Integer and Modular Arithmetic

Jim Carlson

February 2, 2004

Contents

1 Introduction 3

2 Divisibility 5
  2.1 Greatest common divisor ............................. 6
  2.2 Euler’s \( \phi \) function ............................ 7

3 Well-ordering principle 8
  3.1 Existence of prime factorization .................... 8
  3.2 Division algorithm ................................ 9
  3.3 GCD ............................................ 9
  3.4 Uniqueness of prime factorization .................. 10

4 Euclid’s Algorithm 11
  4.1 Implementation .................................. 13
  4.2 Proof that the algorithm works .................... 13
  4.3 The equation \( ax + by = c \) ....................... 14
  4.4 Extended Euclidean algorithm ...................... 16

5 Modular arithmetic 18
  5.1 Congruence .................................... 18
  5.2 Residues ...................................... 21
  5.3 Solving congruences ............................. 22
  5.4 The algebra of residues .......................... 23
  5.5 Sums of squares ................................ 24
  5.6 Rings and fields ................................ 25
1 Introduction

Our main object of study will be the integers: the numbers used for counting, together with their negative companions, and the special number zero:

\[ \ldots -3, -2, -1, 0, 1, 2, 3, \ldots \]

A source of fascination ever since mankind thought about numbers, the integers are rich in problems that are simple to state but sometimes surprisingly difficult to solve. Here are a few such problems:

1. What are the integer solutions to the equation \( x^2 + y^2 = z^2 \)?
2. When can an integer be written as the sum of two squares?
3. What are the integer solutions to the equation \( x^d + y^d = z^d \), where \( d \) is an integer greater than 2?
4. How many primes are there less than a given number \( n \)?
5. How can we tell whether a given number is prime?
6. How can we factor a given number into primes?

The first problem comes from geometry, and was consider by the Greeks and (a millennium earlier) by the Babylonians. One solution corresponds to a right triangle whose sides have lengths \((3, 4, 5)\). Other solutions correspond to other triangles whose side lengths are whole numbers. One is led to ask: how many solutions are there? A few? Many? Infinitely many?

The second problem is clearly related to the first. We will give part of the solution in XX when we introduce the idea of modular arithmetic, and another part in XX when we introduce the ring of Gaussian integers \( a + b\sqrt{-1} \). The third problem, which generalizes the first, is a famous one: Fermat’s last theorem, posed in the mid 1600’s by Pierre de Fermat and solved in 1994 by Andrew Wiles. Fermat wrote in the margin of Diophantus’ *Arithmetica* (ca. 250 AD) that he had found a marvelous argument to show that the only solutions are the obvious ones, with at least one of \( x, y, z \) equal to zero. In honor of Diophantus, integer equations for which we seek integer solutions are called Diophantine equations.

The second group of three problems have to do with prime numbers. These are integers like

\[ 2, 3, 5, 7, 11, 13, \ldots \]

that cannot be factored into smaller numbers. Thus \( 6 = 2 \times 3 \) is not prime. We call numbers like this composite: they are made by putting other numbers together using multiplication. The terminology is Latin: *con* = “with”, *ponere* = “to put.” We don’t count trivial factorizations like \( 2 = 2 \times 1 \).
The primes have also been part of the long fascination with the integers. Already in Euclid’s *Elements*, written in 300 BC, one finds a proof that there are infinitely many primes. But a good answer to (4) did not come until 18XX, when Hadamard (France) and de la Vallée Poussin (Belgium) proved that the number of primes less than \( x \) is approximately \( x/\log x \) — a result conjectured XXX years earlier by Carl Friederich Gauss. Questions like (5) and (6) have gained new importance since 1975, when Rivest, Shamir, and Adleman of MIT invented a way to keep information secret using prime numbers. Their method, known as the RSA algorithm, is used millions of times a day to make credit card purchases on the internet without letting card numbers fall into the hands of criminals. Number theory has become part of applied mathematics in the sense that it is mathematics applied to solve non-mathematical problems. In this case, the problems involve large sums of money.

To answer questions like the ones posed above, we need some mathematical tools — some ideas that go beyond the usual arithmetic and algebra of high school. The most basic tools have to do with the notion of divisibility. We say “\( a \) divides \( b \)” where \( a \) and \( b \) are integers, when \( b/a \) is an integer. Thus 2 divides 6, but it does not divide 7. A key problem is to find the greatest common divisor \( g \) of two positive integers \( a \) and \( b \). This number, often called the GCD, is the largest integer which divides both \( a \) and \( b \). Thus 5 is the GCD of 35 and 55, while the GCD of 12345 and 54321 is 3. There is a very efficient algorithm for finding the GCD that goes back to Euclid. Found in Book IX of the *Elements*, it is nowadays known as the Euclidean algorithm. Closely related to the problem of the GCD is the problem of finding solutions of the Diophantine equation

\[
ax + by = c.
\]

In fact, a slight refinement of the Euclidean algorithm provides solutions whenever they exist. This “extended Euclidean algorithm” is one of the working parts of the RSA system of cryptography mentioned above.

The next set of tools have to do with the notion of congruence, introduced by the German mathematician Carl Friederich Gauss in 1801 to study divisibility problems in a systematic way. We say that integers \( a \) and \( b \) are congruent modulo \( n \) if \( n \) divides \( a - b \). Thus 18 is congruent to 4 modulo 7 because 7 divides 18 - 4, but 18 is not congruent to 5 modulo 7. We write these facts as \( 18 \equiv 4 \mod 7 \) and \( 18 \not\equiv 5 \mod 7 \). In XX we will explain an important congruence due to Fermat:

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

where \( p \) is prime and \( a \not\equiv 0 \mod p \). This congruence turns out to be one of the key pieces of the RSA algorithm — and also the basis of efficient ways to manufacture the needed “industrial primes.”

Closely related to the notion of congruence is the notion of the integers modulo \( n \). These are “new” number systems with “new” operations of addition and multiplication. As an example, to be explained in further XX, the integers
modulo 5 consists of the five numbers

\[ 0, 1, 2, 3, 4 \]

where addition and multiplication are defined by the tables below.

\[
\begin{array}{c|cccc}
+ & 0 & 1 & 2 & 3 \\
\hline
0 & 0 & 1 & 2 & 3 \\
1 & 1 & 2 & 3 & 0 \\
2 & 2 & 3 & 4 & 0 \\
3 & 3 & 4 & 0 & 1 \\
4 & 4 & 0 & 1 & 2 \\
\end{array}
\quad
\begin{array}{c|cccc}
\times & 0 & 1 & 2 & 3 \\
\hline
0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 2 & 3 \\
2 & 0 & 2 & 4 & 1 \\
3 & 0 & 3 & 1 & 4 \\
4 & 0 & 4 & 3 & 2 \\
\end{array}
\]

The operations of modular arithmetic are easily implemented on a computer. For example, in the C or Python computer languages, the expressions

\[(a + b) \mod 5 \]
\[(a \times b) \mod 5 \]

compute the sum and product modulo 5. An important practical use of modular arithmetic is the generation of sequences of integers that “appear” to be random even though they are produced by a formula. These are called pseudorandom sequences, and are used in applications as different as computer games, internet security, and the computation of heat flow in a material body.

A system of numbers like the integers that admits addition, subtraction, multiplication, but not all divisions by nonzero numbers is called a \textit{ring}. If Division by nonzero numbers is permitted, it is called \textit{field}. We wouldn’t bother with this terminology, except that to study the ring of integers, we are naturally led to use ideas that have to do with other rings and fields.

\textbf{Remark}. Mathematicians have given conventional symbols to the most important rings and fields. For the integers we use \( \mathbb{Z} \) and for the integers modulo \( n \) we use \( \mathbb{Z}/n \). For the field of rational, real, and complex numbers we use \( \mathbb{Q}, \mathbb{R}, \mathbb{C} \).

\section{2 Divisibility}

As noted already in the introduction, we say that \( a \) divides \( b \) if the quotient \( a/b \) is an integer. For example 3 divides 6, since \( 6/3 = 2 \), but 4 does not divide 6, since \( 6/4 = 3/2 \) is not an integer. We write these facts as \( 3 \mid 6 \) and \( 4 \notmid 6 \), respectively.

To say that \( b \) divides \( a \) is to say that there is an integer \( q \) such that \( a = bq \). Thus \( 3 \mid 6 \) because \( 6 = 3 \times 2 \), but \( 4 \notmid 6 \) because the only solution of \( 6 = 4q \) is the number \( 3/2 \).
Divisibility is a powerful notion. We can sometimes use it to determine the integer solutions of an equation with integer coefficients. Consider, for example, the equation $x^2 - y^2 = 1$. Does it have integer solutions? Well, we can factor it as $(x + y)(x - y) = 1$. This equation says that $x + y$ divides 1. Since $x + y$ is an integer, it must be $+1$ or $-1$. The same reasoning applies to $x - y$: it is $\pm 1$. Now it is easy to determine $x$ and $y$ by solving linear equations. Notice that you have answered the question, “what are the points on the hyperbola $x^2 - y^2 = 1$ which have integer coordinates.” Points with integer coordinates are sometimes called lattice points.

**Problem 1** Find all lattice points on the hyperbola $x^2 - y^2 = 1$ and $x^2 - y^2 = 2$. Do the same for $x^2 - 2y^2 = 1$.

**Problem 2** Suppose that 7 divides $a$ and $b$. Show that it divides $a + b$. Generalize this result.

**Problem 3** Let $a$, $a + 7$, $a + 14$, ... be an arithmetic progression. Let $d > 1$ be an integer. Show that $d$ divides all the numbers in the progression or none of them.

### 2.1 Greatest common divisor

Given a positive integer $n$, let $D(n)$ be the set of its positive divisors. For example,

$$D(30) = \{ 1; 2, 3, 5; 6, 10, 15; 30 \}$$

and

$$D(43) = \{ 1; 2, 3, 7; 6, 14, 21; 43 \}.$$  

There is, of course, a pattern to the commas and semicolons. The set of common divisors of two positive integers $a$ and $b$ is the set $D(a, b)$ of numbers that divide both $a$ and $b$. In our example $a = 30$, $b = 43$, this is the set

$$D(30, 43) = D(30) \cap D(43) = \{ 1; 2, 3, 6 \}.$$  

The greatest common divisor of $a$ and $b$ is the largest integer that divides both numbers. In our example, the greatest common divisor is 6. We write this as

$$6 = GCD(30, 43).$$

One way of computing the set of divisors of a number is to use its factorization into primes. Another is to list all the numbers between 1 and $a$, then examine each number $j$ in between to see if it divides $a$. If it does not, we discard it. The list of numbers that remains is the list of divisors of $a$. To find the GCD of $a$ and $b$, we examine the list of divisors of $a$ one by one, discarding those that do not divide $b$. From the remaining list we select the largest number. It
is the GCD. Neither of these methods is very efficient: think about how much work (how many divisions) is required. Then estimate the time needed for each division, and use the result to estimate the running time. In section ?? we will discuss a much better (faster, computationally efficient) method for computing the GCD.

Two numbers are said to be relatively prime if they have no common factors besides the obvious one. Thus $a$ and $b$ are relatively prime if $\text{GCD}(a, b) = 1$. The numbers 14 and 33 are relatively prime but the numbers 33 and 43 are not relatively prime.

**Problem 4** Consider the pairs of numbers $m, n$: (a) 12, 21; (b) 123, 321; (c) etc., up to 123456789, 987654321 Find the GCD of $m$ and $n$ in each case.

**Problem 5** (a) Pick two numbers at random and determine whether they are relatively prime. Do this several times. (b) Investigate the question: what is the probability that two numbers chosen at random are relatively prime.

### 2.2 Euler’s $\phi$ function

Let $\text{RP}(n)$ be the set of integers $j$ relatively prime to $n$ which satisfy $1 \leq j \leq n$. For example,

$$\text{RP}(7) = \{ 1, 2, 3, 4, 5, 6 \}.$$  

That was easy since 7 is a prime number. Can you describe $\text{RP}(p)$ for any prime? Here is another example:

$$\text{RP}(15) = \{ 1, 2, 4, 7, 8, 11, 13, 14 \}.$$  

We can compute the set $\text{RP}(n)$ as follows: List the numbers 1, 2, ..., $n$. For the case $n = 15$ we have

$$1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8 \quad 9 \quad 10 \quad 11 \quad 12 \quad 13 \quad 14 \quad 15$$

Find the first number in the list bigger than 1 that divides the last number. Strike out all its multiples. In our example, we strike out all the multiples of 3:

$$1 \quad 2 \quad * \quad 4 \quad 5 \quad * \quad 7 \quad * \quad 8 \quad * \quad 10 \quad 11 \quad * \quad 13 \quad 14 \quad *$$

Find the next number that divides $n$. Strike out all its multiples. In our example we strike out all the multiples of 5 to obtain the list

$$1 \quad 2 \quad * \quad 4 \quad * \quad 7 \quad * \quad * \quad 11 \quad * \quad 13 \quad 14 \quad *$$

Continue in this way until no numbers remain that divide $n$. The numbers that remain are relatively prime to $n$. 

7
Let us now define the quantity
\[ \phi(n) = \#RP(n), \]
where \# denotes the number of elements in a set. Thus we have \( \phi(7) = 6, \phi(15) = 8. \)

The function \( \phi \) is called the Euler phi function. Euler, for the great 18th century Swiss mathematician, phi for the Greek letter \( \phi. \)

**Problem 6** Find \( \phi(11), \phi(9), \) and \( \phi(55). \) Any guesses about general patterns? More data may help.

In number theory it is often a good idea to work out what happens for primes, then compare that to the general case. Let’s do this for a very special case. One has the prime factorization \( 15 = 3 \cdot 5, \) and one computes \( \phi(3) = 2, \phi(5) = 4. \) In this case, \( \phi(15) = \phi(3)\phi(5). \)

**Problem 7** Investigate the question of how generally the identity \( \phi(ab) = \phi(a)\phi(b) \) holds. For now we have only one “data point.”

## 3 Well-ordering principle

An important fact about the integers is the well-ordering principle:

> Every nonempty set of positive integers has a least element.

For example, the least element of the set of positive integers is 1, and the least element of the set of positive integers divisible by 5 is 5 itself. But the set of positive rationals does not have a least element. Indeed, if there were such an element \( L, \) then \( L/2 \) would be even smaller.

We will give several applications of the well-ordering principle. The first two are to the existence of prime factorizations and to fact that the GCD of positive integers \( a \) and \( b \) can be written as \( ax + by \) for some integers \( x \) and \( y. \) The latter fact, which plays a key role in number theory is sometimes called Bezout’s identity.

### 3.1 Existence of prime factorization

Let us first make some important definitions. An integer \( a \) is a unit if there is an integer \( b \) such that \( ab = 1. \) (Which integers are units?) A nonzero integer \( n \) is composite if it can be written as \( n = ab \) where neither \( a \) nor \( b \) are units. A nonzero integer is prime if it is not composite.

**Theorem 1** Every integer \( n > 1 \) can be factored into primes
Proof: Assume the contrary: there is a nonempty set of integers bigger than 1 that cannot be factored into primes. Let $L$ be the least positive integer of this set. Thus $L$ is composite, and so can be written as $L = ab$, where $a$ and $b$ are positive integers bigger than 1. Therefore $a$ and $b$ are less than $L$. Therefore they can be factored into primes. The product of the prime factors of $a$ and the prime factors of $b$ is a prime factorization of $L$. But this contradicts the fact that $L$ is a number that cannot be factored into primes. Q.E.D.

3.2 Division algorithm

A fundamental fact is that given positive integers $a$ and $b$, there is are integers $q \geq 0$ and $0 \leq r < a$ such that

$$b = aq + r \tag{1}$$

This is called the division algorithm. Why? Because long division of $b$ by $a$ provides us with the numbers $q$ and $r$. But there is a pleasing theoretical explanation from the well-ordering principle, which applies to sets of nonnegative numbers as well as to positive ones. Thus, let $S$ be the set of non-negative numbers of the form $b - aq$. It is nonempty because it contains $b - a \cdot 0 = b$. Let $r$ be its least element. By construction $r \geq 0$. Suppose $r \geq a$. Then $r' = r - a$ can also be written as $b - aq'$ for some $q'$. But $r' < r$, a contradiction to the fact that $r$ is the least element of $S$. Q.E.D.

3.3 GCD

Let $S$ be the set of positive integers that can be written as $ax + by$. This set is not empty: it contains $a$ and also $b$. Let $L$ be the least element of $S$. It can be written as $L = ax + by$. To see if $L$ divides $a$, apply the division algorithm: $a = qL + r$, where $0 \leq r < L$. Solve for the remainder, substitute our expression for $L$, and rearrange:

$$r = a - qL = a - q(ax + by) = (1 - qx)a + (-qy)b.$$  

If $r$ is nonzero, it is in the set $S$ and is less than $L$, a contradiction. Therefore $L$ divides $a$. A similar argument shows that $L$ divides $b$.

We have just proved that $L$ is a common divisor of $a$ and $b$. But is $L$ the greatest common divisor? To answer this question, let $C$ be another common divisor of $a$ and $b$. Thus $a = Cq$ and $b = Cq'$. Substitute into $L = ax + by$ to obtain

$$L = ax + by = Cqx + Cq'y = C(qx + q'y).$$

Thus $C$ divides $L$. Consequently $L$ is the greatest of all common divisors of $a$ and $b$. Thus we can assert that

$$\text{GCD}(a, b) = ax + by \tag{2}$$
for some integers $a$ and $b$. This is sometimes called *Bezout’s identity*. It is natural to ask whether there is an algorithm for computing the GCD and also the integers $x$ and $y$. We answer this in the next section. For now, however, let us draw an important theoretical consequence. 

**Corollary 1** Suppose that $p$ is a prime dividing $ab$. Then it divides either $a$ or $b$.

**Proof.** Suppose that $p$ does not divide $a$. Then the GCD of $a$ and $p$ is 1. By Bezout’s lemma, there exist integers $x$ and $y$ such that $px + ay = 1$. Multiply this equation by $b$ to obtain $pxb + aby = b$. The prime $p$ divides both terms on the left-hand side and so divides their sum. Consequently $p$ divides $b$. Q.E.D.

### 3.4 Uniqueness of prime factorization

From the corollary to Bezout’s lemma we can prove a fundamental fact about the integers:

**Theorem 2** A positive integer has only one factorization into primes.

For the idea of the proof, consider the smallest integer $n > 1$ that does not factor uniquely as a product of primes. We may assume those primes are greater than 1, and we may write

$$n = p_1p_2 \cdots p_k = q_1q_2 \cdots q_\ell.$$  

The prime $p_1$ divides the right-hand side, and so, by the corollary, must divide one of the factors. By relabeling the factors, we may assume that $p_1$ divides $q_1$. Since $q_1$ is a positive prime divisible by $p_1$, it equals $p_1$. Thus

$$p_1p_2 \cdots p_k = p_1q_2 \cdots q_\ell.$$  

Cancelling the common factor of $p_1$, we have

$$p_2 \cdots p_k = q_2 \cdots q_\ell.$$  

But this equation represents factorization of $n/p_1$, a number smaller than the smallest number which has two different factorizations. We conclude that $k = \ell$ and that up to changing orders of factors $p_i = q_i$. Thus $n$ has only one factorization, a contradiction. What was contradicted? The assumption that there is a number $n > 1$ which has two factorizations.

It may seem that we are proving the obvious. *Of course* a number can be factored uniquely into primes! But what kind of number. Consider numbers of the form $a + b\sqrt{-5}$ The sum of any two sum numbers is again of the same form, as is the difference and product: they form a *ring*. Thus they are a ring
of numbers rather like the integers. Let’s give this ring a name: \( \mathbb{Z}[\sqrt{-5}] \) — the integers with the square root of \(-5\) adjoined.

Now consider the number 6. It can be written in \( R \) as

\[ 6 = 2 \cdot 3. \]

It can also be written as

\[ 6 = (1 + \sqrt{-5})(1 - \sqrt{-5}). \]

Thus we have two prime factorizations!

Well, let’s reconsider, since our ring \( \mathbb{Z}[\sqrt{-5}] \) is so exotic, maybe we need to rethink the question of whether 2 and 6 are still prime, and whether \((1 \pm \sqrt{-5})\) is prime. This we consider in the problems.

**Problem 8** Define the conjugate of \( z = a + b\sqrt{-5} \) to be \( \bar{z} = a - b\sqrt{-5} \). Define the norm of \( z = a + b\sqrt{-5} \) to be \( N(z) = z\bar{z} \). Compute \( N(2) \), \( N(3) \), and \( N(1 \pm \sqrt{-5}) \).

**Problem 9** List all elements \( z \) of \( \mathbb{Z}[\sqrt{-5}] \) with \( N(z) \leq 5 \). What can you say about the elements of norm 1?

**Problem 10** Show that if \( N(ab) = N(a)N(b) \).

**Problem 11** Show that if \( z \) divides \( w \), then \( N(z) \) divides \( N(w) \). Show that if \( z \) divides \( w \), then \( N(z) \leq N(w) \).

**Problem 12** Define the notion of a prime in \( \mathbb{Z}[\sqrt{-5}] \).

**Problem 13** Show that 2, 3, and \( 1 \pm \sqrt{-5} \) are prime in \( \mathbb{Z}[\sqrt{-5}] \).

**Historical remark.** Lamé, a French mathematician, thought that he had proved Fermat’s last theorem by using special factorizations in rings like \( \mathbb{Z}[\sqrt{-5}] \). The subtle flaw — soon recognized by Kummer?? was that unique factorization into primes does not always hold.

### 4 Euclid’s Algorithm

None of the methods for computing Euclid’s algorithm mentioned so far are practical for computing the GCD of numbers with hundreds of digits — routine work in cryptography Euclid’s algorithm gives such a method. We begin our study with an example: using the table below, we compute the GCD of 2310 and 1547. The answer is given by the value of \( b \) in the last row of the table. We obtain row \( (n) \) from row \( (n-1) \) as follows: (i) let the \( a \) of row \( n-1 \) be the \( a \) of row \( n \). (ii) let the \( b \) of row \( n \) be the \( b \) of row \( n-1 \). (iii) Let the \( r \) of row \( n \) be the remainder as described above. We also need to know when to stop computing. Do this when \( r = 0 \).
<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>2310</td>
<td>1547</td>
<td>763</td>
</tr>
<tr>
<td>(2)</td>
<td>1547</td>
<td>763</td>
<td>21</td>
</tr>
<tr>
<td>(3)</td>
<td>763</td>
<td>21</td>
<td>7</td>
</tr>
<tr>
<td>(4)</td>
<td>21</td>
<td>7</td>
<td>0</td>
</tr>
</tbody>
</table>

Compare the result with the prime factorizations: $2310 = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11$ and $1547 = 7 \cdot 13 \cdot 17$.

Notice how much faster the new algorithm is. In our example it required just four divisions. In the Sunday Morning XX algorithm it would require 1545 divisions.

How was such an algorithm discovered? The answer comes from geometry. The Greeks originally based their theory of proportion on the existence of a common measure. Given line segments $AB$ and $CD$, it was supposed that there was a (perhaps very tiny) line segment $UV$ such that $AB$ is $m$ copies of $UV$ laid out end to end and that $CD$ in $n$ copies of $UV$ laid out end to end. Thus $UV$ is a kind of geometric common divisor of $AB$ and $CD$. The fact that $UV$ can be laid out end to end to fill out the bigger intervals is the statement that it divides each without remainder. See figure XX below.

Figure XX

How can one find the common measure? Well, suppose that $CD$ is the smaller of the two intervals. Lay it out end to end beginning in $AB$ as many times as possible as in figure XX below and let $EF$ be the remainder, as in the figure below. Then lay out $EF$ in $CD$ and let $GH$ be the remainder. Either the process goes on forever or it stops at some point. Assume that it stops, and let $QR$ be the last interval laid out. It is the common measure. If you think about what we have done and translate it into the language of arithmetic, you have the Euclidean algorithm as exemplified in the table above.

Figure XX

What happens if the process never stops? This can’t happen in the arithmetic version of the algorithm with positive integers because the remainders produced get smaller and smaller at each step but can’t be smaller than one without actually being zero. This is why the algorithm — whatever it computes, must terminate after a finite number of steps. But with line segments — or with real numbers instead of integers — the process need not stop. The one says that the two segments or numbers are incommensurable. This is how irrational numbers were discovered, allegedly by Pythagoras. See XX.

**Problem 14** Find the GCD of 12345 and 54321.
4.1 Implementation

To implement the Euclidean algorithm on a computer, we first state it more formally using the notation \( a \% b \) for the remainder of \( a \) upon division by \( b \).

To find the gcd of \( a \) and \( b \):

1. Set \( r = a \% b \)
2. While \( r > 0 \), do the following:
   replace \((a, b, r)\) by \((b, r, b \% r)\)

The gcd is \( b \)

The formal statement can be easily translated into a computer program. Here is the translation into Python:

```python
def gcd(a, b):
    r = a % b
    print a, b, r
    while r > 0:
        a, b, r = b, r, b % r
        print a, b, r
    return b
```

When run as `gcd(2310, 1547)`, it gives exactly the table above. For regular work, remove the `print` statements.

**Problem 15** Compute the GCD of 1234567 and 1683. Then compute the gcd of 7654321 and 1683.

4.2 Proof that the algorithm works

To prove that the Euclidean algorithm works, we need a bit of notation. Let us write a table like XX as follows:

<table>
<thead>
<tr>
<th>( a_0 )</th>
<th>( b_0 )</th>
<th>( r_0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a_1 )</td>
<td>( b_1 )</td>
<td>( r_1 )</td>
</tr>
<tr>
<td>( a_2 )</td>
<td>( b_2 )</td>
<td>( r_2 )</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>( a_{n-2} )</td>
<td>( b_{n-2} )</td>
<td>( r_{n-2} )</td>
</tr>
<tr>
<td>( a_{n-1} )</td>
<td>( b_{n-1} )</td>
<td>( r_{n-1} )</td>
</tr>
<tr>
<td>( a_n )</td>
<td>( d )</td>
<td>( 0 )</td>
</tr>
</tbody>
</table>

Table 1: default
Here \( a_0 = a \) and \( b_0 = b \). Also, \( r_n = 0 \), and we have written \( b_n \) as \( d \), the supposed greatest common divisor. We first show that \( d \) divides \( a \) and \( b \). Start with the fact that
\[
a_{n-1} = q_{n-1}b_{n-1} + r_{n-1}
\]
for some integer \( q_{n-1} \). Since \( d \) divides \( a_n \), which is equal to \( b_{n-1} \), \( d \) divides \( b_{n-1} \). Moreover, \( d \) equals \( r_{n-1} \). Thus \( d \) divides both terms on the right of the above equation. Consequently \( d \) divides the left-hand side of the equation. Thus it divides both \( a_{n-1} \) and \( b_{n-1} \). Repeating the argument with the next row up, we find that \( d \) divides \( a_{n-2} \) and \( b_{n-2} \). Repeating as often as necessary, we see that it divides \( a = a_0 \) and \( b = b_0 \).

We have just seen that \( d \) is a common divisor of \( a \) and \( b \). Is it the greatest common divisor? Suppose \( d' \) divides both \( a \) and \( b \), that is, both \( a_0 \) and \( b_0 \). Then it divides \( r_0 = a_0 - q_0b_0 \). That is, \( d' \) divides all the numbers on line 0. Looking at the next line, we see that \( a_1 = b_0 \) is divisible by \( d' \). Since \( b_1 = r_0, b_1 \) is divisible by \( d' \) as well. And since \( r_1 = a_1 - q_1b_1 \), the remainder is divisible by \( d' \). Consequently the numbers on line 1 are divisible by \( d' \). Continuing in this way, we see that all numbers in the table are divisible by \( d' \). In particular, \( d \) is divisible by \( d' \). And therefore \( d \geq d' \) as required: \( d \) is the greatest of the common divisors.

**Note.** We have proved something more: common divisors of \( a \) and \( b \) divide the greatest common divisor.

### 4.3 The equation \( ax + by = c \)

Our goal in this section is to understand the Diophantine equation
\[
ax + by = c. \tag{3}
\]

Note first that any integer that divides \( a \) and \( b \) must divide the left-hand side of the preceding equation and therefore must also divide the right-hand side. In particular, the GCD of \( a \) and \( b \) divides \( c \). We have just proved the following:

**Proposition 1** If \( ax + by = c \) is solvable, then \( \text{GCD}(a, b) \mid c \).

More difficult is the converse:

**Theorem 3** If \( \text{GCD}(a, b) \mid c \), then \( ax + by = c \) is solvable.

Suppose that the hypothesis of the previous theorem is satisfied. Let \( g = \text{GCD}(a, b) \), and write \( c = gh \). By the theorem, \( ax + by = g \) is solvable. A solution \((x_0, y_0)\) of this equation yields a solution \((hx_0, hy_0)\) of \( ax + by = c \). Thus to prove the theorem in general it is enough to prove the special case in which \( c = g \).
The equation

\[ ax + by = g. \tag{4} \]

can be solved using an extended version of the Euclidean algorithm. We illustrate this with case

\[ 2310x + 1547y = 7. \tag{5} \]

Begin by writing down the usual table that comes from repeated application of the division algorithm, but where we record the quotient \( q \) as well as the remainder \( r \):

<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>r</th>
<th>q</th>
</tr>
</thead>
<tbody>
<tr>
<td>2310</td>
<td>1547</td>
<td>763</td>
<td>1</td>
</tr>
<tr>
<td>1547</td>
<td>763</td>
<td>21</td>
<td>2</td>
</tr>
<tr>
<td>763</td>
<td>21</td>
<td>7</td>
<td>36</td>
</tr>
<tr>
<td>21</td>
<td>7</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>

From this table we extract the following information:

1. \( 763 = 2310 - 1 \cdot 1547 \)
2. \( 21 = 1547 - 2 \cdot 763 \)
3. \( 7 = 763 - 3 \cdot 21 \)

Look at line 3, which we can write as

\[ 7 = 1 \cdot 763 - 36 \cdot 21. \tag{6} \]

The expression on the right is a linear combination of the numbers 763 and 21: an expression of form \( 763x + 21y \). Because this linear combination is equal to 7, we know that \((1, -36)\) is solution of the equation

\[ 763x + 21y = 7. \]

Although this is not the equation we want to solve, it does have the same form: an encouraging sign.

We will now use the rest of the table to solve the original equation. Substitute line 2 of the table in the equation (6) to obtain

\[ 7 = 1 \cdot 763 - 36(1547 - 2 \cdot 763). \]

Rewrite the right-hand side as a linear combination of 1547 and 763:

\[ 7 = (-36) \cdot 1547 + 73 \cdot 763. \tag{7} \]

Finally, substitute line 1 of the table into the preceding equation to obtain

\[ 7 = (-36) \cdot 1547 + 73 \cdot (2310 - 1547). \]
Then put the right-hand side in the form of a linear combination of 2310 and 1547:

\[ 7 = 73 \cdot 2310 + (-109) \cdot 1547. \]  

We have reached our goal: \((x, y) = (73, -109)\) is a solution of equation (5).

The process just illustrated computes both the GCD \(g\) and the unknowns \((x, y)\) in (5). It can be summarized as follows:

To solve \(ax + by = g\):

1. Compute \(g\) by constructing the table \(a, b, r, q\).
2. Use line \(n-1\) of the table to write the remainder \(g\) as a linear combination of \(a\) and \(b\).
3. Move upwards line by line writing \(g\) as a linear combination of \(a\) and \(b\).

The coefficients \((x, y)\) found for line 1 solve the equation.

**Problem 16** Solve the equation \(54321x + 12345y = \text{GCD}\).

### 4.4 Extended Euclidean algorithm

As a first step in automating the process described above for solving \(ax + by = g\), we modify the original \(\gcd\) program so as to produce tables like the one used to solve \(2310x + 1547y = 7\). Here is such a program written in Python:

```python
def gcd(a,b):
    q, r = divmod(a,b)
    print a, b, r, q
    while r > 0:
        qq, rr = divmod(b,r)
        a, b, r, q = b, r, rr, qq
        print a, b, r, q
    return b
```

The expression \(\text{divmod}(a,b)\) produces the pair \((q,r)\) consisting of the quotient and remainder for division of \(a\) by \(b\).

The new GCD program helps, but we should also implement the back-substitution. To do so we would have to store the full table in memory. In fact, this is not necessary. It is possible to work out a forward substitution process where only eight numbers need to be kept in memory at each step, no matter how many lines in the table.
Let us begin with line 1 of table XX. It tells us that

\[ 763 = 2310 + (-1)1547. \]  

(9)

Line 2 yields

\[ 21 = 1547 + (-2)763. \]  

(10)

Substituting (9) into the preceding equation, we have

\[ 21 = 1547 - 2(2310 - 1547) \]

and then

\[ 21 = (-2) \cdot 2310 + 3 \cdot 1547. \]  

(11)

Line 3 yields

\[ 7 = 763 + (-36)21 \]  

(12)

Substituting (9) and (11) into the preceding, we have

\[ 7 = [2310 - 1547] + (-36)[(-2) \cdot 2310 + 3 \cdot 1547] \]

which simplifies to

\[ 7 = 73 \cdot 2310 + (-109)1547 \]  

(13)

We can summarize the results of our computations in the following table.

<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>r</th>
<th>q</th>
<th>x</th>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td>2310</td>
<td>1547</td>
<td>763</td>
<td>1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>1547</td>
<td>763</td>
<td>21</td>
<td>2</td>
<td>-2</td>
<td>3</td>
</tr>
<tr>
<td>763</td>
<td>21</td>
<td>7</td>
<td>36</td>
<td>73</td>
<td>-109</td>
</tr>
</tbody>
</table>

Let us now think about how the last \((x, y)\) was computed. We obtained 73 as \((-36)(-2) + 1\) and \(-109\) as \((-36)(3) - 1\). If \(q\) represents the quotient on line 3, then we have \(73 = 1 - qx\) and \(-109 = -1 - qy\), where \(x\) and \(y\) come from line 2.

A closer look at the table suggests that perhaps the coefficient 1 in \(1 - qx\) and \(-1\) in \(-1 - qy\) are really the \(x\) and \(y\) of line 2. If we give these quantities the names \(u\) and \(v\), then we have the following suggested formulas for computing the new \(x\) and \(y\) from the preceding two \(x\)'s and \(y\)'s.

\[ \text{new } x = u - qx, \quad \text{new } y = v - qy. \]  

(14)

Pure thought or more examples will confirm that the proposed formulas hold in general.

Recall that our goal is to compute \(x\) and \(y\) using only a small number of remembered quantities. The table below shows how to do this. We have added two columns for the quantities \(u\) and \(v\) which are the \(x\) and \(y\) of the previous step. We have also included the final “stop” line.
The question is now: can we compute one line of the table using only data in the previous line? It seems that we can, except for the missing values \( u \) and \( v \) in the first line. But these can be determined in the case of our example from the equations (14) which compute the new \( x \) and \( y \). Writing them down for line 2, we have

\[
-2 = u - 2(1), \quad 3 = v - 2(-1).
\]

These give the values \( u = 0, \) \( v = 1 \). Again, pure thought or more examples confirm that the initial values chosen for \( u \) and \( v \) are correct. It is now a simple matter to translate our ideas into a computer program. Below is one possibility:

```python
def egcd(a,b):
    q,r = divmod(a,b)
    x,y,u,v = 1,-q,0,1
    print a,b,r,q,x,y,u,v
    while r > 0:
        qq,rr = divmod(b,r)
        xx,yy = u - qq*x, v - qq*y
        a,b,r,q,x,y,u,v = b,r,rr,qq,xx,yy,x,y
        print a,b,r,q,x,y,u,v
    return u,v,b
```

5 Modular arithmetic

We will now systematically develop the two related notions of congruence and modular arithmetic. These will give us tools for understanding purely number theoretic problems such as when a number is the sum of two squares as well as important applied problems such as generation of sequences of “random” numbers and cryptography.

5.1 Congruence

Fix a positive integer \( m \) which will shall call the modulus. We say that two integers \( a \) and \( b \) are congruent modulo \( m \) if \( a - b \) is divisible by \( m \). For example, 29 and 5 are congruent modulo 12 since \( 29 - 5 = 2 \cdot 12 \). But 29 and 6 are not congruent. We write these relations as \( 29 \equiv 5 \mod 12 \) and \( 29 \not\equiv 6 \mod 12 \).
Problem 17  (a) List ten consecutive numbers which are congruent to 1 modulo 7. What can you say about this sequence of numbers. (b) Do the same for numbers congruent to zero modulo seven. (c) Consider positive numbers $a$ and $b$. Suppose that $a$ divided by seven has a remainder of 3. Suppose that the same is true of $b$. What can you say about the relation of $a$ to $b$?

Congruence obeys some useful algebraic properties which are rather like those of equality. For example, if $a \equiv b \mod m$ and $c \equiv d \mod m$, then

$$a + c \equiv b + d \mod m \quad \text{and} \quad ac \equiv bd \mod m.$$  

Let us prove the second of these facts. To say that $a \equiv b \mod n$ is to say that $a - b$ is divisible by $m$. Equivalently, $a - b = qm$ for some $q$, which is equivalent to

$$a = b + qm \text{ for some } q.$$  

Likewise, the statement $c \equiv d \mod m$ is equivalent to

$$c = d + q'm \text{ for some } q'.$$  

Multiply the two preceding equations to obtain

$$ac = bd + bq'm + dqm + qq'm^2.$$  

Rearrange to get

$$ac - bd = (bq' + dq + qq'm)m.$$  

Thus $ac - bd$ is divisible by $m$, and so $ac$ is congruent to $bd$ modulo $m$.

Note that by repeated applications of the second congruence identity to $a \equiv b \mod m$ and $a \equiv b \mod m$, we find that

$$a^n \equiv b^n \mod m$$  

for all positive integers $n$.

Problem 18  Prove the following. (i) If $a \equiv b \mod m$ and $c \equiv d \mod m$, then $a + c \equiv b + d \mod m$. (ii) For all $a$, $a \equiv a \mod m$. (reflexivity). (iii) For all $a$, $b$, and $c$, $a \equiv b \mod m$, then $b \equiv a \mod m$. (symmetry). (iv) For all $a$, $b$, and $c$, if $a \equiv b \mod m$ and $b \equiv c \mod m$, then $a \equiv c \mod m$.

As an application of the congruence identities above, we consider the problem:

What is the last digit of $a = 1234567^{100}$

The number $a$ is quite large, and certainly too large to compute exactly on an ordinary calculator. Brute force computation does not seem like a winning strategy. So we try a different approach: we think about the problem. The key
fact is that for any number $a$, the last digit is the remainder upon division by 10. In other words,

$$a \equiv \text{last digit mod } 10.$$

Now the problem becomes easy. First, $1234567 \equiv 7 \mod 10$. Second, $1234567^{100} \equiv 7^{100} \mod 10$.

We seem to be back to the same difficulty: computing something very large, though not so large as before. But we keep thinking. Working modulo 10, we note that $7^2 = 49 \equiv 9$. Therefore $7^3 \equiv 7 \cdot 9 = 63 \equiv 3$. Therefore $7^4 \equiv 7 \cdot 3 = 21 \equiv 1$. This is an important conclusion:

$$7^4 \equiv 1 \mod 10.$$

Let’s keep thinking! Suppose that $x$ and $y$ are positive integers that differ by a multiple of four, with $x > y$. Then $x = y + 4k$ for some positive integer $k$. Therefore

$$7^x = 7^{y+4k} = 7^y7^{4k}.$$ But $7^{4k} = (7^4)^k \equiv 1^k \equiv 1 \mod 10$. We conclude that

$$If \ x \equiv y \mod 4 \ then \ 7^x \equiv 7^y \mod 10.$$  

This was a lot of work, but it is an investment that will pay huge dividends later. But for now, we solve our problem of finding the last digit of $1234567^{100}$, which is the same as computing $7^{100}$ modulo 10. Since 100 is congruent to zero modulo 4, $7^{100}$ is congruent to $7^0 = 1$ modulo 10. The sought-for last digit is 1. Whew!

**Problem 19** Find the last digit of $a = 1779^{100}$. How many digits does the number $a$ have?

The order of a number $a$ modulo $m$ is the least positive integer $k$ such that $a^k \equiv 1 \mod m$. For example, the order of 7 modulo 100 is 4.

**Problem 20** (a) Find the order of 36 modulo 101. (b) Find the order of 37 modulo 101. (c) Find the order of 38 modulo 101.

**Number tricks**

You probably know a few number tricks such as “a number is divisible by 9 if the sum of the digits is divisible by 9.” Consider, for example, the number 10918273644. The sum of the digits is 45, which is divisible by 9. So 10918273644 is divisible by 9. Why does this trick work? Well, consider a number which is decimal form is $abcdef$. What this means is that

$$abcdef = a \cdot 10^5 + b \cdot 10^4 + c \cdot 10^3 + d \cdot 10^2 + e \cdot 10 + f.$$
Since $10 \equiv 1 \mod 9$, $10^2 \equiv 1 \mod 9$ and, more generally, $10^n \equiv 1 \mod 9$ for any $n \geq 0$. Thus
\[ abcd ef \equiv a + b + c + d + f \mod 9. \]

Now a number is divisible by 9 if and only if it is congruent to zero modulo 9. Thus
\[ abcd ef \text{ is divisible by } 9 \text{ if and only if } a + b + c + d + f \text{ is divisible by } 9. \]

**Problem 21** A number $abcd$ is divisible by 11 if the alternating sum of the digits is divisible by 11. That is, if $a - b + c - d$ is divisible by 11. Explain why this is true.

### 5.2 Residues

Every positive integer $a$ is congruent modulo $m$ to one of the numbers
\[ R = \{ 0, 1, 2, \ldots, m - 1 \}. \]

To see this, apply the division algorithm to $a$ and $m$: there are integers $q$ and $r$ with $a = qm + r$, where $0 \leq r < m$. Since $a - r = qm$, $a \equiv r \mod m$. In other words, every positive integer is congruent modulo $m$ to one of the $m$ remainders that occur upon division by $m$.

**Example.** Every positive integer is congruent modulo 2 to an element of the set $\{ 0, 1 \}$. Every even number is congruent to 0 modulo 2, and every odd number is congruent to 1 modulo 2.

Any negative number $b$ is also congruent modulo $m$ to one of the numbers in $R$. Indeed, let $a = b + km$. For $k$ sufficiently large, $a$ is positive. By the result of the previous paragraph, $a \equiv r \mod m$ for some remainder $r$. But $b \equiv a$. Thus $b \equiv r$ by transitivity.

The set of remainders $\{ 0, 1, 2, \ldots, m - 1 \}$ is an example of a system of residues modulo $m$. Such a system $R$ has the following properties: (i) Every integer is congruent modulo $m$ to an element of $R$; (ii) No two distinct elements of $R$ are congruent to one another.

We have proved part (i) needed to show that the remainders form a system of residues. For (ii) suppose that there are such $r$ and $s$ in the list of remainders such that $r \equiv s \mod m$. We may assume without loss of generality that $s < r < m$. Subtract $s$ from this inequality to obtain $0 < r - s < m - s$. Since $m - s < m$, we have $0 < r - s < m$. But $0 < (r - s)/m < m$, so $(r - s)/m$ is not an integer, a contradiction.

**Examples.**

1. $R = \{ 0, 1 \}$ is a system of residues modulo 2.
2. $R = \{0, 1, 2\}$ is a system of residues modulo 3. So is $R' = \{-1, 0, 1\}$.

3. $R = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11\}$ is a system of residues modulo 12. So is $R = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$. Both correspond to the hours on a clock face. Thus modular arithmetic, to be defined later, is a generalization of “clock arithmetic.”

**Problem 22** It is now 7 pm. We have to run an experiment that takes 87 hours. When should we be back in the lab?

**Problem 23** (a) List the residues of the numbers $7k$ as $k$ runs over the residues modulo 12. (b) Do the same for $8k$. (c) Do the same, but now change the modulus to 11. (d) Comment on what you have found. Any patterns that you think hold in general?

A residue $a$ modulo $m$ is called quadratic if it is congruent to the square of a number. In other words, $a$ is a quadratic residue if there is a number $x$ such that $x^2 \equiv a \mod m$. We could then say that $x$ is a square root of $a$ modulo $m$. Let’s look at the example of $\mathbb{Z}/5$. The numbers 0, 1, 2, 3, 4 give a system of residues. The residues of their squares are 0, 1, 4, 4, 1. Thus, 0, 1, and 4 are quadratic residues modulo 5. The numbers 2 and 3 are not quadratic residues modulo 5.

**Problem 24** List the residues modulo 11. Which of these are quadratic residues?

### 5.3 Solving congruences

Congruence is a relation very much like equality. In particular, we can solve congruences like $4x \equiv 1 \mod 11$, just as we can solve equations. For the moment we have no theory to guide us, so we rely on brute force instead. Since every integer is congruent modulo 11 to one of the residues 0, 1, ..., 10, we can try each one in turn to see if satisfies the congruence. If $x = 0$, then $4x = 0 \not\equiv 1$. If $x = 1$, $4x = 4 \not\equiv 1$. If $x = 2$, $4x = 8 \not\equiv 1$. If $x = 3$, $4x = 12 \equiv 1 \not\equiv 1$. Success!

We have a solution. If we keep searching, we find that this is the only solution.

**Problem 25** To the extent possible, solve the following congruences: (a) $5x \equiv 1 \mod 7$, (b) $3x \equiv 1 \mod 6$, (c) $x^2 \equiv 2 \mod 7$, (d) $x^2 \equiv 3 \mod 7$.

So far we have relied on brute force and the fact that there are only finitely many residues modulo $m$. This is not practical when $m$ is large. Is there a better way? Well, for congruences of the form

$$ax \equiv b \mod m$$

there is a better way. This congruence is equivalent to the statement that $m$ divides $ax - b$. Thus $ax - b = \ell m$ for some integer $\ell$. Equivalently,

$$ax + my = b,$$
where \( y = -\ell \). This is a Diophantine equation which we learned to solve in the previous section. For a solution to exist, we require that the GCD of \( a \) and \( m \) divide \( b \). Then we can reduce the equation to \( ax + my = GCD \) and apply the extended Euclidean algorithm.

**Problem 26** Solve the congruences (a) \( 1213x \equiv 1 \mod 12345 \), (b) \( 1213x \equiv 100 \mod 12345 \).

Let us summarize the conclusion reached above as a theorem:

**Theorem 4** Let \( a \) and \( m \) be relatively prime. Then the congruence \( ax \equiv b \mod m \) has a unique solution.

We think of the solution as being constructed in two steps. First, we solve the congruence \( ax \equiv 1 \mod m \). The result, which we write for now as \( a' \), is the multiplicative inverse of \( a \) modulo \( m \). Then we solve \( ay \equiv b \mod m \) by setting \( y \equiv a'b \mod m \).

**Problem 27** Find the multiplicative inverse of 5 modulo 1234567.

### 5.4 The algebra of residues

We can now explain the number system \( \mathbb{Z}/m \), pronounced “the integers modulo \( m \),” mentioned in the introduction. First, the set of such numbers. This is just the set of residues modulo \( m \). Any set of residues will do, but we will usually take the remainders 0, 1, ..., \( m-1 \). Thus

\[
\mathbb{Z}/5 = \{ 0, 1, 2, 3, 4 \}.
\]

Second, the operations of addition and multiplication. To add two residues, add them as integers, and then find the corresponding residue. For example, modulo 5 we have \( 3 + 4 = 7 \) as integers. And \( 7 \equiv 2 \mod 5 \), so the result in \( \mathbb{Z}/5 \) is 2. This is how the addition table in the introduction was constructed. We can do the same for multiplication. Thus \( 2 \times 4 = 8 \), and \( 8 \equiv 3 \mod 5 \). Thus the result in \( \mathbb{Z}/5 \) is 3.

**Remark.** Usually we will be aware of the context in which we do our operations — the integers or the integers mod 5, for example. Thus we will not be confused. However, when we are talking about integers and integers modulo \( m \), we sometimes need a notation to distinguish the two possible interpretations of the symbols. When this is needed we will write \( a \) for an integer and \( [a] \) for its residue modulo \( m \). Thus we could write \( [2] \times [4] = [3] \) in \( \mathbb{Z}/5 \). We could also write the multiplication rule as \( [a] \times [b] = [a \times b] \) — multiply two residues by multiplying them as integers, then taking the residue modulo \( m \).

What about subtraction? We seem to have ignored it. This works in the same way. For example, \( 3 - 4 = -1 \). Since \(-1 \equiv 4 \mod 5\), the result is 4 in \( \mathbb{Z}/5 \).
And what about division? This is tricky. Consider, for example, $3/4$. There is no way to make sense of this — what residue would it be congruent to? We have to approach this problem by a different route. In the rational numbers we can think of the quotient $a/b$ as the product $ab^{-1}$ where $b^{-1}$ is the multiplicative inverse of $b$. The multiplicative inverse is the solution — if it exists — of the equation $bx = 1$. In the rational numbers this solution is $1/b$, the reciprocal of $b$. But in $\mathbb{Z}/m$ we can make sense of the multiplicative inverse whenever the equation $bx = 1$ is solvable. Equivalently, when the congruence $bx \equiv 1 \pmod{m}$ is solvable. For example to find the inverse of 4 in $\mathbb{Z}/5$, we solve $4x \equiv 1 \pmod{5}$. The solution is $x = 4$. This may seem strange, but $4 \times 4 = 16 \equiv 1 \pmod{5}$.

Remark. We didn’t really need the notion of additive inverse because subtraction is defined modulo $m$. However, we can define it as well. The additive inverse of $b$ is the solution of $x + b = 0$.

Problem 28 (a) Construct addition and multiplication tables for $\mathbb{Z}/7$. Construct also tables of additive and multiplicative inverses. (b) Do the same for $\mathbb{Z}/8$.

5.5 Sums of squares

To convince you that the related ideas of congruence and modular arithmetic are indeed very powerful, let us consider the following classical problem in number theory: when is a number the sum of two squares? In other words, given a number $n$, when is the equation

$$x^2 + y^2 = n$$

solvable? One can always try our old work-horse, brute-force search. Thus, to find a solution of $x^2 + y^2 = 13$, we can test the lattice points in some square. Since $y^2$ is positive, $x^2 \leq 13$ and $|x| \leq \sqrt{13} \sim 3.6$. Thus it is enough to assume $0 \leq x, y \leq 3$. There are only 16 lattice points in this square, so we quickly discover that $7 = 2^2 + 3^2$. When we try the same argument on the equation $x^2 + y^2 = 11$, we find no solutions.

So far so good. But why does one equation have a solution while the other one does not? This question is something that the idea of congruence can give an easy partial answer to. The trick is to consider congruence mod 4. Modulo 4, there are just four residues: 0, 1, 2, and 3. The squares of these residues are 0, 1, 0, and 1. Thus the possible values of the expression $x^2 + y^2$ mod 4 are 0 + 0, 1 + 0, 0 + 1, and 1 + 1. Therefore the possible values of $n$ mod 4 are 0, 1, and 2. Since 3 is not on this list, we see that a number congruent to 3 mod 4 cannot be written as the sum of two squares. This is quite remarkable: we have discovered — and proved — a general theorem! Of course, there is more to say: if $n$ is congruent to 0, 1, or 2, is it representable as a sum of squares?

Problem 29 Experiment with numbers congruent to 0, 1, or 2 mod 4. Which
ones can be written as sums of squares? Can you see any general patterns? Or formulate any conjectures?

Remark. The following program is useful in doing brute-force searches for solutions to \( x^2 + y^2 = n \). For example, \( \text{sos}(4937) \) produces the output \( 29 \ 64 \). Thus \( 29^2 + 64^2 = 4937 \).

```python
def sos(n):
    b = int(sqrt(n)) + 1
    for x in range(1,b):
        for y in range(x,b):
            if x*x + y*y == n:
                print x, y
```

5.6 Rings and fields

A ring is a set endowed with operations of addition and multiplication that obey certain laws such as \( a + b = b + a \). Examples are the integers and the integers modulo \( m \). We will give a more formal definition later. A field is a special kind of ring with in which every nonzero element has a multiplicative inverse. Examples are the rationals \( \mathbb{Q} \) and the integers modulo 5. The ring of integers, however, is not a field: the number 2 (for instance) does not have a multiplicative inverse in the integers. In fact, the only integers which have multiplicative inverses are \( \pm 1 \).

Another example of ring which is not a field is \( \mathbb{Z}/6 \). To see why not, observe that \( 2 \cdot 3 = 0 \) in \( \mathbb{Z}/6 \). If 2 had a multiplicative inverse \( b \) then we would have \( b \cdot 2 = 1 \). Multiply the equation before the preceding one to obtain \( b \cdot (2 \cdot 3) = b \cdot 0 = 0 \). Thus \( (b \cdot 2) \cdot 3 = 0 \), and so \( 1 \cdot 3 = 0 \), and, finally, \( 3 = 0 \), which is false. This contradiction shows that 2 cannot have a multiplicative inverse. It is not hard to generalize this observation: whenever a ring has nonzero elements \( a \) and \( b \) such that \( ab = 0 \), the ring is not a field. Consequently, if \( m \) is composite, then \( \mathbb{Z}/m \) is not a field. The converse of the preceding statement is also true:

**Theorem 5** If \( p \) is prime, then \( \mathbb{Z}/p \) is a field.

For the proof, let \( 0 < a < p \) be a nonzero residue modulo \( p \). Consider the Diophantine equation \( ax + py = 1 \). The GCD of \( a \) and \( p \) is 1, so this equation has an integer solution \( (x, y) \). then the residue of \( x \) is the multiplicative inverse of \( a \) in \( \mathbb{Z}/p \).

**Gaussian integers and sums of squares**

We will give two more examples of rings. The first is the ring of numbers of the form \( a + bi \), where \( i = \sqrt{-1} \). This ring has a name, the Gaussian integers, and a symbol, \( \mathbb{Z}[i] \) — the integers with \( i \) adjoined. Why is it a ring? Well,
it is clear how to add and multiply Gaussian integers, since they are complex numbers. The only question is: does the result have the right form? Let’s examine addition.

\[(a + bi) + (c + di) = (a + c) + (b + d)i.\]

The sum is a Gaussian integer. Let’s examine multiplication.

\[(a + bi)(c + di) = (ac - bd) + (ad + bd)i\]

The product is a Gaussian integer. Let \(z = a + bi\) be a Gaussian integer, and let \(\bar{z} = a - bi\) be its conjugate. Then

\[|z|^2 = z \bar{z} = a^2 + b^2.\]

Hmmmmm. This must be a fact useful for studying sums of squares. Indeed, let \(w = c + di\). Then \(|w|^2 = c^2 + d^2\). Now comes the fun. We can write the quantity \(|z|^2|w|^2\) in two ways:

\[(z \bar{z})(w \bar{w}) = (zw)(\bar{z} \bar{w})\]

Equivalently,

\[(a^2 + b^2)(c^2 + d^2) = (ac - bd)^2 + (ad + bd)^2.\]

This identity says something amazing: if two integers \(m\) and \(n\) can be written as sums of squares, then so can their product \(mn\). Indeed, if \(m = a^2 + b^2\) and \(n = c^2 + d^2\), then

\[mn = (ac - bd)^2 + (ad + bd)^2.\]

For example, 101 = 1^2 + 10^2 and 137 = 4^2 + 11^2. Thus 13837 = 101 \cdot 137 is sum of the squares of \(e = 1 \cdot 4 - 10 \cdot 11 = -106\) and \(f = 1 \cdot 11 + 10 \cdot 4 = 51\). That is, 13837 = 106^2 + 51^2.

**Remark.** The expression \(z \bar{z}\) occurs in so many different contexts that we give it its own name — the *norm* of \(z\) — and its own symbol, \(N(z)\). The basic identity is \(N(zw) = N(z)N(w)\). We call this the norm identity.

**Pell’s equation and the ring \(\mathbb{Z}[\sqrt{2}]\)**

The second example is the ring \(\mathbb{Z}[\sqrt{2}]\) obtained by adjoining the square root of 2 to the integers. It consists of expressions of the form \(z = a + b\sqrt{2}\), where \(a\) and \(b\) are integers. It is clear that this ring is closed under addition. Multiplication is almost as easy. If \(w = c + d\sqrt{2}\), then

\[zw = (a + b\sqrt{2})(c + d\sqrt{2}) = (ac + 2bd) + (ad + bd)\sqrt{2}.\]

Define the conjugate of \(z\) by \(\bar{z} = a - b\sqrt{2}\). Then

\[N(z) = z \bar{z} = a^2 - 2b^2.\]
Our interest now is in solutions to Pell’s equation,
\[ x^2 - 2y^2 = 1. \]
Therefore solutions are elements of the ring \( \mathbb{Z}[\sqrt{2}] \) which have norm one. An example is \( z = 3 + 2\sqrt{2} \). A beautiful consequence of the norm identity is that we can multiply numbers corresponding to solutions of Pell’s equation to get new solutions of Pell’s equation. For example, \( (3 + 2\sqrt{2})^2 = 17 + 12\sqrt{2} \). Thus \( (x, y) = (17, 12) \) is also a solution of Pell’s equation. A striking consequence is that Pell’s equation has infinitely many solutions corresponding to the numbers \( (3 + 2\sqrt{2})^n \).

Remark. The solutions \( S \) to Pell’s equation are the units of the ring \( \mathbb{Z}[\sqrt{2}] \): the elements \( z \) that have a multiplicative inverse. Indeed, since \( z\bar{z} = 1 \) for a solution to Pell’s equation, \( \bar{z} \) is the multiplicative inverse. We write \( U(R) \) for the set of units of a ring. Recall that \( U(\mathbb{Z}) = \{ \pm 1 \} \).

Problem 30 Determine the units of (a) the ring of Gaussian integers, (b) the ring \( \mathbb{Z}/7 \), (c) the ring \( \mathbb{Z}/8 \).

5.7 Groups

In the last section we briefly touched on the notion of the set of units of a commutative ring:
\[ U(R) = \{ \text{elements of } R \text{ that have a multiplicative inverse} \}. \]
For example, when \( p \) prime, \( \mathbb{Z}/p \) is a field, and so
\[ U(\mathbb{Z}/p) = \{ \text{nonzero elements} \}. \]
On the other hand,
\[ U(\mathbb{Z}/8) = \{ 1, 3, 5, 7 \}. \]
The set of units forms what is called a group with respect to multiplication. This means that (i) the product of any two elements in \( U \) is in \( U \) (closure); (ii) multiplication is associative, that is, \( a(bc) = (ab)c \); (iii) there is an identity element \( 1 \) such that \( 1a = a1 = a \) for all \( a \); (iv) multiplicative inverses exist. Groups arise in many branches of mathematics and are also important in physics.

The group of units in a ring is abelian, meaning that \( ab = ba \) for all \( a \) and \( b \). Not all groups are abelian. For an example, consider the set \( G \) of two-by-two matrices with nonzero determinant. The condition on the determinant ensures that a multiplicative inverse exists. Consider, for example the matrices
\[ A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \]
Then one checks that \( AB \neq BA \). Quantum mechanics, which is the physics of atoms and subatomic particles, is based on the fact that matrix multiplication

27
is not commutative. Nonetheless, all the groups considered henceforth will be abelian.

A group is called cyclic if each its elements can be expressed as power of a single element $g$. Thus

$$G = \{ \ldots, g^{-2}, g^{-1}, g^0 = 1, g^1, g^2, \ldots \}.$$ 

The element $g$ is a generator of the group.

For an example of a cyclic group, consider the units of $\mathbb{Z}/5$, $U(\mathbb{Z}/5) = \{1, 2, 3, 4\}$.

Compute first the powers of 2. The square is $2^2 = 4$. The cube is $2^3 = 8 \equiv 3$. The fourth power is $2^4 = 2 \cdot 3 = 6 \equiv 1$. Thus all elements of $U$ are powers of 2.

We were lucky: our first choice of element was a generator. Thus, setting $g = 2$, we have

$$U = \{ g^0, g^1, g^2, g^3 \} = \{ 1, 2, 4, 3 \}.$$ 

Is 2 the only generator of $U$? Let’s compute powers of the other elements. $3^2 = 9 \equiv 4$, $3^3 \equiv 3 \cdot 4 \equiv 2$, $3^4 \equiv 3 \cdot 2 \equiv 1$. Thus 3 is also a generator! What about 4? Well, $4^2 = 16 \equiv 1$, so it is not a generator.

**Problem 31** Find a generator of the group of units of $\mathbb{Z}/7$.

For a different kind of abelian group, consider the units of $\mathbb{Z}/8$. The units are the odd residues, 1, 3, 5, and 7. A quick computation shows that $a^2 \equiv 1 \pmod{8}$ for all odd residues modulo 8. Thus no unit modulo 8 can generate the group of units.

Notice that no negative powers of $g$ appeared in the lists of group elements considered above. But the multiplicative inverse of $g^k$ is $g^{-k}$, so negative powers must be there somewhere. This apparent paradox is resolved by the observation that $g^{-k} = 1 \cdot g^{-k} = g^0 g^{-k} = g^{n-k}$. The exponent $n-k$ is positive.

Recall that the (multiplicative) order of an element $a$ is the least positive integer $n$ such that $a^n = 1$. If there is no such element, we say that $a$ has infinite order. The orders of the elements of $U(\mathbb{Z}/5)$ are as follows: 1 has order 1, 2 and 3 have order 4, and 4 has order 2. Notice that elements of order 4 are generators and that the orders of other elements divide 4. The number 4 is also the order of the group — by definition the number of elements. Hmmmmmm. Is there a general pattern here?

**Problem 32** (a) Investigate the orders of elements of $U(\mathbb{Z}/n)$ for several $n$, where $n$ is a small prime. Are these groups cyclic? If so, what fraction of the elements are generators? (b) Investigate the same problem where $n$ is a small composite number.
Let’s look at one more example. We compute the powers of $g = 2$ in $\mathbb{Z}/19$:

$$\{ g^0, g^1, g^2, \ldots, g^{18} \} = \{ 1, 2, 4, 8, 16, 13, 7, 14, 9, 18, 17, 15, 11, 3, 6, 12, 5, 10 \}$$

There doesn’t seem to be a readily discernible pattern to the list of powers $g^k$. In fact, this is the beginning (but only the beginning) of the theory of pseudorandom number generators.

We close with three important theoretical observations.

**Theorem 6** Suppose that $a$ is an element of order $k$ in a group $G$ and that $a^\ell = 1$. Then $k$ divides $\ell$.

**Proof:** To say that $k$ is the order of $a$ is to say that $a^k = 1$ and that no smaller positive integer satisfies this property. Apply the division algorithm to write $\ell = qk + r$, where the remainder $r$ satisfies $0 \leq r < k$. On the one hand, $a^\ell = a^{qk+r} = (a^k)^q a^r = a^r$, since $a^k = 1$. On the other hand, $a^\ell = 1$ by hypothesis. Comparing these two computations of $a^\ell$, we conclude that $a^r = 1$. But $r < k$. Thus we would be forced to conclude that $a$ has order less than $k$, a contradiction, unless we conclude instead that $r = 0$. But this says that $k$ divides $\ell$. Q.E.D.

For a cyclic group of order $n$, $a^n = 1$ for all elements of the group. Thus we have the following.

**Corollary 2** Let $G$ be a cyclic group of order $n$. Let $a$ be an element of $G$. Then the order of $a$ divides $n$.

Finally, we give an example and use it to state prove an amazing theorem about finite abelian groups. The example will be the group of units of $\mathbb{Z}/7$. We list the units:

$$1, 2, 3, 4, 5, 6.$$  \hspace{1cm} (15)

Multiply each element of the list by 3:

$$3 \cdot 1, 3 \cdot 2, 3 \cdot 3, 3 \cdot 4, 3 \cdot 5, 3 \cdot 6.$$  \hspace{1cm} (16)

Then reduce mod 7:

$$3, 6, 2, 5, 1, 4.$$  \hspace{1cm} (17)

The product of all the elements in the first list is $6!$. The product of all the elements in the second list is $3^6 \cdot 6!$. Modulo 7, this product is the same as the product of the elements in the third list. But the third list is the same as the first, except for order. We conclude that 

$$3^6 \cdot 6! \equiv 6! \mod 7.$$
Multiply the preceding equation by the multiplicative inverse of $6!$ to obtain
\[ 3^6 \equiv 1 \pmod{7} \]
But the argument we just have applies to every nonzero element modulo 7. Thus we can say
\[ a^6 \equiv 1 \pmod{7} \text{ if } a \not\equiv 0 \pmod{7}. \]
As a corollary, every nonzero element of $\mathbb{Z}/7$ has order dividing 6. We now prove a general version of this result:

**Theorem 7** Let $G$ be an abelian group of order $n$. Then $a^n = 1$ for every element $a$ of $G$.

**Proof:** List the elements of $G$ as
\[ L = g_1, g_2, \ldots, g_n. \]
Let $g^* = g_1 \cdot g_2 \cdots g_n$ be the product of these elements. Let $a$ be an element of $G$, and consider the list
\[ L' = ag_1, ag_2, \ldots, ag_n. \]
On the one hand, the product of the elements in this list is $a^n g^*$. On the other hand we claim that second list is just the first list written in a different order, as in the example above. If this is so, then
\[ a^n g^* = g^*. \]
Multiplying both sides of this equation by the inverse of $g^*$, we have
\[ a^n = 1, \]
as desired. It remains to prove the claim. To this end, suppose that $ag_i = ag_j$. Multiply by the inverse of $a$ to conclude that $g_i = g_j$, and hence $i = j$. The contrapositive of this statement is that if $i \neq j$ then $ag_i \neq ag_j$. Therefore the list $L'$ consists of $n$ distinct elements. Each element of $L'$ is an element of $L$. Since both lists have the same size, they must be the same, except for the order in which the elements are listed. **Q.E.D.**

**Corollary 3** If $n$ is a prime, then $a^{n-1} \equiv 1 \pmod{n}$ for all $a \not\equiv 0 \pmod{p}$.

For the proof, simply apply the previous result to the group of nonzero elements of $\mathbb{Z}/n$. There is another form of this result:

**Corollary 4** If $n$ is prime, then $a^n \equiv a \pmod{n}$ for all $a$.

The two corollaries just stated are versions of Fermat’s little theorem. Fermat (1601 - 1665), who lived in Toulouse, France, is one of the founders of the theory of numbers.
5.8 Primality tests

There is one obvious test for whether a number is prime: factor it into primes. If there is just one factor, the number is prime. We can do this by trial division. Thus, to determine whether 437 is prime, we divide it by 2 and find a remainder of 1, then divide by 3 and find a remainder of 2, and so on. Finally, when we divide by 19, we find a remainder of zero. Thus 437 is not prime. By contrast, trial division applied to 439 yields no divisors. Thus 439 is prime.

Note that to factor $n$, it is not necessary to use trial divisors $d$ larger than the square root of $n$. To see why this is true, suppose that $n = ab$, and suppose that $a \leq b$. Multiply this inequality by $a$ to obtain $a^2 \leq ab = n$. Take the square root to conclude that $a \leq \sqrt{n}$. Thus to find factors of 439, we need consider factors no larger than 20, which is the integer part of $\sqrt{439} = 20.952 \cdots$.

Even with this improvement, trial division is not an efficient algorithm for testing to see whether a number is prime. Suppose, for example, that we need to factor a 200 digit number — a typical cryptographic task. The largest trial divisor is a 100 digit number. Thus roughly $10^{100}$ divisors must be tested. Suppose that ten billion divisors can be tested per second. In a worst-case situation trial division would require $10^{90}$ seconds. A year is about $3 \times 10^7$ seconds. Thus trial division requires $10^{83}$ years — far too many billions of billions of years to wait.

Let’s look for other ways of determining whether a number is prime. Fermat’s little theorem says that if $n$ is prime then $a^{n-1} \equiv 1 \mod n$ if $a \neq 0 \mod n$. The contrapositive of this statement is that for $a \neq 0 \mod n$, if $a^{n-1} \not\equiv 1 \mod n$, then $n$ is not prime. Consider, for example, the number 35. We readily find that

$$2^{34} = 17179869184 \equiv 9 \mod 35.$$  

Thus 35 is not prime. This is no surprise, but what is surprising is that we did not have to factor 35 to reach this conclusion. Indeed, we know for certain that 35 is composite even though we made no use of its factorization as $5 \times 7$. On the other hand, for the prime 37, we have

$$2^{36} = 68719476736 \equiv 1 \mod 37.$$  

Perhaps computations of this kind can lead to an efficient primality test?

To investigate this question, let’s compute $2^{n-1} \mod n$ for all the odd numbers in the range 3 to 41. The results, given below, look pretty good! Below every prime is the number 1, and above every 1 sits a prime. Perhaps we are on to something?

For our method to be useful, we must have an efficient way of computing powers modulo a given integer. This is not hard using the method of “repeated squares.” We illustrate it by showing how to compute $11^{105} \mod 113$. The first step is to find the binary expansion of the exponent. The binary expansion of
105 is 1101001, which means that
\[ 105 = 2^6 + 2^5 + 2^3 + 1. \]

Therefore
\[ 11^{105} = 11^{2^6} \cdot 11^{2^5} \cdot 11^{2^3} \cdot 11. \]

Now observe that the square of \( 11^{2^k} \) is \( 11^{2^{k+1}} \). Thus we can construct the needed powers of 11 by repeatedly squaring 11 and reducing modulo 113. The result is recorded in the table below. Note that the binary expansion of 105 is written backwards.

Once a table like the one above is constructed, just multiply together the entries in the middle row that sit above a 1, always working modulo 113. The result is \( 11^{105} \mod 113 \).

To find the binary expansion of a number, we also construct a little table. The idea is to repeatedly divide by 2 and record the remainder, as illustrated below. Thus the entry below 105 is 1, since when we divide 2 into 105, the remainder is 1. The quotient 52 is recorded to the right of 105. In other words: record the quotient to the right, the remainder below.

**Problem 33** Compute (a) \( 321^{571} \mod 991 \), (b) \( 123^{4567} \mod 891011 \)

**Problem 34** Consider the 38-digit number
\[ 15140955104265152459705939852879368553. \]
Problem 35 Find the first number \( n \) > \( 10^{100} \) which is “likely to be prime.”

It is convenient to have a computer program for calculating modular powers. Below is one such program. It implements the algorithm you used for hand computations: to compute \( b^e \mod m \), use the binary expansion of the exponent to control which of the numbers \( a^{2^k} \) should be multiplied together.

```python
def modpower(b,e,m):
    """modpower(b,e,m) = b^e mod m"
    P = 1
    S = b
    while e > 0:
        r = e%2
        e = e/2
        if r == 1:
            P = P*S % m
            S = S*S % m
    return P
```

The variable \( P \) accumulates the quantity \( b^e \). The remainder \( r \) is the current digit in the binary expansion of \( e \). Each time we compute the remainder, we also replace \( e \) by its quotient upon division by 2. The variable \( S \) holds the current value of the repeated square of \( b \).

Once a function like `modpower` is in hand, it is easy to construct a function that carries out the Fermat test:

```python
ft = lambda n: modpower(2,n-1,n)
```

With `ft` it is easy to search for primes. Here is one short example, which applies the Fermat test to all the odd numbers in the range \([10000, 10100]\).

```python
for k in range(10000, 10100):
    if k % 2 == 1:
        if ft(k) == 1:
            print k,
```

The output of the program is

```
10007 10009 10037 10039 10061 10067 10069 10079 10091 10093 10099
```
Out of fifty odd numbers in the given interval, eleven pass the Fermat test, and all eleven are actually prime, as is shown by trial division.

We now have to face up to the fact that the Fermat test, despite its good success rate so far, is not backed up by a theorem. Fermat’s theorem says that if \(a^{n-1} \not\equiv 1 \pmod{n}\) for some \(a\) satisfying \(0 < a < n\), then \(a\) is composite. It does not say that if \(a^{n-1} \equiv 1 \pmod{n}\) for some \(a\) in this range, then \(a\) is prime. So we will make another experiment. Suppose given a function \texttt{factor}\ that accepts an integer and produces a list of its prime factors, e.g., the crude trial division program given below. For example, we have the following:

\[
\texttt{factor(1234567)}
\]
\[
[127, 9721]
\]

Now we can run an extensive test for numbers that pass Fermat’s test but are not prime:

```python
def test(a,b):
    for n in range(a,b):
        if n % 2 == 1:
            if ft(n) == 1:
                ff = factor(n)
                if len(ff) > 1:
                    print n, ff
```

Running \texttt{test(2,100)} we find no failures of the Fermat test. But here is the output of \texttt{test(2,1000)}:

```
>>> test(2,1000)
341 [11, 31]
561 [3, 11, 17]
645 [3, 5, 43]
```

This is three false results out of a total of 170 numbers accepted by the Fermat test — a failure rate of about 1.76 percent.

**Problem 36** Define the failure rate of the Fermat test in an interval \([a, b]\) to be the number of odd integers in this interval that are accepted but are not prime, divided by the number that are accepted. Study the failure rate as a function of \([a, b]\).

**Problem 37** Can the Fermat test be modified so as to lower the failure rate?

```python
def factor(n):
    """factor(n): return list of prime factors of n""
    d = 2
```
factors = [ ]
while n > 1:
    if n % d == 0:
        n = n/d
        factors.append(d)
    else:
        d = d + 1
if n % d == 0:
    factors.append(d)
return factors