be factored in \( \mathbb{Z} \) by having one of its factors be a unit, in this case \( \pm 1 \).

For instance, \( 2 \) is a prime in \( \mathbb{Z} \), but \( 2 \) is not a Gaussian prime because \( 2 = (1 + i)(1 - i) \), and neither \( 1 + i \) nor \( 1 - i \) is a unit. Likewise, \( 5 \) is not a Gaussian prime since \( 5 = (2 + i)(2 - i) \).

To practice using our norm, let’s show \( 7 \) remains a Gaussian prime. Supposing \( 7 = zw \) where \( z, w \in \mathbb{Z}[i] \), we get the following factorization in \( \mathbb{Z} \):

\[
49 = N(7) = N(zw) = N(z)N(w).
\]

Hence \( N(z) = 1, 7 \) or \( 49 \). If \( N(z) = 1 \), then \( z \) is a unit, by Proposition 7.9. If \( N(z) = 49 \), then \( N(w) = 1 \), and \( w \) is a unit. Could \( N(z) = 7 \)? Write \( z = a + bi \) where \( a, b \in \mathbb{Z} \). If \( N(z) = 7 \), then \( a^2 + b^2 = 7 \). But a simple inspection of the possibilities for \( a, b \) shows this can never happen. Therefore, \( N(z) \neq 7 \). Consequently, \( 7 \) is a Gaussian prime, since the only way to factor it inside \( \mathbb{Z}[i] \) is by having one of its factors be a unit.

For more practice, let’s show \( 2 + i \) is a Gaussian prime. Put \( 2 + i = zw \) for \( z, w \in \mathbb{Z}[i] \). Apply the norm to get

\[
5 = 2^2 + 1^2 = N(2 + i) = N(zw) = N(z)N(w).
\]

Since \( N(z) \) and \( N(w) \) are ordinary integers, it follows that \( N(z) = 1 \) or \( 5 \). If \( N(z) = 1 \), then \( z \) is a unit. If \( N(z) = 5 \), then \( N(w) = 1 \), and \( w \) is a unit.

The remainder theorem for Gaussian integers

A crucial property of primes \( p \) in \( \mathbb{Z} \) is that if \( p \) divides \( ab \), then \( p \) divides \( a \) or \( p \) divides \( b \). This was demonstrated back in Proposition 2.4. We have to wonder if Gaussian primes enjoy such a property. They do, but to see it, we first need a remainder theorem for \( \mathbb{Z}[i] \).

**Proposition 7.11.** If \( z, w \) are Gaussian integers and \( z \neq 0 \), then there exist Gaussian integers \( q, r \) such that

\[
w = zq + r \text{ where } N(r) < N(z).
\]

**Proof.** The quotient \( w/z \) is a complex number somewhere in \( \mathbb{C} \) as shown. This \( w/z \) need not be a Gaussian integer.
Let $q$ be a Gaussian integer whose distance to $\frac{w}{z}$ is as small as possible. By inspection, we see that

$$\left| \frac{w}{z} - q \right| \leq \frac{1}{\sqrt{2}}$$

because $q$ has to be in one of the four boxes as shown, and the diagonal of each box has length

$$\sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2} = \frac{1}{\sqrt{2}}.$$

Thus

$$\left| \frac{w}{z} - q \right|^2 \leq \frac{1}{2} < 1.$$
7.2. **SUMS OF SQUARES IN GENERAL**

From this we get

\[
\left| \frac{w - zq}{z} \right|^2 < 1,
\]

and so

\[|w - zq|^2 < |z|^2.\]

In other words

\[N(w - zq) < N(z).\]

Put \(r = w - zq\), and obtain \(w = zq + r\) where \(N(r) < N(z).\)

To illustrate Proposition 7.11, let us find a suitable quotient and remainder when \(z = 3 - 2i\) is divided into \(w = 4 + 5i\). According to the recipe in the proof of Proposition 7.11, a suitable quotient comes from taking the Gaussian integer that is closest to the complex number \(w/z\). We calculate that

\[
\frac{w}{z} = \frac{4 + 5i}{3 - 2i} = \frac{4 + 5i}{3 - 2i} \cdot \frac{3 + 2i}{3 + 2i} = \frac{2 + 23i}{13} = \frac{2}{13} + \frac{23}{13}i.
\]

By sketching the position of \(\frac{2}{13} + \frac{23}{13}i\) in the complex plane, we see that the closest Gaussian integer to it is \(2i\). Thus we take \(q = 2i\), and for the remainder we take

\[r = w - zq = 4 + 5i - (3 - 2i)2i = -i.
\]

Thus we have

\[4 + 5i = (3 - 2i)2i - i,
\]

and indeed,

\[N(-i) = 1 < 13 = N(3 - 2i).
\]

**The greatest common divisor of two Gaussian integers**

The next thing we need is the notion of greatest common divisors in \(\mathbb{Z}[i]\). If \(z, w \in \mathbb{Z}[i]\), a Gaussian integer of type \(zt + ws\) where \(t, s \in \mathbb{Z}[i]\) is called a **Gaussian linear combination** of \(z\) and \(w\). If at least one of \(z\) or \(w\) is non-zero, it is easy to make many Gaussian linear combinations that are not zero. For example, if \(w\) is non-zero the combination \(w = z \cdot 0 + w \cdot 1\) is non-zero, and so is the combination

\[z\overline{z} + w\overline{w} = |z|^2 + |w|^2 > 0.
\]
CHAPTER 7. SUMS OF SQUARES AND OTHER POWERS

Of all possible Gaussian linear combinations of \( z \) and \( w \) there has to be one which is non-zero and whose positive norm is as small as possible. We refer to such a combination as a non-zero combination of minimal norm. Here is a result that should remind us of Proposition 1.4.

**Proposition 7.12.** Let \( z, w \in \mathbb{Z}[i] \) with at least one of them \( \neq 0 \). If \( u = zt + ws \) is a Gaussian linear combination such that \( u \neq 0 \) and \( N(u) \) is minimal, then \( u \mid z \) and \( u \mid w \) in \( \mathbb{Z}[i] \).

**Proof.** To prove \( u \mid z \), Proposition 7.11 gives

\[
z = uq + r \quad \text{where } q, r \in \mathbb{Z}[i] \text{ and } N(r) < N(u).
\]

Then notice that

\[
r = z - uq = z - (zt + ws)q = z(1 - tq) + w(-sq),
\]

which is Gaussian combination of \( z \) and \( w \) with \( N(r) < N(u) \). By the minimality of \( N(u) \) we deduce that \( N(r) = 0 \), and thus \( r = 0 \). So \( z = uq \). In other words, \( u \mid z \). Similarly, \( u \mid w \). \( \Box \)

Any Gaussian integer \( u = zt + ws \) where \( u \mid z \) and \( u \mid w \) is called a **greatest common divisor** of \( z \) and \( w \). We just saw in Proposition 7.12 that any pair of Gaussian integers, not both zero, does possess a greatest common divisor. Just take a non-zero Gaussian linear combination of the pair of Gaussian integers having minimal norm. A simple inspection shows that if \( x \) is a Gaussian integer and \( x \mid z \) and \( x \mid w \), then \( x \mid u \). So a greatest common divisor \( u \) is indeed a common divisor of \( z \) and \( w \) that is divisible by every other common divisor of \( z \) and \( w \).

A little exercise will show that \( u, x \) are both greatest common divisors of \( z \) and \( w \) if and only if \( x = yu \) where \( y \) is one of the units \( \pm 1, \pm i \).

**The signature property of Gaussian primes**

Here is a major property that Gaussian primes have in common with ordinary integer primes. The proof is a bit tricky. See also Proposition 2.4.

**Proposition 7.13.** If \( p \) is a Gaussian prime and \( p \mid zw \) for some Gaussian integers \( z, w \), then \( p \mid z \) or \( p \mid w \).
Proof. Supposing \( p \nmid z \), let’s deduce that \( p \mid w \). Let \( u \) be a greatest common divisor of \( p \) and \( z \). So

\[
u = pt + zs \text{ for some } t, s \text{ in } \mathbb{Z}[i], \text{ and } u \mid p, u \mid z.
\]

Write \( p = uk \) for some \( k \) in \( \mathbb{Z}[i] \). Since \( p \) is a Gaussian prime, one of \( u \) or \( k \) is a unit in \( \mathbb{Z}[i] \).

If \( k \) is a unit, then \( u = pk^{-1} \) where \( k^{-1} \in \mathbb{Z}[i] \), and we see that \( p \mid u \). Since \( u \mid z \) we get \( p \mid z \) contrary to assumption. Thus \( u \) is a unit with inverse \( u^{-1} \) in \( \mathbb{Z}[i] \).

Now multiply \( u = pt + zs \) by the unit \( w \) to get

\[
w = pwt + wzs
\]

and then

\[
w = pwtu^{-1} + wzsu^{-1}.
\]

Clearly \( p \mid pwtu^{-1} \) by just looking, and we are given that \( p \mid wz \). Thus, \( p \mid w \), and we are done.

Proposition 7.13 deserves to be contrasted to the situation with that strange ring \( \mathbb{Z}[\sqrt{-5}] \) back in Chapter 1.

Fermat’s Christmas Theorem

We now possess enough knowledge about \( \mathbb{Z}[i] \) to prove a very interesting result, known as Fermat’s Christmas Theorem. He sent it in a letter to Mersenne on December 25, 1640. It makes a major breakthrough in understanding which integers are sums of two squares. As with all major discoveries, the proof does not just fall into one’s lap.

Theorem 7.14 (Fermat’s Christmas Theorem). If \( p \) is an odd prime in \( \mathbb{Z} \) and \( p \equiv 1 \mod 4 \), then \( p = a^2 + b^2 \) for some \( a, b \) in \( \mathbb{Z} \).

Proof. The proof offered here is due to Richard Dedekind, a German mathematician, circa 1894. Among other things, it brings in the theory of quadratic residues and illustrates how interconnected mathematics can be.

Since \( p \equiv 1 \mod 4 \), we know from Proposition 6.10 that \(-1\) is a quadratic residue modulo \( p \). So \(-1 \equiv x^2 \mod p \) for some \( x \) in \( \mathbb{Z} \). Thus \( p \mid x^2 + 1 \) in \( \mathbb{Z} \). So \( p \mid (x + i)(x - i) \) in \( \mathbb{Z}[i] \). But \( p \mid x + i \). Indeed, the condition \( x + i = p(c + di) \) for some \( c + di \) in \( \mathbb{Z}[i] \) forces \( 1 = pd \), and we know \( p \nmid 1 \). Likewise \( p \nmid x - i \).
Since $p$ divides a product without dividing either of the factors, Proposition 7.13 tells us that $p$ is not a Gaussian prime. Thus $p = uv$ where $u, v \in \mathbb{Z}[i]$, and they are not units. Then

$$p^2 = N(p) = N(uv) = N(u)N(v).$$

So

$$N(u) = 1, \ p \text{ or } p^2,$$

If $N(u) = 1$, then $u$ is a unit, which is not so. If $N(u) = p^2$, then $N(v) = 1$, which would make $v$ a unit. Hence $N(u) = p$. Writing $u = a + bi$, what this tells us is that $p = a^2 + b^2$, a sum of two squares. □

The sum of squares problem for primes

Now we know which primes are sums of squares.

- If $p = 2$, then obviously $2 = 1^2 + 1^2$.
- If $p \equiv 1 \text{ mod } 4$, then $p$ is a sum of two squares, by Fermat’s Christmas Theorem.
- If $p \equiv 3 \text{ mod } 4$, then $p$ is not a sum of two squares, by Proposition 7.6.

And we are left with the question for arbitrary integers $n$.

Sums of squares for arbitrary integers

It’s easy to see that a product of integers that are themselves sums of two squares is again a sum of two squares. Indeed, if $m, n$ are sums of two squares, then $n = N(z), m = N(w)$ for some Gaussian integers $z, w$. And then $nm = N(z)N(w) = N(zw)$, which shows $mn$ is a sum of squares.

Thus, any $n$ that is the product of $2$’s and of primes congruent to $1$ modulo $4$ is a sum of two squares. For instance,

$$72778279720 = 2^3 \cdot 5^1 \cdot 13^7 \cdot 29^1$$

is a sum of squares.
Could the above be the only way to obtain sums of two squares? Unfortunately, it gets a bit more complicated than that. For instance, $45 = 3^2 \cdot 5$ has the prime 3 in its factorization, and yet $45 = 6^2 + 3^2$.

To see better what is going on, we can present the unique factorization of any integer $n$ as follows:

$$n = 2^d p_1^{e_1} \cdots p_k^{e_k} q_1^{c_1} \cdots q_{\ell}^{c_{\ell}},$$

where the primes $p_j \equiv 1 \mod 4$, the primes $q_j \equiv 3 \mod 4$, and the exponents $d, e_j, c_j \geq 0$. We know that the factor $2^d p_1^{e_1} \cdots p_k^{e_k}$ is a sum of two squares. But if all $c_j$ are even, the factor $q_1^{c_1} \cdots q_{\ell}^{c_{\ell}}$ is a sum of two squares as well. Indeed, if all $c_j = 2b_j$, then

$$q_1^{c_1} \cdots q_{\ell}^{c_{\ell}} = (q_1^{b_1} \cdots q_{\ell}^{b_{\ell}})^2 + 0^2.$$

Since a product of two integers that are themselves sums of two squares is again a sum of two squares, we have discovered the following result.

**Proposition 7.15.** Let

$$n = 2^d p_1^{e_1} \cdots p_k^{e_k} q_1^{c_1} \cdots q_{\ell}^{c_{\ell}},$$

where the primes $p_j \equiv 1 \mod 4$, the primes $q_j \equiv 3 \mod 4$, and the exponents $d, e_j, c_j \geq 0$. If all $c_j$ are even exponents, then $n$ is a sum of two squares.

It turns out that the converse of the above proposition holds, and so we have the definitive condition for an integer to be a sum of two squares. To prove the converse, more work needs to be done.

At this stage one can be in awe of the demands that advanced mathematics places upon our powers of reasoning and memory. It might also help give us a moment to empathize with the difficulties that our own students may be confronting.

**Another consequence of the signature property**

From the signature property of Proposition 7.13, it can be shown that every non-zero Gaussian integer has a unique factorization into Gaussian primes. But rather than go into that longer story, let’s prove the generalization of Proposition 7.13 which we will directly need.

**Proposition 7.16.** If $p$ in $\mathbb{Z}[i]$ is a Gaussian prime and $p^k \mid uv$ for some $u, v$ in $\mathbb{Z}[i]$ and some exponent $k \geq 1$, then there are exponents $j, \ell = 0, 1, \ldots, k$ such that $p^j \mid u, p^\ell \mid v$ and $j + \ell = k$. 
Proof. Let’s use induction on $k$.

For $k = 1$, the result says that if $p | uv$, then either $p^1 | u$, $p^0 | v$ or $p^0 | u$, $p^1 | v$, which is nothing but the restatement of Proposition 7.13.

Suppose the result is good for $k - 1$, and now let $p^k | uv$. Then $p | u$ or $p | v$. Say $p | v$. Write $v = wp$ for some $w$ in $\mathbb{Z}[i]$. Then $p^k | uwp$, whence $p^{k-1} | uw$ (For those who don’t see this, we have $uwp = p^k s$ and so $uw = p^{k-1}s$). By the inductive hypothesis, we have integers $j, m$ between 0 and $n - 1$ such that

$$p^j | u, \ p^m | w \text{ and } j + m = k - 1.$$  

Then $p^{m+1} | wp = u$. So we have $j$ and $m + 1$ between 0 and $n$ such that $p^j | u, p^{m+1} | u$ and $j + m + 1 = k$. With $\ell = m + 1$ we have our desired $j, \ell$. \qed

Primes congruent to 3 modulo 4

By Fermat’s Christmas Theorem every prime $p$ in $\mathbb{Z}$, congruent to 1 modulo 4, can be expressed as a sum of two squares:

$$p = a^2 + b^2.$$  

Then $p = (a + bi)(a - bi)$. Since $p$ is prime, neither $a$ nor $b$ can equal 0. Hence $a + bi$ and $a - bi$ are not units in $\mathbb{Z}[i]$. So we have that such primes $p$ no longer remain Gaussian primes.

On the other hand, primes in $\mathbb{Z}$, congruent to 3 modulo 4, remain Gaussian primes.

Proposition 7.17. If $p \in \mathbb{Z}$ and $p$ is prime and $p \equiv 3 \mod 4$, then $p$ remains a Gaussian prime.

Proof. Suppose $p = uv$ for some $u, v \in \mathbb{Z}[i]$. Then

$$p^2 = N(p) = N(uv) = N(u)N(v).$$  

If $N(u) = 1$ or $N(v) = 1$, we see that $u$ or $v$ is a unit, proving $p$ is a Gaussian prime.

The only other option is that $N(u) = N(v) = p$. But that would mean $p$ is a sum of two squares, contrary to the fact that integers congruent to 3 modulo 4 are not the sum of two squares. \qed
7.2. SUMS OF SQUARES IN GENERAL

The end of the story

**Proposition 7.18.** Let \( n \) be an integer such that \( n = a^2 + b^2 \) and let \( p \) be a prime in \( \mathbb{Z} \) congruent to 3 modulo 4. If \( k \) is the exponent such that 
\[
 n = p^k m \text{ and } p \nmid m, 
\]
then \( k \) has to be even.

**Proof.** We see that \( p^k | a^2 + b^2 \). So 
\[
 p^k | (a + bi)(a - bi) \text{ in } \mathbb{Z}[i].
\]
By Proposition 7.17, \( p \) is a Gaussian prime and by Proposition 7.16 we have integers \( j, \ell \) between 0 and \( k \) such that 
\[
 p^j | a + bi \text{ and } p^\ell | a - bi \text{ and } j + \ell = k.
\]
Say \( j \geq \ell \). Then \( a + bi = p^j(c + di) \) for some \( c, d \) in \( \mathbb{Z} \). So \( a + bi = p^j c + p^j di \). Then \( a = p^j c, b = p^j d \), and then 
\[
 n = a^2 + b^2 = p^{2j} c^2 + p^{2j} d^2 = p^{2j}(c^2 + d^2).
\]
Since \( j \geq \ell \) we get \( 2j = j + \ell \geq j + \ell = k \), and since \( p^k \) is the highest power of \( p \) that divides \( n \) and \( p^{2j} | n \), we must conclude \( k = 2j \), which is an even number. \( \square \)

Here is the way to interpret what we have just proven. If 
\[
 n = 2^d p_1^{e_1} \cdots p_k^{e_k} q_1^{e_1} \cdots q_l^{e_l}
\]
where \( p_j \) are primes congruent to 1 modulo 4, and \( q_j \) are primes congruent to 3 modulo 4, and if \( n = a^2 + b^2 \), then all \( c_j \) are even. This is exactly the converse of Proposition 7.15.

Thus, for example, we see that 
\[
 2 \cdot 5^3 \cdot 13^5 \cdot 101^7 \cdot 31^2 \cdot 11^6 \cdot 7^{10}
\]
is a sum of two squares because \( 5, 13, 101 \equiv 1 \mod 4 \), and while \( 31, 11, 7 \equiv 3 \mod 4 \), their exponents 2, 6, 10 are even. But 
\[
 2 \cdot 5^3 \cdot 13^5 \cdot 101^7 \cdot 31^2 \cdot 11^6 \cdot 7^9
\]
is not a sum of two squares because \( 7 \equiv 3 \mod 4 \) and its exponent 9 is odd.
CHAPTER 7. SUMS OF SQUARES AND OTHER POWERS

What is number theory for?

Having struggled to understand the solution to the sums of squares problem, we might be tempted to pause and ask if it was all worth it. Or, for that matter, if the whole exploration of number theory is worth the sacrifice. Certainly, for some people, the effort brings them immense pleasure. For the majority, probably not so much.

We seem driven by our cerebral cortex to understand how things work. Some pursuits are of a practical nature, such as the functioning of an electric motor. Others are born of pure curiosity, such as the origins of the universe, with no obvious benefit in sight. Then there are these Platonic objects called numbers, which we seem to either love or hate. It’s up to each of us to decide. Yet, once in a while, even they yield something the whole world finds useful, such as cryptography.

We don’t know what we might come to wish we had known.

7.3 Some remarks on Fermat’s Last Theorem

The result that the Diophantine equation

\[ x^n + y^n = z^n \]

has no positive integer solution \( x, y, z \) when \( n \geq 3 \) has been recognized, even in the everyday press, as one of the remarkable mathematical discoveries of the twentieth century. Fermat wrote his theorem in the margins of his notebook, claiming to have a proof but lacking room in which to write it. The general consensus is that he did not have a proof, but we still must marvel at how he could have known such a thing was true.

The proof of Fermat’s Last Theorem is far beyond our scope. But let us say something when \( n = 4 \).

The equation \( x^4 + y^4 = z^4 \)

Before proving that the equation

\[ x^4 + y^4 = z^4 \]

has no solution in positive integers \( x, y, z \), a couple of observations may be worth the trouble.
If \( x, y, z \) are positive integers that solve \( x^4 + y^4 = z^4 \), and if \( p \) is a common factor of \( x, y, z \), then
\[
\left( \frac{x}{p} \right)^4 + \left( \frac{y}{p} \right)^4 = \left( \frac{z}{p} \right)^4.
\]

By taking out all such common factors we come up with a solution \( x, y, z \) of the degree-four Fermat equation in which \( x, y, z \) are coprime. As with Pythagorean triples we should call that a \textit{primitive} solution.

Next, if \( x, y, z \) is a primitive solution to \( x^4 + y^4 = z^4 \), then \( x, y, w = z^2 \) are positive, coprime integers that solve \( x^4 + y^4 = w^2 \). So we might as well prove that \( x^4 + y^4 = w^2 \) has no solution using coprime, positive integers \( x, y, w \).

**Theorem 7.19.** There are no positive integers \( x, y, z \) such that \( x^4 + y^4 = z^4 \).

**Proof.** As discussed already, it suffices to prove that there are no coprime, positive integers \( x, y, w \) such that
\[
x^4 + y^4 = w^2.
\]

Assuming such a triplet \( x, y, w \) exists, there must also be such a triplet in which \( w \) is as small as possible. In other words a \textit{minimal}, positive, coprime solution. We can now use the characterization of Pythagorean triples given in Theorem 7.4 to force a contradiction to pop out.

Obviously, \((x^2, y^2, w)\) is now a Pythagorean triple, and it is primitive. According to Proposition 7.3, \( w \) is odd while \( x^2 \) and \( y^2 \) have different parity. We might as well say that \( x^2 \) (and thereby \( x \)) is odd and \( y^2 \) (and thereby \( y \)) is even. According to Theorem 7.4, there exist integers \( r, s \) such that
\[
\begin{align*}
&\text{\( r, s \) have different parity} \\
&\text{\( r, s \) are coprime} \\
&0 \leq r < s \\
&x^2 = s^2 - r^2 \\
&y^2 = 2rs \\
&w = s^2 + r^2.
\end{align*}
\]

Note that \( r > 0 \), in fact. For otherwise, \( y^2 = 2rs \) would be 0, which it is not.

The equation
\[
x^2 + r^2 = s^2
\]
tells us that \((x, r, s)\) is once more a primitive Pythagorean triple, and \(x\) remains odd. So \(r\) is even and \(s\) is odd, by Proposition 7.3. By Theorem 7.4 one more time, there exist integers \(t, u\) such that

\[
\begin{align*}
&\text{\(t, u\) have different parity} \\
&\text{\(t, u\) are coprime} \\
&0 \leq t < u \\
&x = u^2 - t^2 \\
r = 2tu \\
s = u^2 + t^2.
\end{align*}
\]

Since \(r > 0\) the equation \(r = 2tu\) causes \(t > 0\) as well.

Now there is an abundance of equations to play with.

Substitute \(r = 2tu\) into \(y^2 = 2rs\) to obtain

\[
y^2 = 4tus \quad \text{and thus} \quad \left(\frac{y}{2}\right)^2 = tus.
\]

So the positive integer \(tus\) is a perfect square.

Since \(t, u\) are coprime and since \(s = u^2 + t^2\) it follows that \(t, s\) are coprime. Otherwise, some prime \(p\) would appear in the unique factorization of \(t\) and of \(u^2 + t^2\). Then \(p\) would appear in the factorization of \(u^2\), and thereby in that of \(u\), which contradicts the coprimeness of \(t\) and \(u\). Likewise \(u, s\) are coprime.

Because the pairs \(t, u\) and \(t, s\) and \(u, s\) are each coprime and because \(tus\) is a perfect square, each of \(t, u\) and \(s\) is a perfect square. The quickest way to see this is to observe that an integer is a perfect square if and only if every prime in its unique factorization has even multiplicity. The primes in the factorizations of \(t, u\) and \(s\) never overlap. If, in one of these integers, a prime \(p\) had odd multiplicity, then that odd multiplicity would persist in the factorization of \(tus\).

So we can write

\[
t = a^2, \ u = b^2, \ s = c^2,
\]

for some positive integers \(a, b, c\). And since \(t, u, s\) are pairwise coprime, so are \(a, b, c\). Then the equation

\[
s = u^2 + t^2
\]

becomes

\[
b^4 + a^4 = c^2.
\]
We have come up with another positive solution \((b, a, c)\) to the Diophantine equation \(x^4 + y^4 = w^2\).

But now, since \(c \geq 1\) and \(r \geq 1\), we have the inequalities

\[
c \leq c^2 = \frac{s}{s} < s^2 + r^2 = w.
\]

This contradicts the fact that \((x, y, w)\) was a positive, coprime solution in which \(w\) was minimal.

The only way out is to conclude that \(x^4 + y^4 = w^2\) has no positive solution. Then neither does \(x^4 + y^4 = z^4\). 

The rather subtle argument in Theorem 7.19 is due to (not surprisingly) Fermat. Because it shows us how to construct a lesser solution from any hypothetical, given solution, it is known as Fermat’s method of descent. It appears often in the solution of various Diophantine equations.

**Was that satisfying?**

It’s a peculiarity of mathematicians that they relish in showing something can’t be done, as much as they relish showing how to find solutions. In a way, some may think of Fermat’s Last Theorem as rather useless because it’s a statement of failure to solve. Yet, it should be satisfying because it’s a discovery of truth. Just as it must have been satisfying to learn that the so called ether, thought to pervade space, is actually not there.

Furthermore, we cannot predict where one piece of knowledge will take us. Let’s now go to such a place.

### 7.4 A peek at an elliptic curve

One is tempted to think that a proof of the non-solvability of \(x^4 + y^4 = z^4\) marks the end of a story. Happily, with mathematics, stories rarely end. There is always a new spinoff that emerges from an old solution. Just to illustrate this point, and hoping we still retain some stamina, let us see how Fermat’s result about the sum of fourth powers leads to a piece of geometry. We will need the fact, as seen in the proof of Theorem 7.19, that \(x^4 + y^4 = w^2\) has no positive integer solution.

We want to discuss the curve

\[
y^2 = x^3 + x,
\]
whose graph is as shown.

![Graph of an elliptic curve](image)

Such a curve is an example of an **elliptic curve**. Elliptic curves are now famous for their major applications in cryptography, in the final solution of Fermat’s Last Theorem, in practical ways to factor huge integers having small factors, and, mysteriously enough, in advanced theories of physics. It would seem that elliptic curves have little or nothing to do with classical ellipses. So why the name? It arose from attempts to compute the circumference of an actual ellipse. In that problem, integrals involving the square root of a cubic polynomial tend to pop up. Admittedly it’s not the best of names. Mathematical concepts can have obscure origins and thereby end up with names that are hard to fathom.

A **rational point** on our elliptic curve is a point \((r, s)\) where \(r, s\) are rational numbers and \(s^2 = r^3 + r\). In other words, a point that sits on the curve and has rational coordinates.

For the unit circle \(x^2 + y^2 = 1\), we have seen in Proposition 7.2 a full description of all possible rational points. So, what are the rational points on \(y^2 = x^3 + x\)? For sure, the rational point \((0, 0)\) is on the curve. Surprisingly, that’s all there are.

**Theorem 7.20.** The only rational point on the curve \(y^2 = x^3 + x\) is \((0, 0)\).

**Proof.** As might be expected, we suppose that \((r, s)\) is a rational point on our curve, other than \((0, 0)\), and see if a contradiction can be fished out.

By looking at the graph of \(y^2 = x^3 + x\), we can see that \(r > 0\) and \(s \neq 0\).
Thus $r = \frac{a}{b}$ and $s = \frac{c}{d}$, where

- $a, b, c, d$ are integers
- $a, b$ are coprime
- $c, d$ are coprime
- $a > 0$, $b > 0$, $d > 0$ and $c \neq 0$
- \[
    \left(\frac{c}{d}\right)^2 = \left(\frac{a}{b}\right)^3 + \left(\frac{a}{b}\right).
\]

By multiplying out the denominators, the Diophantine equation

\[b^3c^2 = d^2a^3 + b^2d^2a\]

pops out. And now we can milk this until a contradiction emerges, but we might get dizzy keeping track of the steps.

Factor the right side of the equation to get

\[b^3c^2 = d^2a(a^2 + b^2).\]

Thus $d^2 | b^3c^2$, and since $d$ is coprime with $c$ it follows that $d^2 | b^3$. Next, rearrange the original Diophantine equation and factor to get

\[d^2a^3 = b^2(bc^2 - ad^2).\]

Now $b^2 | d^2a^3$, and since $a$ is coprime with $b$ it follows that $b^2 | d^2$. A little reflection using unique factorization reveals that $b | d$. So,

- $b | d$
- $d^2 | b^3$.

Next write $d = be$ and $b^3 = d^2f$ for some positive integers $e, f$. (They are positive because $b, d$ are positive.) Substitute the first of these equations into the second to obtain

\[b^3 = (be)^2f, \quad \text{and then} \quad b = e^2f, \quad \text{as well as} \quad d = e^3f.\]

Put these versions of $b$ and $d$ into the original Diophantine equation to get

\[(e^2f)^3c^2 = (e^3f)^2a^3 + (e^2f)^2(e^3f)^2a.\]
Look carefully to see a common factor of $e^6f^2$ in all three terms just above, and then cancel it to arrive at

$$f c^2 = a^3 + e^4 f^2 a,$$

which can be rewritten as $a^3 = f(c^2 - e^4fa)$.

Now it can be seen that $f \mid a^3$. Since $a, b$ are coprime and since $f \mid b$, the positive integer $f$ has to equal 1. Thus

$$b = e^2 \text{ and } d = e^3,$$

and then $f e^2 = a^3 + e^4 f^2 a$ becomes

$$e^2 = a^3 + e^4 a = a(a^2 + e^4).$$

The next thing is to notice that $a$ and $a^2 + e^4$ are coprime. Indeed, if some prime $p$ divided both $a$ and $a^2 + e^4$, then that $p$ would also divide $e$. Since $e = b^2$, our $p$ would divide $b$, as well as $a$. This would contradict the fact $a$ and $b$ are coprime.

So, just above, we have the product of two coprime integers forming a perfect square. As seen already in previous arguments based on unique factorization, each of the factors in the product is a perfect square. That is,

$$a = g^2 \text{ for some positive integer } g, \text{ and } a^2 + e^4 = h^2 \text{ for some positive integer } h.$$ 

Substitute the first of these equations into the second to arrive at

$$g^4 + e^4 = h^2.$$ 

We have discovered a positive solution $(g, e, h)$ to the equation $x^4 + y^4 = w^2$. But, as can be seen from the remark at the start of Theorem 7.19, this equation does not have solutions in which all unknowns are positive integers. This contradiction shows that the only rational point on $y^2 = x^3 + x$ is $(0, 0)$. \hfill $\Box$

The congruence $y^2 \equiv x^3 + x \pmod p$

Having solved one problem, a committed mathematician keeps coming back for more. We just established that the only rational point on the curve $y^2 = x^3 + x$ is $(0, 0)$. What can happen if we work modulo some prime $p$? Let’s briefly discuss the number of solutions to the congruence

$$y^2 \equiv x^3 + x \pmod p.$$
7.4. A PEEK AT AN ELLIPTIC CURVE

A solution, of course, is any pair of integers \((a, b)\) such that \(b^2 \equiv a^3 + a \mod p\). Since the replacement principle is in effect and we do not wish to be counting a solution \((c, d)\) twice when \(c \equiv a \mod p\) and \(d \equiv b \mod p\), we really should only look at solutions \((a, b)\) where \(a, b\) run from 0 to \(p - 1\). The number of possible non-redundant solutions is thereby at most \(p^2\), but what is the actual number? This is an intriguing problem, which touches on the forefront of research in number theory.

Roughly we might expect that as \(x\) varies from 0 to \(p - 1\), the reductions modulo \(p\) of \(x^3 + x\) will appear randomly between 0 and \(p - 1\). Since half of all integers from 1 to \(p - 1\) have quadratic residues, we might expect that about half of the \(x\) from 1 to \(p - 1\) will cause \(x^3 + x\) to hit a quadratic residue. So roughly \(\frac{p - 1}{2}\) integers \(x\) between 1 and \(p - 1\) will cause \(y^2 \equiv x^3 + x \mod p\) to have a solution. But for each such \(x\) there will be two solutions, namely \(\pm y\). So roughly we might expect that for \(x\) from 1 to \(p - 1\) there will be about \(2 \cdot \frac{p - 1}{2} = p - 1\) solutions. Throwing in the obvious solution \((0, 0)\), we might expect \(y^2 \equiv x^3 + x\) to have roughly \(p\) solutions. Now, it turns out that if \(p \equiv 3 \mod 4\), then there will be exactly \(p\) solutions to the congruence \(y^2 \equiv x^3 + x \mod p\). Number theory, just like the world, is full of surprises.

**Theorem 7.21.** If \(p\) is a prime and \(p \equiv 3 \mod 4\), then the number of solutions \((x, y)\) to the congruence \(y^2 \equiv x^3 + x \mod p\), where \(x, y\) are between 0 and \(p - 1\), equals exactly \(p\).

**Proof.** Notice that \((0, 0)\) is an obvious solution to our congruence. After that, in counting the solutions \((x, y)\) of \(y^2 \equiv x^3 + x \mod p\), start by counting those \(x\), from 1 to \(p - 1\), that cause \(x^3 + x\) to have a quadratic residue.

First check that, when \(x\) runs from 1 to \(p - 1\), the integer \(x^3 + x \not\equiv 0 \mod p\). Well, since \(p \equiv 3 \mod 4\), Proposition 6.10 says that the integer \(-1\) does not have a quadratic residue modulo \(p\). This means that \(x^2 \not\equiv -1 \mod p\) no matter what \(x\) is. In other words \(x^2 + 1 \not\equiv 0 \mod p\). Remembering that \(p\) is a prime, it follows that if \(x \not\equiv 0 \mod p\), then also \(x^3 + x = x(x^2 + 1) \not\equiv 0 \mod p\).

In counting those \(x\) from 1 to \(p - 1\), which cause \(x^3 + x\) to have a quadratic residue modulo \(p\), it helps to work with the Legendre symbol for \(x^3 + x\). Notice that,

\[
\left( \frac{x^3 + x}{p} \right) = \left( \frac{x(x^2 + 1)}{p} \right) = \left( \frac{x}{p} \right) \left( \frac{x^2 + 1}{p} \right).
\]

This shows that \(x^3 + x\) will have a quadratic residue modulo \(p\) precisely when \(\left( \frac{x}{p} \right)\) and \(\left( \frac{x^2 + 1}{p} \right)\) are both +1 or both −1. In other words, precisely when these
Let us now see that the equation
\[
\left( \frac{x^2 + 1}{p} \right) = \left( \frac{x}{p} \right)
\]
happens for exactly \( \frac{p-1}{2} \) of the \( x \)'s between 1 and \( p - 1 \). Since \( p - x \equiv -x \mod p \), the rules in Proposition 6.15 for working with the Legendre symbols give
\[
\left( \frac{p - x}{x} \right) = \left( \frac{-x}{p} \right) = -\left( \frac{x}{p} \right),
\]
as well as
\[
\left( \frac{(p - x)^2 + 1}{p} \right) = \left( \frac{(-x)^2 + 1}{p} \right) = \left( \frac{x^2 + 1}{p} \right).
\]
This reveals that \( \left( \frac{x^2 + 1}{p} \right) = \left( \frac{x}{p} \right) \) if and only if \( \left( \frac{(p-x)^2+1}{p} \right) \neq \left( \frac{p-x}{p} \right) \). So, every time an \( x \) between 1 and \( p - 1 \) causes \( \left( \frac{x^2+1}{p} \right) = \left( \frac{x}{p} \right) \) to hold or fail, there is a compensating \( p - x \) for which the same equation fails or holds, respectively. This shows that, half the time, this equation holds, and, half the time, it fails.

This proves that there are exactly \( \frac{p-1}{2} \) integers \( x \) from 1 to \( p - 1 \) that cause \( x^3 + x \) to have a quadratic residue modulo \( p \). For every one of these integers \( x \), the congruence \( y^2 \equiv x^3 + x \) has, not one but two, solutions for \( y \). Indeed, for every solution \( y \), there is the companion solution \( p - y \), which is also between 1 and \( p - 1 \). Hence, the number of pairs \((x, y)\) where \( x, y \) run from 1 to \( p - 1 \) and \( y^2 \equiv x^3 + x \mod p \) equals
\[
2 \cdot \frac{p - 1}{2} = p - 1.
\]
Throwing in the trivial solution \((0, 0)\) gives a grand total of \( p \) solutions to our elliptic curve congruence. \( \square \)

For a truly enduring project we could ask to find the rational points, or maybe count the points modulo \( p \), for a general elliptic curve:
\[
y^2 = x^3 + ax^2 + bx + c,
\]
where \( a, b, c \) are integers. But the time has come to pause.
7.5 EXERCISES

1. Which of the following integers are sums of two squares?

1960, 21560, 131^9, 2011, 10!, 2^3 \cdot 3^2 \cdot 7^4 \cdot 13 \cdot 89 \cdot 101^3, 2^3 \cdot 3^2 \cdot 7^4 \cdot 13 \cdot 89 \cdot 107^3?

2. If \( z, w \) are Gaussian integers, and \( u, x \) are both greatest common divisors of \( z \) and \( w \), show that \( x = yu \) where \( y \) is one of the units \( \pm 1, \pm i \).

Conversely, suppose \( u \) is a greatest common divisor of \( z \) and \( w \) and \( x = yu \) for some unit \( y \). Prove that \( x \) is also a greatest common divisor of \( z \) and \( w \).

3. Find all primitive non-negative Pythagorean triples \((a, b, c)\) for which \( a \) is odd and \( c \leq 40 \). Then find all Pythagorean non-negative triples \((a, b, c)\) primitive or not, for which \( c \leq 40 \).

4. (a) Find all primitive Pythagorean triples \((a, b, c)\) such that one of \( a, b, c \) equals 185.

(b) Find all primitive, non-negative Pythagorean triples \((a, b, c)\) for which \( c = a + 2 \)

(c) Show that primitive Pythagorean triples \((a, b, c)\) for which \( c = a + 4 \) do not exist.

5. Show that a prime \( p \) in \( \mathbb{Z} \) remains a Gaussian prime if and only if \( p \equiv 3 \mod 4 \).

Hint. This is already buried in the notes of this chapter.

6. Prove that \( 2 + 7i \) is a Gaussian prime.

7. Show that a Gaussian integer \( z \) is a Gaussian prime if and only if it is in one of the following two categories:

- \( z \) is an prime integer already in \( \mathbb{Z} \), congruent to 3 modulo 4
- its norm \( N(z) \) is an prime in \( \mathbb{Z} \)

8. Suppose that a Gaussian integer \( z \) has the property that whenever \( z \mid uv \) then \( z \mid u \) or \( z \mid v \). Prove that such \( z \) must be a Gaussian prime.

Note. All you really need is that \( \mathbb{Z}[i] \) is a ring.
9. Show that the Diophantine equations $x^8 + y^8 = z^8$ and $x^{12} + y^{12} = z^{12}$ have no positive solutions.

10. If $p$ is a prime and $p \equiv 1 \mod 4$, we know that $p = a^2 + b^2$ for some integers $a, b$. That’s what Fermat’s Christmas Theorem said. When this happens show that the pair $(a, b)$ is unique if we insist that $0 \leq a \leq b$. The latter condition eliminates the possibility of changing the signs of $a$ and $b$ or interchanging $a$ and $b$, which would produce solutions that, in essence, are not new.

11. Show that there is no right angled triangle with all sides of integer length and such that its area is a perfect square. This is not so easy.
Chapter 8
Multiplicative functions

In this chapter we discuss something with a rather combinatorial flavour.

8.1 Examples and properties

Recall that \( \mathbb{N} \) is the set of positive integers.

**Definition 8.1.** A function \( f : \mathbb{N} \to \mathbb{C} \) is called *multiplicative* provided \( f \) is not the zero function and

\[
f(nm) = f(n)f(m) \text{ whenever } n, m \text{ are coprime.}
\]

The target space \( \mathbb{C} \) of multiplicative functions is chosen to be as versatile as possible, but for our main examples of multiplicative functions, their values will be integers. Note that we only require the multiplicative property to work on a product of coprime integers.

The key examples

Here are some of the important multiplicative functions.

1. The *indicator function* for the integer 1:

\[
I(n) = \begin{cases} 
1 & \text{if } n = 1 \\
0 & \text{if } n \neq 1.
\end{cases}
\]
A bit of reflection reveals that

\[ I(mn) = I(m)I(n) \] for all \( m, n \), not just for coprime \( m, n \).

2. The constant function with value 1:

\[ 1(n) = 1 \] for all \( n \).

This one is obviously multiplicative.

3. The identity function:

\[ i(n) = n \] for all \( n \).

Again obviously multiplicative, but the next one is a bit more subtle.

4. The number of divisors function:

\[ \tau(n) = \text{the number of positive divisors of } n. \]

For example, \( n = 60 \) has the divisors 1, 2, 3, 4, 5, 6, 10, 12, 15, 20, 30, 60. Since 60 has 12 divisors, \( \tau(60) = 12 \).

Let’s find a formula for \( \tau(n) \). After that we can verify \( \tau \) is multiplicative. Write

\[ n = p_1^{e_1} \cdot p_2^{e_2} \cdots p_k^{e_k}, \]

where the \( p_j \) are distinct primes and \( e_j \geq 1 \). In accordance with Proposition 2.6, a positive integer \( a \) divides \( n \) if and only if

\[ a = p_1^{d_1} \cdot p_2^{d_2} \cdots p_k^{d_k} \]

where \( 0 \leq d_j \leq e_j \). Each \( d_j \) has \( e_j + 1 \) possibilities. Since \( \tau(n) \) counts the total number of possibilities for getting a divisor \( a \) of \( n \), we see that

\[ \tau(n) = (e_1 + 1)(e_2 + 1) \cdots (e_k + 1). \]

For instance,

\[ \tau(3^4 \cdot 5^3 \cdot 101^2) = 5 \cdot 4 \cdot 3 = 60. \]

We can readily see now that \( \tau \) is multiplicative. Indeed, say \( m, n \) are coprime. So

\[ m = p_1^{e_1} \cdots p_k^{e_k} \text{ and } n = q_1^{d_1} \cdots q_\ell^{d_\ell} \]
where the primes \( p_j, q_j \) are all distinct. The unique factorization of \( mn \) is:

\[
mn = p_1^{e_1} \cdots p_k^{e_k} q_1^{d_1} \cdots q_\ell^{d_\ell}
\]

And from our formula for this function,

\[
\tau(mn) = (e_1 + 1)(e_2 + 1) \cdots (e_k + 1)(d_1 + 1) \cdots (d_\ell + 1) = \tau(m)\tau(n).
\]

5. The sum of the positive divisors function:

\[
\sigma(n) = \sum_{d \mid n} d = \text{the sum of the positive divisors of } n.
\]

For example,

\[
\sigma(60) = 1 + 2 + 3 + 4 + 5 + 6 + 10 + 12 + 15 + 20 + 30 + 60 = 168.
\]

It’s not so obvious that \( \sigma \) is multiplicative, but we shall soon see why it is.

6. The (by now familiar) **Euler \( \phi \)-function**:

\[
\varphi(n) = \text{the number of units of } \mathbb{Z}_n
\]

\[
= \text{the number of integers from 1 to } n \text{ that are coprime with } n.
\]

A tiny technicality. When \( n > 1 \), whether we say that \( \varphi(n) \) counts the integers from 1 to \( n - 1 \) or counts the integers from 1 to \( n \) that are coprime with \( n \), we always get the same number \( \varphi(n) \), because \( n \) is not coprime with \( n \). And for \( n = 1 \), we see that \( \varphi(1) = 1 \) because 1 is coprime with 1.

We checked back in Proposition 4.8 that \( \varphi \) is multiplicative by using the Chinese Remainder Theorem. We saw also back in Proposition 4.9 that if \( n = p_1^{e_1} \cdots p_k^{e_k} \) with distinct primes \( p_j \), then

\[
\varphi(n) = (p_1^{e_1} - p_1^{e_1 - 1}) \cdots (p_k^{e_k} - p_k^{e_k - 1}).
\]

7. The **Legendre symbol** for a fixed odd prime \( p \):

\[
\lambda_p(n) = \left( \frac{n}{p} \right).
\]

We checked in Proposition 6.15 that \( \lambda_p(mn) = \lambda_p(m)\lambda_p(n) \) for all \( m, n \), even when \( m, n \) are not coprime.
8. The \textit{Möbius function}:

\[
\mu(n) = \begin{cases} 
1 & \text{if } n = 1 \\
0 & \text{if a prime repeats in the factorization of } n \\
(-1)^k & \text{if } n \text{ factors into } k \text{ distinct primes.}
\end{cases}
\]

Another way to say it is that \(\mu(n) = 1\) when \(n\) is the product of an even number of distinct primes, and \(\mu(n) = -1\) when \(n\) is the product of an odd number of distinct primes. And if any prime repeats itself inside \(n\), then \(\mu(n) = 0\).

We shall come to appreciate that this strange beast is quite important for the study of all multiplicative functions, and for counting some unusual things.

To show that \(\mu\) is multiplicative, suppose that \(m, n\) are coprime integers. If \(m\) or \(n\) equal 1, say \(m = 1\), then \(mn = n\) and

\[\mu(mn) = \mu(n) = 1\mu(n) = \mu(1)\mu(n) = \mu(m)\mu(n).\]

If one of \(m\) or \(n\), say \(m\), has a repeated prime, then that prime continues to repeat in the product \(mn\), maybe even more than before. So,

\[\mu(mn) = 0 = 0\mu(n) = \mu(m)\mu(n).\]

The last case to handle is where both \(m\) and \(n\) are products of distinct primes. Say \(m\) has \(k\) distinct primes, and \(n\) has \(\ell\) distinct primes, in their respective unique factorizations. Then the number of distinct primes in the factorization of \(mn\) is \(k + \ell\), because \(m\) and \(n\) coprime, which tells us that the primes in \(m\) are different than the primes in \(n\). Therefore in this case

\[\mu(mn) = (-1)^{k+\ell} = (-1)^k(-1)^\ell = \mu(m)\mu(n).\]

9. The \textit{Liouville function}:

\[\lambda(n) = (-1)^{\omega(n)},\]

where \(\omega(n)\) is the total number of primes needed to factor \(n\), counting repeats.

For example, if \(n = 2 \cdot 2 \cdot 3 \cdot 7 \cdot 7 \cdot 13 \cdot 53 \cdot 71\) we get \(\omega(n) = 9\) and \(\lambda(n) = (-1)^9 = -1.\)
Clearly, if \( m = p_1 \cdots p_k \) and \( n = q_1 \cdots q_l \) where \( p_j, q_j \) are primes, then \( mn = p_1 \cdots p_k \cdot q_1 \cdots q_l \) and then
\[
\lambda(mn) = (-1)^{k+l} = (-1)^k(-1)^l = \lambda(m)\lambda(n).
\]

So, \( \lambda \) is completely multiplicative, even when \( m, n \) are not coprime.

There will be more multiplicative functions to come, but first we need a brief look at some of their properties.

**Two easy properties of multiplicative functions**

Every multiplicative function \( f \) is such that \( f(1) = 1 \). Indeed, \( f \) is non-zero, so \( f(n) \neq 0 \) for some \( n \). Since 1 and \( n \) are coprime,
\[
f(n) = f(1 \cdot n) = f(1)f(n).
\]

Cancel \( f(n) \) to get \( 1 = f(1) \).

Here is something fundamental. If \( f \) is a multiplicative function, then \( f \) is completely determined by what it does to prime powers. Indeed, write
\[
n = p_1^{e_1}p_2^{e_2} \cdots p_k^{e_k} \text{ where the } p_j \text{ are distinct primes and } e_j \geq 1.
\]

Then
\[
f(n) = f(p_1^{e_1})f(p_2^{e_2}) \cdots f(p_k^{e_k}) \quad \text{since } p_1^{e_1} \text{ and } p_2^{e_2} \cdots p_k^{e_k} \text{ are coprime}
\]
\[
= f(p_1^{e_1})f(p_2^{e_2})f(p_3^{e_3}) \cdots f(p_k^{e_k}) \quad \text{since } p_2^{e_2} \text{ and } p_3^{e_3} \cdots p_k^{e_k} \text{ are coprime}
\]
\[
\cdots
\]
\[
= f(p_1^{e_1})f(p_2^{e_2}) \cdots f(p_k^{e_k}) \quad \text{by repeated use of the coprimeness of the } p_j^{e_j}.
\]

If we know what \( f \) does to prime powers \( p^e \), then we know what \( f \) must do to every integer \( n \).

**A formula for \( \sigma(n) \)**

Now we can get a formula for \( \sigma(n) \), the sums of the divisors of \( n \), discussed in item 5 just above, assuming \( \sigma \) is multiplicative and assuming that \( n \) is factored...
into primes. If \( n = p^e \) where \( p \) is prime and \( e \geq 1 \), then the divisors of \( n \) are 1, \( p, p^2, \ldots, p^e \). By adding up a geometric series we obtain

\[
\sigma(n) = 1 + p + p^2 + \cdots + p^{e-1} + p^e = \frac{p^{e+1} - 1}{p - 1}
\]

So if

\[ n = p_1^{e_1} \cdots p_k^{e_k} \] with distinct primes \( p_j \) and \( e_j \geq 1 \),

the sum of the divisors of \( n \) comes out to be the rather messy

\[
\sigma(n) = \left( \frac{p_1^{e_1+1} - 1}{p_1 - 1} \right) \cdot \left( \frac{p_2^{e_2+1} - 1}{p_2 - 1} \right) \cdots \left( \frac{p_k^{e_k+1} - 1}{p_k - 1} \right).
\]

But we still need to know that \( \sigma \) is multiplicative. What comes next will help a lot.

**Building multiplicative functions**

Our first more subtle fact in this chapter is a little, but tricky, outcome of unique factorization.

**Proposition 8.2.** If \( m, n \) are coprime integers, then every divisor \( d \) of their product \( mn \) comes from a unique pair of integers \( a, b \) such that

\[ a \mid m, \ b \mid n \text{ and } ab = d. \]

**Proof.** If the unique factorizations of \( m, n \) are given by

\[ m = p_1^{e_1} \cdots p_k^{e_k} \text{ and } n = q_1^{d_1} \cdots q_\ell^{d_\ell}, \]

where the primes \( p_i, q_j \) are all distinct and \( e_i \geq 1, d_j \geq 1 \), the unique factorization of \( mn \) is given by

\[ mn = p_1^{e_1} \cdots p_k^{e_k} \cdot q_1^{d_1} \cdots q_\ell^{d_\ell}. \]

A divisor \( d \) of \( mn \) takes the form

\[ d = p_1^{r_1} \cdots p_k^{r_k} q_1^{s_1} \cdots q_\ell^{s_\ell} \text{ where } 0 \leq r_i \leq e_i \text{ and } 0 \leq s_j \leq d_j. \]

With

\[ a = p_1^{r_1} \cdots p_k^{r_k} \text{ and } b = q_1^{s_1} \cdots q_\ell^{s_\ell}, \]
it becomes obvious that
\[a \mid m, \ b \mid n \ \text{and} \ d = ab.\]

We need to confirm that the above \(a, b\) form the only pair of integers with the above properties. To that end, suppose \(c, e\) is another such pair. Thus
\[c \mid m, \ e \mid n \ \text{and} \ d = ce.\]

Clearly \(ce = ab\). Since \(c \mid m\) while \(b \mid n\) and \(m, n\) are coprime, the primes appearing in \(c\) are different than the primes appearing in \(b\). Therefore \(c, b\) are coprime too. Thus \(c \mid a\), by the now ancient Proposition 1.7. By a symmetrical argument, \(a \mid c\), whence \(a = c\), and then \(b = e\).

In the next important result there is an unusual summation notation. We have encountered it briefly in example 5 defining the function \(\sigma\), which we have yet to prove is multiplicative. If \(f\) is any function defined on the set of positive integers, the notation
\[\sum_{d \mid n} f(d)\]
stands for the sum of all values \(f(d)\) as \(d\) runs over the positive divisors of \(n\). For example, \(\sum_{d \mid n} 1\) stands for the sum of the number 1 taken once for every divisor of \(n\), which, of course, is the total number of divisors of \(n\), also known as \(\tau(n)\). And for another example, \(\sum_{d \mid n} d\) gives us the sum of the divisors of \(n\), which we had called \(\sigma(n)\).

**Proposition 8.3.** If \(f\) is a multiplicative function, then the function \(g\) given by
\[
g(n) = \sum_{d \mid n} f(d),
\]
is multiplicative too.

**Proof.** Let \(m, n\) be coprime integers. We want \(g(mn) = g(m)g(n)\). Well, by a
close look at the index of our summations and using Proposition 8.2, we see that
\[
g(mn) = \sum_{d \mid mn} f(d)
\]
\[
= \sum_{a \mid m, b \mid n} f(ab) \quad \text{by Proposition 8.2}
\]
\[
= \sum_{a \mid m, b \mid n} f(a)f(b) \quad \text{since } a, b \text{ are coprime and } f \text{ is multiplicative}
\]
\[
= \left( \sum_{a \mid m} f(a) \right) \left( \sum_{b \mid n} f(b) \right)
\]
\[
= g(m)g(n).
\]

And we are done. \(\square\)

A few illustrations of building multiplicative functions

For instance, take the indicator function \(I\) that maps 1 to 1 and every other positive integer to 0. We see that \(\sum_{d \mid n} I(d) = 1\) for all \(n\), since the only summand that gives a non-zero value is \(I(1) = 1\). The summation gives the constant function 1 where \(1(n) = 1\) for all \(n\). Proposition 8.3 predicts 1 is multiplicative, but we already knew that.

Now apply Proposition 8.3 to the constant function 1. Here we obtain that \(\sum_{d \mid n} 1(d) = \sum_{d \mid n} 1\) gives a multiplicative function. In fact, this is the multiplicative function \(\tau\) that counts the number of divisors of \(n\).

What if we apply Proposition 8.3 to \(i\), the identity function that maps every \(n\) to \(n\)? Then we get that \(\sum_{d \mid n} i(d) = \sum_{d \mid n} d\) is multiplicative. But as we have seen, this is the function \(\sigma\) where \(\sigma(n)\) is the sum of the divisors of \(n\). As promised earlier on, we have just proven that \(\sigma\) is multiplicative.

An unexpected property of Euler’s \(\varphi\)-function

Proposition 8.3, applied to Euler’s \(\varphi\)-function, gives a nice curiosity. Let \(g(n) = \sum_{d \mid n} \varphi(d)\). Let’s do a couple of experiments to see what the multiplicative function \(g(n)\) might yield. For instance,
\[
g(10) = \varphi(1) + \varphi(2) + \varphi(5) + \varphi(10) = 1 + 1 + 4 + 4 = 10.
\]
8.2. THE $\sigma$-FUNCTION AND PERFECT NUMBERS

and

$$g(60) = \varphi(1) + \varphi(2) + \varphi(3) + \varphi(4) + \varphi(5) + \varphi(6) + \varphi(10) + \varphi(12) + \varphi(15) + \varphi(20) + \varphi(30) + \varphi(60) = 1 + 1 + 2 + 2 + 4 + 2 + 4 + 4 + 8 + 8 + 8 + 16 = 60.$$  

It’s no coincidence that, on two trys, we got $g(n) = n$. Here is a proof of this rather non-obvious fact.

**Proposition 8.4.** For every positive integer $n$:

$$\sum_{d|n} \varphi(n) = n.$$

**Proof.** Let $g(n) = \sum_{d|n} \varphi(d)$. We know by Proposition 8.3 that $g$ is multiplicative.

First consider the case where $n$ is a prime power, say $n = p^e$ where $p$ is prime and $e \geq 1$. Expand the summation for $g$ in longhand and use $\varphi(p^e) = p^e - p^{e-1}$ to see that

$$g(p^e) = \varphi(1) + \varphi(p) + \varphi(p^2) + \varphi(p^3) + \cdots + \varphi(p^n) = 1 + (p - 1) + (p^2 - p) + (p^3 - p^2) + \cdots + (p^e - p^{e-1}) = p^e$$  

by the cancellation effect in the sum.

For general $n$, write $n = p_1^{e_1} \cdot p_2^{e_2} \cdots p_k^{e_k}$, where the $p_j$ are distinct primes and $e_j \geq 1$. The pairwise coprimeness of the $p_j^{e_j}$ and the fact $g$ is multiplicative reveals

$$g(n) = g(p_1^{e_1}) \cdot g(p_2^{e_2}) \cdots g(p_k^{e_k}) = p_1^{e_1} \cdot p_2^{e_2} \cdots p_k^{e_k} = n,$$

and the proof is done. \(\square\)

One could now be tempted to explore the effects of Proposition 8.3 on other known multiplicative functions, in order to see what might pop out. But let’s turn to a special class of numbers in connection with the function $\sigma$.

8.2 The $\sigma$-function and perfect numbers

Recall $\sigma(n) = \sum_{d|n} d$, the sum of the divisors of $n$. Occasionally we bump into numbers $n$ such that $\sigma(n) = 2n$. These are called **perfect numbers**. Another way
to say it is that $n$ is perfect when the sum of its proper divisors equals $n$, i.e.

$$n = \sum_{d|n, d<n} d.$$ 

For example, 6 is perfect since $1 + 2 + 3 + 6 = 2 \cdot 6$. For another example, take 28. The proper divisors of 28 are 1, 2, 4, 7, 14, which, by inspection, add up to 28. It would be desirable to have a way of building all perfect numbers. For the even perfect numbers a lot is known, but the odd ones remain a big mystery. What follows is an outline some of the progress that has been made for those that are even.

**Mersenne primes**

It turns out that even perfect numbers are intimately related to so-called Mersenne primes. Any number of the form $2^q - 1$ is called a *Mersenne number*, and if such a number happens to be a prime, we call it, naturally, a *Mersenne prime*. Marin Mersenne (1588-1648), a disciple of Descartes, made contributions to philosophy, the theory of music and, of course, mathematics.

Here is a table of the first few Mersenne numbers.

<table>
<thead>
<tr>
<th>$2^q - 1$</th>
<th>$2^3 - 1$</th>
<th>$2^4 - 1$</th>
<th>$2^5 - 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>7</td>
<td>15</td>
<td>31</td>
</tr>
<tr>
<td>63</td>
<td>127</td>
<td>255</td>
<td>511</td>
</tr>
<tr>
<td>1023</td>
<td>2047</td>
<td>4095</td>
<td>8191</td>
</tr>
</tbody>
</table>

After a bit of staring at this table, we might notice that when the exponent $q$ is not prime, then the Mersenne number $2^q - 1$ does not seem to be prime either. For instance $63 = 9 \cdot 7$ and $511 = 73 \cdot 7$ and $1023 = 341 \cdot 3$.

Let’s verify that if $q$ is not prime, then neither is $2^q - 1$. If $q$ is not prime, then $q = ab$ where $a \geq 2$ and $b \geq 2$. Then $2^q - 1 = (2^a)^b - 1$. One of the most important factorization formulas in mathematics is the following:

$$X^b - 1 = (X - 1) \left(1 + X + X^2 + X^3 + \cdots + X^{b-1}\right).$$

Indeed, this is the basis for the sum of all geometric series, and from that the entire theory of compound interest, not to mention its many uses in calculus and advanced mathematics. From this formula it follows that

$$2^q - 1 = (2^a - 1) \left(1 + 2^a + 2^{2a} + \cdots + 2^{a(b-1)}\right).$$
Since $a \geq 2$, so is $2^a - 1 \geq 2$ (in fact it’s at least 3). And since $b \geq 2$, the sum in the right bracket has at least two terms, which makes it at least 2 as well. Thus, a proper factorization of $2^q - 1$ is achieved when $q$ is not prime.

One might hope for a converse of the preceding observation. Namely, if $q$ is prime, is $2^q - 1$ always prime as well? In our little table we can see that $2^q - 1$ is prime when $q = 2, 3, 5, 7$, and 13, but we are disappointed to discover that

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

So the question becomes: for which primes $q$ is $2^q - 1$ also prime? We do not know. At this point in the history of mathematics its not even known if the are infinitely many primes $q$ that cause $2^q - 1$ to be a Mersenne prime. One way to become famous is to prove such a claim, or disprove it. Yet some huge Mersenne primes have been discovered, especially with the advent of computers. Until the 1950’s the largest known Mersenne prime was $2^{127} - 1$ which is quite gigantic. But in 2005 it was discovered that the truly unfathomable

$$2^{30402457} - 1$$ 

is a Mersenne prime.

For a quite some time now, there has been a collaborative, internet-based search for ever larger Mersenne primes, called the Great Internet Mersenne Prime Search (GIMPS). Cash prizes are offered for success in finding them. To join the search, just Google in “GIMPS”.

But let’s get back to perfect numbers and their relationship to Mersenne primes.

**The link between perfect numbers and Mersenne primes**

Here is an interesting result.

**Theorem 8.5.** An even integer $n$ is a perfect number if and only if

$$n = 2^{q-1} \left(2^q - 1\right),$$

where $q$ is a prime that makes $2^q - 1$ be a Mersenne prime.

**Proof.** For the easy part, suppose that $n = 2^{q-1} \left(2^q - 1\right)$ where $2^q - 1$ is prime. Let $p = 2^q - 1$, just for brevity. Here is the complete list of divisors of $n$.

$$1, 2, 2^2, 2^3, \ldots, 2^{q-1} \text{ and } p, 2p, 2^2p, \ldots, 2^{q-1}p.$$
Add them up using the formula for the sum of a geometric series to get

\[
\sigma(n) = (1 + 2 + 2^2 + 2^3 + \cdots + 2^{q-1}) + (p + 2p + 2^2p + \cdots + 2^{q-1}p)
\]
\[
= (2^q - 1) + p(2^q - 1)
\]
\[
= (1 + p)(2^q - 1)
\]
\[
= 2^q \cdot (2^q - 1)
\]
\[
= 2 \cdot 2^{q-1} \cdot (2^q - 1)
\]
\[
= 2n,
\]
as desired.

The converse is more of a challenge, requiring virtuosity in substitution within equations. Suppose \(n\) is even and perfect. By unique factorization it is possible to write

\[
n = 2^{q-1}m
\]
for some \(q \geq 2\) and for some odd \(m\). The fact \(n\) is even allows us to say that \(q \geq 2\). This will matter at some point in the proof.

Our job is to show that \(m = 2^q - 1\), and that \(m\) is prime.

We have that \(n\) is perfect, and so

\[
\sigma(n) = 2n = 2^q m.
\]

Because \(2^{q-1}\) and \(m\) are coprime and \(\sigma\) is multiplicative, the first equation yields

\[
\sigma(n) = \sigma(2^{q-1})\sigma(m).
\]

By adding up the divisors of \(2^{q-1}\) we obtain

\[
\sigma(2^{q-1}) = 1 + 2 + 2^2 + \cdots + 2^{q-1} = 2^q - 1.
\]

Therefore,

\[
\sigma(n) = (2^q - 1)\sigma(m).
\]

Equate the two different expressions for \(\sigma(n)\) to get

\[
2^q m = (2^q - 1)\sigma(m).
\]

Since \(2^q\) and \(2^q - 1\) are coprime, \(2^q - 1 \mid m\). So

\[
m = (2^q - 1)d,
\]
for some $d$.

To get $m = 2^q - 1$, we need to show $d = 1$. To that end we should notice that, since $n$ is even, our $q \geq 2$, and thus $2^q - 1 \geq 3$. This tells us that $d$ is a proper divisor of $m$, i.e. $d < m$.

Now put $m = (2^q - 1)d$ into the equation $2^q m = (2^q - 1)\sigma(m)$ to obtain

$$2^q(2^q - 1)d = (2^q - 1)\sigma(m),$$

and thus

$$2^q d = \sigma(m).$$

From $m = (2^q - 1)d$ and $2^q d = \sigma(m)$ we come to

$$m + d = 2^q d = \sigma(m).$$

Hoping that $d = 1$, let’s see what happens if $d > 1$. In that case (and remembering that $d < m$) $m$ gets at least three divisors: $1, d$ and $m$. So $\sigma(m) \geq m + d + 1$, in contradiction to the fact $\sigma(m) = m + d$. Thus, $d = 1$, which tells us that $m = 2^q - 1$.

We also need to know $m$ is prime. This is true because $\sigma(m) = m + d = m + 1$. Since the divisors of $m$ add up to $m + 1$, our $m$ can have only 1 and $m$ as divisors, which makes $m$ a prime. Hence our perfect even number $n$ takes the form $2^{q-1}(2^q - 1)$, where $2^q - 1$ is a Mersenne prime. $\square$

For an illustration of Theorem 8.5, take the Mersenne prime $2^7 - 1 = 127$. We know by our theorem that $2^6 \cdot 127 = 8128$ is a perfect number.

Since only finitely many Mersenne primes are known to exist, the same remark applies to perfect even numbers. The search for Mersenne primes is simultaneously a search for even perfect numbers. And then there remains the issue of odd perfect numbers. As of now none has been found, and no one has proven that there are none to be found. About the best we can say so far is that if any odd perfect numbers do exist, they must be bigger than $10^{50}$.

### 8.3 M"{o}bius inversion

The M"{o}bius function plays a major role in number theory and in combinatorics. While we are not in a position to exploit its usefulness in advanced number theory, it remains worthwhile to explore its seemingly magical powers. Recall the
CHAPTER 8. MULTIPLICATIVE FUNCTIONS

definition of the Möbius function:

\[
\mu(n) = \begin{cases} 
1 & \text{if } n = 1 \\
0 & \text{if a prime repeats itself in the factorization of } n \\
(-1)^k & \text{if } n \text{ factors into } k \text{ distinct primes.}
\end{cases}
\]

We have verified already that this somewhat strange function is multiplicative.

What multiplicative function pops out when we apply Proposition 8.3 to \(\mu\)? We get back the indicator function \(I\) for the integer 1.

**Proposition 8.6.** For every \(n \geq 1\),

\[
\sum_{d \mid n} \mu(d) = I(n) = \begin{cases} 
1 & \text{if } n = 1 \\
0 & \text{if } n > 1.
\end{cases}
\]

**Proof.** For \(n = 1\) we have

\[
\sum_{d \mid 1} \mu(d) = \mu(1) = 1 = I(1).
\]

Next consider \(n \geq 2\). We need to show \(\sum_{d \mid n} \mu(d) = 0\). We offer a couple of proofs for this.

This first proof is more direct, but also a bit more challenging. Let the unique factorization of \(n\) be \(n = p_1^{e_1} \cdots p_k^{e_k}\), where the \(p_j\) are distinct primes. The only divisors \(d\) that contribute a non-zero value to the sum \(\sum_{d \mid n} \mu(d)\) are those \(d\) that divide the product \(p_1 \cdot p_2 \cdots p_k\), without repetition of the primes. The other divisors of \(n\) (if any) have a repeated prime in their factorization, and thereby contribute a 0 to the sum. Now, a divisor \(d\) of \(p_1 \cdot p_2 \cdots p_k\) comes from choosing a subset of the set of indices \(\{1, \ldots, k\}\) and taking \(d\) to be the product of the primes that go with that subset. If \(d\) is the product of an even number of primes, we have \(\mu(d) = 1\). If \(d\) is the product of an odd number of primes, we have \(\mu(d) = -1\). Since every finite, non-empty set has exactly as many subsets with an even number of elements as it has with an odd number of elements, the number of \(+1\)’s in the sum \(\sum_{d \mid n} \mu(d)\) cancels exactly the number of \(-1\)’s in the sum. Hence, the sum \(\sum_{d \mid n} \mu(d) = 0\).

This next proof is a bit easier. Let \(g(n) = \sum_{d \mid n} \mu(n)\). Since \(\mu\) is multiplicative, so is \(g\), by Proposition 8.3. Because multiplicative functions are completely
determined by what they do to prime powers, it suffices to see that \( g(n) = 0 \) when \( n = p^e \), a prime power. Well,

\[
g(p^e) = \sum_{d|p^e} \mu(d)
\]

\[
= \mu(1) + \mu(p) + \mu(p^2) + \cdots + \mu(p^e)
\]

\[
= 1 - 1 + 0 + \cdots + 0 = 0,
\]

and we are done.

\[
\square
\]

The Möbius inversion formula

The Möbius function is important because it gives us a way to solve the equation

\[
g(n) = \sum_{d|n} f(d)
\]

for \( f \) in terms of \( g \). The solution is called Möbius inversion, and here’s what it says.

**Theorem 8.7.** If \( f, g \) are any functions defined on the set of positive integers and

\[
g(n) = \sum_{d|n} f(d),
\]

then

\[
f(n) = \sum_{d|n} g(d) \mu \left( \frac{n}{d} \right) = \sum_{d|n} g \left( \frac{n}{d} \right) \mu(d).
\]

**Proof.** In this theorem, \( f, g \) do not need to be multiplicative.

It should be clear that the two summations give the same thing, since, as \( d \) runs over the divisors of \( n \), the quotients \( \frac{n}{d} \) also run over the very same divisors of \( n \).

For a positive integer \( n \) and a pair of positive integers \( d, e \) we first convince ourselves that

\[
d e | n \text{ if and only if } d | n \text{ and } e | \frac{n}{d}.
\]
With that in mind, keep a very close look at the indices of summation in each line to obtain:

\[ \sum_{d \mid n} g \left( \frac{n}{d} \right) \mu(d) \]

\[ \begin{align*}
&= \sum_{d \mid n} \left( \sum_{e \mid \frac{n}{d}} f(e) \right) \mu(d) \quad \text{by definition of } g \\
&= \sum_{d \mid n, e \mid \frac{n}{d}} f(e) \mu(d) \quad \text{by multiplying in } \mu(d) \\
&= \sum_{e \mid n, d \mid \frac{n}{e}} f(e) \mu(d) \quad \text{by the conviction of our observation} \\
&= \sum_{e \mid n, d \mid \frac{n}{e}} f(e) \mu(d) \quad \text{by the symmetry between } d \text{ and } e \\
&= \sum_{e \mid n} \left( \sum_{d \mid \frac{n}{e}} \mu(d) \right) f(e) \quad \text{by factoring out } f(e) \\
&= \sum_{e \mid n} I \left( \frac{n}{e} \right) f(d) \quad \text{by Proposition 8.6} \\
&= f(n) \quad \text{by the definition of the indicator function } I.
\]

When all is said and done, Möbius inversion is just a way altering the order of a complicated summation.

Illustrations of Möbius inversion

We have seen in Proposition 8.4 that

\[ n = \sum_{d \mid n} \varphi(d) \]

By Möbius inversion,

\[ \varphi(n) = \sum_{d \mid n} d \mu \left( \frac{n}{d} \right) , \]
which is an unusual formula for $\varphi(n)$ in terms of the Möbius function. For instance,

$$\varphi(28) = \sum_{d\mid 28} d\mu\left(\frac{28}{d}\right)$$

$$= 1\mu(28) + 2\mu(14) + 4\mu(7) + 7\mu(4) + 14\mu(2) + 28\mu(1)$$

$$= 1 \cdot 0 + 2 \cdot 1 + 4 \cdot (-1) + 7 \cdot 0 + 14 \cdot (-1) + 28 \cdot 1$$

$$= 12$$

This corroborates with our calculation of $\varphi(28)$ using the multiplicativity of $\varphi$. Indeed,

$$\varphi(28) = \varphi(4) \cdot \varphi(7) = 2 \cdot 6 = 12.$$

For another illustration of Möbius inversion, take the function $\sigma$ defined by $\sigma(n) = \sum_{d\mid n} d$. Invert to get another funny formula:

$$n = \sum_{d\mid n} \sigma(d)\mu\left(\frac{n}{d}\right).$$

For instance, with $n = 28$ we obtain

$$\sum_{d\mid 28} \sigma(d)\mu\left(\frac{n}{d}\right)$$

$$= \sigma(1)\mu(28) + \sigma(2)\mu(14) + \sigma(4)\mu(7) + \sigma(7)\mu(4) + \sigma(14)\mu(2) + \sigma(28)\mu(1)$$

$$= 1 \cdot 0 + 2 \cdot 1 + 4 \cdot (-1) + 7 \cdot 0 + 14 \cdot (-1) + 28 \cdot 1$$

$$= 28$$

Next look at $\tau(n) = \sum_{d\mid n} 1$, which is the number of divisors of $n$. Möbius gives

$$1 = \sum_{d\mid n} \tau(d)\mu\left(\frac{n}{d}\right).$$

For instance, with $n = 18$ we get

$$\tau(1)\mu(18) + \tau(2)\mu(9) + \tau(3)\mu(6) + \tau(6)\mu(3) + \tau(9)\mu(2) + \tau(18)\mu(1)$$

$$= 1 \cdot 0 + 2 \cdot 0 + 2 \cdot 1 + 4 \cdot (-1) + 3 \cdot (-1) + 6 \cdot 1$$

$$= 1.$$
An application of Möbius inversion to a combinatorial problem

A binary string of length \( n \) is a list

\[ w = a_1 a_2 \cdots a_n \]

where \( a_j = 0 \) or 1. For example, 1001101 is a string of length 7.

To form a binary string of length \( n \) we have to fill in \( n \) slots with one of two possible entries. So, the total number of binary strings of length \( n \) is \( 2^n \).

We say a string \( w \) is the \( k \)'th power of a string \( z \) of length \( d \) when \( n = kd \) and

\[ w = z z z \cdots z, \text{ where } z \text{ is repeated } k \text{ times.} \]

In that case we may write \( w = z^k \). For example, in the string \( w = 001001001001 \) we see that \( z = 001 \) is repeated four times to get \( w \). That is \( w = z^4 \). It’s also true that \( w = (001001)^2 \). On the other hand, while \( w = 001101 \) is the first power of itself, it is not a higher power of any shorter string.

We say a string \( w \) is irreducible when the only way to get \( w = z^k \) is by having \( z = w \) and \( k = 1 \). For instance, \( w = 010010010010 = (010)^4 \) is reducible. But \( w = 11011100 \) is irreducible.

Here comes a key observation. Every string \( w \) of length \( n \) determines a unique string \( z \) of length \( d \) such that

- \( d \mid n \)
- \( z \) is irreducible
- \( w = z^\frac{n}{d} \)

We may call \( z \) the irreducible root of \( w \). For example, the irreducible root of 1111111 is \( z = 1 \). The irreducible root of 101101101101 is \( z = 101 \). Note that the length of the irreducible root of a string \( w \) always divides the length of \( w \).

Here’s the combinatorial problem.

If \( f(n) \) is the number of irreducible strings of length \( n \), calculate \( f(n) \).

The answer comes from Möbius inversion, as we now explain.

For any finite set \( A \), let the notation \( \# A \) stand for the number of elements in \( A \), also known as the cardinality of \( A \). If \( X \) is the set of all strings of length \( n \), we observed already that \( \# X = 2^n \).

For every divisor \( d \) of \( n \) let \( X_d \) be the set of strings of length \( n \) which have an irreducible root of length \( d \). Here are some key observations.
• Every string of length $n$ had a unique irreducible root $z$ of some length $d$ that divides $n$. Thus every string of length $n$ belongs to one and only one $X_d$. In other words, the $X_d$ do not overlap and they exhaust all of $X$.

• For every divisor $d$ of $n$ and every irreducible string $z$ of length $d$, the power $z^n$ lies in $X_d$.

• Different strings $z$ of length $d$ give different powers $z^n$ in $X_d$.

In light of the above observations we see that

$$X = \bigcup_{d|n} X_d,$$

and this union is disjoint (no overlaps in the sets $X_d$). Also, the last two of our bullets reveal that

$$\#X_d = f(d) = \text{the number of irreducible strings of length } d.$$

Putting these facts together gives

$$2^n = \#X = \sum_{d|n} \#X_d = \sum_{d|n} f(d).$$

Now Möbius inversion is at hand to yield

$$f(n) = \sum_{d|n} 2^d \mu \left( \frac{n}{d} \right),$$

which is quite a remarkable formula.

For instance, the number of irreducible strings of length 12 equals

$$f(12) = 2^1 \mu(12) + 2^2 \mu(6) + 2^3 \mu(4) + 2^4 \mu(3) + 2^6 \mu(2) + 2^{12} \mu(1)$$

$$= 2 \cdot 0 + 4 \cdot 6 + 8 \cdot 0 + 16 \cdot (-1) + 64 \cdot (-1) + 4096 \cdot 1$$

$$= 4020.$$

It would have been quite a chore to look at every one of the 4096 possible strings of length 12 and discard those 76 that are reducible.
And the number of irreducible strings of length 100 is

\[ f(100) = 2^1 \mu(100) + 2^2 \mu(50) + 2^4 \mu(25) + 2^5 \mu(20) + 2^{10} \mu(10) + 2^{20} \mu(5) + 2^{25} \mu(4) + 2^{50} \mu(2) + 2^{100} \mu(1) \]

\[ = 2^1 \cdot 0 + 2^2 \cdot 0 + 2^4 \cdot 0 + 2^5 \cdot 0 + 2^{10} \cdot 1 + 2^{20} \cdot (-1) + 2^{25} \cdot 0 + 2^{50} \cdot (-1) + 2^{100} \cdot 1 \]

\[ = 2^{10} - 2^{20} - 2^{50} + 2^{100} \]

\[ = 1\,267\,650\,600\,228\,228\,275\,596\,795\,315\,200, \]

which is quite a lot.

**Where might we go from here?**

After seeing some number theory we become more aware of that much remains to be learned. There is the huge matter involving the distribution of primes, in particular the prime number theorem which tells us that the number of primes up to a give number \(x\) is asymptotic to \(\frac{\ln x}{x}\). There is the exciting matter of **continued fractions**, which not only solve some significant Diophantine equations but also approximate real numbers in ways that are far superior to decimal expansions. Then there is the matter of telling if a real number is algebraic or transcendental. If are given a number \(x\), for example \(\pi\) or \(e\) or \(\ln 2\), is there polynomial with integer coefficients that \(x\) is root of? Just for starters, is there even one real number that is not the root of some polynomial with integer coefficients? The entire subject of elliptic curves, and algebraic curves in general remains untouched, not to mention its widely used application to cryptography.

Finally there are all the problems that have yet to be solved. Lately, the twin primes problem seems to be surrendering it secrets. It asks if there are infinitely many prime pairs, such as \((3, 5), (5, 7), (11, 13), (17, 19), (29, 31), (51, 53)\), where the second prime is two more than the first. Maybe someone will tell us one day soon.

But, as it is with all things, there also comes a time at which to stop.

### 8.4 Exercises

1. Evaluate \(\mu(100), \mu(231), \mu(462), \sigma(100), \sigma(231), \sigma(462)\).

2. Let \(X\) be a finite set with \(n\) elements. There are \(2^n\) subsets of \(X\). Show
that the number of subsets of X with an even number of elements equals
the number of subsets with an odd number of elements.

3. Find all n such that \( \sigma(n) = 28 \).

4. Prove that for every n there exist n successive integers \( x, x+1, x+2, \ldots, x+(n-1) \) such that \( \mu(x+j) = 0 \) for all \( j = 0, \ldots, n-1 \). This tells us that
there are arbitrarily long runs of successive integers each of which has a
repeated prime factor.

Hint. Pick any n distinct primes \( p_0, p_1, \ldots, p_{n-1} \). Then explain why the
congruences

\[
\begin{align*}
x & \equiv 0 \pmod{p_0^2} \\
x + 1 & \equiv 0 \pmod{p_1^2} \\
x + 2 & \equiv 0 \pmod{p_2^2} \\
& \vdots \\
x + n - 1 & \equiv 0 \pmod{p_{n-1}^2}
\end{align*}
\]

have a solution \( x \).

5. If \( b \) is a positive integer and \( p \) is an odd prime and \( p \nmid b \), prove that
\[
\sum_{k=1}^{p-1} \left( \frac{kb}{p} \right) = 0.
\]

6. Let \( f(n) = \sum_{d|n} \frac{1}{d} \).

(a) Explain why \( f \) is multiplicative.

(b) An integer \( n \) is perfect when \( \sigma(n) = 2n \). Show that \( n \) is perfect if and
only if \( f(n) = 2 \).

7. Find a multiplicative function \( f \) such that
\[
\mu(n) = \sum_{d|n} f(d),
\]
where \( \mu \) is the Móbius function.
8. A pair of numbers \(m, n\) is called amicable when
\[\sigma(m) = \sigma(n) = m + n.\]

(a) Verify that the pairs 220, 284 is amicable. Verify that 17296 = \(2^4 \cdot 23 \cdot 47\) and 18416 = \(2^4 \cdot 1151\) are amicable.

(b) Show that a prime \(p\) is not part of an amicable pair.

(c) Show that the square \(p^2\) of a prime \(p\) is not part of an amicable pair.

9. Let \(a, b, c\) represent three symbols. A string on three letters, of length \(n\), is any expression
\[w = x_1 x_2 x_3 \cdots x_n,\]
where each \(x_j\) is one of \(a, b, c\). For instance, \(babbcaacb\) is a string on three letters, of length 9. These are also called ternary strings.

(a) How many ternary strings on three letters of length \(n\) are there?

A ternary string of length \(n\) is the \(k\)'th power of a string of length \(d\), provided
\[n = kd\text{ and }w = zzz\cdots z\]
where \(z\) is a string of length \(d\) and \(z\) is repeated \(k\) times. In this case write \(w = z^k\). The string \(w\) is called irreducible, when the only way to get \(w = z^k\) is by having \(z = w\) and \(k = 1\).

(b) By imitating our reasoning on binary strings, determine the number of irreducible ternary strings of lengths 5, 10 and 20.

10. Find all multiplicative functions \(f\) such that \(f\) is strictly increasing and \(f(2) = 2\). For example, the identity function is such a function. Are there any others?
Index

absolute value, 175
associative law, 59

binary string, 214
box notation, 49

cancellation, 51
cardinality, 91, 214
Carmichael number, 121
Cartesian product, 88
Chinese Remainder Theorem, 84
Chinese remainder theorem, 85, 89
commutative law, 59
complex conjugate, 175
complex number, 3
composite integer, 30
congruence class, 49
congruent, 43
coprime, 20
degree of a polynomial, 113
Diophantine equation, 21
discriminant, 135
dividend, 6
divisibility, 3
divisor, 6
divisor of a Gaussian integer, 174
联络 curve, 190, 194
Euclidean Algorithm, 10
Euler ϕ-function, 62, 199

Euler’s test for quadratic residues, 140
Euler’s theorem, 62, 63
Extended Euclidean Algorithm, 16

factor, 3
factor of a Gaussian integer, 174
Fermat number, 80
Fermat prime, 80
Fermat’s Christmas Theorem, 181
Fermat’s Last Theorem, 21
Fermat’s Little Theorem, 64
field, 3, 58, 113
floor, 157

Gauss’ Lemma, 153
Gaussian integer, 40, 173
Gaussian linear combination, 179
Gaussian prime, 40, 176
Gaussian unit, 40
golden ratio, 28
greatest common divisor, 9, 180
Greatest common divisor theorem, 9
group, 60

index of an integer, 124, 139
indicator function, 197
integer, 1
integer combination, 7
integer lattice point, 157
integer part of a real number, 157
inverse, 58
Korselt’s criterion, 121
least common multiple, 39
Legendre symbol, 137, 199
linear Diophantine equation, 21
Liouville function, 200
Möbius function, 200, 210
Möbius inversion, 211
Mersenne number, 206
Mersenne prime, 41, 206
modular arithmetic, 51
multiplicative function, 92, 197
multiplicity, 33
natural number, 5
non-negative Pythagorean triple, 169
norm, 36
norm of a Gaussian integer, 175
order of a residue, 108
order of an integer modulo \( n \), 108
perfect number, 41, 205
polynomial over \( \mathbb{Z}_p \), 113
prime, 29, 37
primitive, 187
primitive Pythagorean triple, 169
primitive root in \( \mathbb{Z}_n^* \), 107
primitive root modulo an integer, 107, 118
Primitive root theorem, 117
pseudo-primes, 74, 131
Pythagorean triple, 165
quadratic reciprocity, 146
quadratic residue, 81, 136
quotient, 6
rational number, 3
real number, 3
reduced form of a rational number, 25
reduction modulo \( n \), 47
reduction modulo an integer, 66
relatively prime, 20
remainder, 6
Remainder theorem, 5
Replacement principle, 46, 47
representative, 49
residue, 49
residue class, 49
ring, 2, 174
root of a polynomial, 114
RSA-scheme, 93
signature property of primes, 31
square and multiply, 66
square and multiply algorithm, 66
symmetric reduction, 151
symmetric system of remainders, 151
Unique Factorization Theorem, 32
unit, 37
unit Gaussian integer, 175
unit in \( \mathbb{Z}_n \), 57
unit modulo \( n \), 57
well-defined, 52
Wilson’s theorem, 82
witness for non-primality, 72, 119