MTH 505: Number Theory
Course Notes

Spring 2017
Drew Armstrong

Contents

Introduction

This is a course about the system of integer numbers. We denote the set of these numbers by a blackboard bold letter Z (for *Zahlen*):

$$\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}.$$

In some sense this is the oldest kind of mathematics that is still actively studied. The subject is full of open problems that are easy to state but have resisted solution for hundreds or even thousands of years. In this course I will focus on the period of activity that began with Fermat in the 1600s and ended roughly with Dedekind at the end of the 19th century. This was the period in which the modern language of number theory was developed.

There are two major branches of modern number theory, called *algebraic* and *analytic*. At higher levels the two branches are combined but in the beginning they look quite different. In this class I will focus on the *algebraic* branch of number theory, but I may mention some analytic ideas from time to time. Most of our time will be spent looking for integer solutions to polynomial equations, such as

$$x^n + y^n = z^n. \tag{1}$$

It is a famous conjecture of Fermat that this equation has **no integer solutions** when $xyz \neq 0$ and $n \geqslant 3$.[1] After hundreds of years of intense work, a proof of this conjecture was finally announced by Andrew Wiles of Princeton in 1993. There is no way that we can discuss Wiles' proof in this class, but we will examine the special cases of (1) when $n = 2, 3, 4$. Such polynomial equations of integers are called *Diophantine equations* after Diophantus of Alexandria (ca. 201–299 AD), but they were also studied in the *Sunzi Suanjing*, a Chinese mathematical work from the 3rd to 5th century AD.

We will begin by looking at the so-called **linear Diophantine equations** such as

$$ax + by + c = 0,$$

and we will work our way up to **quadratic Diophantine equations** such as

$$ax^2 + bxy + cy^2 + dx + ey + f = 0.$$

Along the way we will learn about the algebraic techniques that were developed throughout history to solve such equations. Cubic equations and above can become impossibly difficult. In fact, it was proved by Yuri Matyasevich and Julia Robinson in the 1970s that the general problem is technically undecidable (in a certain sense).[2] This is just as well, because the linear and quadratic cases are more than enough to keep us busy for a full semester.

[1] In fact, Fermat claimed that he could prove this, but most people today believe that his proof was wrong.

[2] They gave a negative answer to Hilbert's 10th Problem by showing that there does not exist an algorithm to determine whether a given Diophantine equation has a solution.

1 The Definition of the Integers

This is a class in which we will prove things, and it is impossible to prove anything without some technical definitions to work with. Therefore we will begin by stating the technical definition of the integers.

Logically speaking, the modern definition of $\mathbb{Z}$ begins with the definition of the *natural numbers*

$$\mathbb{N} = \{0, 1, 2, \ldots\}.$$

It turns out that the only properties of $\mathbb{N}$ that are really fundamental are the properties of "successor" and "induction". These are captured with Giusuppe Peano's system of axioms.

1.1 Peano's Axioms for $\mathbb{N}$

Let $\mathbb{N}$ be a set equipped with an equivlence relation "=" and a "successor" function $\sigma : \mathbb{N} \to \mathbb{N}$ satisfying the following four properties:

(P1) There exists a special element called $0 \in \mathbb{N}$.

(P2) The element 0 is not the successor of any number, i.e.,

$$\forall n \in \mathbb{N}, \sigma(n) \neq 0.$$

(P3) Every number has a unique successor, i.e.,

$$\forall m, n \in \mathbb{N}, (\sigma(m) = \sigma(n)) \Rightarrow (m = n).$$

(P4) *The Induction Principle.* If a set of natural numbers $S \subseteq \mathbb{N}$ contains 0 and is closed under succession, then we must have $S = \mathbb{N}$. In other words, if we have

- $0 \in S$,
- $\forall n \in \mathbb{N}, (n \in S) \Rightarrow (\sigma(n) \in S)$,

then it follows that $S = \mathbb{N}$.

//

I'll admit that this definition doesn't look much like the integers we know and love. For example, where are the arithmetic operations of addition/subtraction and multiplication? It turns out that these structures are inherent in the Peano axioms but it takes some work to get them out. [You will investigate this on HW1.] After all the work is done, we could rephrase the definition in a much friendlier way.

1.2 A Friendly Definition of $\mathbb{Z}$

Let $\mathbb{Z}$ be a set equipped with

- an *equivalence relation* "=" defined by
  - $\forall\, a \in \mathbb{Z}$, $a = a$ (reflexive)
  - $\forall\, a, b \in \mathbb{Z}$, $(a = b) \Rightarrow (b = a)$ (symmetric)
  - $\forall\, a, b, c \in \mathbb{Z}$, $(a = b \ \wedge \ b = c) \Rightarrow (a = c)$ (transitive),
- a *total ordering* "$\leqslant$" defined by
  - $\forall\, a, b \in \mathbb{Z}$, $(a \leqslant b \ \wedge \ b \leqslant a) \Rightarrow (a = b)$ (antisymmetric)
  - $\forall\, a, b, c \in \mathbb{Z}$, $(a \leqslant b \ \wedge \ b \leqslant c) \Rightarrow (a \leqslant c)$ (transitive)
  - $\forall\, a, b \in \mathbb{Z}$, $(a \leqslant b \ \vee \ b \leqslant a)$ (total)
- and two binary operations
  - $\forall\, a, b \in \mathbb{Z}$, $\exists\, a + b \in \mathbb{Z}$ (addition)
  - $\forall\, a, b \in \mathbb{Z}$, $\exists\, ab \in \mathbb{Z}$ (multiplication)

which satisfy the following twelve properties:

Axioms of Addition.

(A1) $\forall\, a, b \in \mathbb{Z}$, $a + b = b + a$ (commutative)

(A2) $\forall\, a, b, c \in \mathbb{Z}$, $a + (b + c) = (a + b) + c$ (associative)

(A3) $\exists\, 0 \in \mathbb{Z}$, $\forall a \in \mathbb{Z}$, $0 + a = a$ (additive identity exists)

(A4) $\forall\, a \in \mathbb{Z}$, $\exists\, b \in \mathbb{Z}$, $a + b = 0$ (additive inverses exist)

These four properties tell us that $\mathbb{Z}$ is an *additive group*. It has a special element called 0 that acts as an "identity element" for addition, and every integer a has an "additive inverse". If b and c are two such additive inverses then by applying axioms (A1)–(A3) we obtain

$$b = b + 0 = b + (a + c) = (b + a) + c = 0 + c = c.$$

Thus additive inverses are unique; we will denote **the** additive inverse of a by "$-a$".

Axioms of Multiplication.

(M1) $\forall\, a, b \in \mathbb{Z}$, $ab = ba$ (commutative)

(M2) $\forall\, a, b, c \in \mathbb{Z}$, $a(bc) = (ab)c$ (associative)

(M3) $\exists\, 1 \in \mathbb{Z} \backslash 0, \forall\, a \in \mathbb{Z}$, $1a = a$ (multiplicative identity exists)

Notice that elements of $\mathbb{Z}$ do **not** have "multiplicative inverses". That is, we can't divide in $\mathbb{Z}$. So $\mathbb{Z}$ is not quite a group under multiplication. We also need to say how addition and multiplication behave together.

Axiom of Distribution.

(D) $\forall\, a, b, c \in \mathbb{Z},\ a(b + c) = ab + ac$

We can paraphrase these first eight properties by saying that $\mathbb{Z}$ is a *(commutative) ring*. Next we will describe how arithmetic and order interact.

Axioms of Order.

(O1) $\forall\, a, b, c \in \mathbb{Z},\ (a \leqslant b) \Rightarrow (a + c \leqslant b + c)$

(O2) $\forall\, a, b, c \in \mathbb{Z},\ (a \leqslant b\ \wedge\ 0 \leqslant c) \Rightarrow (ac \leqslant bc)$

(O3) $0 < 1$ (this means that $0 \leqslant 1\ \wedge\ 0 \neq 1$)

These first eleven properties tell us that $\mathbb{Z}$ is an *ordered ring*. However, we have not yet defined $\mathbb{Z}$ because there exist other ordered rings, for example the real numbers $\mathbb{R}$. To distinguish $\mathbb{Z}$ among the ordered rings we need one final axiom. This last axioms is equivalent to the Induction Principle but we will state it in a more convenient way.

The Well-Ordering Principle.

Let $S \subseteq \mathbb{Z}$ be any *non-empty* set of integers that is *bounded below*. That is, assume that there exists some $s \in S$ and assume that there exists some $b \in \mathbb{Z}$ such that we have $b \leqslant s$ for all $s \in S$. In this case we conclude that the set S **contains a least element**, i.e., an element $\ell \in S$ such that $\ell \leqslant s$ for all $s \in S$. Formally, we have the following:

(WO) $\forall\, S \in \wp(\mathbb{N}) \backslash \varnothing,\ \exists\, \ell \in S,\ \forall\, s \in S,\ \ell \leqslant s$

You can see from the formal statement that this axioms is logically the most complicated. It took quite a while for people to realize that this is an axiom and not a theorem. This was essentially the contribution of Giuseppe Peano in 1889, following earlier work of Richard Dedekind. //

Exercise: Convince yourself that the rational numbers $\mathbb{Q}$ do **not** satisfy the Well-Ordering Principle. [Don't worry about the formal definition of $\mathbb{Q}$; just use your intuition.]

Again, it is quite a bit of work to prove that these axioms characterize the system of integers uniquely. We won't bother.

1.3 Discussion: Multiplicative Cancellation

There are two further properties of the integers that are useful in proofs, called *additive* and *multiplicative cancellation*. Additive cancellation can be proved easily from the Friendly Axioms of Addition.

Additive Cancellation in $\mathbb{Z}$. For all integers $a, b, c \in \mathbb{Z}$ we have

$$(a + c = b + c) \Rightarrow (a = b)$$

//

Proof. Consider any $a, b, c \in \mathbb{Z}$ such that $a + c = b + c$. From axiom (A4) we know that there exists an integer $d \in \mathbb{Z}$ with the property $c + d = 0$. By adding d to both sides we obtain

$$\begin{aligned} a + c &= b + c \\ (a + c) + d &= (b + c) + d \\ a + (c + d) &= b + (c + d) \\ a + 0 &= b + 0 \\ a &= b. \end{aligned}$$

□

In other words, additive cancellation holds in $\mathbb{Z}$ because we can "subtract". However, we can also prove that additive cancellation holds in $\mathbb{N}$, where subtraction in not always possible. On HW1 you will use the Peano axioms to show that the following recursively defined operations $+, \cdot : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ are commutative and associative:

$$\begin{aligned} a + 0 &= a, \\ a + \sigma(b) &= \sigma(a + b), \\ a \cdot 0 &= 0, \\ a \cdot \sigma(b) &= (a \cdot b) + a. \end{aligned}$$

Additive Cancellation in $\mathbb{N}$. For all natural numbers $a, b, c \in \mathbb{N}$ we have

$$(a + c = b + c) \Rightarrow (a = b)$$

//

Proof. From the definition of "+" we already know that $a + 0 = b + 0$ implies $a = a + 0 = b + 0 = b$. Now fix $a, b \in \mathbb{N}$ and let $S \subseteq \mathbb{N}$ be the set of natural numbers $c \in \mathbb{N}$ with the property $(a + c = b + c) \Rightarrow (a = b)$. We will use induction to prove that $S = \mathbb{N}$. We just saw

that $0 \in S$. So consider any $n \in \mathbb{N}$ and assume for induction that $n \in S$. That is, assume that $(a+n = b+n) \Rightarrow (a = b)$. In this case we want to show that $(a+\sigma(n) = b+\sigma(n)) \Rightarrow (a = b)$ is also true. Indeed, if $a + \sigma(n) = b + \sigma(n)$ then we must have

$$\begin{aligned} \sigma(a+n) &= a + \sigma(n) \\ &= b + \sigma(n) \\ &= \sigma(b+n). \end{aligned}$$

Then axiom (P3) says that $a + n = b + n$ and the assumption $n \in S$ implies that $a = b$ as desired. We have show that $0 \in S$ and that $(n \in S) \Rightarrow (\sigma(n) \in S)$ for all $n \in \mathbb{N}$. It follows from the Principle of Induction (P4) that $S = \mathbb{N}$ as desired. □

The issue of multiplicative cancellation in $\mathbb{Z}$ is much more subtle. We would like to prove that for all $a, b \in \mathbb{Z}$ we have $(ac = bc) \Rightarrow (a = b)$, but we quickly note that something is wrong with this statement because we have

$$2 \cdot 0 = 0 = 3 \cdot 0, \quad \text{but} \quad 2 \neq 3.$$

The correct statement goes as follows.

Multiplicative Cancellation in $\mathbb{Z}$. For all $a, b, c \in \mathbb{Z}$ we have

$$(ac = bc \wedge c \neq 0) \Rightarrow (a = b)$$

//

To mimic the proof of additive cancellation in $\mathbb{Z}$ we could consider some $d \in \mathbb{Z}$ with the property that $cd = 1$ and then multiply both sides of $ac = bd$ by d to obtain

$$\begin{aligned} ac &= bc \\ (ac)d &= (bc)d \\ a(cd) &= b(cd) \\ a \cdot 1 &= b \cdot 1 \\ a &= b. \end{aligned}$$

But this is a false proof because in general **there is no such integer** $d \in \mathbb{Z}$.[3]

What can we do? It turns out that it is **impossible** to prove multiplicative cancellation from just the axioms of addition (A1)–(A4) and the axioms of multiplication (M1)–(M3). These seven axioms define a structure called a "commutative ring" and it turns out that there are plenty of interesting commutative rings in which multiplicative cancellation **does not hold**.[4]

[3] We will prove this in the next section as a consequence of Division With Remainder.

[4] We will see some of these rings in the chapter on Modular Arithmetic

So it must be that multiplicative cancellation in $\mathbb{Z}$ has some mysterious connection to the order structure $(\mathbb{Z}, \leqslant)$ and the principle of induction.

Many authors choose to ignore this subtlety and they just take Multiplicative Cancellation as a friendly axiom. You can do this if you want but it's not necessary. The following sequence of exercises will guide you through a proof of multiplicative cancellation from first principles.

TO BE CONTINUED

2 Linear Diophantine Equations

As mentioned in the introduction, the central problem of number theory is to solve Diophantine equations. The general problem of Diophantine equations is in some sense impossible, but there is one case that we understand completely: the case of **linear Diophantine equations**. In this chapter I will present the complete solution of linear Diophantine equations and I will use this as motivation to introduce the basic definitions and algorithms of number theory.

2.1 Division With Remainder

The natural numbers $(\mathbb{N}, +, 0)$ are called a *commutative monoid* because they have a commutative and associative binary operation $+ : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ with an identity element $0 \in \mathbb{N}$. We can formally enlarge this to a commutative *group* $(\mathbb{Z}, +, 0)$ by adjoining "negative numbers" [see HW1], and this commutative group also carries a commutative monoid structure $(\mathbb{Z}, \times, 1)$ in which the multiplication operation $\times$ distributes over addition $+$. Putting all of this together gives us a *commutative ring* structure:

$$(\mathbb{Z}, +, \times, 0, 1).$$

Here the operation $+$ is invertible (we can subtract) but the operaiton $\times$ is not (we can not divide by an arbitrary integer). [For example, 2 is an integer, but $1/2$ is not.] This can be fixed by formally adjoining multiplicative inverses (called "fractions"), to obtain the system of rational numbers $(\mathbb{Q}, +, \times, 0, 1)$. But we don't **want** to do that in this course because it kills all the interesting properties of number theory.

Instead, we will investigate the subtle properties of "divisibility" for integers.

Definition of Divisibility. Consider two integers $a, b \in \mathbb{Z}$. We say that b *divides* a or that a *is divisible by* b if there exists an integer $q \in Z$ such that $a = qb$, and this case we will write "$b|a$". In symbols, we have

$$b|a \quad \iff \quad \exists q \in \mathbb{Z}, a = qb.$$

//

Observe that we have $1|a$ and $a|0$ for all $a \in \mathbb{Z}$. In other words,

1 *divides everything and everything divides* 0.

If b does not divide a we will write $b \nmid a$. But there is something much more specific we can say in this case. This is the first "theorem" of number theory on which everything else is based.

Theorem (Division With Remainder). Given integers $a, b \in \mathbb{Z}$ with $b \neq 0$, there exists a unique pair of integers $q, r \in \mathbb{Z}$ satisfying the following two simultaneous properties:

$$\begin{cases} a = qb + r \\ 0 \leqslant r < |b| \end{cases}$$

We say that q is the *quotient* and is the *remainder* of a modulo b. //

Proof: Consider $a, b \in \mathbb{Z}$ with $b \neq 0$. First we will show that the quotient and remainder *exist*. To do this we consider the set of integers of the form $a - nb$ for various $n \in \mathbb{Z}$:

$$S := \{a - nb : n \in \mathbb{Z}\}.$$

Since $b \neq 0$ this set must contain a non-negative integer. So let $S_{\geqslant 0}$ be the subset of S consisting of its non-negative elements. Since the set $S_{\geqslant}0$ is not empty, the Well-Ordering Principle (i.e., the Principle of Induction) says that it has a least element. Let us call this least element $r \in S_{\geqslant 0}$. Since $r \in S$ we have by definition that $r = a - qb$ for some $q \in \mathbb{Z}$. Thus we have obtained specific integers $q, r \in \mathbb{Z}$ with the property

$$a = qb + r.$$

Furthermore, since $r \in S_{\geqslant 0}$ we know that $0 \leqslant r$. It only remains to show that the remainder satisfies $r < |b|$. To prove this, let us **assume for contradiction** that $|b| \leqslant r$. Then we have

$$\begin{aligned} |b| &\leqslant r \\ |b| - |b| &\leqslant r - |b| \\ 0 &\leqslant r - |b|. \end{aligned}$$

Then since

$$\begin{aligned} r - |b| &= (a - qb) - |b| \\ &= a - qb - (\pm b) \\ &= a - (q \pm 1)b \quad \in S \end{aligned}$$

we conclude that $r - |b| \in S_{\geqslant 0}$. On the other hand, since $b \neq 0$ we have

$$\begin{aligned} 0 &< |b| \\ r &< r + |b| \\ r - |b| &< r, \end{aligned}$$

and thus we have found an element of $S_{\geqslant 0}$ that is smaller than r. This is the desired contradiction.//

Next we will show that the quotient and remainder of a mod b are *unique*. To do this, suppose that we have some integers $q_1, q_2, r_1, r_2 \in \mathbb{Z}$ satisfying the simultaneous properties

$$\begin{cases} a = q_1 b + r_1 \\ 0 \leqslant r_1 < |b| \end{cases} \qquad \text{and} \qquad \begin{cases} a = q_2 b + r_2 \\ 0 \leqslant r_2 < |b| \end{cases}$$

In this case I claim that we must have $q_1 = q_2$ and $r_1 = r_2$. To see this, first observe that the simultaneous equations

$$a = q_1 b + r_1 \qquad \text{and} \qquad a = q_2 b + r_2$$

imply that

$$\begin{aligned} q_1 b + r_1 &= q_2 b + r_2 \\ q_1 b - q_2 b &= r_2 - r_1 \\ (q_1 - q_2) b &= (r_2 - r_1). \end{aligned} \tag{2}$$

This equation is certainly true when $(q_1 - q_2) = 0 = (r_2 - r_1)$; we want to show that this is the **only** possible solution.

So let us **assume for contradiction that** $(r_2 - r_1) \neq 0$. Since $b \neq 0$ we conclude from equation (2) that $(q_1 - q_2) \neq 0$, and then since $(q_1 - q_2)$ is a whole number we conclude that $1 \leqslant |q_1 - q_2|$. Now we use equation (2) and the multiplicative property of the absolute value to obtain

$$\begin{aligned} 1 &\leqslant |q_1 - q_2| \\ |b| &\leqslant |q_1 - q_2||b| \\ |b| &\leqslant |(q_1 - q_2)b| \\ |b| &\leqslant |r_2 - r_1|. \end{aligned} \tag{3}$$

Now I claim that this inequality contradicts the assumptions

$$0 \leqslant r_1 < |b| \qquad \text{and} \qquad 0 \leqslant r_2 < |b|.$$

There are two cases to deal with: since $(r_1 - r_2) \neq 0$ we must have either $r_1 < r_2$ or $r_2 < r_1$. For the purpose of this proof we will assume that $r_1 < r_2$ (the proof of the other case is similar). In this case we have $0 < (r_2 - r_1) = |r_2 - r_1|$ so that the inequality (3) becomes $|b| \leqslant (r_2 - r_1)$. Then the assumption $0 \leqslant r_1$ gives

$$\begin{aligned} 0 &\leqslant r_1 \\ -r_1 &\leqslant 0 \\ |b| - r_1 &\leqslant |b| \end{aligned}$$

and the assumption $r_2 < |b|$ gives

$$\begin{aligned} r_2 &< |b| \\ r_2 - r_1 &< |b| - r_1. \end{aligned}$$

Putting these together gives $(r_2 - r_1) < (|b| - r_1) \leqslant |b|$, which **contradicts** (3).

We have shown that $(r_2 - r_1) = 0$. Finally, from equation (2) we have

$$(q_1 - q_2)b = (r_2 - r_1) = 0$$

and then since $b \neq 0$ we conclude that $(q_1 - q_2) = 0$ as desired. □

Remarks:

- I skipped some steps in there, mostly involving the absolute value function. It is formally defined by
$$|a| := \begin{cases} a & \text{if } 0 \leqslant a \\ -a & \text{if } a < 0 \end{cases}$$
and then one can prove that $|ab| = |a||b|$ for all $a, b \in \mathbb{Z}$. It is not entirely trivial to prove this from the axioms but I'm going to skip those details. [You can find them in my old Course Notes for MTH 230.]
- In the past I have seen many students write "$b|a = q$" when $a = qb$. This is wrong. The symbol "$b|a$" on the left is **not** a number; it is a logical statement meaning that **there exists an integer $q \in \mathbb{Z}$ with the property $a = qb$.** I advise you to avoid the use of fractional notation when proving theorems about $\mathbb{Z}$ since it can cause confusion.

I claimed above that is it generally not possible to divide by an integer (that is, without introducing the formal concept of "fractions"). Now we can prove it.

Example. I claim that $2 \nmid 1$. In other words, there does not exist an integer $n \in \mathbb{Z}$ with the property $2n = 1$. In other other words, there does not exist an integer that deserves to be called "1/2".

Proof. Suppose for contradiction that such an integer does exist. This would mean that the quotient of 1 mod 2 is n and the remainder is zero:

$$\begin{cases} 1 = n \cdot 2 + 0 \\ 0 \leqslant 0 < |2| \end{cases}$$

On the other hand, the following two properties are also true:

$$\begin{cases} 1 = 0 \cdot 2 + 1 \\ 0 \leqslant 1 < |2| \end{cases}$$

But this says that the quotient of 1 mod 2 is zero and the remainder is 1. Since $0 \neq 1$ this contradicts the uniqueness of remainders which we proved above. □

Now it is time to start solving equations. Here is an easy one.

Problem (Linear Diophantine Equation in One Unknown). Given integers $a, b \in \mathbb{Z}$, find all integers $x \in \mathbb{Z}$ satisfying

$$ax = b.$$

//

Solution. This equation has a solution if and only if a divides b, i.e., if and only if the remainder of b mod a is zero. In this case there is a unique solution $x \in \mathbb{Z}$, which is the quotient of b mod a. //

2.2 Greatest Common Divisor

Okay, now here's a harder one.

Problem (Linear Diophantine Equation in Two Unknowns). Given integers $a, b, c \in \mathbb{Z}$, find all integers $x, y \in \mathbb{Z}$ satisfying

(LDE) $$ax + by = c.$$

//

We will work up to the full solution of this problem but it will take some time to get there. There are a few separate issues involved in the solution:

- Determine whether a solution exists.
- Find one particular solution $x, y \in \mathbb{Z}$.
- Classify all possible solutions $x, y \in \mathbb{Z}$.

We'll first deal with the **non-existence** of solutions since this is easiest. Suppose that $d \in \mathbb{Z}$ is a *common divisor* of a and b. That is, suppose that there exist integers $a', b' \in \mathbb{Z}$ such that $a = da'$ and $b = db'$. Now suppose that the equation (LDE) has a solution, i.e., assume that there exist integers $x, y \in \mathbb{Z}$ such that

$$ax + by = c.$$

Then we must have

$$\begin{aligned} c &= ax + by \\ &= (da')x + (db')y \\ &= d(a'x) + d(b'y) \\ &= d(a'x + b'y), \end{aligned}$$

which implies that d also divides c.

Conclusion. If a and b have a common divisor that does not divide c, then (LDE) has no solution. For example, if a and b are both **even** (i.e., if they have the common divisor 2) and if c is **odd** (i.e., if it is not divisible by 2) then there is no solution.

This suggests that we should investigate the common divisors of a and b in more detail. For this purpose we will denote the set of all common divisors by

$$\mathrm{Div}(a,b) := \{d \in \mathbb{Z} : d|a \wedge d|b\}.$$

If $a = b = 0$ then we have $\mathrm{Div}(a,b) = \mathbb{Z}$ (every integer divides zero) which is not very interesting. So let's assume that a and b are not both zero.

Theorem/Definition. Given two integers $a, b \in \mathbb{Z}$ with a, b not both zero, the set $\mathrm{Div}(a,b)$ of common divisors is non-empty and bounded above. Thus, by the Well-Ordering Principle it must have a greatest element. We call this element the *greatest common divisor* of a and b, and we denote it by $\gcd(a,b) \in \mathrm{Div}(a,b)$. //

Proof. Without loss of generality, let's assume that $a \neq 0$. Then I claim that each common divisor $d \in \mathrm{Div}(a,b)$ satisfies $d \leqslant |a|$. So consider any $d \in \mathrm{Div}(a,b)$. Since $d|a$ and $a \neq 0$ we must have $d \neq 0$. Since $d|a$ we also have $a = da'$ for some $a' \in \mathbb{Z}$ and since a and d are both nonzero we must have $a' \neq 0$. Then since a' is a nonzero integer we must have

$$\begin{aligned} 1 &\leqslant |a'| \\ |d| &\leqslant |d||a'| \\ |d| &\leqslant |da'| \\ |d| &\leqslant |a|, \end{aligned}$$

which implies that $d \leqslant |d| \leqslant |a|$ as desired. We conclude that the set $\mathrm{Div}(a,b)$ is bounded above by $|a|$. We also know that $\mathrm{Div}(a,b)$ is non-empty because $1 \in \mathrm{Div}(a,b)$ (1 divides everything).

The usual statement of the Well-Ordering Principle says that every nonempty set of integers that is bounded **below** has a **least** element. By multiplying everything by -1 one can show that this is equivalent to the statement that every non-empty set of integers that is bounded **above** has a **greatest** element. Thus the greatest common divisor exists. □

This allows us to be more precise about the solvability of (LDE).

Theorem (Reduction of LDE). Consider integers $a, b, c \in \mathbb{Z}$ with a, b not both zero and let $d = \gcd(a,b)$. If $d \nmid c$ then the equation $ax + by = c$ has **no integer solution** $x, y \in \mathbb{Z}$. On the other hand, if $d|c$ then we have integers $a', b', c' \in \mathbb{Z}$ such that $a = da'$, $b = db'$, and $c = dc'$. In this case I claim that the integer solutions of $ax + by = c$ coincide with the solutions of the *reduced equation*:

$$a'x + b'y = c'.$$

//

Proof. We already proved the first statement. To prove the second statement we will denote the set of solutions of of the equation by

$$V_{a,b,c} := \{(x, y) \in \mathbb{Z}^2 : ax + by = c\}.$$

We want to prove that $V_{a,b,c} = V_{a',b',c'}$. To show that $V_{a',b',c'} \subseteq V_{a,b,c}$ consider any solution $(x, y) \in V_{a',b',c'}$, i.e., consider any ordered pair of integers $(x, y) \in \mathbb{Z}^2$ such that $a'x + b'y = c'$. Now multiply both sider of this equation by d to obtain

$$\begin{aligned} a'x + b'y &= c' \\ d(a'x + b'y) &= dc' \\ (da')x + (db')y &= dc' \\ ax + by &= c. \end{aligned}$$

We conclude that $(x, y) \in V_{a,b,c}$ and hence $V_{a',b',c'} \subseteq V_{a,b,c}$. To show that $V_{a,b,c} \subseteq V_{a',b',c'}$ consider any solution $(x, y) \in V_{a',b',c'}$, i.e., any ordered pair of integers $(x, y) \in \mathbb{Z}^2$ such that $ax + by = c$. Then we must have

$$\begin{aligned} ax + by &= c \\ (da')x + (db')y &= (dc') \\ d(a'x + b'y) &= dc'. \end{aligned}$$

Then since $d \neq 0$ we can multiplicatively cancel d from both sides to obtain $a'x + b'y = c'$. We conclude that $(x, y) \in V_{a',b',c'}$ and hence $V_{a,b,c} \subseteq V_{a',b',c'}$. □.

The process of dividing out by the greatest common divisor is called *reduction*. It is also convenient to talk about reduction in the language of *coprimality*. Given integers $a, b \in \mathbb{Z}$ recall that we have $1 \in \mathrm{Div}(a, b)$ because the integer 1 divides every other integer. This tells us that the **greatest** element of $\mathrm{Div}(a, b)$ must satisfy $1 \leqslant \gcd(a, b)$ by definition.

Definition of Coprimality. Given two integers $a, b \in \mathbb{Z}$, with a, b not both zero, we have seen that there exists a greatest common divisor $d = \gcd(a, b)$ and that this greatest common divisor satisfies

$$1 \leqslant \gcd(a, b) \leqslant \min\{|a|, |b|\}.$$

In the extreme case that $\gcd(a, b) = 1$ we say that the integers a and b are *coprime*.

The utility of this concept is that **any** pair of integers $a, b \in \mathbb{Z}$ (not both zero) can be reduced to a coprime pair of integers as follows. Let $d = \gcd(a, b)$ such that $a = da'$ and $b = db'$. In this case I claim that $\gcd(a', b') = 1$ and hence the pair $a', b' \in \mathbb{Z}$ is coprime.

Proof. Let $d' \in \mathrm{Div}(a', b')$ be any common divisor of a' and b', so that we have $a' = d'a''$ and $b' = d'b''$ for some integers $a'', b'' \in \mathbb{Z}$. Then we can substitute $a = d'a''$ into the equation $a = da'$ to obtain

$$a = da' = d(d'a'') = (dd')a'',$$

which implies that dd' divides a. Similarly we find that dd' divides b and hence that $dd' \in \mathrm{Div}(a, b)$. But d is by definition the **greatest** element of $\mathrm{Div}(a, b)$ so we must have $dd' \leqslant d$. Now I claim that $d' \leqslant 1$. Indeed, if $d' > 1$ then multiplying both sides by d yields the contradiction $dd' > d$.

We have shown that every element $d' \in \mathrm{Div}(a', b')$ satisfies $d' \leqslant 1$, which implies that $1 \in \mathrm{Div}(a', b')$ is the **greatest** element of this set. In other words, $\gcd(a', b') = 1$. □

In summary, we can restate the problem of linear Diophantine equations as follows.

Problem' (Linear Diophantine Equations in Two Unknowns). Given integers $a, b, c \in \mathbb{Z}$ with $\gcd(a, b) = 1$, find all integers $x, y \in \mathbb{Z}$ satisfying

$$ax + by = c.$$

Indeed, if $\gcd(a, b) = 1$ then we automatically have $\gcd(a, b)|c$. If $1 \neq d = \gcd(a, b)$ then we can divide both sides of the equation by d to obtain the reduced equation $a'x + b'y = c'$ which has the same solution. Note that the reduced equation satisfies $\gcd(a', b') = 1$. We will assume from now on that all linear Diophantine equations are reduced in this way.

2.3 A Bit of Linear Algebra

Suppose for the moment that we are able to *compute* the greatest common divisor in an efficient way.[5] Then we can restrict our attention to linear Diophantine equations

(LDE) $$ax + by = c$$

in which a and b are **coprime integers**, i.e., in which $\gcd(a, b) = 1$. In this section we will reduce the problem even further. The ideas will be familir to you if you have already taken linear algebra.

The case of (LDE) in which $c = 0$ is called a *homogeneous linear Diophantine equation* and it turns out that this case is much easier to solve. Furthermore, it turns out that solving homogeneous equations is almost enough to solve the full problem.

Theorem (Reduction to the Homogeneous Case). Consider any integers $a, b, c \in \mathbb{Z}$. (For this theorem is doesn't matter if a and b are coprime.) Now consider the solution sets to

[5] We will give an algorithm for this in the next section.

the equation $ax + by = c$ and its homogeneous version $ax + by = 0$:

$$V_0 := \{(x, y) \in \mathbb{Z}^2 : ax + by = 0\},$$
$$V_c := \{(x, y) \in \mathbb{Z}^2 : ax + by = c\}.$$

I claim that these sets are "almost the same" in the following sense: if $(x', y') \in V_c$ is any **one specific solution** then the full solution is given by

$$V_c = V_0 + (x', y') := \{(x + x', y + y') : ax + by = 0\}.$$

//

In other words, the complete solution of the non-homogeneous equation coincides with the complete solution of the homogeneous equation after translation by one particular solution.

Proof. To prove that $V_0 + (x', y') \subseteq V_c$ consider any element $(x + x', y + y') \in V_0 + (x', y')$. Then we have

$$\begin{aligned} a(x + x') + b(y + y') &= (ax + by) + (ax' + by') \\ &= 0 + c \\ &= c, \end{aligned}$$

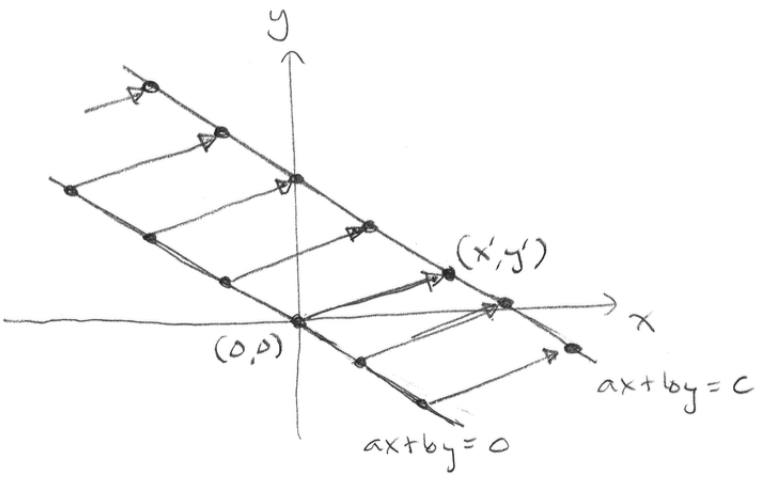
and hence $(x' + x, y' + y) \in V_c$ as desired. Conversely, consider any element $(u, v) \in V_c$ and define the vector $(x, y) := (u, v) - (x', y') = (u - x', v - y')$. Then we must have

$$\begin{aligned} ax + by &= a(u - x') + b(v - y') \\ &= (au + bv) - (ax' + by') \\ &= c - c \\ &= 0, \end{aligned}$$

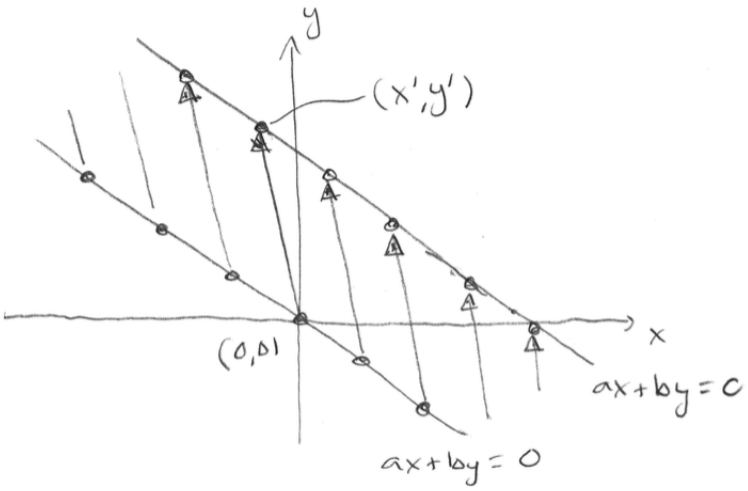
and it follows that $(x, y) \in V_0$. But then we have $(u, v) = (x, y) + (x', y') \in V_0 + (x', y')$ as desired. □

We can visualize the situation as follows. If we temporarily allow x and y to be *real* numbers then the equation $ax + by = c$ defines a line in the plane $\mathbb{R}^2$. The integer solutions (x, y) can be thought of as the "integer points" on this line; there may be none or there may be infinitely many integer points. The associated homogeneous equation $ax + by = 0$ defines a **parallel line** passing through the origin $(0, 0)$; thus it always has an integer point. Since a and b are integers we will shortly see that the line $ax + by = 0$ has infinitely many integers points.

The above theorem says that if we can find **just one integer point** (x', y') on the line $ax + by = c$ then we will obtain a one-to-one correspondence between the integer points on $ax + by = 0$ and the integer points on $ax + by = c$ as in the following picture:



This one-to-one correspondence is of course not unique. If we chose a different integer point (x', y') on $ax + by = c$ then we would obtain a different picture:



For this reason, there is no one correct way to express the solution of a linear Diophantine equation. Two people with the correct solution might have answers that look slightly different. That won't bother us in this class. If you're implementing the problem on a computer then you might want to choose a standard format for the output; there are several available, such as the *Hermite normal form*.

Thus we have broken down the problem into two steps:

- Find one particular solution $ax' + by' = c$.

- Find the general homogeneous solution $ax + by = 0$.

We will deal with these two problems in the next two sections.

2.4 The Euclidean Algorithm

First we will deal with the problem of actually *computing* the greatest common divisor of two numbers, for the purpose of reducing the Diophantine equation. As a side effect of the computation we will obtain an efficient method to compute a single solution to the equation $ax + by = c$ when $\gcd(a, b)$ divides c.

For relatively small numbers we can simply compute the set of common divisors by hand and then select the greatest element of this set. For example, the set of common divisors of -18 and 30 is

$$\mathrm{Div}(-18, 30) = \{-6, -3, -2, -1, 1, 2, 3, 6\},$$

from which we conclude that $\gcd(-18, 30) = 6$. In this section I will present a beautiful method, called the *Euclidean Algorithm*, that can compute the solution in logarithmic time. To be precise, if $0 \leqslant |a| < |b|$ then the Euclidean Algorithm will compute $\gcd(a, b)$ in less than $2 \cdot \log_2 |b|$ steps.

First I'll show you an example of the algorithm and then I'll prove why it works. To compute $\gcd(3094, 2513)$ we first note that $3094 > 2513$ and then we find the quotient and remainder of 3094 mod 2513:

$$3094 = 1 \cdot 2513 + 581.$$

Then we replace the number 3094 by the remainder 581 to obtain the new pair of numbers $2513 > 581$. Now we compute the quotient and remainder of 2513 mod 581:

$$2513 = 4 \cdot 581 + 189.$$

This results in the new pair of numbers $581 > 189$. Now we repeat the process until a remainder of 0 is reached:

$$\begin{aligned} 3094 &= 1 \cdot 2513 + 581 \\ 2513 &= 4 \cdot 581 + 189 \\ 581 &= 3 \cdot 189 + 14 \\ 189 &= 13 \cdot 14 + 7 \\ 14 &= 2 \cdot 7 + 0. \end{aligned}$$

I claim that the last nonzero remainder in this sequence is the greatest common divisor:

$$\gcd(3094, 2513) = 7.$$

Now let me justify the claim. It all depends on the following lemma.

Lemma. Consider any integers $a, b, q, r \in \mathbb{Z}$ such that $a = qb + r$. (This q, r need not be the quotient and remainder of a mod b.) Then we have

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

//

Proof. We will show more generally that the sets of common divisors are equal:

$$\mathrm{Div}(a, b) = \mathrm{Div}(b, r).$$

Then since the greatest common divisors are the greatests elements of these sets, they must also be equal. To show that $\mathrm{Div}(a, b) \subseteq \mathrm{Div}(b, r)$, consider any common divisor $d \in \mathrm{Div}(a, b)$. By definition there exist integers $a', b' \in \mathbb{Z}$ such that $a = da'$ and $b = db'$. Then we must have

$$\begin{aligned} r &= a - qb \\ &= (da') - q(db') \\ &= d(a' - qb), \end{aligned}$$

which implies that d divides r, and it follows that $d \in \mathrm{Div}(b, r)$. Conversely, to show that $\mathrm{Div}(b, r) \subseteq \mathrm{Div}(a, b)$ consider any common divisor $d \in \mathrm{Div}(b, r)$. By definition there exist integers $b', r' \in \mathbb{Z}$ such that $b = db'$ and $r = dr'$. Then we must have

$$\begin{aligned} a &= qb + r \\ &= q(db') + (dr') \\ &= d(qb' + r'), \end{aligned}$$

which implies that d divides a, and it follows that $d \in \mathrm{Div}(a, b)$ as desired. □

Theorem (The Euclidean Algorithm). Consider any integers $a, b \in \mathbb{Z}$ with $b \neq 0$. To compute $\gcd(a, b)$ we first apply Division With Remainder to obtain

$$\begin{cases} a = q_1 \cdot b + r_1 \\ 0 \leqslant r_1 < |b| \end{cases}$$

If $r_1 \neq 0$ then we continue to compute the quotient and remainder of b mod r_1:

$$\begin{cases} a = q_2 \cdot r_1 + r_2 \\ 0 \leqslant r_2 < r_1 \end{cases}$$

And if $r_2 \neq 0$ we compute the quotient and remainder of r_1 mod r_2:

$$\begin{cases} a = q_3 \cdot r_2 + r_3 \\ 0 \leqslant r_3 < r_2 \end{cases}$$

Thus we obtain a strictly descending sequence of non-negative integers:

$$0 \neq |b| =: r_0 > r_1 > r_2 > \cdots \geqslant 0.$$

I claim that the sequence must terminate. That is, I claim that there exists an integer $n \geqslant 1$ such that $r_n = 0$ and $r_{n-1} \neq 0$. Furthermore, I claim in this case that r_{n-1} is the greatest common divisor of a and b:

$$\gcd(a, b) = r_{n-1}.$$

//

Proof. Assume for contradiction that we have $r_n \neq 0$ for all $n \geqslant 1$. Then we obtain an infinite strictly decreasing sequence of positive integers:

$$0 \neq |b| =: r_0 > r_1 > r_2 > \cdots > 0.$$

Now consider the set $S := \{r_0, r_1, r_2, \ldots, \}$. This set is non-empty because $|b| \in S$ and it is bounded below by 0. Thus the Well-Ordering Principle says that S must have a least element of the form $r_\ell \in S$. But this is impossible because $r_{\ell+1} < r_\ell$ is also an element of S. We conclude that there exists $n \geqslant 1$ with $r_n = 0$ and by another application of Well-Ordering we can assume that $r_{n-1} \neq 0$.

To prove that this numbe r_{n-1} is the greatest common divisor of a and b we repeatedly use the previous Lemma to obtain

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

This last gcd exists because $r_{n-1} \neq 0$. Furthermore, the common divisors of r_{n-1} and 0 are just the divisors of r_{n-1} because everything divides zero. Since r_{n-1} is positive we conclude that $\gcd(r_{n-1}, 0) = r_{n-1}$ and hence

$$\gcd(a, b) = \gcd(r_{n-1}, 0) = r_{n-1}.$$

□

Remarks:

- I'll ask you to compute the complexity of the Euclidean Algorithm on HW2.
- The algorithm can be implemented very simply without even mentioning the words "quotient" and "remainder". Given two non-negative integers $a, b \in \mathbb{N}$ not both zero perform the following steps:

  `while` $a \neq b$ `do`

```
        if a > b then replace a by a − b
        else replace b by b − a
    return a
```

I claim that this is just the Euclidean Algorithm in disguise. [Why?]

Thus the Euclidean Algorithm is an efficient way to compute the greatest common divisor of two integers. However, I claim that the same algorithm can also be used to compute solutions to linear Diophantine equations. Before describing the general method I'll illustrate the ideas behind it by considering the equation

$$3094x + 2513y = 21.$$

We saw above that $\gcd(3094, 2513) = 7$. Then since $7|21$ we know that this equation might possibly have an integer solution $(x, y) \in \mathbb{Z}^2$. To find such a solution the trick is to broaden our scope and consider the following homogeneous Diophantine equation in **three unknowns** x, y, z:

$$3094x + 2513y = z.$$

The reason we do this is because there are **two obvious solutions** to this equation:

$$3094(1) + 2513(0) = (3094)$$
$$3094(0) + 2513(1) = (2513).$$

And once we have two solutions we can combine them in various ways to get infinitely many solutions. We will borrow a principle from linear algebra.

The Principle of Linear Combination. Fix two integers $a, b \in \mathbb{Z}$ and suppose that the vectors $(x', y', z') \in \mathbb{Z}^3$ and $(x'', y'', z'') \in \mathbb{Z}^3$ are two solutions of the equation

$$ax + by = z.$$

Then for any integers $u, v \in \mathbb{Z}$ I claim that the *linear combination vector*

$$u(x', y', z') + v(x'', y'', z'') = (ux' + vx'', uy' + vy'', uz' + vz'') \in \mathbb{Z}^3$$

is another solution. In other words, the **set** of solution vectors

$$V = \{(x, y, z) \in \mathbb{Z}^3 : ax + by = z\}$$

is closed under *vector addition* and *scalar multiplication* by integers. //

Proof. Suppose that (x', y', z') and (x'', y'', z'') are in V and consider any two integers $u, v \in \mathbb{Z}$. Then we have

$$a(ux' + vx'') + b(uy' + vy'') = u(ax' + by') + v(ax'' + by'') = uz' + vz'',$$

and it follows that $u(x', y', z') + v(x'', y'', z'')$ is also in V. $\square$

Let's apply this idea to our problem. If V is the set of solutions $(x, y, z) \in \mathbb{Z}^3$ to the equation $3094x + 2513y = z$ then of course we must have the trivial solution $(0, 0, 0) \in V$. But we also saw above that there two "obvious but non-trivial solutions":

$$(1, 0, 3094) \in V \quad \text{and} \quad (0, 1, 2513) \in V.$$

Now we can apply the Principle of Linear Combination to create as many new solutions as we want. In the end we are looking for a solution of the form $(x, y, 7)$; is there some sequence of linear combinations that will achieve this? Certainly. We can just apply the steps of the Euclidean Algorithm to the third coordinates and let the first two coordinates come along for the ride. To keep track of the steps I will use the vector notation

$$\mathbf{x}_1 := (1, 0, 3094) \quad \text{and} \quad \mathbf{x}_2 := (0, 1, 2513).$$

Then the first step of the Euclidean Algorithm says that we should divide 3094 by 2513 to obtain $3094 = 1 \cdot 2513 + 581$. In terms of vectors we compute

$$\begin{aligned} \mathbf{x}_3 &:= \mathbf{x}_1 - 1 \cdot \mathbf{x}_2 \\ &= (1, 0, 3094) - 1 \cdot (0, 1, 2513) \\ &= (1, -1, 581). \end{aligned}$$

Observe that the resulting vector is another solution of the equation because

$$3094(1) + 2513(-1) = (581).$$

So far this is not very interesting, but then we continue the process:

$$\begin{array}{lcll} \mathbf{x}_1 & & = & (1, \quad 0, \quad 3094) \\ \mathbf{x}_2 & & = & (0, \quad 1, \quad 2513) \\ \mathbf{x}_3 & := \quad \mathbf{x}_1 - 1 \cdot \mathbf{x}_2 & = & (1, \quad -1, \quad 581) \\ \mathbf{x}_4 & := \quad \mathbf{x}_2 - 4 \cdot \mathbf{x}_3 & = & (-4, \quad 5, \quad 189) \\ \mathbf{x}_5 & := \quad \mathbf{x}_3 - 3 \cdot \mathbf{x}_4 & = & (13, \quad -16, \quad 14) \\ \mathbf{x}_6 & := \quad \mathbf{x}_4 - 13 \cdot \mathbf{x}_5 & = & (-173, \quad 213, \quad 7) \\ \mathbf{x}_7 & := \quad \mathbf{x}_5 - 2 \cdot \mathbf{x}_6 & = & (359, \quad -442, \quad 0) \end{array}$$

At each step the Principle of Linear Combination guarantees that $\mathbf{x}_n$ is a new solution of the original equation. In the second-to-last step, the Euclidean Algorithm guarantees that the third coordinate is the gcd, and we obtain the (non-trivial!) equation:

$$3094(-173) + 2513(213) = 7 = \gcd(3094, 2513).$$

Finally, we can return to our motivating equation

$$3094x + 2513y = 21.$$

Since $7|21$ we believed that there **might** be a solution, and now we can find one easily. Since $21 = 3 \cdot 7$ we just "scalar multiply" the solution $\mathbf{x}_6 = (-173, 213, 7)$ by 3 to obtain

$$3094(-519) + 2513(639) = (21).$$

This example illustrates that the following general method is correct.

The Vector Euclidean Algorithm. Consider any integers $a, b \in \mathbb{Z}$ with $b > 0$. As in the usual Euclidean Algorithm we define a sequence of quotients and remainders $(q_i, r_i) \in \mathbb{Z}^2$ by repeated division as follows:

$$\begin{aligned} a &= q_1 \cdot b + r_1 \\ b &= q_2 \cdot r_1 + r_2 \\ r_1 &= q_3 \cdot r_2 + r_3 \\ &\vdots \\ r_{n-3} &= q_{n-1} \cdot r_{n-2} + r_{n-1} \\ r_{n-2} &= q_n \cdot r_{n-1} + 0. \end{aligned}$$

Recall that $r_{n-1} = \gcd(a, b)$. Now consider the two obvious solutions $\mathbf{x}_1 := (1, 0, a)$ and $\mathbf{x}_2 = (0, 1, b)$ of the equation $ax + by = z$. If we recursively define the vectors

$$\mathbf{x}_{i+2} := \mathbf{x}_{i+1} - q_i \cdot \mathbf{x}_i$$

then the vector $\mathbf{x}_n$ has the form $(x', y', \gcd(a, b))$ for some integer $x', y' \in \mathbb{Z}$ such that

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

//

Here is a summary of our progress in this section.

- Consider integers $a, b, c \in \mathbb{Z}$ with a, b not both zero. We have shown that the linear Diophantine equation

  (LDE) $$ax + by = c$$

  has a solution $(x, y) \in \mathbb{Z}^2$ **if and only if** $\gcd(a, b)$ divides c.

- More specifically, if $c = n \cdot \gcd(a, b)$ for some $n \in \mathbb{Z}$ then we can use the Vector Euclidean Algorithm to obtain specific integers $x', y' \in \mathbb{Z}$ such that

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

  and then we can multiply both sides by n to obtain a specific solution to (LDE):

  $$a(nx') + b(ny') = c.$$

To complete the solution of (LDE) it remains to find the complete solution of the associated homogeneous equation: $ax + by = 0$. The answer is easy to guess but a bit tricky to prove. We will do this in the next section.

2.5 Euclid's Lemma

Consider two integers $a, b \in \mathbb{Z}$ not both zero. Our goal in this section is to find the complete solution of the **homogeneous** linear Diophantine equation

(HLDE) $$ax + by = 0$$

If $d = \gcd(a, b) \neq 1$ with $a = da'$ and $b = db'$ then recall from section 2.2 that we can "multplicatively cancel" d from both sides of (HLDE) to obtain a new equation

$$a'x + b'y = 0$$

which has the same solutions and where $\gcd(a', b') = 1$. Thus it is sufficient to solve (HLDE) in the case that a and b are coprime. The complete solution is given by the following theorem.

Theorem (Homogeneous Linear Diophantine Equations). Consider two integers $a, b \in \mathbb{Z}$ with $\gcd(a, b) = 1$. Then the complete solution of the Diophantine equation

$$ax + by = 0$$

is given by $(x, y) = k(b, -a) := (kb, -ka)$ for all $k \in \mathbb{Z}$. //

It is easy to verify that every pair of the form $(x, y) = (kb, -ka)$ is a solution. Indeed, we have

$$\begin{aligned} ax + by &= a(kb) + b(-ka) \\ &= k(ab) + k(-ab) \\ &= k(ab - ab) \\ &= k \cdot 0 \\ &= 0. \end{aligned}$$

But proving that **every** solution has this form is a bit harder. To do this we will need a lemma whose proof depends on the Euclidean Algorithm. This lemma is important enough to deserve a special name.

Euclid's Lemma. Consider integers $a, b, c \in \mathbb{Z}$ with $\gcd(a, b) = 1$. Then we have

$$a|(bc) \quad \Rightarrow \quad a|c.$$

Proof of the Lemma. Assume that $a|(bc)$ so there exists an integer k with the property $ak = (bc)$. Since $\gcd(a, b) = 1$ it follows from the Vector Euclidean Algorithm that there exists a pair of integers $x', y' \in \mathbb{Z}$ with the property

$$1 = ax' + by'.$$

By multiplying both sides of this equation by c we obtain

$$\begin{aligned} c &= c(ax' + by') \\ &= cax' + (bc)y' \\ &= cax' + (ak)y' \\ &= a(cx' + ky'), \end{aligned}$$

and it follows that $a|c$ as desired. □

[Remark: That was a very good trick; never forget it.]

Proof of the Theorem. Assume that $\gcd(a, b) = 1$. We have already seen that all vectors of the form $(x, y) = k(a, -b) = (ka, -kb)$ with $k \in \mathbb{Z}$ are solutions of the equation $ax + by = 0$.

Conversely, let $(x, y) \in \mathbb{Z}^2$ be any vector satisfying $ax + by = 0$. In this case we want to prove that $(x, y) = k(b, -a) = (kb, -ka)$ for some $k \in \mathbb{Z}$. If $a = 0$ or $b = 0$ then one can check that the solution has the correct form. [Maybe you should check this.] **So let us assume that a and b are both nonzero**. Now we will rewrite the equation in two ways:

$$\begin{aligned} ax + by &= 0 \\ ax &= b(-y) \\ a(-x) &= by. \end{aligned}$$

The equation $ax = b(-y)$ says that $a|b(-y)$ and then since $\gcd(a, b) = 1$ Euclid's Lemma says that a divides $-y$. In other words, there exists an integer k with the property

$$\begin{aligned} -y &= ka \\ y &= -ka. \end{aligned}$$

Similarly, the equation $a(-x) = by$ says that $b|a(-x)$ and then Euclid's Lemma implies that $b|(-x)$ so there exists an integer $\ell \in \mathbb{Z}$ with the property

$$\begin{aligned} -x &= \ell b \\ x &= -\ell b. \end{aligned}$$

Now we substitute these expressions for x and y into the original equation:

$$\begin{aligned} ax + by &= 0 \\ a(-\ell b) + b(-ka) &= 0 \\ ab(-\ell - k) &= 0. \end{aligned}$$

Since we have assumed that a and b are both nonzero we must have $ab \neq 0$ and then we can multiplicatively cancel ab to obtain

$$\begin{aligned} -\ell - k &= 0 \\ -\ell &= k. \end{aligned}$$

It follows that $(x, y) = (-\ell b, -ka) = (kb, -ka) = k(b, -a)$ as desired. $\square$

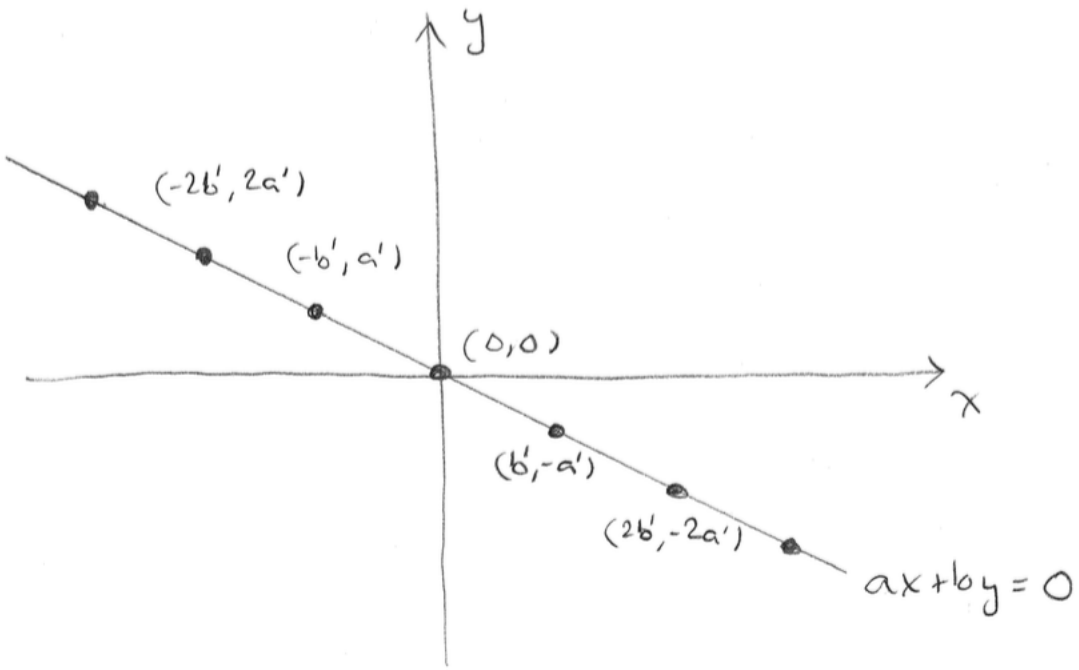
In summary, consider two integers $a, b \in \mathbb{Z}$, not both zero, and let $d = \gcd(a, b)$ with $a = da'$ and $b = db'$. We have shown that the homogeneous Diophantine equation

$$ax + by = 0$$

has the complete solution

$$V = \{(x, y) \in \mathbb{Z}^2 : ax + by = 0\} = \{(kb', -ka') : k \in \mathbb{Z}\}.$$

Geometrically we can think of V as the family of integer points on the line $ax + by = 0$ in the Cartesian plane. This line contains the integer point $(0, 0)$ because the equation is homogeneous. Then from the above result we see that the rest of the integer points are equally spaced with distance $\sqrt{(a')^2 + (b')^2}$ between them:



2.6 Summary and Discussion

We have now completely solved the linear Diophantine equation

(LDE) $$ax + by = c.$$

Here is a point-form summary of our results:

- If $a = b = 0$ and $c \neq 0$ then there is no solution. If $a = b = c = 0$ then every point $(x, y) \in \mathbb{Z}^2$ is a solution.
- If a, b are not both zero then there exists a greatest common divisor $d := \gcd(a, b)$ with $a = da'$ and $b = db'$ for some unique integers $a', b' \in \mathbb{Z}$. If $d \nmid c$ then there is no solution.
- If $d|c$, i.e., if there exists an integer $c' \in \mathbb{Z}$ with $c = dc'$ then the equation (LDE) is equivalent to the "reduced" Diophantine equation

  (RLDE) $$a'x + b'y = c',$$

  where now we have $\gcd(a', b') = 1$.
- By applying the Vector Euclidean Algorithm we can find a **specific** pair of integers $x', y' \in \mathbb{Z}$ such that
  $$a'x' + b'y' = \gcd(a', b') = 1,$$
  and then multiplying both sides by c' gives us a **specific** solution to (RLDE):
  $$a'(c'x') + b'(c'y') = c'.$$
- On the other hand, we can use Eucild's Lemma to prove that the associated **homogeneous** Diophantine equation

  (HLDE) $$a'x + b'y = 0$$

  has the **complete** solution given by
  $$V_0 := \{(x, y) \in \mathbb{Z} : a'x + b'y = 0\} = \{(kb', -ka') : k \in \mathbb{Z}\}.$$
- Finally, let $V := \{(x, y) \in \mathbb{Z}^2 : ax + by = c\}$ denote the complete solution to the original equation (LDE), which is the same as the solution to (RLDE). By combining the complete solution V_0 to (HLDE) with the specific solution $(c'x', c'y')$ to (LDE), a Bit of Linear Algebra shows us that
  $$\begin{aligned} V &= V_0 + (c'x', c'y') \\ &= \{(kb', -ka') : k \in \mathbb{Z}\} + (c'x', c'y') \\ &= \{(c'x' + kb', c'y' - ka') : k \in \mathbb{Z}\}. \end{aligned}$$

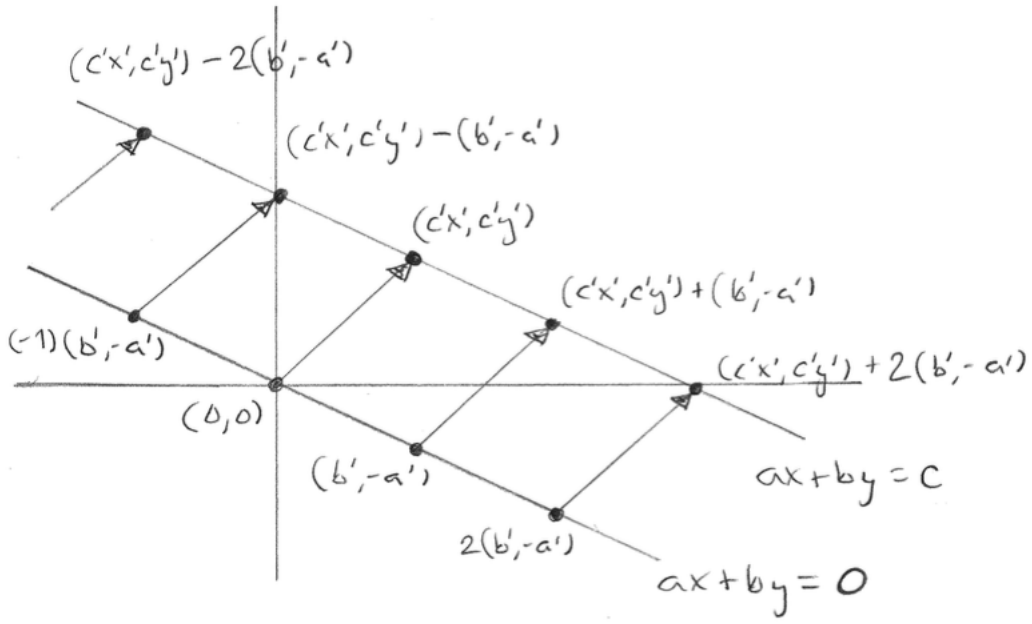
//

Geometrically, we can think of (LDE) as the equation of a general line in the real Cartesian plane $\mathbb{R}^2$. If $a, b, c \in \mathbb{Z}$ then the Diophantine problem is to find all of the "integer points" on this line. If $\gcd(a, b) \nmid c$ then we find that the line contains **no integer points** (amazing as that may be), and when $\gcd(a, b)|c$ we find that the line contains **infinitely many equally spaced integer points**. The displacement between any two consecutive points is the vector

$(b', -a')$, which has length $\sqrt{(a')^2 + (b')^2}$. If we can find just one integer point $(c'x', c'y')$ on the line, then the rest of the integer points are "parametrized" by

$$(c'x', c'y') + k(b', -a').$$

There are infinitely many choices for the specific integers x', y'. Unfortunately, no choice is better than any other, so there is no one "correct" way to parametrize the solution. Here is a picture of the situation:



Finally, let's complete our running example.

Example: Find the complete solution to the Diophantine equation

$$3094x + 2513y = 21.$$

Solution: We consider the associated equation $3094x + 2513y = z$ and the set of integer vectors $V = \{(x, y, z) \in \mathbb{Z}^3 : 3094x + 2513y = z\}$ solving this equation. We run the Vector Euclidean Algorithm starting with the two "basis vectors":

$$\mathbf{x}_1 = (1, 0, 3094) \quad \text{and} \quad \mathbf{x}_2 = (0, 1, 2513).$$

By omitting unnecessary symbols we obtain the following table of solution vectors:

x	y	z	vector (x, y, z)	
1	0	3094	$\mathbf{x}_1$	
0	1	2513	$\mathbf{x}_2$	
1	−1	581	$\mathbf{x}_3$	$= \mathbf{x}_1 - 1 \cdot \mathbf{x}_2$
−4	5	189	$\mathbf{x}_4$	$= \mathbf{x}_3 - 4 \cdot \mathbf{x}_3$
13	−16	14	$\mathbf{x}_5$	$= \mathbf{x}_3 - 3 \cdot \mathbf{x}_4$
−173	213	7	$\mathbf{x}_6$	$= \mathbf{x}_4 - 13 \cdot \mathbf{x}_5$
359	−442	0	$\mathbf{x}_7$	$= \mathbf{x}_5 - 2 \cdot \mathbf{x}_6$

I claim that the final two vectors $\mathbf{x}_6$ and $\mathbf{x}_7$ contain the solution to our problem. Indeed, by the Principle of Linear Combination we know that all linear combinations $\ell\mathbf{x}_6 + k\mathbf{x}_7$ with $\ell, k \in \mathbb{Z}$ are in the set V. In other words, we have

$$3094(-173\ell + 359k) + 2513(213\ell - 442k) = (7\ell + 0k) \quad \forall \ell, k \in \mathbb{Z}.$$

Finally, since $21 = 3 \cdot 7 = 3 \cdot \gcd(3094, 2513)$ we specify $\ell = 3$ to obtain the equation

$$3094(-519 + 359k) + 2513(639 - 442k) = 21 \quad \forall k \in \mathbb{Z}.$$

This is the complete solution to the problem.

To end the chapter, I will show how the problem of solving the linear Diophantine equation $ax + by = c$ can be extended in two directions.

Extension 1: Systems of linear Diophantine equations. We have seen the complete solution to a single linear Diophantine equation in two unknowns. More generally, we might consider a system of m linear Diophantine equations in n unknowns:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad\qquad\qquad\qquad\qquad\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

If you have taken a course in linear algebra then you know that this system of equations can be written as a single matrix equation:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Then to simplify notation we can abbreviate this matrix equation as

$$A\mathbf{x} = \mathbf{b}$$

where A is the $m \times n$ matrix of coefficients, $\mathbf{x}$ is the $n \times 1$ column vector of unknowns and $\mathbf{b}$ is the $m \times 1$ column vector of constants. We assume that A and $\mathbf{b}$ have **integer** entries and the problem is to find all **integer** vectors $\mathbf{x}$ satisfying the matrix equation.

Now I will sketch the complete solution to this problem. The main difficulty is to generalize the Euclidean Algorithm from a pair of integers to a matrix of integers. The algorithm was first written down by Henry J. Stephen Smith in 1861 and the general theory was developed by Weierstrass and Frobenius. I will state their result without proof.[6]

Theorem (Smith Normal Form). Let A be an $m \times n$ matrix with integer entries. For convenience we will write $\ell := \min\{m, n\}$. Then there exists a unique sequence of natural numbers $d_1, d_2, \ldots, d_\ell \in \mathbb{N}$ with the following properties:

- We have $1 =: d_0 | d_1 | d_2 | \cdots | d_\ell | d_{\ell+1} := 0$. This implies that there exists a unique number $0 \leqslant r \leqslant \ell$ such that $d_0, d_1, \ldots, d_r$ are positive and $d_{r+1} = \cdots = d_{\ell+1} = 0$. The number r is called the *rank* of the matrix A.
- There exists an invertible $m \times m$ matrix P and an invertible $n \times n$ matrix Q such that the matrices P, P^{-1}, Q, Q^{-1} all have integer entries and such that

$$PAQ = D := \left(\begin{array}{ccc|c} d_1 & & & \\ & d_2 & & \\ & & \ddots & 0_{r,n-r} \\ & & & d_r \\ \hline & 0_{m-r,r} & & 0_{m-r,n-r} \end{array}\right)$$

  where $0_{a,b}$ denotes the $a \times b$ matrix with all zero entries.

The numbers $(d_1, d_2, \ldots, d_r)$ are called the *invariant factors* of the integer matrix A. We can think of them as a matrix generalization of the "greatest common divisor". (Indeed, the first invariant factor d_1 is the greatest common divisor of all the entries of A and for all i the product $d_1 d_2 \cdots d_i$ is the greatest common divisor of all the $i \times i$ minors of A.) //

Assuming this result we can solve the linear Diophantine system $A\mathbf{x} = \mathbf{b}$ as follows. If $PAQ = D$ is the Smith Normal Form then we can invert this equation to obtain $A = P^{-1}DQ^{-1}$. Then we substitute into the equation $A\mathbf{x} = \mathbf{b}$ to obtain

$$\begin{aligned} \text{(LDS)} \qquad A\mathbf{x} &= \mathbf{b} \\ P^{-1}DQ^{-1}\mathbf{x} &= \mathbf{b} \\ DQ^{-1}\mathbf{x} &= P\mathbf{b} \\ \text{(RLDS)} \qquad D\mathbf{y} &= \mathbf{c} \end{aligned}$$

where $\mathbf{y} := Q^{-1}\mathbf{x}$ and $\mathbf{c} := P\mathbf{b}$. We conclude that $\mathbf{x}$ is a solution of the linear Diophantine system (LDS) if and only if $\mathbf{y} = Q^{-1}\mathbf{x}$ is a solution of the **reduced** linear Diophantine system

[6] The proof is quite involved.

(RLDS). But the reduced system is easy to solve. If we write $\mathbf{y} = (y_1, y_2, \ldots, y_n)^T$ and $\mathbf{c} = (c_1, c_2, \ldots, c_m)^T$ then the matrix equation $D\mathbf{y} = \mathbf{c}$ translates into the following system of linear Diophantine equations:

$$\begin{cases} d_1 y_1 &= c_1 \\ d_2 y_2 &= c_2 \\ &\vdots \\ d_r y_r &= c_r \\ 0 &= c_{r+1} \\ &\vdots \\ 0 &= c_m \end{cases}$$

Observe that this system has a solution if and only if

- for all $1 \leqslant i \leqslant r$ we have $d_i | c_i$, say $c_i = q_i d_i$ with $q_i \in \mathbb{Z}$,
- for all $r < i$ we have $c_i = 0$,

in which case the complete solution to (RLDS) is

$$\mathbf{y} = \begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_r \\ k_1 \\ k_2 \\ \vdots \\ k_{n-r} \end{pmatrix} \quad \text{for all} \quad k_1, k_2, \ldots, k_{n-r} \in \mathbb{Z}.$$

Then to obtain the complete solution of (LDS) we just compute $\mathbf{x} = Q\mathbf{y}$.

Let's test this on our running example to see if it makes sense. We can express the linear equation $3094x_1 + 2513x_2 = 21$ as the matrix equation

$$\begin{pmatrix} 3094 & 2513 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 21 \end{pmatrix}.$$

My computer tells us that the Smith Normal Form of the coefficient matrix is

$$\begin{pmatrix} 1 \end{pmatrix} \begin{pmatrix} 3094 & 2513 \end{pmatrix} \begin{pmatrix} 186 & 359 \\ -229 & -422 \end{pmatrix} = \begin{pmatrix} 7 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 3094 & 2513 \end{pmatrix} = \begin{pmatrix} 1 \end{pmatrix} \begin{pmatrix} 7 & 0 \end{pmatrix} \begin{pmatrix} 442 & 359 \\ -229 & -186 \end{pmatrix}.$$

(You might recognize some of these numbers from our Euclidean Algorithm computations above.) Substituting this into the original equation gives

$$\begin{aligned}
(1)\begin{pmatrix}7 & 0\end{pmatrix}\begin{pmatrix}442 & 359\\ -229 & -186\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix} &= \begin{pmatrix}21\end{pmatrix}\\
\begin{pmatrix}7 & 0\end{pmatrix}\begin{pmatrix}442 & 359\\ -229 & -186\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix} &= (1)\begin{pmatrix}21\end{pmatrix}\\
\begin{pmatrix}7 & 0\end{pmatrix}\begin{pmatrix}y_1\\ y_2\end{pmatrix} &= \begin{pmatrix}21\end{pmatrix}.
\end{aligned}$$

This last equation is the reduced form of the system; it obviously has the complete solution $(y_1, y_2) = (3, k)$ for all $k \in \mathbb{Z}$. Finally, we invert the process to obtain the complete solution of the original system:

$$\begin{aligned}
\begin{pmatrix}y_1\\ y_2\end{pmatrix} &= \begin{pmatrix}3\\ k\end{pmatrix}\\
\begin{pmatrix}442 & 359\\ -229 & -186\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix} &= \begin{pmatrix}3\\ k\end{pmatrix}\\
\begin{pmatrix}x_1\\ x_2\end{pmatrix} &= \begin{pmatrix}186 & 359\\ -229 & -422\end{pmatrix}\begin{pmatrix}3\\ k\end{pmatrix}\\
\begin{pmatrix}x_1\\ x_2\end{pmatrix} &= \begin{pmatrix}558+359k\\ -687-442k\end{pmatrix}.
\end{aligned}$$

This is not in the same form as our solution above, but we can see that they are the same by making the substitution $k \mapsto (k-3)$ to obtain

$$\begin{pmatrix}x_1\\ x_2\end{pmatrix} = \begin{pmatrix}558+359(k-3)\\ -687-442(k-3)\end{pmatrix} = \begin{pmatrix}-519+359k\\ 639-442k\end{pmatrix} \quad \forall k \in \mathbb{Z}.$$

The non-uniqueness of the parametrization comes from the non-uniqueness of the matrices P and Q in the Smith Normal Form $PAQ = D$. For example, our original solution corresponds to the equally valid factorizations

$$\begin{aligned}
(1)\begin{pmatrix}3094 & 2513\end{pmatrix}\begin{pmatrix}-173 & 359\\ 213 & -422\end{pmatrix} &= \begin{pmatrix}7 & 0\end{pmatrix}\\
\begin{pmatrix}3094 & 2513\end{pmatrix} &= (1)\begin{pmatrix}7 & 0\end{pmatrix}\begin{pmatrix}442 & 359\\ 213 & 173\end{pmatrix},
\end{aligned}$$

but this is not what my computer spit out when I asked it for the Smith Normal Form.

Extension 2: Positive solutions to linear Diophantine equations. Another way to extend the problem is to require that the coefficients and solutions to a linear Diophantine equation be **non-negative**. We can state the problem as follows:

Given natural numbers $a_1, a_2, \ldots, a_n \in \mathbb{N}$ and $b \in \mathbb{N}$, find all natural numbers $x_1, x_2, \ldots, x_n \in \mathbb{N}$ such that

$$\text{(Syl)} \qquad a_1x_1 + a_2x_2 + \cdots + a_nx_n = b.$$

This problem goes by a few different names:

- **The Frobenius Coin Problem.** You live in a country where the coins come in n different denominations: $\$a_1, \$a_2, \ldots, \$a_n$. For which values $\$b$ can you make change? (Debt is not allowed.) If you can make change for $\$b$, in how many different ways can you do it?
- **Sylvester's Postage Stamp Problem.** The numbers $a_1, \ldots, a_n$ are the values of stamps. What amounts of postage can you obtain?
- **The Chicken McNugget Problem.** Chicken McNuggets come in boxes of size $a_1, a_2, \ldots, a_n$. What quantities of Chicken McNuggets can you purchase?

If the numbers $a_1, a_2, \ldots, a_n$ are pairwise coprime then it is known that there exists number

$$g(a_1, a_2, \ldots, a_n),$$

called the *Frobenius number*, such that every integer greater than this can be represented as a non-negative combination of $a_1, \ldots, a_n$. In general it is an *NP-hard* problem to compute the Frobenius number. In the case $n = 2$, Sylvester (1884) proved that

$$g(a_1, a_2) = a_1a_2 - a_1 - a_2$$

and moreover, he proved that the total number of non-representable positive integers is

$$\frac{(a_1 - 1)(a_2 - 1)}{2}.$$

Homework 2 will guide you through proofs of Sylvester's theorems. The equation (Syl) for general n is still an active area of research. For example, let

$$E(a_1, a_2, \ldots, a_n)(b)$$

denote the total number of solutions to the equation. This function is called *Sylvester's denumerant*. It was known to Sylvester and Cayley that this function can be expressed as a polynomial in b with non-constant coefficients:

$$E(a_1, a_2, \ldots, a_n)(b) = \sum_{i=1}^{n} E_i(b) \cdot b^i.$$

Furthermore, each coefficient $E_i(b)$ is a periodic function of b with period that divides the least common multiple of $a_1, \ldots, a_n$. Today these kinds of investigations go under the name "Ehrhart Theory".[7]

[7]For an introduction to Ehrhart Theory I recommend the book *Computing the Continuous Discretely* by Beck and Robins. For the specific problem of Sylvester's denumerant see the recent paper: `https://arxiv.org/abs/1312.7147`

3 Modular Arithmetic

In the previous chapter we saw that in general it is not possible to "divide" by an integer. For example, if there exists an integer $n \in \mathbb{Z}$ with the property $2n = 1$ then we observe that the following properties are true:

$$\begin{cases} 1 = n \cdot 2 + 0 \\ 0 \leqslant 0 < |2| \end{cases} \qquad \text{and} \qquad \begin{cases} 1 = 0 \cdot 2 + 1 \\ 0 \leqslant 1 < |2| \end{cases}$$

The properties on the left say that 1 has quotient n and remainder 0 mod 2, while the properties on the right say that 1 has quotient 0 and remainder 1 mod 2. Since $0 \neq 1$ this contradicts the uniqueness of remainders, so we conclude that there is no such integer $n \in \mathbb{Z}$. In other words, it is not possible to "divide by 2" in the system of integers.

One way to fix this situation is to introduce the formal concept of "fractions". To do this we consider the following set of abstract symbols:

$$\{[a/b] : a, b \in \mathbb{Z}, b \neq 0\}.$$

We think of the abstract symbol "$[a/b]$" as the integer a divided by the nonzero integer b, even though such a number does not necessarily exist within $\mathbb{Z}$. Based on this intuition we should have an equivalence relation on symbols defined by

$$[a/b] \sim [c/d] \quad \iff \quad ad = bc.$$

You checked on HW1 that this does indeed define an equivalence relation. Then we define a *rational number* as an *equivalence class* of abstract symbols. For example, the rational number "1/2" corresponds to the equivalence class

$$\text{"1/2"} = \{[1/2], [(-1)/(-2)], [2/4], [(-2)/(-4)], [3/6], [(-3)/(-6)], \ldots\}.$$

Our intuition also tells us that it should be possible to add and multiply rational numbers using the following rules:

$$\begin{aligned} [a/b] \cdot [c/d] &= [(ac)/(bd)] \\ [a/b] + [c/d] &= [(ad + bc)/(bd)]. \end{aligned}$$

Note that the symbols on the right exist because $b \neq 0$ and $d \neq 0$ implies $bd \neq 0$. But there is still a subtle issue here: each rational number has many different representations; we need to check that the definitions of addition and multiplication of fractions do not depend on the choice of representation. For example, our definition of addition says that

$$[1/2] + [5/8] = [(1 \cdot 8 + 2 \cdot 5)/(2 \cdot 8)] = [18/16].$$

But we can rewrite these fractions as $[1/2] \sim [(-3)/(-6)]$ and $[5/8] \sim [10/16]$ and then the definition of addition gives

$$[(-3)/(-6)] + [10/16] = [((-3) \cdot 16 + (-6) \cdot 10)/((-6) \cdot 16)] = [(-108)/(-96)].$$

If these notions are to make any sense then it must be the case that

$$[18/16] \sim [(-108)/(-96)],$$

and indeed this is true because $18 \cdot (-96) = -1728 = 16 \cdot (-108)$. On HW1 you checked that all of these definitions fit together to create a new number system, the ordered commutative ring of *rational numbers*:

$$(\mathbb{Q}, \leqslant, +, \cdot, 0, 1).$$

This ring satisfies all of the friendly axioms of $\mathbb{Z}$, **except for the Well-Ordering Principle**, but it has the advantage that we can now "divide" by any non-zero number. In abstract-algebraic terminology we say that $\mathbb{Q}$ is a *field*.

You are probably so familiar with fractions that you forgot how abstract they are.[8] Once upon a time, someone had to invent the concept of a "fractions" and then it took quite a while before everyone was comfortable calling them "numbers".

In this chapter we will follow the pattern just described, to define a new family of extensions of $\mathbb{Z}$ out of thin air. These are less familiar than the rational numbers but they have more number-theoretic interest. At the end of the chapter I'll explain how these new number systems are central to modern cryptography.

3.1 Equivalence Mod n

To define the rational numbers we considered an equivalence relation on a set of abstract symbols. To define our new number system we will consider an unusual equivalence on the usual set of integers.

Throughout this section we fix a positive integer $n > 0$.

Definition. Given integers $a, b \in \mathbb{Z}$ we will define the relation $\sim_n$ by

$$a \sim_n b \quad \iff \quad n|(a-b).$$

When $a \sim_n b$ holds we say that a and b are *equivalent modulo* n. //

Before doing anything else let's check that $\sim_n$ is indeed an equivalence relation:

(1) **Reflexive.** For all $a \in \mathbb{Z}$ we have $n|(a-a)$ because $(a-a) = 0 = n \cdot 0$. Thus by definition we have $a \sim_n a$.

(2) **Symmetric.** Consider $a, b \in \mathbb{Z}$ and assume that $a \sim_n b$ so that $n|(a-b)$. By definition this means that there exists an integer $q \in \mathbb{Z}$ such that $a - b = nq$. But then we have

$$\begin{aligned} a - b &= nq \\ b - a &= n(-q), \end{aligned}$$

[8] Indeed, the concept of "adding fractions" signals the end of most people's mathematical careers.

and hence $n|(b-a)$. It follows that $b \sim_n a$ as desired.

(3) **Transitive.** Consider $a, b, c \in \mathbb{Z}$ and assume that we have $a \sim_n b$ and $b \sim_n c$. By definition this means that there exist integers q, q' such that $a - b = nq$ and $b - c = nq'$. But then we have

$$\begin{aligned} a - c &= (a - b) + (b - c) \\ &= nq + nq' \\ &= n(q + q'), \end{aligned}$$

and hence $n|(a-c)$. It follows that $a \sim_n c$ as desired. //

We conclude that the relation $\sim_n$ is an equivalence on the set of integers. This gives us infinitely many new equivalence relations on $\mathbb{Z}$ in addition to our favorite equivalence "=". Now recall that an equivalence relation on a set determines a partition of the set into *equivalence classes.*

Definition. For each $a \in \mathbb{Z}$ consider the set of elements $b \in \mathbb{Z}$ that are equivalent to a mod n:

$$[a]_n := \{b \in \mathbb{Z} : a \sim_n b\}.$$

//

We can be more explicit here; for all $a \in \mathbb{Z}$ I claim that

$$\begin{aligned} [a]_n = (n\mathbb{Z} + a) &:= \{nk + a : k \in \mathbb{Z}\} \\ &= \{\ldots, a - 2n, a - n, a, a + n, a + 2n, \ldots\}. \end{aligned}$$

To see that $[a]_n \subseteq (n\mathbb{Z} + a)$, consider any $b \in [a]_n$. By definition this means that $n|(b-a)$ and hence we have $(b-a) = nk$ for some $k \in \mathbb{Z}$. But then we have $b = nk + a$ and hence $b \in (n\mathbb{Z} + a)$ as desired. Conversely, to see that $(n\mathbb{Z} + a) \subseteq [a]_n$, consider any element $b \in (n\mathbb{Z} + a)$. By definition this means that $b = nk + a$ for some $k \in \mathbb{Z}$ and then we have $(b - a) = nk$. It follows that $n|(b - a)$ and hence $b \in [a]_n$ as desired.

A nice thing about this notation is that we can replace the abstract concept of *equivalence of integers mod n* with the concrete concept of *equality of equivalence classes.*

Equality of Equivalence Classes. For all $a, b \in \mathbb{Z}$ we have

$$a \sim_n b \quad \Longleftrightarrow \quad [a]_n = [b]_n.$$

Proof. Assume that $a \sim_n b$ so that $a = nk + b$ for some $k \in \mathbb{Z}$. Then we have

$$\begin{aligned} c \in [a]_n &\Rightarrow c = n\ell + a && \text{for some } \ell \in \mathbb{Z} \\ &\Rightarrow c = n\ell + (nk + b) \\ &\Rightarrow c = n(\ell + k) + b \\ &\Rightarrow c \in [b]_n, \end{aligned}$$

and conversely,

$$\begin{aligned} c \in [b]_n &\Rightarrow c = nm + b && \text{for some } m \in \mathbb{Z} \\ &\Rightarrow c = nm + (a - nk) \\ &\Rightarrow c = n(m-k) + a \\ &\Rightarrow c \in [a]_n. \end{aligned}$$

We conclude that $[a]_n = [b]_n$ as desired. Next assume that $[a]_n = [b]_n$. Then, in particular, since b is equivalent to itself we must have $b \in [b]_n = [a]_n$ and it follows that $a \sim_n b$ as desired.

□

Now let $\sim$ be a general equivalence relation on a general set S and for each element $x \in S$ let $[x]_\sim := \{x' \in S : x' \sim x\}$ denote the equivalence class of x. If S is not empty then we can choose an element $x_1 \in S$ and then we can express S as a disjoint union

$$S = [x_1]_\sim \amalg S',$$

where $S' \subseteq S$ is the set of elements that are **not** equivalent to x_1. Now if S' is not empty then we can choose an element $x_2 \in S'$ to obtain a disjoint union

$$S = [x_1]_\sim \amalg [x_2]_\sim \amalg S'',$$

where $S'' \subseteq S$ are the elements that are equivalent to **neither** of x_1 and x_2. Continuing in this way, we obtain a partition of the set

$$S = \coprod_{i \in I} [x_i]_\sim,$$

where I is an indexing set and the elements x_i are some arbitrary choice of class representatives. If S is an infinite set then the sets I and $[x_i]_\sim$ might be infinite or finite; there is not much we can say in general.

But now let us return to the equivalence relation $\sim_n$ on the set of integers $\mathbb{Z}$. In this case there is a lot we can say.

Theorem (Division With Remainder, Fancy Version). Let n be a positive integer. Then I claim that $\mathbb{Z}$ decomposes as a disjoint union of the following n equivalence classes:

$$\mathbb{Z} = [0]_n \amalg [1]_n \amalg [2]_n \amalg \cdots \amalg [n-1]_n.$$

Proof. Consider any $a \in \mathbb{Z}$. Since $n > 0$ there exist integers $q, r \in \mathbb{Z}$ such that

$$\begin{cases} a = qn + r \\ 0 \leqslant r < n \end{cases}$$

Then since $(a - r) = nq$ we see that $a \sim_n r$ and hence $a \in [r]_n$. We have shown that every integer is contained in some equivalence of the form $[r]_n$ for some $0 \leqslant r < n$. In other words we can express $\mathbb{Z}$ as a union of n equivalence classes:

$$\mathbb{Z} = [0]_n \cup [1]_n \cup [2]_n \cup \cdots \cup [n-1]_n.$$

To show that this union is **disjoint**, assume for contradiction that two of the classes overlap: say that $a \in [r]_n \cap [r']_n$ for some $0 \leqslant r < r' < n$. The fact that $a \in [r]_n$ tells us that r is the remainder of a mod n and the fact that $a \in [r']_n$ tells us that r' is the remainder of a mod n. But this contradicts the uniqueness of remainders because $r \neq r'$. □

The key to this proof is to express each equivalence class $[a]_n$ in the "standard form" $[a]_n = [r]_n$ where r is the remainder of a mod n. You should compare this to the concept of "lowest terms" for fractions: for each fraction $[a/b] \in \mathbb{Q}$ there is a unique way to write

$$[a/b] \sim [a'/b']$$

where $\gcd(a', b') = 1$ and b' is strictly positive. When computing with fractions we know that we can reduce to lowest terms at any time without affecting the result of the computation. In the next section we will show that the same idea holds for computations with remainders mod n.

3.2 Addition and Multiplication of Remainders

In the previous section we showed that $\mathbb{Z}$ can be written as a disjoint union of n equivalence classes mod n. Using the alternate notation $[a]_n = (n\mathbb{Z} + a)$ we can write this as

$$\mathbb{Z} = (n\mathbb{Z}) \amalg (n\mathbb{Z} + 1) \amalg \cdots \amalg (n\mathbb{Z} + n - 1).$$

The class $n\mathbb{Z}$ is called a *subgroup* of $(\mathbb{Z}, +, 0)$ because it is closed under addition and subtraction, and for a general $a \in \mathbb{Z}$ the equivalence class $(n\mathbb{Z} + a)$ is called the *coset of $n\mathbb{Z}$ generated by a*. We will use that standard abstract-algebraic notation "$\mathbb{Z}/n\mathbb{Z}$" for the set of all cosets of the subgroup $n\mathbb{Z} \subseteq \mathbb{Z}$. Then from the previous theorem we obtain the following.

Definition. Given an integer $n > 0$ we denote the set of equivalence classes mod n by

$$\mathbb{Z}/n\mathbb{Z} := \{[0]_n, [1]_n, [2]_n, \ldots, [n-1]_n\}.$$

//

We can think of $\mathbb{Z}/n\mathbb{Z}$ as the set of possible remainders upon division by n. Indeed, when the context is very clear we might shorten the notation to And if the context is very clear, we might occasionally shorten this to

$$\mathbb{Z}/n\mathbb{Z} = \{0, 1, 2, \ldots, n-1\}.$$

The problem for this section is whether we can add and multiply remainders mod n. For example, consider the numbers 9 and 10 as remainders mod 12. Their sum $9 + 10 = 19$ and product $9 \cdot 10 = 90$ are **not** valid remainders mod 12 but the following equations are valid:

$$[9 + 10]_{12} = [19]_{12} = [7]_{12}$$
$$[9 \cdot 10]_{12} = [90]_{12} = [6]_{12}.$$

Thus we would **like** to say that "$9 + 10 = 7$" and "$9 \cdot 10 = 6$" mod 12, and because of the First Amendment we are free to say this.

Definition(?). Fix an integer $n > 0$. Then for all integers $a, b \in \mathbb{Z}$ we define the *sum* and *product* of equivalence classes as follows:

$$[a]_n + [b]_n := [a + b]_n$$
$$[a]_n \cdot [b]_n := [ab]_n.$$

//

However, just because we can say it doesn't mean that it makes any sense. To turn this definition(?) into a real definition we have to show that it does not logically contradict itself.

Theorem. For all $a, a', b', b' \in \mathbb{Z}$ with $[a]_n = [a']_n$ and $[b]_n = [b']_n$ we have

$$[a + b]_n = [a' + b']_n$$
$$[ab]_n = [a'b']_n.$$

In other words, we say that addition and multiplication of remainders is *well-defined*. //

Proof. Assume that $[a]_n = [b]_n$ and $[a']_n = [b']_n$ so there exist integers $k, \ell \in \mathbb{Z}$ such that $(a - a') = nk$ and $(b - b') = n\ell$. Then we have

$$\begin{aligned}(a + b) - (a' + b') &= (a - a') + (b - b') \\ &= nk + n\ell \\ &= n(k + \ell),\end{aligned}$$

from which it follows that $[a + b]_n = [a' + b']_n$, and we have

$$\begin{aligned}(ab) - (a'b') &= ab - a'b + a'b - a'b' \\ &= (a - a')b + a'(b - b') \\ &= nkb + an\ell \\ &= n(kb + a\ell),\end{aligned}$$

from which it follows that $[ab]_n = [a'b']_n$. □

We have obtained a finite set $\mathbb{Z}/n\mathbb{Z}$ with two (well-defined) binary operations

$$+, \cdot : (\mathbb{Z}/n\mathbb{Z}) \times (\mathbb{Z}/n\mathbb{Z}) \to \mathbb{Z}/n\mathbb{Z}.$$

Since these operations are "inherited" from the integers, it is easy to check that both "$+$" and "$\cdot$" commutative and associative and that "$\cdot$" distributes over "$+$". Note that we have an identity element for addition,

$$[a]_n + [0]_n = [a]_n \quad \forall a \in \mathbb{Z},$$

and for each element $[a]_n \in \mathbb{Z}/n\mathbb{Z}$ we have an additive inverse:

$$[a]_n + [-a]_n = [0]_n.$$

Since this inverse is unique we will write "$-[a]_n$" $= [-a]_n$. Furthermore, if $n \geqslant 2$ then we also have an identity element for multiplication:

$$[a]_n \cdot [1]_n = [a]_n \quad \forall a \in \mathbb{Z}.$$

In summary, for each integer $n \geqslant 2$ we have obtained a new commutative ring, which we call the *ring of integers mod n*:

$$(\mathbb{Z}/n\mathbb{Z}, +, \cdot, [0]_n, [1]_n).$$

This ring shares some properties in common with $\mathbb{Z}$ and $\mathbb{Q}$ but it is also quite different, the key difference being that $\mathbb{Z}/n\mathbb{Z}$ is a **finite ring**. Here is one consequence of finiteness.

Fact. It is impossible to give $(\mathbb{Z}/n\mathbb{Z}, +, \cdot, [0]_n, [1]_n)$ the structure of an *ordered ring*. //

Proof. Assume for contradiction that $\mathbb{Z}/n\mathbb{Z}$ carries an order structure "$\leqslant$". One of the axioms of order says that

$$[0]_n < [1]_n.$$

Another axiom of order says that inequalites are preserved by addition, so we must also have

$$\begin{aligned} [0]_n + [1]_n &< [1]_n + [1]_n \\ [0+1]_n &< [1+1]_n \\ [1]_n &< [2]_n. \end{aligned}$$

By successively adding $[1]_n$ to both sides we eventually obtain the inequality

$$[n-1]_n < [n]_n = [0]_n,$$

and then by transitivity we conclude that $[1]_n < [0]_n$, which is a contradiction. □

Additive cancellation holds in $\mathbb{Z}/n\mathbb{Z}$, as it does in any ring. But recall that multiplicative cancellation in $\mathbb{Z}$ was a consequence of its order structure. You will not be surprised, then, to find out that multiplicative cancellation does not generally hold in $\mathbb{Z}/n\mathbb{Z}$. Even worse, we the ring $\mathbb{Z}/n\mathbb{Z}$ may contain *zero divisors*.

For example:

- Let $n = 6$ and consider the elements $[2]_6$ and $[3]_6$ of the ring $\mathbb{Z}/6\mathbb{Z}$. By uniqueness of remaindes mod 6 we know that $[2]_6 \neq [0]_6$ and $[3]_6 \neq [0]_6$. On the other hand, we have

$$[2]_6 \cdot [3]_6 = [2 \cdot 3]_6 = [6]_6 = [0]_6.$$

- It follows from this that we cannot "multiplicatively cancel $[2]_6$" in the ring $\mathbb{Z}/6\mathbb{Z}$. Indeed, multiplicative cancellation would imply that we have

$$[2]_6 \cdot [x]_6 = [2]_6 \cdot [y]_6 \qquad \Rightarrow \qquad [x]_6 = [y]_6$$

  for all elements $[x]_6, [y]_6 \in \mathbb{Z}/6\mathbb{Z}$. But $[x]_6 = [3]_6$ and $[y]_6 = [0]_6$ is a counterexample.

- Finally, this implies that there is no element in $\mathbb{Z}/6\mathbb{Z}$ that deserves to be called $[1/2]_6$. Indeed, suppose for contradiction that there exists an element $[x]_6 \in \mathbb{Z}/6\mathbb{Z}$ with the property $[2]_6 \cdot [x]_6 = [1]_6$. Then by multiplying both sides of the equation $[2]_6 \cdot [3]_6 = [0]_6$ by $[x]_6$ we would obtain

$$\begin{aligned} [2]_6 \cdot [3]_6 &= [0]_6 \\ [x]_6 \cdot [2]_6 \cdot [3]_6 &= [x]_6 \cdot [0]_6 \\ [1]_6 \cdot [3]_6 &= [x]_6 \cdot [0]_6 \\ [3]_6 &= [0]_6, \end{aligned}$$

  which is a contradiction.

In the next section we will investigate the full story behind this example.

3.3 Euler's Totient Function

We have seen that the element $[2]_6$ has no "multiplicative inverse" in the ring $\mathbb{Z}/6\mathbb{Z}$. This means that there is no element $[x]_6 \in \mathbb{Z}/6\mathbb{Z}$ with the property

$$[2]_6 \cdot [x]_6 = [1]_6.$$

However, if we work modulo 7 then we have

$$[2]_7 \cdot [4]_7 = [8]_7 = [1]_7,$$

which says that the element $[4]_7$ behaves like the number "1/2 modulo 7". Furthermore, this element is unique. Indeed, suppose that we had another multiplicative inverse $[2]_7 \cdot [x]_7 = [1]_7$. Then the associative property gives

$$[x]_7 = [1]_7 \cdot [x]_7 = ([4]_7 \cdot [2]_7) \cdot [x]_7 = [4]_7 \cdot ([2]_7 \cdot [x]_7) = [4]_7 \cdot [1]_7 = [4]_7.$$

Then since the element is unique we can give is a special name. For abstract-algebraic reasons we prefer to use negative exponents instead of fractional notation.

Definition. If the element $[a]_n \in \mathbb{Z}/n\mathbb{Z}$ has a (necessarily unique) multiplicative inverse then we will denote this inverse by $[a^{-1}]_n$, so that

$$[a]_n \cdot [a^{-1}]_n = [1]_n.$$

For all natural numbers $k \in \mathbb{N}$ we will denote k-th power of the inverse by

$$[a^{-k}]_n := ([a^{-1}]_n)^k = \underbrace{[a^{-1}]_n \cdot [a^{-1}]_n \cdot \cdots \cdot [a^{-1}]_n}_{k \text{ times}}$$

Thus the notation $[a^k]_n$ makes sense for all **integers** $k \in \mathbb{Z}$. One can check that this notation satisfies the usual properties of exponents; in particular, we see that every power of $[a]_n$ is invertible because

$$[a^k]_n \cdot [a^{-k}]_n = [a^0]_n = [1]_n.$$

//

Generalizing the above example, we can show that 2 is invertible mod n for any **odd number** n. Indeed, suppose that $n = 2k - 1$ for some integer $k \in \mathbb{Z}$. Then we have

$$[2]_n \cdot [k]_n = [2k]_n = [n + 1]_n = [1]_n$$

and hence $[2^{-1}]_n = [k]_n$. On the other hand, if n is an **even number** then 2 is **not** invertible mod n. Indeed, if $n = 2\ell$ for some $\ell \in \mathbb{Z}$ then we have

$$[2]_n \cdot [\ell]_n = [2\ell]_n = [n]_n = [0]_n,$$

and it follows from the above arguments that $[2]_n$ can have no multiplicative inverse. Here is the general situation.

Theorem (Existence of Multiplicative Inverses Mod n). Let n be a fixed positive integer. Then the element $[a]_n \in \mathbb{Z}/n\mathbb{Z}$ has a multiplicative inverse if and only if $\gcd(a, n) = 1$. Moreover, this inverse can be computed efficiently using the Euclidean Algorithm. //

Proof. First assume that $\gcd(a, n) = 1$. Then the Euclidean Algorithm gives us (non-unique) integers $x, y \in \mathbb{Z}$ with the property

$$ax + ny = 1.$$

By rearranging this equation we obtain

$$\begin{aligned}(ax - 1) = n(-y) &\Longrightarrow n|(ax - 1) \\ &\Longrightarrow (ax) \sim_n 1 \\ &\Longrightarrow [ax]_n = [1]_n \\ &\Longrightarrow [a]_n \cdot [x]_n = [1]_n,\end{aligned}$$

and we conclude that $[x]_n$ is the multiplicative inverse of a mod n. Conversely, suppose that the multiplicative inverse $[x]_n = [a^{-1}]_n$ exists. Then the equation $[a]_n \cdot [x]_n = [1]_n$ gives

$$\begin{aligned}
[ax]_n = [1]_n &\Longrightarrow (ax) \sim_n 1 \\
&\Longrightarrow n|(ax-1) \\
&\Longrightarrow (ax-1) = nq \text{ for some } q \in \mathbb{Z} \\
&\Longrightarrow ax + n(-q) = 1 \text{ for some } q \in \mathbb{Z}.
\end{aligned}$$

Now we want to show that $\gcd(a,n) = 1$. So let d be **any** common divisor of a and n, with $a = da'$ and $n = dn'$. Substituting these into the previous equation gives

$$\begin{aligned}
ax + n(-q) &= 1 \\
(da'x + (dn')(-q) &= 1 \\
d(a'x - n'q) &= 1,
\end{aligned}$$

which implies that d divides 1. But the only integers that divide 1 are $d = \pm 1$ and so we conclude that the **greatest** common divisor of a and n is $d = 1$. □

For example, let us try to compute the multiplicative inverse of 71 modulo 1024. We consider the collection of integer triples $(x, y, z) \in \mathbb{Z}^3$ satisfying $1024x + 71y = z$. Then the Vector Euclidean Algorithm gives:

x	y	z
1	0	1024
0	1	71
1	−14	30
−2	29	11
5	−72	8
−7	101	3
19	−274	2
−26	375	1
71	−1024	0

The second to last row says that

$$1024(-26) + 71(375) = 1,$$

from which we conclude that 71 is invertible mod 1024 with inverse

$$[71^{-1}]_{1024} = [375]_{1024}.$$

[If the algorithm had stopped with $\gcd(71, 1024) \neq 1$ then we would have concluded that 71 is not invertible mod 1024.] //

In a commutative ring $(R, +, \cdot, 0, 1)$, the element 0 never has a multiplicative inverse [why not?] and the element 1 always has a multiplicative inverse; namely itself. In general we denote the collection of invertible elements by

$$R^\times := \{r \in R : \text{there exists a (unique) element } s \in R \text{ with the property } rs = 1\}$$

and we call this the *group of units* of the ring R. The name is meant to indicate that the triple $(R^\times, \cdot, 1)$ has the structure of a *group*. [It is a set with an associative binary operation and an identity element, in which each element has a (necessarily unique) inverse.] Now we can rephrase the previous theorem as follows.

Theorem (The Group of Units of $\mathbb{Z}/n\mathbb{Z}$). Fix a positive integer $n > 0$ and consider the ring $\mathbb{Z}/n\mathbb{Z}$ of integers modulo n. Its group of units is given by

$$(\mathbb{Z}/n\mathbb{Z})^\times = \{[a]_n : \gcd(a, n) = 1\}.$$

//

For example, we have

$$\begin{aligned}
(\mathbb{Z}/6\mathbb{Z})^\times &= \{[1]_6, [5]_6\}, \\
(\mathbb{Z}/7\mathbb{Z})^\times &= \{[1]_7, [2]_7, [3]_7, [4]_7, [5]_7, [6]_7\}, \\
(\mathbb{Z}/8\mathbb{Z})^\times &= \{[1]_8, [3]_8, [5]_8, [7]_8\}.
\end{aligned}$$

Observe that the sum of two elements in $(\mathbb{Z}/n\mathbb{Z})^\times$ is not necessarily in $(\mathbb{Z}/n\mathbb{Z})^\times$ so this is just a group; not a ring. The remainder of this chapter will be devoted to studying the structure of this group. To begin we will look at "exponentiation mod n".

For example, consider the element $[71]_{1024} \in (\mathbb{Z}/1024\mathbb{Z})^\times$. We saw above that $[71^{-1}]_{1024} = [375]_{1024}$. Here are the first few **positve** powers of 71 mod 1024:

$$\begin{aligned}
[71^2]_{1024} &= [5041]_{1024} = [945]_{1024}, \\
[71^3]_{1024} &= [71^2]_{1024} \cdot [71]_{1024} = [945]_{1024} \cdot [71]_{1024} = [67095]_{1024} = [535]_{1024} \\
[71^4]_{1024} &= [71^3]_{1024} \cdot [71]_{1024} = [535]_{1024} \cdot [71]_{1024} = [37987]_{1024} = [97]_{1024} \\
[71^5]_{1024} &= [743]_{1024} \\
[71^6]_{1024} &= [529]_{1024} \\
[71^7]_{1024} &= [695]_{1024} \\
&\vdots
\end{aligned}$$

The sequence of powers

$$1, 71, 945, 535, 97, 743, 529, 695, \ldots$$

looks pretty random.[9] However, because $(\mathbb{Z}/1024\mathbb{Z})^\times$ is a **finite set** we do know that the sequence must contain some repeated element. That is, there must exist two integers $0 < k < \ell$ with the property that

$$[71^k]_{1024} = [71^\ell]_{1024}.$$

[9] It is called *pseudo-random*.

Then multiplying both sides of this equation by the inverse element

$$[71^{-k}]_{1024} = ([71^{-1}]_{1024})^k = ([375]_{1024})^k = [375^k]_{1024}$$

gives

$$\begin{aligned} [71^k]_{1024} &= [71^\ell]_{1024} \\ [71^k]_{1024} \cdot [71^{-k}]_{1024} &= [71^\ell]_{1024} \cdot [71^{-k}]_{1024} \\ [1]_{1024} &= [71^{\ell-k}]_{1024}. \end{aligned}$$

We conclude that there exists some natural number $m \geqslant 1$ with the property that $[71^m]_{1024} = [1]_{1024}$. By the Well-Ordering principle there must be a smallest such number.

Definition. For any element $[a]_n \in (\mathbb{Z}/n\mathbb{Z})^\times$ let $\mathrm{ord}_n(a)$ be the **smallest positive integer** with the property

$$[a^{\mathrm{ord}_n(a)}]_n = [1]_n.$$

We call $\mathrm{ord}_n(a)$ the *multiplicative order of a mod n.* //

The numbers $\mathrm{ord}_n(a)$ are unpredictable in general but they do satisfy some important restrictions. For example, here are the multiplicative orders for the elements of $(\mathbb{Z}/7\mathbb{Z})^\times$:

$[a]_7$	$[1]_7$	$[2]_7$	$[3]_7$	$[4]_7$	$[5]_7$	$[6]_7$
$\mathrm{ord}_7(a)$	1	3	6	3	6	2

Note that all of these numbers divide the size of the group: $6 = (\mathbb{Z}/7\mathbb{Z})^\times$. Leonhard Euler proved in 1750 that this phenomenon holds in general.

Euler's Totient Theorem. Fix a positive integer $n > 0$ and let

$$\varphi(n) := \#(\mathbb{Z}/n\mathbb{Z})^\times,$$

i.e., let $\varphi(n)$ is the number of integers $0 < a < n$ that are coprime to n.[10] Then for all integers $a \in \mathbb{Z}$ satisfying $\gcd(a, n) = 1$ we have

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

and it follows from this that the multiplicative order $\mathrm{ord}_n(a)$ divides $\varphi(n)$. //

Proof. We don't know exactly what the elements of $(\mathbb{Z}/n\mathbb{Z})^\times$ are, but at least we know that there are $\varphi(n)$ of them. Thus we can write

$$(\mathbb{Z}/n\mathbb{Z})^\times = \{[c_1]_n, [c_2]_n, \ldots, [c_{\varphi(n)}]_n\}$$

[10] J.J. Sylvester in 1879 called this *Euler's totient function.* Sylvester was always coming up with ridiculous mathematical terminology, some of which has stuck.

for some distinct class representatives $0 < c_1, c_2, \ldots, c_{\varphi(n)} < n$. Now consider any integer $a \in \mathbb{Z}$ with $\gcd(a, n) = 1$. In this case I claim that we also have

$$(\mathbb{Z}/n\mathbb{Z})^\times = \{[ac_1]_n, [ac_2]_n, \ldots, [ac_{\varphi(n)}]_n\}.$$

Indeed, for each index i we must have $[ac_i]_n \in (\mathbb{Z}/n\mathbb{Z})^\times$ and hence we must have $[ac_i]_n = [c_j]_n$ for some index j. But since $[a]_n$ is invertible (and hence cancellable) we know that

$$\begin{aligned}[ac_i]_n = [ac_j]_n &\iff [a]_n \cdot [c_i]_n = [a]_n \cdot [c_j]_n \\ &\iff [c_i]_n = [c_j]_n.\end{aligned}$$

Now we will multiply all of the elements of $(\mathbb{Z}/n\mathbb{Z})^\times$ together. I don't know which element of $(\mathbb{Z}/n\mathbb{Z})^\times$ this gives me but I do have two different ways to express it:

$$\begin{aligned}[c_1]_n \cdot [c_2]_n \cdot \cdots \cdot [c_{\varphi(n)}]_n &= [ac_1]_n \cdot [ac_2]_n \cdot \cdots \cdot [ac_{\varphi(n)}]_n \\ [c_1c_2 \cdots c_{\varphi(n)}]_n &= [ac_1ac_2 \cdots ac_{\varphi(n)}]_n \\ [c_1c_2 \cdots c_{\varphi(n)}]_n &= [a^{\varphi(n)}c_1c_2 \cdots c_{\varphi(n)}]_n \\ [c_1c_2 \cdots c_{\varphi(n)}]_n &= [a^{\varphi(n)}]_n \cdot [c_1c_2 \cdots c_{\varphi(n)}]_n.\end{aligned}$$

Now we can multiply both sides by the inverse of $[c_1c_2 \cdots c_{\varphi(n)}]_n$ (whatever it is) to obtain

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

Finally, recall that $\mathrm{ord}_n(a)$ is the smallest positive integer satisfying $[a^{\mathrm{ord}_n(a)}]_n = [1]_n$. Now divide $\varphi(n)$ by $\mathrm{ord}_n(a)$ to obtain a quotient and remainder:

$$\begin{cases} \varphi(n) = q \cdot \mathrm{ord}_n(a) + r \\ 0 \leqslant r < \mathrm{ord}_n(a) \end{cases}$$

The first equation tells us that

$$\begin{aligned}[1]_n &= [a^{\varphi(n)}]_n \\ &= [a^{q \cdot \mathrm{ord}_n(a) + r}]_n \\ &= ([a^{\mathrm{ord}_n(a)}]_n)^q \cdot [a^r]_n \\ &= ([1]_n)^q \cdot [a^r]_n \\ &= [a^r]_n.\end{aligned}$$

If $0 < r$ then this contradicts the minimality of $\mathrm{ord}_n(a)$, so we conclude that $r = 0$ and hence $\mathrm{ord}_n(a) | \varphi(n)$ as desired. □

For example, let's compute $\varphi(1024)$. Since $1024 = 2^{10}$ is a power of 2 we have $\gcd(a, 1024)$ if and only if a is **odd**. In other words, we have

$$(\mathbb{Z}/1024\mathbb{Z})^\times = \{[1]_{1024}, [3]_{1024}, [5]_{1024}, \ldots, [1023]_{1024}\}.$$

Since exactly half of the numbers are odd we conclude that

$$\varphi(1024) = \#(\mathbb{Z}/1024\mathbb{Z})^\times = 1024/2 = 512 = 2^9,$$

and then it follows from Euler's Totient Theorem that the multiplicative order of any element $[a]_{1024}$ satisfies:

$$\mathrm{ord}_{1024}(a) \in \{d \in \mathbb{N} : d|512\} = \{1, 2, 4, 8, 16, 32, 64, 128, 356, 512\}.$$

This cuts down on the work necessary to compute $\mathrm{ord}_{1024}(71)$, but it's still not trivial. My computer used a brute-force method to find that

$$\mathrm{ord}_{1024}(71) = 128.$$

A Party Trick. Have you ever looked at a sequence of powers and noticed that the final digit repeats? For example, consider the powers of 3 and note that the final digits cycle through the sequence $1, 3, 9, 7$:

$$\underline{1}, \underline{3}, \underline{9}, 2\underline{7}, 8\underline{1}, 24\underline{3}, 72\underline{9}, 218\underline{7}, \ldots.$$

This phenomenon is explained by Euler's Theorem. Indeed, note that

$$(\mathbb{Z}/10\mathbb{Z})^\times = \{[1]_{10}, [3]_{10}, [7]_{10}, [9]_{10}\}$$

and hence we have $\varphi(10) = 4$. Then for any integer $a \in \mathbb{Z}$ coprime to 10, Euler's Theorem says that

$$[a^4]_{10} = [1]_{10}.$$

Furthermore, for any integer $n = 4q + r$ we have

$$[a^n]_{10} = [a^{4q+r}]_{10} = ([a^4]_{10})^q \cdot [a^r]_{10} = ([1]_{10})^q \cdot [a^r]_{10} = [a^r]_{10}.$$

When q and r are the quotient and remainder of n mod 4 then we conclude that

$$[a^n]_{10} = \begin{cases} [1]_{10} & \text{if } n \in [0]_4 \\ [a]_{10} & \text{if } n \in [1]_4 \\ [a^2]_{10} & \text{if } n \in [2]_4 \\ [a^3]_{10} & \text{if } n \in [3]_4 \end{cases}$$

Since 3 is coprime to 10 this explains our observation about the powers of 3. In fact, you will show on HW3 that for $q \geqslant 0$ and $r \geqslant 1$ we still have $[a^{4q+r}]_{10} = [a^r]_{10}$ even when $\gcd(a, 10) \neq 1$.

You can use this trick at a party to impress people by calculating the final digit of a large power by hand. However, if your friends notice the "mod 4 repetition" then they might not be very impressed. To be safe you should learn how to compute the **final two digits**. For this trick we need to know that

$$\varphi(100) = 40.$$

I'll show you a quick way to compute this in the next section; for now we'll just take it as given. Now for any power a^n, observe that the final two digits are given by the reduced form of $[a^n]_{100}$. If $\gcd(a, 100) = 1$ and if $n = 40q + r$ then the same argument as above tells us that

$$[a^n]_{100} = [a^r]_{100}.$$

And when $q \geqslant 0$ and $r \geqslant 2$ then the result from HW3 says that the same equation still holds for $\gcd(a, 100) \neq 1$. This guarantees that you will never have to compute a higher exponent than 39. For optimum effect you should arrange for the exponent to be 2 more than than a multiple of 40. For example, you could say:

Give me any number "a" and I'll compute the final two digits of "a^{42}".

Then since $[a^{42}]_{100} = [a^2]_{100}$ you just need to compute a^2 (which isn't so hard) and tell them the final two digits. //

In the next two sections we will develop a general formula for the Euler totient function. Then in the final section of the chapter we will apply this formula to cryptography.

3.4 Unique Prime Factorization

To compute the totient function $\varphi(n)$ we first need to compute the "prime factorization" of the integer $n \in \mathbb{Z}$. So far we have only discussed **coprimality** in this class; now it is finally time to discuss **primality**. I postponed the concept of primality until now because it's more subtle than you might think.

What is a prime number? Observe that every integer $n \in \mathbb{Z}$ has two *trivial factorizations*:

$$n = 1 \cdot n \quad \text{and} \quad n = (-1)(-n).$$

Any other factorization $n = ab$ with $a, b \in \mathbb{Z}\backslash\{\pm 1\}$ is called *non-trivial*. We want to say that $n \in \mathbb{Z}$ is prime when it has **no non-trivial factorization**, but there are a few issues here:

- Are we allowed to have negative prime numbers?
- Are the numbers $+1$ and -1 prime?
- What about 0?

There are no completely satisfying answers to these questions and you will find books with differing opinions. I will base my definition of primality on two considerations:

- Aesthetics: I want the statements of big theorems to be as simple as possible.
- Generality: I want my definition to generalize correctly to other commutative rings.

For these reasons I will first state the definition of primality for a general commutative ring R and then we will restrict this definition to the integers $\mathbb{Z}$. Recall that the collection of invertible elements in a ring is called the "group of units" $R^\times$. If $u \in R^\times$ (i.e., if there exists

a multiplicative inverse u^{-1}) then u necessarily divides every element of the ring. Indeed, for all $r \in R$ we have

$$r = 1r = (uu^{-1})r = u(u^{-1}r).$$

Thus, from the point of view of divisibility we should **ignore the units** of the ring. And what about the zero element $0 \in R$? From a sophisticated point of view I would say that 0 is prime if and only if the ring R contains no zero-divisors. But that's a bit too sophisticated for this course, so here I will just say that 0 is **not prime**.

Defininition (Primality in a Commutative Ring). Let R be a general commutative ring. We say that an element $p \in R$ is *prime* when:

- p is not zero,
- p is not a unit,
- if $p = rs$ for some $r, s \in R$ then either r or s is a unit (but not both).

//

We will return to this definition in the next chapter when we study primality in the ring of "Gaussian integers" $\mathbb{Z}[\sqrt{-1}]$. For now we restrict our attention to the "plain old integers" $\mathbb{Z}$. Recall that the invertible integers are just ± 1:

$$\mathbb{Z}^\times = \{-1, +1\}.$$

Definition (Primality in $\mathbb{Z}$). Let p be an integer. We say that p is *prime* when:

- $p \notin \{-1, 0, 1\}$,
- if $p = ab$ for some $a, b \in \mathbb{Z}$ then either $a = \pm 1$ or $b = \pm 1$ (but not both).

//

The notion of primality is not affected by multiplication by units. Thus the prime integers come in positive-negative pairs:

$$\pm 2, \pm 5, \pm 7, \pm 11, \text{etc.}$$

The possibility of negative primes makes the following proofs cleaner, but you can ignore the negative primes when it comes to applications.

For the rest of this section I will present the *Fundamental Theorem of Arithmetic*, which says that every (non-zero, non-unit) integer can be written as a product of prime integers in an (essentially) unique way. These results were originally proved in Books VII and IX of Euclid's *Elements* (c. 300 BC).[11]

[11] This is also where we get the Euclidean Algorithm and Euclid's Lemma.

Theorem (Existence of Prime Factors in $\mathbb{Z}$). Every integer $n \notin \{-1, 0, 1\}$ is divisible by a prime integer. //

Proof. Note that for all $n, p \in \mathbb{Z}$ we have $p|n$ if and only if $p|(-n)$. Thus we can restrict our attention to positive integers. So assume for contradiction that there exists an integer $n \geqslant 2$ with **no prime factor**. Then by the Well-Ordering Principle there exists a smallest such integer; call it $m \geqslant 2$. Since m divides itself (i.e., $m = 1m$) and since by assumption m has no prime factor, it must be the case that m is **not prime**. By definition this means that there exists a "non-trivial" factorization

$$m = ab$$

in which neither of a or b is a unit. Since $ab = (-a)(-b)$ and since a is not a unit we can assume without loss of generality that $a \geqslant 2$. Since $a|m$ we must also have $a \leqslant m$, but if $a = m$ then $m = ab$ implies that $b = 1$, which contradicts the fact that b is not a unit. Thus we conclude that $2 \leqslant a \leqslant m - 1$. Since m was the **smallest** positive integer with no prime factor, this implies that a **has** a prime factor, say $a = pa'$. Finally, we conclude that m itself has a prime factor since

$$m = ab = (pa')b = p(a'b),$$

and this is the desired contradiction. □

For the next theorem I will use a common mathematical convention: Let $S \subseteq \mathbb{Z}$ be any finite collection of integers and let $n = \prod_{s \in S} s$ denote the product of these integers. For $|S| \geqslant 2$ this product is well-defined because of the commutative and associative laws of multiplication. In the cases $S = \{s\}$ or $S = \varnothing$ we say by convention that $n = s$ or $n = 1$, respectively. That is:

a product of no numbers equals 1.

Theorem (Existence of Prime Factorization in $\mathbb{Z}$). Every non-zero integer can be expressed as a unit times a product of prime numbers. //

Proof. Consider $0 \neq n \in \mathbb{Z}$. If n is a unit or a prime then we are done. Otherwise, we know from the previous theorem that there exists a prime factor, say $n = pn'$. If n' is a unit or a prime then we are done. Otherwise, n' has a prime factor, say $n' = p'n''$. If n'' is a prime or a unit then we are done; otherwise we continue. When this process stops we will obtain the desired factorization.

To prove that the process does stop, observe that the integers n, n', n'' from above satisfy $|n| > |n'| > |n''|$. If the process continues forever then we will obtain an infinite decreasing sequence of positive integers

$$|n| > |n'| > |n''| > |n'''| > \cdots > 0,$$

which violates the Well-Ordering Principle. □

We have now proved that every non-zero integer can be written as a unit times a product of primes. For example, the number -30 can be written as

$$\begin{aligned} -30 &= (-1) \cdot 2 \cdot 3 \cdot 5 \\ &= 1 \cdot (-2) \cdot 3 \cdot 5 \\ &= 1 \cdot 2 \cdot (-3) \cdot 5 \\ &= 1 \cdot 2 \cdot 3 \cdot (-5) \\ &= (-1) \cdot (-2) \cdot (-3) \cdot 5 \\ &\vdots \\ &= 1 \cdot (-5) \cdot (-3) \cdot (-2) \end{aligned}$$

There are lots of ways (48 ways, in fact) to write this factorization, but the the differences are only cosmetic; all I have done is rearranged the units and permuted the prime factors. Our final theorem says that prime factorization is unique except for these trivial rearrangements.

Theorem (Uniqueness of Prime Factorization in $\mathbb{Z}$). Consider a non-zero integer $n \in \mathbb{Z}$ and suppose that we have

$$n = \pm p_1 p_2 \cdots p_k = \pm q_1 q_2 \cdots q_\ell$$

where the integers $p_1, \ldots, p_k, q_1, \ldots, q_k \in \mathbb{Z}$ are all prime. In this case I claim that $k = \ell$, and furthermore I claim that we can permute the indices so that

$$p_1 = \pm q_1, \quad p_2 = \pm q_2, \quad \ldots, \quad p_k = \pm q_k.$$

//

To prove this we need a lemma. This is the original version of Euclid's Lemma from Euclid's *Elements* (Proposition VII.30).

Euclid's Lemma (Prime Version). Let $p \in \mathbb{Z}$ be prime. Then for all $a, b \in \mathbb{Z}$ we have

$$(p|ab) \Rightarrow (p|a \vee p|b).$$

//

Proof of the Lemma. Let p be prime. We will assume that $p|ab$ and $p \nmid a$ and this case we will show that $p|b$. Recall from our original version of Euclid's Lemma that

$$(p|ab \wedge \gcd(a,p) = 1) \Rightarrow (p|b).$$

Thus we will be done if we can show that $d := \gcd(a,p)$ equals 1. Recall that the gcd satisfies $1 \leqslant d \leqslant |p|$. Since $d|p$ we have $p = dp'$ for some $p' \in \mathbb{Z}$. Then since p is prime it must be the case that d is a unit (i.e., $d = 1$) or that p' is a unit (i.e., $d = |p|$). On the other hand, the case $d = |p|$ is impossible because we have $d|p$ and $a \nmid p$. We conclude that $d = 1$ as desired. $\square$

Proof of the Theorem. Suppose that we have

$$p_1 p_2 \cdots p_k = \pm q_1 q_2 \cdots q_\ell \tag{UF}$$

for some primes $p_1, \ldots, p_k, q_1, \ldots, q_\ell \in \mathbb{Z}$ and assume without loss of generality that $k \leqslant \ell$. Since p_1 divides the left hand side it also divides the right hand side: $p_1 | (q_1 q_2 \cdots q_\ell)$. Since p_1 is prime, by Euclid's Lemma this means that p_1 divides q_i for some i. After relabeling the primes q_i we can assume without loss of generality that $p_1 | q_1$, say $q_1 = p_1 u$. Since q_1 is prime and since p_1 (being prime) is not a unit, this implies that u **is a unit** and we conclude that $p_1 = \pm q_1$. Now we apply multiplicative cancellation to the equation (UF) to obtain

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

By repeating the argument and relabeling the primes q_i as necessary we will find that $p_2 = \pm q_2$, $\ldots$, $p_k = \pm q_k$. Finally, we assume for contradiction that $\ell \geqslant k + 1$. After canceling the first k factors we obtain the equation

$$1 = \pm q_{k+1} \cdots q_\ell,$$

which implies that $q_{k+1} | 1$. But then we must have $q_{k+1} = \pm 1$ which contradicts the fact that q_{k+1} (being prime) is not a unit. □

In summary, each non-zero integer $n \in \mathbb{Z}$ has a *unique prime factorization*. It is often convenient to express this in the following form:

> Denote the positive primes by $2 = p_1 < p_2 < p_3 < \cdots$. Then for all $0 \neq n \in \mathbb{Z}$ there exists a unique sequence of non-negative exponents $e_1, e_2, e_3, e_4, \ldots$ (all but finitely many equal to zero) such that
>
> $$n = \pm p_1^{e_1} p_2^{e_2} p_3^{e_3} p_4^{e_4} \cdots$$

Exercises. The language of unique factorization gives us a new way to think about divisibility. For these exercise we will fix two non-zero integers $a, b \in \mathbb{Z}$ with unique prime factorizations

$$a = \pm p_1^{a_1} p_2^{a_2} p_3^{a_3} \cdots,$$
$$b = \pm p_1^{b_1} p_2^{b_2} p_3^{b_3} \cdots.$$

(a) Prove that $a | b$ if and only if $a_i \leqslant b_i$ for all i.

(b) Prove that the greatest common divisor is given by

$$\gcd(a, b) = p_1^{\min(a_1, b_1)} p_2^{\min(a_2, b_2)} p_3^{\min(a_3, b_3)} \cdots.$$

(c) Find a similar formula for the *least common multiple* $\operatorname{lcm}(a, b)$ and use it to prove that

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

3.5 Chinese Remainder Theorem

Finally, we will use the unique prime factorization of a positive integer n to compute the value of the Euler totient function $\varphi(n)$. Let me state the result right away and then we will work up to the proof.

Theorem (Value of the Totient Function). Let n be a positive integer. Then the totient function is given by

$$\varphi(n) = n \cdot \prod_{p|n} \frac{p-1}{p},$$

where the product is taken over the distinct positive prime factors of n. //

For example, our Party Trick used the fact that $\varphi(100) = 40$. Now we can see why this is true. The prime factorization $100 = 2^2 \cdot 5^2$ shows us that the distinct prime factors of 100 are 2 and 5. Then the formula gives

$$\varphi(100) = 100 \cdot \frac{1}{2} \cdot \frac{4}{5} = 40.$$

The theorem depends on two lemmas. The first one is straightforward.

Lemma 1 (Totient of a Prime Power). Consider two positive integers p, n where p is prime. Then we have

$$\varphi(p^n) = (p^n - p^{n-1}) = p^n \left(1 - \frac{1}{p}\right) = p^n \cdot \frac{p-1}{p}.$$

//

Proof. I claim that for all $a \in \mathbb{Z}$ we have

$$\gcd(a, p^n) = 1 \quad \Longleftrightarrow \quad p \nmid a.$$

To give a quick[12] proof of this we will apply unique prime factorization. Suppose that some non-zero integer a has the prime factorization

$$a = \pm p_1^{e_1} p_2^{e_2} p_3^{e_3} \cdots$$

and suppose that p is the k-th prime so that p^n has prime factorization

$$p^n = p_1^0 \cdots p_{k-1}^0 p_k^n p_{k+1}^0 \cdots$$

[12] A slow proof is also possible.

Then from the exercises in the previous section we find that the greatest common divisor has prime factorization given by

$$\gcd(a, p^n) = p_1^0 \cdots p_{k-1}^0 p_k^{\min(e_k, n)} p_{k+1}^0 \cdots .$$

From this factorization we observe that $\gcd(a, p^n) = 1$ if and only if $\min(e_k, n) = 0$, i.e., if and only if $e_k = 0$, i.e., if and only if a is not divisible by $p = p_k$. In other words, we have

$$(\mathbb{Z}/p^n\mathbb{Z})^\times = \{[a]_{p^n} : 1 \leqslant a \leqslant p^n \text{ and } p \nmid a\}.$$

To count the elements of this group, observe that the multiples of p between 1 and p^n are

$$1, 2p, 3p, \ldots, (p^{n-1})p = p^n,$$

and there are precisely p^{n-1} of these. Finally, we have

$$\begin{aligned}
\varphi(p^n) &= \#(\mathbb{Z}/p^n\mathbb{Z})^\times \\
&= \#(\text{integers from 1 to } p^n \text{ not divisible by } p) \\
&= \#(\text{integers from 1 to } p^n) - \#(\text{multiples of } p) \\
&= p^n - p^{n-1}.
\end{aligned}$$

□

The second lemma depends on a significant trick, so it deserves a name. This result was called the "Chinese Remainder Theorem" by Leonard Dickson in 1929. Apparently it became known in the West after Wylie's 1953 article *Jottings on the Science of the Chinese Arithmetic*. We now know that the result was discovered by the mathematician Sun Zu in the 3rd century AD. Some authors have tried to change the name to *Sun Zu's Theorem* but it might be too late.

Lemma 2 (The Chinese Remainder Theorem). For any coprime integers $\gcd(m, n) = 1$ there exists a one-to-one correspondence between elements of the ring $\mathbb{Z}/mn\mathbb{Z}$ and pairs of elements from $\mathbb{Z}/m\mathbb{Z}$ and $\mathbb{Z}/n\mathbb{Z}$:

$$(\mathbb{Z}/mn\mathbb{Z}) \longleftrightarrow (\mathbb{Z}/m\mathbb{Z}) \times (\mathbb{Z}/n\mathbb{Z}).$$

This correspondence restricts to the invertible elements

$$(\mathbb{Z}/mn\mathbb{Z})^\times \longleftrightarrow (\mathbb{Z}/m\mathbb{Z})^\times \times (\mathbb{Z}/n\mathbb{Z})^\times,$$

and it follows from this that the totient function satisfies $\varphi(mn) = \varphi(m)\varphi(n)$.

Proof. The map from $(\mathbb{Z}/mn\mathbb{Z})$ to pairs $(\mathbb{Z}/m\mathbb{Z}) \times (\mathbb{Z}/n\mathbb{Z})$ is easy to define: for all integers $a \in \mathbb{Z}$ we send the equivalence class $[a]_{mn}$ to the pair of equivalence classes $([a]_m, [a]_n)$. To show that this is a one-to-one correspondence, there are three things to check:

(1) The map is "well-defined". Assume that $[a]_{mn} = [a']_{mn}$, so we have $(a - a') = mnk$ for some $k \in \mathbb{Z}$. In particular, since $(a - a') = m(nk)$ we have $[a]_m = [a']_n$, and since $(a - a') = n(mk)$ we have $[a]_n = [a']_n$. Thus the pairs $([a]_m, [a]_n)$ and $([a']_m, [a']_n)$ are equal as desired. //

(2) The map is "one-to-one". Assume that the pairs $([a]_m, [a]_n)$ and $([b]_m, [b]_n)$ are equal. In this case we want to show that $[a]_{mn} = [b]_{mn}$. By assumption we have $[a]_m = [b]_m$ so that $m|(a - b)$ and we have $[a]_n = [b]_n$ so that $n|(a - b)$. Then a result from HW3 tells us that $(mn)|(a - b)$ and hence $[a]_{mn} = [b]_{mn}$ as desired. //

(3) The map is "onto". This is the part where we need a trick. For any two integers $a, b \in \mathbb{Z}$ we need to show that the pair $([a]_m, [b]_n)$ has the form $([c]_m, [c]_n)$ for some common integer $c \in \mathbb{Z}$. And here's the trick: Since $\gcd(m, n) = 1$ we know from the Euclidean Algorithm that there exist some integers $x, y \in \mathbb{Z}$ such that $mx + ny = 1$. Then we define

$$c := any + bmx.$$

To check that $[c]_m = [a]_m$ we note that

$$\begin{aligned}
[c]_m &= [any + bmx]_m \\
&= [any]_m + [m(bx)]_m \\
&= [any]_m + [0]_m \\
&= [any]_m \\
&= [a(1 - mx)]_m \\
&= [a]_m - [m(ax)]_m \\
&= [a]_m - [0]_m \\
&= [a]_m.
\end{aligned}$$

The proof that $[c]_n = [b]_n$ is similar. //

To complete the proof we need to show that this one-to-one correspondence matches the inverticle elements $[a]_{mn} \in (\mathbb{Z}/mn\mathbb{Z})^\times$ with pairs of invertible elements $([a]_m, [a]_n) \in (\mathbb{Z}/m\mathbb{Z})^\times \times (\mathbb{Z}/n\mathbb{Z})^\times$.[13] In other words, we need to show that for all integers $a \in \mathbb{Z}$ we have

$$\gcd(a, mn) = 1 \quad \Longleftrightarrow \quad \gcd(a, m) = 1 \wedge \gcd(a, n) = 1.$$

For this we don't even need the assumption $\gcd(m, n) = 1$. We will use the fact that two integers $p, q \in \mathbb{Z}$ are coprime **if and only if** there exist integers $x, y \in \mathbb{Z}$ such that $px + qy = 1$. [Remind yourself why this is true.] First assume that $\gcd(a, mn) = 1$ so there exist integers $x, y \in \mathbb{Z}$ such that $ax + mny = 1$. Then since $ax + m(ny) = 1$ we have $\gcd(a, m) = 1$ and since $ax + n(my) = 1$ we have $\gcd(a, n) = 1$. Conversely, assume that we have $\gcd(a, m) = 1$ and

[13] We could give an abstract proof by showing that the correspondence preserves ring operations and then by showing that the group of units of a "product ring" $R \times S$ satisfies $(R \times S)^\times = R^\times \times S^\times$, but that would be too abstract for this class.

$\gcd(a,n)=1$, so there exist integers $x,y,x',y'\in\mathbb{Z}$ such that $ax+my=1$ and $ax'+ny'=1$. Multiplying these two equations gives

$$\begin{aligned}(ax+my)(ax'+ny')&=1\\ a(xax'+xny'+myx')+mn(yy')&=1,\end{aligned}$$

and hence $\gcd(a,mn)=1$ as desired. In conclusion, we have a one-to-one correspondence between the sets $(\mathbb{Z}/mn\mathbb{Z})^\times$ and $(\mathbb{Z}/m\mathbb{Z})^\times\times(\mathbb{Z}/n\mathbb{Z})^\times$. By comparing cardinalities we obtain

$$\begin{aligned}\#(\mathbb{Z}/mn\mathbb{Z})^\times&=\#\left[(\mathbb{Z}/m\mathbb{Z})^\times\times(\mathbb{Z}/n\mathbb{Z})^\times\right]\\ \#(\mathbb{Z}/mn\mathbb{Z})^\times&=\#(\mathbb{Z}/m\mathbb{Z})^\times\cdot\#(\mathbb{Z}/n\mathbb{Z})^\times\\ \varphi(mn)&=\varphi(m)\varphi(n).\end{aligned}$$

□

Proof of the Theorem. Suppose that a positive integer $n\geqslant 2$ has prime factorization

$$n=p_1^{e_1}p_2^{e_2}\cdots p_k^{e_k}$$

for some distinct primes $1<p_1<p_2<\cdots<p_k$. One can easily check that the factors $p_i^{e_i}$ and $p_j^{e_j}$ are coprime for all $i\neq j$. Thus from the two previous lemmas we have

$$\begin{aligned}\varphi(n)&=\varphi(p_1^{e_1}p_2^{e_2}\cdots p_k^{e_k})\\ &=\varphi(p_1^{e_1})\varphi(p_2^{e_k})\cdots\varphi(p_k^{e_k}) && \text{Lemma 2}\\ &=p_1^{e_1}\cdot\frac{p_1-1}{p_1}\cdot p_2^{e_2}\cdot\frac{p_2-1}{p_2}\cdots p_k^{e_k}\cdot\frac{p_k-1}{p_k} && \text{Lemma 1}\\ &=p_1^{e_1}p_2^{e_2}\cdots p_k^{e_k}\cdot\frac{p_1-1}{p_1}\cdot\frac{p_2-1}{p_2}\cdots\frac{p_k-1}{p_k}\\ &=n\cdot\prod_{i=1}^{k}\frac{p_i-1}{p_i}\\ &=n\cdot\prod_{p\mid n}\frac{p-1}{p}.\end{aligned}$$

□

To end the section I will give a probabilistic interpretation of this theorem. For example, consider our favorite number $100=2^2\cdot 5^2$ and consider any integer $1\leqslant a\leqslant 100$. We know that $\gcd(a,100)=1$ if and only if a is not a multiple of 2 and a is not a multiple of 5. To remove the multiples of 2 we can multiply by $1/2$ to get

$$100\cdot\frac{1}{2}=50,$$

and to remove the multiples of 5 we can multiply by 4/5 to get

$$100 \cdot \frac{4}{5} = 80.$$

It seems plausible that we could remove **both** kinds of numbers by multiplying by both fractions to get

$$100 \cdot \frac{1}{2} \cdot \frac{4}{5} = 40.$$

In other words, we are assuming that for integers $1 \leqslant a \leqslant 100$ the two events

$$\text{“}a \text{ is not a multiple of 2”} \quad \text{and} \quad \text{“}a \text{ is not a multiple of 5”}$$

are probabilistically *independent*. The theorem above guarantees that this is correct.

Epilogue (Sun Zu Suan Jing). The original purpose of the Chinese Remainder Theorem was to solve systems of simultaneous linear “congruences”. For example, here is a problem from the fourth-century text *Sun Zu Suan Jing* (Master Sun's Mathematical Manual):

> There are certain things whose number is unknown. If we count them by threes, we have two left over; by fives, we have three left over; and by sevens, two are left over. How many things are there?

In modern terms we can phrase the problem as follows: Find all integers $c \in \mathbb{Z}$ such that

$$\begin{cases} [c]_3 &= [2]_3 \\ [c]_5 &= [3]_5 \\ [c]_7 &= [2]_7. \end{cases}$$

We will solve this by dealing with the equations two at a time. Let's begin with the first two equations:

$$\text{(SunZu)} \qquad \begin{cases} [c]_3 &= [2]_3 \\ [c]_5 &= [3]_5. \end{cases}$$

Now let's recall what the Chinese Remainder Theorem says. If $\gcd(m,n) = 1$ then there exists a unique element $[c]_{mn} \in \mathbb{Z}/mn\mathbb{Z}$ with the property $([c]_m, [c]_n) = ([a]_m, [b]_n)$, and this element is given explicitly by

$$[c]_{mn} = [any + bmx]_{mn},$$

where $x, y \in \mathbb{Z}$ are any integers satisfying $mx + ny = 1$. In our case we have $(a,b) = (2,3)$, $(m,n) = (3,5)$ and I found $(x,y) = (-3,2)$ by trial-and-error. Thus the pair of equations (SunZu) has a unique solution mod $3 \cdot 5 = 15$ which is given by

$$\begin{aligned} [c]_{15} &= [any + bmx]_{15} \\ &= [2 \cdot 5 \cdot 2 + 3 \cdot 3 \cdot (-3)]_{15} \\ &= [20 - 27]_{15} \\ &= [-7]_{15} \\ &= [8]_{15}. \end{aligned}$$

In other words, the pair of equations (SunZu) is equivalent to the single equation $[c]_{15} = [8]_{15}$ and the original system of three equations is equivalent to the following system of two equations:

$$\begin{cases} [c]_{15} &= [8]_{15} \\ [c]_7 &= [2]_7. \end{cases}$$

We can solve this with the same method; this time we have $(a, b) = (8, 2)$, $(m, n) = (15, 7)$ and I found $(x, y) = (1, -2)$ by trial-and-error. Finally, the Chinese Remainder Theorem tells us that the original system has a unique solution mod $15 \cdot 7 = 105$, which is given by

$$\begin{aligned} [c]_{105} &= [any + bmx]_{105} \\ &= [8 \cdot 7 \cdot (-2) + 2 \cdot 15 \cdot 1]_{105} \\ &= [-112 + 30]_{105} \\ &= [-82]_{105} \\ &= [23]_{105} \end{aligned}$$

In other words, the complete solution of the problem is $c = 23 + 105k$ for all integers $k \in \mathbb{Z}$. Sun Zu used a similar method, but he solved all three equations at the same time. First he (somehow) found the integers $(x, y, z) = (2, 1, 1)$ such that $([x^{-1}]_{5\cdot7}, [y^{-1}]_{3\cdot7}, [z^{-1}]_{3\cdot5}) = ([1]_{5\cdot7}, [1]_{3\cdot7}, [1]_{3\cdot5})$ and then he computed the solution

$$\begin{aligned} [c]_{105} &= [2(x \cdot 5 \cdot 7) + 3(3 \cdot y \cdot 7) + 2(3 \cdot 5 \cdot z)]_{105} \\ &= [2(70) + 3(21) + 2(15)]_{105} \\ &= [233]_{105} \\ &= [23]_{105}. \end{aligned}$$

Apparently this solution was even recorded in a folk song called "The Song of Master Sun":

> Not in every third person is there one aged three score and ten,
> On five plum trees only twenty-one boughs remain,
> The seven learned men meet every fifteen days,
> We get our answer by subtracting one hundred and five over and over again.[14]

Here is the general statement Sun Zu's method in modern language. Suppose that the sequence of moduli $m_1, m_2, \ldots, m_n \in \mathbb{N}$ are pairwise coprime. Then for any integers $a_1, a_2, \ldots, a_n \in \mathbb{Z}$ the system of congruences $[c]_{m_i} = [a_i]_{m_i}$ has a unique solution $[c]_M$ modulo $M := m_1 m_2 \cdots m_n$, which can be computed as follows. For each index $1 \leqslant i \leqslant n$, use the Euclidean Algorithm to find an integer $x_i \in \mathbb{Z}$ such that

$$[x_i^{-1}]_{m_i} = [m_1 \cdots m_{i-1} m_{i+1} \cdots m_n]_{m_i}.$$

Then the complete solution is given by

$$[c]_M = [a_1(x_1 m_2 \cdots m_n) + a_2(m_1 x_2 m_2 \cdots m_n) + \cdots + a_n(m_1 m_2 \cdots m_{n-1} x_n)]_M.$$

[14] Quoted from *The Crest of the Peacock* by George Ghereghese Joseph.

The later work *Shu Shu Jiu Zhang* (1247) by the mathematician Qin Jiushao describes algorithms (*da yan shu*) for solving linear systems that were unknown in Europe until over 500 years later. Eventually theses methods were rediscovered by Euler (1743) in his work on linear differential equations and by Gauss (1801) in his work on least-squares regression.

3.6 Applications to Cryptography

The dividing line between "arithmetic" and "higher arithmetic" (i.e., number theory) was traditionally placed at the point where arithmetic ceases being useful. From recreational problems such as Sylvester's postage stamp problem, to significant challenges such as Fermat's Last Theorem, a common feature of all types of number theory was its lack of applications.[15]

This all changed in the 1960s and 70s, when researchers working in academia and behind the scenes at US and British intelligence agencies came up with a new kind of cryptography, called

asymmetric cryptography.

To understand asymmetric cryptography we first have to discuss its precursor, symmetric cryptography. Suppose that Alice and Bob[16] want to send secret messages to each other. This traditionally involved two steps:

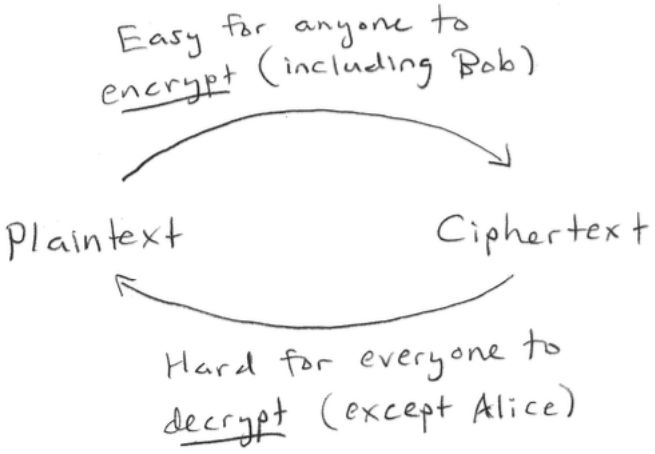
> **Key Exchange.** Alice and Bob meet in secret or establish a *secure channel* to exchange the keys for a symmetric cryptosystem.
>
> **Message Exchange.** Now Alice and Bob can exchange encrypted messages from over an *insecure channel*.

The term "symmetric cryptosystem" means that Alice and Bob will both use the same process for encryption and decryption; I will assume that relatively good schemes are available. The real difficulty of symmetric cryptography is that is seems to require a secure channel in order to exchange the keys. As the US Department of Defense developed the ARPANET in the 1960s, **pressure mounted to find some way to perform this key exchange over an insecure channel**.

The desire to solve this problem forced people to consider the possibility of an **asymmetric cryptosystem**. In brief, this is a scheme in which the encryption key is public and only the decryption key needs to be private. (For this reason it is also called "public-key cryptography".) Thus, in order to send messages back and forth, Alice and Bob must set up two separate systems. Here is what Alice's system looks like:

[15] The number theorist G.H. Hardy wrote an essay in 1940 called *A Mathematician's Apology* in which he celebrated the fact that his science was "gentle and clean" and could never be applied to military purposes.

[16] I am legally obligated to use these names.



Presumably we will encode both the plaintext and the ciphertext as numbers. Thus the technical problem is to find a mathematical "encryption function" $m \mapsto e(m)$ with the following properties:

- Given some plaintext message m it is easy to compute the ciphertext $e(m)$.
- The function e is one-to-one, so that a given ciphertext $e(m)$ corresponds to a **unique** plaintext message m.
- Given the ciphertext $e(m)$ it is generally difficult to compute the plaintext message m (i.e., to "decipher" the ciphertext), however there is a secret key that allows Alice to compute it quickly.

Such a function $m \mapsto e(m)$ is called a *trapdoor function* because of the secret door that allows only Alice to compute the inverse $e(m) \mapsto m$ quickly. However, it is not clear whether any practical trapdoor functions exist. William Stanley Jevons tried to apply number theory to the problem in his book *The Principles of Science* (1874):

> Can the reader say what two numbers multiplied together will produce the number 8616460799? I think it unlikely that anyone but myself will ever know.

Here he is referring to the fact that it is relatively easy to multiply two numbers (in this case the prime numbers 89681 and 96079) but that it is relatively difficult to factor them apart again. Unfortunately Jevons was not able to incorporate a trapdoor into the prime factorization problem.

The breakthrough came in the early 1970s when researchers working at GCHQ (the British counterpart of the NSA) and at American universities independently came up with the same solutions to the problem. The government discoveries were first, but since the work was classified the algorithms were named after the public discoverers. For the rest of this section I will discuss the two key protocols of asymmetric cryptography, called *Diffie-Hellman Key Exchange* and the *RSA Cryptosystem*.

Protocol 1: Diffie-Hellman and ElGamal. Before we discussed Euler's Totient Theorem I wrote down the sequence of powers of 71 mod 1024:

$$1, 71, 945, 535, 97, 743, 529, 695, \ldots$$

Since $\varphi(1024) = \varphi(2^9) = 2^9(1 - 1/2) = 512$, Euler's theorem tells us that the multiplicative order $\mathrm{ord}_{1024}(71)$ divides 512; in fact, my computer tells me that $\mathrm{ord}_{1024}(71) = 128$. In other words, the sequence of powers will repeat after 128 steps. However, other than this repetition mod 128 there seems to be no discernable pattern in the sequence. I will phrase this as an assumption:

Assumption: computing discrete logarithms is hard.

By a "discrete logarithm" I mean that we are given an element of $(\mathbb{Z}/1024\mathbb{Z})^\times$ of the form $[71^\ell]_{1024}$ and we are asked to find the exponent ℓ. This ℓ (which is well-defined modulo the order $\mathrm{ord}_{1024}(71) = 128$) is something like a "logarithm to the base 71" modulo 1024. More generally, if $\gcd(a, n) = 1$ then we will assume that it is difficult to compute the exponent ℓ given an element of the form $[a^\ell]_n \in (\mathbb{Z}/n\mathbb{Z})^\times$.

On the other hand, we have the following fact:

Fact: computing discrete exponentials is easy.

That is, given integers a, ℓ and n it is easy to compute the element $[a^\ell]_n \in (\mathbb{Z}/n\mathbb{Z})$; in fact we can do it in less than $4 \cdot \log_2(\ell)$ operations by the method of "repeated squaring". The trick is to repeatedly use the formula

$$x^n = \begin{cases} x \cdot (x^2)^{(n-1)/2} & \text{for } n \text{ odd} \\ (x^2)^{n/2} & \text{for } n \text{ even} \end{cases}$$

so that we only have to compute binary products and squares. To see how this works, let's compute the reduced form of the element $[71^{43}]_{1024}$. First we repeatedly apply the above formula to obtain

$$\begin{aligned}
[71^{43}]_{1024} &= [71]_{1024} \cdot \left([71^2]_{1024}\right)^{21} = [71]_{1024} \cdot [945^{21}]_{1024}, \\
[945^{21}]_{1024} &= [945]_{1024} \cdot \left([945^2]_{1024}\right)^{10} = [945]_{1024} \cdot [97^{10}]_{1024}, \\
[97^{10}]_{1024} &= \left([97^2]_{1024}\right)^{5} = [193^5]_{1024}, \\
[193^5]_{1024} &= [193]_{1024} \cdot \left([193^2]_{1024}\right)^{2} = [193]_{1024} \cdot [385^2]_{1024}, \\
[385^2]_{1024} &= [769]_{1024}.
\end{aligned}$$

Then we back-substitute to obtain

$$\begin{aligned}
[71^{43}]_{1024} &= [71]_{1024} \cdot [945]_{1024} \cdot [193]_{1024} \cdot [769]_{1024}, \\
&= [71]_{1024} \cdot [945]_{1024} \cdot [961]_{1024}, \\
&= [71]_{1024} \cdot [881]_{1024}, \\
&= [87]_{1024}.
\end{aligned}$$

In total we computed 5 squares and 3 binary products, and for each of these we performed a single reduction mod 1024. If we regard each multiplication, squaring and reduction mod 1024 as a single operation, then we used a total of 16 operations, which is indeed less than $4 \cdot \log_2(43) \approx 4 \cdot (5.42) = 21.7$.

In contrast, suppose that someone tells you that $[87]_{1024}$ is a power of $[71]_{1024}$. No one has yet found a method to compute the logarithm that is significantly faster than the brute force method of computing each element the sequence $[1]_{1024}, [71]_{1024}, [71^2]_{1024}, \ldots$ and waiting until we hit $[87]_{1024}$.

The idea of public-key cryptography was proposed by Whitfield Diffie and Martin Hellman in 1976. In this paper they proposed a method that allows two people (not yet called Alice and Bob) to agree on a shared secret number over an insecure channel. This method is now called the **Diffie-Hellman Key Exchange**. Here's how it works:

- Alice and Bob agree publicly on a large prime number p and an invertible element $[g]_p \in (\mathbb{Z}/p\mathbb{Z})^\times$ such that the multiplicative order $\mathrm{ord}_p(g)$ is as large as possible.[17]
- Alice chooses a secret number a and Bob chooses a secret number b.
- Alice transmits $[A]_p = [g^a]_p$ to Bob and Bob transmits $[B]_p = [g^b]_p$ to Alice.
- Alice computes $[K_1]_p = [B^a]_p$ in standard form and Bob computes $[K_2]_p = [A^b]_p$ in standard form.

But now observe that

$$[K_1]_p = [B^a]_p = [(g^b)^a]_p = [(g^a)^b]_p = [A^b]_p = [K_2]_p.$$

Since the elements are in standard form, the uniqueness of remainders implies that $K_1 = K_2$. This number $K := K_1 = K_2$ is the "secret key" that Alice and Bob can now use as the foundation for a symmetric cryptosystem.

Let's investigate why this system is secure. If Eve the eavesdropper is listening to all transmissions between Alice and Bob then she will know the numbers p, g, $A = g^a$ and $B = g^b$ (reduced mod p). To break the system Eve needs to use these numbers to somehow compute $K = g^{ab}$. At present it seems that the only way to do this is to compute the discrete logs of A and B to obtain the exponents a and b, and computing discrete logs is assumed to be computationally expensive.

One weakness of the Diffie-Hellman Key Exchange is that neither of Alice or Bob gets to choose the secret number K in advance, thus it cannot be used to directly transmit messages. Instead, Alice and Bob can use the secret number K as a "key" to set up a symmetric cryptosystem. In the same paper (1976) Diffie and Hellman proposed the idea of "public key cryptography" and "trapdoor functions", but they didn't provide any explicit examples. The Diffie-Hellman Key Exchange was upgraded to a full cryptosystem in 1985 by Taher ElGamal. Here is the **ElGamal Protocol**, which allows everyone (including Bob) to send secret messages to Alice:

[17] We know from Euler's Totient Theorem that $\mathrm{ord}_p(g)$ always divides $\varphi(p) = p-1$. Moreover, one can prove that there always exists an element g with $\mathrm{ord}_p(g) = p-1$; this is called the "primitive root theorem".

- Alice chooses a large prime p and and element $[g]_p \in (\mathbb{Z}/p\mathbb{Z})^\times$ of order $p-1$, just as in the Diffie-Hellman protocol.
- Alice chooses a secret number $0 < a < p$ and computes $[A]_p = [g^a]_p$ in reduced form. She publishes the numbers (p, g, A) as her *public key*. She retains a as her *private key*.
- Bob converts his message to a number $0 < m < p$.[18] To encrypt the message he chooses a secret number $0 < b < p$ and computes the numbers $[B]_p = [g^b]_p$ and $[K]_p = [A^b]_p$ in reduced form. He sends the pair of numbers $([B]_p, [mK]_p)$ to Alice.
- To decrypt the message, Alice first computes the shared secret number $[K]_p = [B^a]_p$, just as in the Diffie-Hellman protocol. Then she uses the Euclidean Algorithm to compute the inverse $[K^{-1}]_p$ and multiplies with the encrypted message $[mK]_p$ to obtain

$$[mK]_p \cdot [K^{-1}]_p = [m]_p \cdot \big([K]_p \cdot [K^{-1}]_p\big) = [m]_p.$$

The ElGamal Protocol is slower than some other public-key cryptosystems (see the RSA Cryptosystem below), however it has the advantage that it can be generalized to other mathematical situations. That is, instead of choosing an element $[g]_p$ in the group of units $(\mathbb{Z}/p\mathbb{Z})^\times$ one can use any element of any group, as long as the group computations can be encoded efficiently in a computer. One popular choice is the "group of rational points on an elliptic curve", which is unfortunately a bit too advanced for this course.

Protocol 2: The RSA Cryptosystem. The most popular public-key cryptosystem was discovered in 1977 by Ron Rivest, Adi Shamir, and Leonard Adelman of MIT. It was also discovered in 1973 by Clifford Cocks working for the UK intelligence agency GCHQ. However, since Cocks' work was classified until 1997, the system is known as RSA.

The security of the RSA Cryptosystem is based on the following assumption, which is the same idea that was proposed by William Stanley Jevons in 1874:

Assumption: factoring integers is hard.

Specifically, if p and q are large prime numbers then it is much easier to multiply them to obtain $n = pq$ than it is to factor n back into p and q. Now let me describe the **RSA Protocol**, which allows everyone (including Bob) to send secret messages to Alice. It is mathematically a bit more sophisticated than ElGamal but it turns out to be more efficient in practice.

- Alice chooses two large prime numbers p and q and computes their product $n = pq$. Then she chooses a random number e that is coprime to $(p-1)(q-1)$ and she publishes the numbers (n, e) as her *public key*.
- Next Alice uses the Euclidean Algorithm to compute the inverse of e mod $(p-1)(q-1)$,

$$[d]_{(p-1)(q-1)} = [e^{-1}]_{(p-1)(q-1)},$$

  and she keeps the secret number d as her *private key*. The individual primes p and q must also be kept secret.

[18] If th message is long he can repeat the process several times.

- Bob converts his message to a number $0 \leqslant m < n$ (or a sequence of numbers of this form) and then he computes the number

$$[c]_n = [m^e]_n$$

in standard form. [Recall that modular exponentiation can be done in logarithmic time.] He sends the "ciphertext" number c to Alice.

- To decode the message, Alice uses her private key d to compute the number

$$[m']_n = [c^d]_n$$

in standard form. I claim that the number $m' = m$ and hence Alice has recovered Bob's secret message.

Proof that $m' = m$. Since e and d are inverses mod $(p-1)(q-1)$ we know that there exists some integer $k \in \mathbb{Z}$ such that

$$de = (p-1)(q-1)k + 1.$$

Now we compute

$$\begin{aligned}
[m']_n &= [c^d]_n \\
&= [c^d]_n \\
&= ([c]_n)^d \\
&= ([m^e]_n)^d \\
&= [m^{de}]_n \\
&= [m^{(p-1)(q-1)k+1}]_n.
\end{aligned}$$

In the likely case that Bob's message m is coprime to $n = pq$ then since $\varphi(n) = (p-1)(q-1)$, Euler's Totient Theorem tells us that

$$\begin{aligned}
[m^{(p-1)(q-1)k+1}]_n &= \left([m^{(p-1)(q-1)}]_n\right)^k \cdot [m]_n \\
&= \left([m^{\varphi(n)}]_n\right)^k \cdot [m]_n \\
&= ([1]_n)^k \cdot [m]_n \\
&= [m]_n,
\end{aligned}$$

and hence $[m']_n = [m]_n$. Then since $0 \leqslant m < n$ and $0 \leqslant m' < n$, the uniqueness of remainders implies that $m' = m$ as desired.

In the unlikely case that Bob's message m is **not** coprime to $n = pq$,[19] then the generalization of Euler's Totient Theorem proved on HW3.6 tells us that the equation

$$[m^{(p-1)(q-1)k+1}]_n = [m]_n$$

[19] Bob doesn't know the individual primes p and q so he has no way to guarantee that this does not happen.

is true anyway. So the RSA Cryptosystem works even when Bob is unlucky. □

Finally, let's discuss why the RSA Cryptosystem is secure. If Eve the eavesdropper is listening to all transmissions between Alice and Bob she knows the numbers n, e and c and she wants to somehow combine these numbers to compute the secret message m. At present it seems that the only way to do this is to first compute the secret number d and then compute $[d^d]_n$, just as Alice does, and since d is the inverse of e mod $(p-1)(q-1)$, Eve will be able to do this if she can find the number $(p-1)(q-1)$. Thus, here is the problem Eve needs to solve:

$$compute\ (p-1)(q-1)\ given\ pq.$$

At present it seems that the only way to do this is to compute the prime factors p and q of pq, which is assumed to be computationally expensive.

Remark: The security of the Diffie-Hellman/ElGamal and RSA systems is based on the assumption that the problems of computing discrete logs and factoring integers are computationally expensive. We do not yet have any mathematical theorems to justify these assumptions.[20] However, armies of well-paid mathematicians have been working on the problem now for a few decades with little success; perhaps that's just as good as a mathematical theorem.

4 Interlude on Quadratic Forms

In Chapter 2 we considered the linear Diophantine equation

(LDE) $$ax + by + c = 0$$

as motivation for the basic concepts of number theory, including the greatest common divisor and the Euclidean Algorithm. We dressed up the topic in modern clothing (using a bit of linear algebra) but the ideas here were completely classical, going back to Euclid's *Elements* (c. 200 BC).

Then in Chapter 3 we extended the discussion to the **finite** number systems $\mathbb{Z}/n\mathbb{Z}$. The fundamental theorems here were Euler's Totient Theorem and the Chinese Remainder Theorem, which gave us the tools to discuss applications of number theory to public-key cryptography. After preliminary work by Fermat and Euler, the theory of "modular arithmetic" was given its modern form in Gauss' *Disquisitiones Arithmeticae* (1798).[21]

The results in Chapters 2 and 3 can be regarded as the essential core of an undergraduate number theory course; after this there is a bit of freedom in the selection of topics. In this course we will spend the rest of our time investigating the general *quadratic Diophantine equation*

(QDE) $$ax^2 + bxy + cy^2 + dx + ey + f = 0$$

[20] There do exist efficient "quantum algorithms" to solve both problems, so both systems will be broken when (and if) quantum computers are developed. Quantum cryptography is a completely different subject.

[21] written when he was 21 and published when he was 24

with integer coefficients $a, b, c, d, e, f \in \mathbb{Z}$. This is a nice topic because in principle the solution is completely understood.[22] However, the solution of (QDE) is much more involved than the solution of (LDE). In particular, it will not fit in a single chapter.

So I will break up the discussion into three separate phases:

The current chapter is mostly algebraic. First we will prove a theorem relating rational solutions of Diophantine equations to "primitive" integer solutions of "homogeneous" Diophantine equations. Then we will discuss how to reduce the solution of the general quadratic equation (QDE) to a collection of standard cases. We will see that the geometric picture of (QDE) is a "conic section" in the real x, y-plane, thus our problem is to find all integer points on a given conic section.

This problem is too difficult to approach directly, so in Chapter 5 we will retreat temporarily to consider **rational points** $(x, y) \in \mathbb{Q}^2$ on the conic section (QDE). It turns out that if we can find **one rational point** then a geometric method (called the Diophantus chord method) will give us an explicit one-to-one correspondence between rational solutions of (QDE) and rational numbers $t \in \mathbb{Q}$. The question of whether any rational points exist is completely solved by *Legendre's Theorem* and *Quadratic Reciprocity*, however it more difficult to actually **find** a rational point.

Finally, in Chapter 6 we return to the problem of integer points on conics. Lagrange showed that in the worst case scenario (QDE) can be reduced to a pair of equations of the form

$$x^2 - \Delta y^2 = k \tag{PE}$$

where $\Delta = b^2 - 4ac$ and $k \in \mathbb{Z}$. The equation (PE) was solved several times throughout history. The last European to rediscover the solution was Fermat in 1657 and he posed it as a challenge to other mathematicians. Lord Brouncker described a solution in terms of *continued fractions*, which Euler misattributed to John Pell, hence we know (PE) as *Pell's equation*. Lagrange gave the first rigorous proof that Brouncker's algorithm always terminates and then Dirichlet gave a shorter (nonconstructive) proof of existence of solutions using his "pigeonhole principle" (German: *Schubfachprinzip*). The search for a deeper understanding of (PE) leads to *class field theory* but I'm sure we won't get that far.

4.1 Rational Versus Integer Solutions

Our first task is to distinguish clearly between **rational** and **integer** solutions of Diophantine equations. To begin, let's recall from Chapter 2 how we dealt with the linear Diophantine equation in two variables:

$$ax + by + c = 0. \tag{LDE}$$

If a or b is zero then (LDE) is an equation in one variable whose solution is trivial. Therefore we will assume that a and b are both nonzero with greatest common divisor $d := \gcd(a, b)$.

[22] Diophantine equations of degree $\geqslant 3$ are a different matter.

If $d \nmid c$ then there is no solution so we will assume that $d|c$. When $c \neq 0$ then the equation (LDE) is called *inhomogeneous*. In order to find the complete solution of the inhomogeneous equation we reduce it to a pair of related *homogeneous* equations:

- We can use a bit of linear algebra to show that the complete solution of $ax + by + c = 0$ is deteremined by **one particular solution** $ax' + by' + c = 0$ together with the complete solution of the associated homogeneous equation in two variables:

  (HLDE) $$ax + by = 0.$$

  The general solution of $ax + by = 0$ is easy to find using Euclid's Lemma.

- Finding a particular solution $ax' + by' + c = 0$ is a bit more challenging. In order to do this we consider the associated homogeneous equation in **three** variables:

  (HLDE') $$ax + by + cz = 0.$$

  Since the collection of solutions $(x, y, z) \in \mathbb{Z}^3$ to (HDE') is closed under vector addition and scalar multiplication by integers, we are able to combine simple solutions via the Euclidean Algorithm in order to obtain a solution of the desired form $(x', y', 1)$.

The main benefit of reducing the inhomogeneous equation (LDE) to the pair of homogeneous equations (HLDE) and (HLDE') is that homogeneous equations are amenable to the techniques of **linear algebra**, and linear algebra is something that we humans are good at.

The same general strategy can be applied to Diophantine equations of arbitrary degree, with varying degrees of success. For example, in order to solve the quadratic Diophantine equation

(QDE) $$ax^2 + bxy + cy^2 + dx + ey + f = 0$$

we will focus our attention on the associated homogeneous equation in two variables

(HQDE) $$ax^2 + bxy + cy^2 = 0$$

and the associated homogeneous equation in three variables

(HQDE') $$ax^2 + bxy + cy^2 + dxz + eyz + fz^2 = 0,$$

both of which are amenable to techniques of linear algebra.

Before moving on to the analysis of (HQDE) and (HQDE'), let me define some terminology for the general situation. A general Diophantine equation has the form

(DE) $$f(x_1, \ldots, x_n) = 0$$

where $f(x_1, \ldots, x_n) \in \mathbb{Z}[x_1, \ldots, x_n]$ is a polynomial in n variables with integer coefficients. What does it mean to say that f is a "homogeneous" polynomial?

Definition of Homogeneous Polynomials. Consider a polynomial in n variables with integer coefficients[23]

$$f(x_1, \ldots, x_n) \in \mathbb{Z}[x_1, \ldots, x_n].$$

We say that f is *homogeneous of degree* d if for any number λ we have

$$f(\lambda x_1, \ldots, \lambda x_n) = \lambda^d f(x_1, \ldots, x_n).$$

In older books, homogeneous polynomials are called *forms*. Thus you might see homogeneous polynomials of degrees 1, 2 and 3 referred to *linear forms*, *quadratic forms* and *ternary forms*, respectively. //

For example, consider the polynomial $f(x, y) = ax^2 + bxy + cy + d$. By making the substitution $(x, y) \mapsto (\lambda x, \lambda y)$ we obtain

$$\begin{aligned} f(\lambda x, \lambda y) &= a(\lambda x)^2 + b(\lambda x)(\lambda y) + c(\lambda y) + d \\ &= \lambda^2(ax^2 + bxy) + \lambda^1(cy) + \lambda^0(d). \end{aligned}$$

This suggests that we should define the auxiliary polynomials

$$\begin{aligned} f_2(x, y) &:= ax^2 + bxy, \\ f_1(x, y) &:= cy, \\ f_0(x, y) &:= d, \end{aligned}$$

where each $f_i(x, y)$ is homogeneous of degree i. Then we can express $f(x, y)$ as the sum of its homogeneous parts:

$$f(x, y) = f_2(x, y) + f_1(x, y) + f_0(x, y)$$

In this case we say that the non-homogeneous polynomial $f(x, y)$ has degree 2 and we say that the homogeneous polynomial (quadratic form) $f_2(x, y)$ is its *leading form*.

More generally, given any polynomial $f(\mathbf{x}) := f(x_1, \ldots, x_n) \in \mathbb{Z}[x_1, \ldots, x_n]$, we can express $f(\mathbf{x})$ uniquely as a sum of homogeneous polynomials

$$f(\mathbf{x}) = f_d(\mathbf{x}) + f_{d-1}(\mathbf{x}) + \cdots + f_1(\mathbf{x}) + f_0(\mathbf{x}),$$

where $f_i(\mathbf{x}) \in \mathbb{Z}[x_1, \ldots, x_n]$ is a homogeneous polynomial of degree i called the *i-th homogeneous part of f*. If $f_d(\mathbf{x})$ is not the *zero polynomial* then we say that $f(\mathbf{x})$ has degree d and we call $f_d(\mathbf{x})$ the leading form. Furthermore, for any number λ, the change of variables $\mathbf{x} = (x_1, \ldots, x_n) \mapsto (\lambda x_1, \ldots, \lambda x_n) = \lambda \mathbf{x}$ has the folloing result:

$$\begin{aligned} f(\lambda \mathbf{x}) &= f_d(\lambda \mathbf{x}) + f_{d-1}(\lambda \mathbf{x}) + \cdots + f_1(\lambda \mathbf{x}) + f_0(\lambda \mathbf{x}) \\ &= \lambda^d f_d(\mathbf{x}) + \lambda^{d-1} f_{d-1}(\mathbf{x}) + \cdots + \lambda^1 f_1(\mathbf{x}) + \lambda^0 f_0(\mathbf{x}). \end{aligned}$$

[23] We could also define homogeneous polynomials with coefficients in an arbitrary commutative ring.

Now there are two obvious ways we can convert the Diophantine equation (DE) into a homogeneous equation. On the one hand, we could just focus on the leading form:

$$f_d(\mathbf{x}) = 0. \tag{HDE}$$

This equation is useful, but it is not sufficient to solve (DE) because it loses information. On the other hand, we can preserve more information by promoting (DE) to a homogeneous Diophantine equation with **one extra variable**, which we call x_{n+1}. To be precise, we substitute $\lambda = 1/x_{n+1}$ into the above equation to obtain

$$\begin{aligned} f(\lambda\mathbf{x}) &= \lambda^d f_d(\mathbf{x}) + \lambda^{d-1} f_{d-1}(\mathbf{x}) + \cdots + \lambda^1 f_1(\mathbf{x}) + \lambda^0 f_0(\mathbf{x}) \\ f\left(\frac{x_1}{x_{n+1}}, \ldots, \frac{x_n}{x_{n+1}}\right) &= \frac{1}{x_{n+1}^d} f_d(\mathbf{x}) + \frac{1}{x_{n+1}^{d-1}} f_{d-1}(\mathbf{x}) + \cdots + \frac{1}{x_{n+1}} f_1(\mathbf{x}) + f_0(\mathbf{x}) \end{aligned}$$

and then after multiplying both sides by x_{n+1}^d we obtain a homogeneous degree d polynomial in the variables $x_{n+1}, x_1, \ldots, x_n$ called the *homogenization* of f:

$$\begin{aligned} F(x_1, \ldots, x_n, x_{n+1}) &:= x_0^d \cdot f\left(\frac{x_1}{x_0}, \ldots, \frac{x_n}{x_0}\right) \\ &= f_d(x_1, \ldots, x_n) + x_0 \cdot f_{d-1}(x_1, \ldots, x_n) + \cdots + x_0^d \cdot f_0(x_1, \ldots, x_n). \end{aligned}$$

Exercise: Check that $F(x_1, \ldots, x_n, x_{n+1}) \in \mathbb{Z}[x_1, \ldots, x_n, x_{n+1}]$ is indeed homogeneous.

To illustrate this definition, consider the non-homogeneous polynomial $f(x, y) = f_2(x, y) + f_1(x, y) + f_0(x, y) = (ax^2 + bxy) + (cy) + (d)$ from our example above. Then its "homogenization" is given by

$$F(x, y, z) = (ax^2 + bxy) + z(cy) + z^2(d) = ax^2 + bxy + cyz + dz^2,$$

which is homogeneous of degree 2. To recover the original polynomial (that is, to *dehomogenize*) we just substitute $z = 1$ to obtain

$$F(x, y, 1) = f(x, y).$$

The main goal of this section is to describe a close relationship between the following concepts:

- Rational solutions of non-homogeneous Diophantine equations.
- Integer solutions of homogeneous Diophantine equations.

We begin with the following definition and theorem on integer solutions of homogeneous equations.

Definition of GCD and Primitive Vectors. Consider a nonzero integer vector

$$(0, \ldots, 0) \neq (a_1, \ldots, a_n) \in \mathbb{Z}^n$$

and let $\mathrm{Div}(a_1, \ldots, a_n) = \{d \in \mathbb{Z} : \forall i, d|a_i\}$ be the set of common divisors. Observe that $1 \in \mathrm{Div}(a_1, \ldots, a_n)$ and, since the integers a_i are not all zero, the set $\mathrm{Div}(a_1, \ldots, a_n)$ is bounded above by the minimum of the absolute values $|a_1|, \ldots, |a_n|$. Thus by Well-Ordering there exists a greatest element of the set which we call the *greatest common divisor*:

$$1 \leqslant \gcd(a_1, \ldots, a_n) \leqslant \min\{|a_1|, \ldots, |a_n|\}.$$

We say that the vector $(a_1, \ldots, a_n) \in \mathbb{Z}^n$ is *primitive* when

$$\gcd(a_1, \ldots, a_n) = 1.$$

//

Theorem (Unique Primitive Reduction of Homogeneous Equations). Consider a homogeneous polynomial

$$F(x_1, \ldots, x_n) \in \mathbb{Z}[x_1, \ldots, x_n]$$

of degree d. Note that the zero vector is always a solution because we have

$$F(0, \ldots, 0) = F(0 \cdot x_1, \ldots, 0 \cdot x_n) = 0^d \cdot F(x_1, \ldots, x_n) = 0$$

for any values of $x_1, \ldots, x_n \in \mathbb{Z}$. Now consider an arbitrary nonzero integer vector $(0, \ldots, 0) \neq (a_1, \ldots, a_n) \in \mathbb{Z}^n$ such that

$$F(a_1, \ldots, a_n) = 0.$$

I claim that the exists a unique positive integer $1 \leqslant \lambda \in \mathbb{Z}$ and a unique primitive vector $(a'_1, \ldots, a'_n) \in \mathbb{Z}^n$ such that

- $(a_1, \ldots, a_n) = \lambda \cdot (a'_1, \ldots, a'_n)$
- $F(a'_1, \ldots, a'_n) = 0$.

That is, every integer solution of $F(\mathbf{x}) = 0$ can be written uniquely as a positive multiple of a **primitive** integer solution. //

The proof will require the following Lemma, which it interesting enough to have its own name.

Lemma (Vector Bézout Identity). For every nonzero integer vector $(0, \ldots, 0) \neq (a_1, \ldots, a_n) \in \mathbb{Z}^n$ there exist integers $x_1, \ldots, x_n \in \mathbb{Z}$ such that

$$\gcd(a_1, \ldots, a_n) = a_1 x_1 + \cdots a_n x_n.$$

//

Proof of the Lemma. We already know that this statement is true for $n = 2$ because of the Vector Euclidean Algorithm. Now we will **assume for induction** that the statement is true for some $n = k \geqslant 2$, and in this case we will prove that the statement is true for $n = k + 1$.

So consider an arbitrary nonzero integer vector

$$(0,\ldots,0) = (a_1,\ldots,a_{k+1}) \in \mathbb{Z}^{k+1}.$$

In this case I claim that we have

$$(*) \qquad \gcd(a_1,\ldots,a_k) = \gcd(\gcd(a_1,\ldots,a_k),a_k).$$

To see this we will show that the following sets of common divisors are equal, since then their **greatest elements** will also be equal:

$$\mathrm{Div}(a_1,\ldots,a_{k+1}) = \mathrm{Div}(\gcd(a_1,\ldots,a_k),a_{k+1}).$$

So consider any integer $d \in \mathbb{Z}$. If d is in the right hand set then we have $d|a_{k+1}$ and $d|\gcd(a_1,\ldots,a_k)$. Since d divides some common divisor of $a_1,\ldots,a_k$ it must divide each a_i individually, and we conclude that d is in $\mathrm{Div}(a_1,\ldots,a_{k+1})$ as desired. Conversely, suppose that d is a common divisor of $a_1,\ldots,a_{k+1}$. We have assumed for induction that there exist integers $y_1,\ldots,y_k \in \mathbb{Z}$ such that

$$(**) \qquad \gcd(a_1,\ldots,a_k) = a_1y_1 + \cdots a_ky_k.$$

Then since $d|a_i$ for all $1 \leqslant i \leqslant k$ we see from equation $(**)$ that $d|\gcd(a_1,\ldots,a_k)$ and it follows that d is in the right hand set as desired.

Finally, from equation $(*)$ and the Euclidean Algorithm (EA) there exist integers $x,y \in \mathbb{Z}$ such that

$$\begin{aligned}
\gcd(a_1,\ldots,a_{k+1}) &= \gcd(\gcd(a_1,\ldots,a_k),a_{k+1}) && (*)\\
&= \gcd(a_1,\ldots,a_k)x + a_{k+1}y && (\mathrm{EA})\\
&= (a_1y_1 + \cdots + a_ky_k)x + a_{k+1}y && (**)\\
&= a_1(y_1x) + \cdots a_k(y_kx) + a_{k+1}y.
\end{aligned}$$

We conclude that the statement of the theorem is true for $n = k+1$, which completes the proof by induction. □

Proof of the Theorem. Let $F(x_1,\ldots,x_n) \in \mathbb{Z}[x_1,\ldots,x_n]$ be homogeneous of degree d and consider a nonzero integer solution

$$F(a_1,\ldots,a_n) = 0.$$

Let $1 \leqslant \lambda = \gcd(a_1,\ldots,a_n)$ so that we have $a_i = \lambda a_i'$ for some integers a_i'. Note that the vector $(a_1',\ldots,a_n')$ is primitive since if $\varepsilon > 1$ is any common divisors of $a_1',\ldots,a_n'$ then $\lambda\varepsilon > \lambda$ is a common divisor of $a_1,\ldots,a_n$, which contradicts the fact that λ was the **greatest** common divisor. Furthermore, since $\lambda \neq 0$ (and also $\lambda^d \neq 0$) we have

$$\begin{aligned}
F(a_1,\ldots,a_n) &= F(\lambda a_1',\ldots,\lambda a_n')\\
0 &= \lambda^d \cdot F(a_1',\ldots,a_n')\\
0 &= F(a_1',\ldots,a_n').
\end{aligned}$$

We have shown that the integer solution $(a_1, \ldots, a_n) \in \mathbb{Z}^n$ can be expressed as

$$(a_1, \ldots, a_n) = \lambda \cdot (a'_1, \ldots, a'_n)$$

where $1 \leqslant \lambda \in \mathbb{Z}$ and whenre $(a'_1, \ldots, a'_n) \in \mathbb{Z}^n$ is a **primitive** integer solution.

It only remains to show that this expression is unique. So suppose that we have another positive integer $1 \leqslant \mu \in \mathbb{Z}$ and another primitive vector $(a''_1, \ldots, a''_n) \in \mathbb{Z}^n$ such that

$$(\lambda a'_1, \ldots, \lambda a'_n) = (a_1, \ldots, a_n) = (\mu a''_1, \ldots, \mu a''_n)$$

Since $\gcd(a'_1, \ldots, a'_n) = 1$ it follows from the previous lemma that there exist integers $y_1, \ldots, y_n \in \mathbb{Z}$ such that $1 = a'_1 y_1 + \cdots a'_n y_n$. Then multiplying both sides by λ gives

$$\begin{aligned} \lambda &= \lambda(a'_1 y_1 + \cdots a'_n y_n) \\ &= (\lambda a'_1) y_1 + \cdots (\lambda a'_n) y_n \\ &= (\mu a''_1) y_1 + \cdots (\mu a''_n) y_n \\ &= \mu(a''_1 y_1 + \cdots + a''_n y_n). \end{aligned}$$

It follows that $\mu | \lambda$, and a similar argument shows that $\lambda | \mu$. In other words, there exist integers $k, \ell \in \mathbb{Z}$ such that $\lambda = k\mu$ and $\mu = \ell\lambda$. Since $\lambda \neq 0$ this implies that

$$\begin{aligned} \lambda &= k\mu \\ \lambda &= k\ell\lambda \\ (1 - k\ell)\lambda &= 0 \\ (1 - k\ell) &= 0 \\ 1 &= k\ell, \end{aligned}$$

and hence we have either $k = \ell = 1$ or $k = \ell = -1$. But since $\lambda = k\mu$ and since λ and μ are both **positive** we must have $k = \ell = 1$ and hence $\lambda = \mu$. Finally, by cancelling the non-zero factor λ in the equations

$$\begin{aligned} \lambda a'_1 = \mu a''_1 &= \lambda a''_1 \\ &\vdots \\ \lambda a'_n = \mu a''_1 &= \lambda a''_n \end{aligned}$$

we conclude that $(a'_1, \ldots, a'_n) = (a''_1, \ldots, a''_n)$ as desired. □

We have shown that the complete **integer** solution of a homogeneous Diophantine equation is determined by its **primitive** solutions. The next theorem shows that **rational** solutions of **non-homogeneous** Diophantine equations are determined by **primitive** solutions of the homogenized equation.

Theorem (Rational Versus Integer Solutions). Consider a polynomial in n variables with integer coefficients

$$f(x_1, \ldots, x_n) \in \mathbb{Z}[x_1, \ldots, x_n].$$

If f (which is not necessarily homogeneos) has degree d then we define its homogenization by

$$F(x_1,\ldots,x_n,x_{n+1}) := x_{n+1}^d \cdot f\left(\frac{x_1}{x_{n+1}},\ldots,\frac{x_n}{x_{n+1}}\right) \in \mathbb{Z}[x_1,\ldots,x_n,x_{n+1}].$$

I claim that **rational solutions** $(x_1,\ldots,x_n) \in \mathbb{Q}^n$ of the non-homogeneous equation

$$f(x_1,\ldots,x_n) = 0 \tag{DE}$$

are in one-to-one correspondence with **primitive integer solutions** $(x_1,\ldots,x_n,x_{n+1}) \in \mathbb{Z}^{n+1}$ of the homogenized equation

$$F(x_1,\ldots,x_n,x_{n+1}) = 0 \tag{HDE'}$$

in which $x_{n+1} \geqslant 1$. //

Proof. Consider an arbitrary integer solution $F(a_1,\ldots,a_n,b) = 0$ with $\gcd(a_1,\ldots,a_n,b) = 1$ and $b \geqslant 1$. Since $b \neq 0$ (and $b^d \neq 0$) it follows that

$$\begin{aligned} F(a_1,\ldots,a_n,b) &= b^d \cdot f(a_1/b,\ldots,a_n/b) \\ 0 &= c^d \cdot f(a_1/b,\ldots,a_n/b) \\ 0 &= f(a_1/b,\ldots,a_n/b) \end{aligned}$$

and hence we have found a rational solution $(x_1,\ldots,x_n) = (a_1/b,\ldots,a_n/b)$ of the equation $f(x_1,\ldots,x_n) = 0$. I claim that the mapping $\mathbb{Z}^{n+1} \to \mathbb{Q}^n$ defined by

$$(a_1,\ldots,a_n,b) \mapsto (a_1/b,\ldots,a_n/b)$$

is the desired one-to-one correspondence. There are two things to show:

(1) The map is "onto". Consider an arbitrary rational solution $f(x_1,\ldots,x_n) = 0$. By finding a common denominator we can write $(x_1,\ldots,x_n) = (a_1/b,\ldots,a_n/b)$ for **some** integers $(a_1,\ldots,a_n,b) \in \mathbb{Z}^{n+1}$ with $b \geqslant 1$. Then since $b \neq 0$ (and $b^d \neq 0$) we have

$$F(a_1,\ldots,a_n,b) = b^d \cdot f(a_1/b,\ldots,a_n/b) = b^d \cdot f(x_1,\ldots,x_n) = 0$$

so that $(a_1,\ldots,a_n,b) \in \mathbb{Z}^{n+1}$ is an integer solution of $F(a_1,\ldots,a_n,b) = 0$. It follows from the previous theorem that there exists an expression $(a_1,\ldots,a_n,b) = \lambda \cdot (a_1',\ldots,a_n')$ with $1 \leqslant \lambda \in \mathbb{Z}$ and $\gcd(a_1,\ldots,a_n,b) = 1$ such that $F(a_1',\ldots,a_n',b') = 0$, and since $b = \lambda b'$ with $b \geqslant 1$ and $\lambda \geqslant 1$ we must have $b' \geqslant 1$. Finally, observe that the primitive integer solution $(a_1',\ldots,a_n',b')$ gets sent under our map to $(a_1'/b',\ldots,a_n'/b') = (a_1/b,\ldots,a_1/b) = (x_1,\ldots,x_n)$ as desired.

(2) The map is "one-to-one". Suppose that we can write

$$(a_1/b,\ldots,a_n/b) = (x_1,\ldots,x_n) = (a_1'/b',\ldots,a_n'/b')$$

for some integers $a_i, b, a_i', b' \in \mathbb{Z}$ satisfying

- $\gcd(a_1, \ldots, a_n, b) = \gcd(a'_1, \ldots, a'_n, b') = 1$,
- $b \geqslant 1$ and $b' \geqslant 1$.

To show that $(a_1, \ldots, a_n, b) = (a'_1, \ldots, a'_n, b')$ we will follow a similar strategy to the proof of uniqueness in the previous theorem. Since $\gcd(a_1, \ldots, a_n, b) = 1$ the Vector Bézout Identity says that there exist integers $x_1, \ldots, x_n, y \in \mathbb{Z}$ such that

$$\begin{aligned}
1 &= a_1 x_+ \cdots a_n x_n + by \\
b' &= b'(a_1 x_1 + \cdots a_n x_n + by) \\
b' &= (b' a_1) x_1 + \cdots + (b' a_n) x_n + b'(by) \\
b' &= (b a'_1) x_1 + \cdots + (b a'_n) x_n + b(b' y) \\
b' &= b(a'_1 x_1 + \cdots a'_n x_n + b' y),
\end{aligned}$$

and we conclude that $b | b'$. A similar argument shows that $b' | b$ then since b and b' are both **positive** we must have $b = b'$. Finally, since $b \neq 0$ and since $b a'_i = b' a_i = b a_i$ for all i we conclude that $a_i = a'_i$ for all i as desired. □

Thus we have shown that the problem of finding **rational** solutions to **general** Diophantine equations is equivalent to the problem of finding **integer** solutions fo **homogeneous** Diophantine equations. We can approach both of these problems with linear algebra. Unfortunately, the problem of finding **integer** solutions to **general** Diophantine equations is more difficult because it it not as susceptible to linear algebraic techniques.

4.2 A Moderate Amount of Linear Algebra

In the previous section we proved that rational solutions of the general quadratic Diophantine equation

(QDE) $$ax^2 + bxy + cy^2 + dx + ey + f = 0$$

are equivalent to primitive integer solutions of the homogenized equation

(HQDE') $$ax^2 + bxy + cy^2 + dxz + eyz + fz^2 = 0$$

in which $z \geqslant 1$. In this section we will use a moderate amount of linear algebra in order to reduce the equation (QDE) to a small number of standard forms and in the next section we will deal with (QDE").

The first step is to express the equation (QDE) in the language of matrix multiplication. Observe that we have

$$\begin{aligned}
ax^2 + bxy + cy^2 + dx + ey + f &= 0 \\
\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} d & e \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + f &= 0 \\
\mathbf{x}^T A \mathbf{x} + \mathbf{d}^T \mathbf{x} + f &= 0,
\end{aligned}$$

where the 2×1 column vectors $\mathbf{d}, \mathbf{x}$ and the 2×2 matrix A are defined by

$$\mathbf{d} = \begin{pmatrix} d \\ e \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}, \quad \text{and} \quad A = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}.$$

There is some non-uniqueness in the choice of the matrix A. Indeed, it would be equally correct to choose

$$A = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \qquad \text{or} \qquad A = \begin{pmatrix} a & 0 \\ b & c \end{pmatrix}.$$

I made the choice I did because I want A to be a **symmetric matrix**, i.e., $A^T = A$. You might worry about the entry $b/2$ if we are working over the integers, but we will see later that this is not a big problem.

Now that we have expressed the equation (QDE) in terms of matrices and vectors, it makes sense to look for a change of variables $\mathbf{x} = (x, y) \mapsto (x', y') = \mathbf{x}'$ that can also be expressed in this language. In general we will consider so-called *affine transformations*, which have the form

(AT)
$$\mathbf{x} = P\mathbf{x}' + \mathbf{u}$$
$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} p & q \\ r & s \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} + \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} px' + qy' + u \\ rx' + sy' + v \end{pmatrix}.$$

Later we may require the numbers p, q, r, s, u, v to be integers or rational numbers but for now they can be arbitrary. For the next step, recall that the matrix transpose satisfies $(M+N)^T = M^T + N^T$ and $(MN)^T = N^T M^T$ whenever the matrix sum and product are defined. Then substituting the change of variables $\mathbf{x} = P\mathbf{x}' + \mathbf{u}$ into the equation (QDE) and using a moderate amount of matrix arithmetic yields

$$\begin{aligned}
\mathbf{x}^T A\mathbf{x} + \mathbf{d}^T\mathbf{x} + f &= 0 \\
(P\mathbf{x}' + \mathbf{u})^T A(P\mathbf{x}' + \mathbf{u}) + \mathbf{d}^T(P\mathbf{x}' + \mathbf{u}) + f &= 0 \\
((\mathbf{x}')^T P^T + \mathbf{u}^T)A(P\mathbf{x}' + \mathbf{u}) + \mathbf{d}^T(P\mathbf{x}' + \mathbf{u}) + f &= 0 \\
(\mathbf{x}')^T(P^T AP)\mathbf{x}' + (\mathbf{x}')^T P^T A\mathbf{u} + \mathbf{u}^T AP\mathbf{x}' + \mathbf{d}^T P\mathbf{x}' + \mathbf{u}^T A\mathbf{u} + \mathbf{d}^T\mathbf{u} + f &= 0 \\
(\mathbf{x}')^T(P^T AP)\mathbf{x}' + \left[(\mathbf{x}')^T P^T A\mathbf{u}\right]^T + \mathbf{u}^T AP\mathbf{x}' + \mathbf{d}^T P\mathbf{x}' + \mathbf{u}^T A\mathbf{u} + \mathbf{d}^T\mathbf{u} + f &= 0 \qquad (!) \\
(\mathbf{x}')^T(P^T AP)\mathbf{x}' + \mathbf{u}^T A^T P\mathbf{x}' + \mathbf{u}^T AP\mathbf{x}' + \mathbf{d}^T P\mathbf{x}' + \mathbf{u}^T A\mathbf{u} + \mathbf{d}^T\mathbf{u} + f &= 0 \\
(\mathbf{x}')^T(P^T AP)\mathbf{x}' + \mathbf{u}^T AP\mathbf{x}' + \mathbf{u}^T AP\mathbf{x}' + \mathbf{d}^T P\mathbf{x}' + \mathbf{u}^T A\mathbf{u} + \mathbf{d}^T\mathbf{u} + f &= 0 \qquad (!!) \\
(\mathbf{x}')^T(P^T AP)\mathbf{x}' + \left[\left(2\mathbf{u}^T A + \mathbf{d}^T\right)P\right]\mathbf{x}' + \left(\mathbf{u}^T A\mathbf{u} + \mathbf{d}^T\mathbf{u} + f\right) &= 0 \\
(\mathbf{x}')^T(P^T AP)\mathbf{x}' + \left[P^T(2A\mathbf{u} + \mathbf{d})\right]^T\mathbf{x}' + \left(\mathbf{u}^T A\mathbf{u} + \mathbf{d}^T\mathbf{u} + f\right) &= 0 \qquad (!!) \\
(\mathbf{x}')^T A'\mathbf{x}' + (\mathbf{d}')^T\mathbf{x}' + f' &= 0 \qquad \text{(QDE')}
\end{aligned}$$

where the 2×2 matrix A', the 2×1 column vector $\mathbf{d}'$ and the 1×1 number f' are defined by

$$A' = P^T AP, \quad \mathbf{d}' = P^T(2A\mathbf{u} + \mathbf{d}), \qquad \text{and} \qquad f' = \mathbf{u}^T A\mathbf{u} + \mathbf{d}^T\mathbf{u} + f.$$

Observe that in step (!) I used the fact that $(\mathbf{x}')^T P^T A\mathbf{u} = \left[(\mathbf{x}')^T P^T A\mathbf{u}\right]^T$, which is true because $(\mathbf{x}')^T P^T A\mathbf{u}$ is just a 1×1 matrix (i.e., a "number") and every 1×1 matrix is equal to its own transpose. The steps labeled (!!) are true because $A^T = A$.

In summary, we began with the quadratic equation

$$\text{(QDE)} \qquad \mathbf{x}^T A\mathbf{x} + \mathbf{d}^T\mathbf{x} + f = 0.$$

Then we made the affine transformation

$$\text{(AT)} \qquad \mathbf{x} = P\mathbf{x}' + \mathbf{u}$$

and substituted this into (QDE) to obtain the transformed equation

$$\text{(QDE')} \qquad (\mathbf{x}')^T A'\mathbf{x}' + (\mathbf{d}')^T\mathbf{x}' + f' = 0$$

where

$$A' = P^T AP, \quad \mathbf{d}' = P^T(2A\mathbf{u} + \mathbf{d}), \quad \text{and} \quad f' = \mathbf{u}^T A\mathbf{u} + \mathbf{d}^T\mathbf{u} + f.$$

If the matrix P is invertible then the affine transformation (AT) is also invertible, with inverse given by

$$\text{(AT')} \qquad \mathbf{x}' = P^{-1}(\mathbf{x} - \mathbf{u}),$$

and this case we observe that

$$\mathbf{x} \text{ is a solution of (QDE)} \quad \Longleftrightarrow \quad \mathbf{x}' \text{ is a solution of (QDE')}.$$

The goal now is to choose an invertible matrix P and a vector $\mathbf{u}$ so that the equation (QDE') is as simple as possible. If we can find the complete solution of (QDE') then we will obtain the complete solution of the original (QDE) after applying the change of variables (AT).

So far all of this is pure algebra that holds over any commutative ring. However, if we are looking for solutions in a specific ring (such as $\mathbb{R}$, $\mathbb{Q}$ or $\mathbb{Z}$) then we will need to place restrictions on the matrix P and the vector $\mathbf{u}$.

Observations:

- *Real Case.* Suppose that P and $\mathbf{u}$ have entries in $\mathbb{R}$. If P is invertible then its inverse is given by

  $$P^{-1} = \begin{pmatrix} p & q \\ r & s \end{pmatrix}^{-1} = \frac{1}{ps - qr}\begin{pmatrix} s & -q \\ -r & p \end{pmatrix},$$

  which also has entries in $\mathbb{R}$. In this case we see that

  $$\mathbf{x} \in \mathbb{R}^2 \quad \Leftrightarrow \quad \mathbf{x}' \in \mathbb{R}^2,$$

  so that the real solutions of (QDE) correspond to the real solutions of (QDE'). We will examine this case below to get a feeling for the geometry of the problem.

- *Rational Case.* Suppose that P and $\mathbf{u}$ have entries in $\mathbb{Q}$. If P is invertible then we see from the formula above that P^{-1} also has entries in $\mathbb{Q}$ and we conclude that

$$\mathbf{x} \in \mathbb{Q}^2 \quad \Leftrightarrow \quad \mathbf{x}' \in \mathbb{Q}^2.$$

  Thus the rational solutions of (QDE) correspond to the rational solutions of (QDE'). The reduction in this case is only a bit trickier than the real case.

- *Integer Case.* Suppose that P and $\mathbf{u}$ have entries in $\mathbb{Z}$. Suppose furthermore that P is invertible and that the inverse P^{-1} has entries in $\mathbb{Z}$. (By the formula above this is equivalent to having $ps - qr = \pm 1$.) In this case we see that

$$\mathbf{x} \in \mathbb{Z}^2 \quad \Leftrightarrow \quad \mathbf{x}' \in \mathbb{Z}^2,$$

  so the integer solutions of (QDE) correspond to the integer solutions of (QDE'). This case is much trickier and we will postpone a full discussion until the next chapter.

//

General Strategy: With these observations in mind, here is the general strategy that we will pursue. Consider a ring $K \in \{\mathbb{R}, \mathbb{Q}, \mathbb{Z}\}$. To find solutions $\mathbf{x} \in K^2$ of the equation (QDE) we perform the following steps:

- First we search for a matrix P such that P and P^{-1} both have entries in K and such that the matrix A' is "diagonal":

$$A' = P^T A P = \begin{pmatrix} a' & 0 \\ 0 & c' \end{pmatrix}.$$

  This has the effect of eliminating the xy-term from (QDE'):

$$a'(x')^2 + c'(y')^2 + (\mathbf{d}')^T \mathbf{x}' + f = 0.$$

  We will find that this is **always** possible when $K \in \{\mathbb{R}, \mathbb{Q}\}$. In the case $K = \mathbb{R}$ the desired matrix P is just a rotation of the plane. In the case $K = \mathbb{Z}$ it is not always possible.

- Then we search for a vector $\mathbf{u}$ with entries in K such that

$$\mathbf{d}' = P^T(2A\mathbf{u} + \mathbf{d}) = \mathbf{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

  This has the effect of eliminating the x-term and y-term from (QDE'):

(QDE") $$a'(x')^2 + c'(y')^2 + f' = 0.$$

  We will find that this is **almost always** possible when $K \in \{\mathbb{R}, \mathbb{Q}\}$. Algebraically we can think of it as "completing the squares". Geometrically it is a translation of the plane.

- Finally, we attempt to characterize the full solution $\mathbf{x} \in K^2$ of the equation (QDE"). If we can do this then we obtain the full solution of (QDE) by inverting the change of variables.

//

The rest of the section is devoted to carrying out strategy, first for $K = \mathbb{R}$ and then for $K = \mathbb{Q}$.

Reduction of (QDE) for real numbers. There is a general theorem of linear algebra that says the following:

> *Principal Axes Theorem.* Let A be a real $n \times n$ matrix. If A is symmetric (i.e., if $A^T = A$) then there exists a real $n \times n$ matrix P with the following properties
>
> - P is invertible with inverse equal to its transpose (i.e., $P^{-1} = P^T$)
> - $P^T AP$ is diagonal.

We will not use this theorem, but at least it tells us what kind of solution to look for. The real 2×2 matrices P satisfying $P^{-1} = P^T$ are just the *reflections* and *rotations* of the plane that leave the origin fixed. In particular, for any angle θ the matrix

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

has the effect of rotating every point around the origin, counterclockwise by θ. So let us assume that $P = R_\theta$ and see if we can find an angle θ such that $P^T AP$ is diagonal. I will temporarily use the notation $C := \cos\theta$ and $S := \sin\theta$ to save space.[24] Then we have

$$\begin{aligned} P^T AP &= \begin{pmatrix} C & S \\ -S & C \end{pmatrix} \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \begin{pmatrix} C & -S \\ S & C \end{pmatrix} \\ &= \begin{pmatrix} aC^2 + bSC + cS^2 & (c-a)SC + \frac{b}{2}(C^2 - S^2) \\ (c-a)SC + \frac{b}{2}(C^2 - S^2) & aS^2 - bSC + cC^2 \end{pmatrix}. \end{aligned}$$

Observe that this matrix is diagonal if and only if

$$\begin{aligned} (c-a)SC + \frac{b}{2}(C^2 - S^2) &= 0 \\ (c-a)\sin\theta\cos\theta + \frac{b}{2}(\cos^2\theta - \sin^2\theta) &= 0 \\ \frac{c-a}{2}\sin(2\theta) + \frac{b}{2}\cos(2\theta) &= 0 \\ (a-c)\sin(2\theta) &= b\cos(2\theta). \end{aligned}$$

If $b = 0$ then $\theta = 0$ is a solution. Indeed, in this case the matrix A is already diagonal so we only need to "rotate by zero". If $(a - c) = 0$ then $\theta = \pi/4$ is a solution since in this case we have

$$(a-c)\sin(\pi/2) = 0 = b\cos(\pi/2).$$

[24] Trigonometry is beautiful except for the notation.

Finally, if we have $b \neq 0$ and $(a-c) \neq 0$ then we must have either $\sin(2\theta) = \cos(2\theta) = 0$ (which is **impossible**) or we must have $\sin(2\theta) \neq 0$ and $\cos(2\theta) \neq 0$. In this last case we can solve the equation to obtain

$$\frac{\sin(2\theta)}{\cos(2\theta)} = \frac{b}{a-c}$$
$$\tan(2\theta) = \frac{b}{a-c}.$$

And since the tan function takes on all real values, we can always find a solution θ. In summary, there exists an invertible real matrix P with the property

$$P^T A P = \begin{pmatrix} a' & 0 \\ 0 & c' \end{pmatrix},$$

and thus the transformed equation (QDE') takes the form

$$a'(x')^2 + c'(y')^2 + \mathbf{d}^T \mathbf{x}' + f' = 0,$$

where

$$\mathbf{d}' = P^T(2A\mathbf{u} + \mathbf{d}) \qquad \text{and} \qquad f' = \mathbf{u}^T A \mathbf{u} + f.$$

Our next goal is to choose the vector $\mathbf{u} = \begin{pmatrix} u & v \end{pmatrix}^T$ so that $\mathbf{d}' = \mathbf{0}^T = \begin{pmatrix} 0 & 0 \end{pmatrix}^T$. Since the matrix P (and hence also P^T) is invertible, we observe that this will happen if and only if

$$(*) \qquad 2A\mathbf{u} + \mathbf{d} = \mathbf{0}.$$

The question is whether this equation $(*)$ has a solution. The rest of the discussion depends on a specific integer $\Delta \in \mathbb{Z}$, which is called the *discriminant* of the equation (QDE):

$$\Delta := b^2 - 4ac.$$

We observe that the determinant of the matrix A is

$$\det \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} = ac - \left(\frac{b}{2}\right)^2 = ac - \frac{b^2}{4} = \frac{4ac - b^2}{4} = -\frac{\Delta}{4},$$

and hence the matrix A is invertible if and only if $\Delta \neq 0$. Furthermore, recall that we have $\det(M^T) = \det(M)$ and $\det(MN) = \det(M)\det(N)$ for all matrices M and N for which these expressions make sense. Now observe that

$$\begin{aligned} a'c' &= \det \begin{pmatrix} a' & 0 \\ 0 & c' \end{pmatrix} \\ &= \det(P^T A P) \\ &= \det(P^T)\det(A)\det(P) \\ &= \det(P)\det(A)\det(P) \\ &= \det(A)\det(P)^2 \\ &= -\frac{\Delta}{4}\det(P)^2. \end{aligned}$$

Since $\det(P) \neq 0$ (because P is invertible) we must have $\det(P)^2 > 0$ and this implies that $a'c'$ is negative/positive/zero precisely when Δ is positive/negative/zero.[25] There are three cases:

• If $\Delta = b^2 - 4ac = 0$ then we also have $a'c' = 0$, so at least one of a' and c' is zero. If they are both zero then we get $A' = 0$ and hence $A = 0$, so that (QDE) and (QDE') are each the equation of a **line**. Otherwise, by switching x and y if necessary, we can assume that $a' \neq 0$ and $c' = 0$. Then the equation (QDE') takes the form

$$a'(x')^2 + d'x' + e'y' + f' = 0$$

and by "completing the square" in x' we obtain

$$\begin{aligned} (x')^2 + \frac{d'}{a'}x' + \frac{e'}{a'}y' + \frac{f'}{a'} &= 0 \\ \left(x' + \frac{d'}{2a'}\right)^2 + \frac{e'}{a'}y' + \frac{f'}{a'} - \left(\frac{d'}{2a'}\right)^2 &= 0 \\ (x'')^2 + e''y' + f'' &= 0. \end{aligned}$$

If $e'' = 0$ then this is the equation of a **line** (when $f'' = 0$), **two parallel lines** (when $f'' < 0$) or has **no real solution** (when $f'' > 0$). If $e'' \neq 0$ then we identify this as the equation of a **parabola**.

• If $\Delta = b^2 - 4ac < 0$ then we also have $a'c' > 0$ so that a' and c' have the same parity. Since $\Delta \neq 0$, the matrix A is invertible and we can solve the equation $(*)$ for $\mathbf{u}$ to obtain

$$\begin{aligned} 2A\mathbf{u} + \mathbf{d} &= \mathbf{0} \\ \mathbf{u} &= -\frac{1}{2}A^{-1}\mathbf{d}. \end{aligned}$$

This forces $\mathbf{d}' = \mathbf{0}$ and so equation (QDE') takes the form

(QDE") $$a'(x')^2 + c'(y')^2 + f' = 0.$$

If f' has the same parity as a' and c' then there is **no real solution**. If $f' = 0$ then the solution is the **single point** $(x', y') = (0, 0)$ and if f' has opposite parity from a' and c' then we identify (QDE") as the equation of an **ellipse**.

• If $\Delta = b^2 - 4ac > 0$ then we also have $a'c' < 0$ so that a' and c' have opposite parity. Since $\Delta \neq 0$ we can again force $\mathbf{d}' = \mathbf{0}$ by choosing $\mathbf{u} = -A^{-1}\mathbf{d}/2$ so equation (QDE') takes the form

(QDE") $$a'(x')^2 + c'(y')^2 + f' = 0.$$

[25] In the current situation we have $\det(P) = 1$, but this will not be the case below so I wanted to keep the discussion as general as possible.

If $f' = 0$ then we can solve this to obtain $y'/x' = \pm\sqrt{-a'/c'}$, which is the equation of **two lines meeting at the point** $(x', y') = (0, 0)$. If $f' \neq 0$ then we identify (QDE") as the equation of a **hyperbola**.

In summary, we can choose a rotation matrix P and a translation vector $\mathbf{u}$ so that (QDE') is the equation of a conic section in standard position in the x', y'-plane. This tells us that the original (QDE) is the equation of a conic section in non-standard position in the x, y-plane. See below for an explanation of the words "conic section".

//

Reduction of (QDE) for rational numbers. The general outline here is the same as the computation over the real numbers, however we will not be able to choose the matrix P so that $P^{-1} = P^T$. Indeed, the rotation matrix R_θ **almost never** has rational entries.

To find a suitable matrix P we will use a different method called *Hermite reduction*. You may remember that we can perform an invertible "row operation" on a 2×2 matrix A by multiplying on the left by one of the *elementary matrices*

$$E = \begin{pmatrix} 1 & \ell \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ \ell & 1 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

The goal of *Gaussian row-reduction* is to multiply on the left by a sequence of elementary matrices until we reach a diagonal matrix, or at least an upper-triangular matrix. This is often expressed as an algorithm to compute the **inverse** of a matrix. First we place A next to an identity matrix:

$$\left(\begin{array}{c|c} A & I \end{array} \right).$$

Then we perform a sequence of row operations $E_1, E_2, \ldots, E_k$ on the whole matrix to obtain

$$\begin{array}{c} \left(\begin{array}{c|c} A & I \end{array} \right) \\ \left(\begin{array}{c|c} E_1 A & E_1 \end{array} \right) \\ \left(\begin{array}{c|c} E_2 E_1 A & E_2 E_1 \end{array} \right) \\ \vdots \\ \left(\begin{array}{c|c} E_k \cdots E_2 E_1 A & E_k \cdots E_2 E_1 \end{array} \right) \\ \left(\begin{array}{c|c} PA & P \end{array} \right) \end{array}$$

where $P = E_k \cdots E_2 E_1$ is the product of the elementary matrices. If it is possible to reduce A to the identity matrix in this way then we will eventually reach $PA = I$ and the inverse matrix $P = A^{-1}$ will appear on the right hand side:

$$\left(\begin{array}{c|c} PA & P \end{array} \right) = \left(\begin{array}{c|c} I & A^{-1} \end{array} \right).$$

There is also a variant of this method called *column-reduction* which performs invertible column operations by multiplying on the **right** by elementary matrices:

$$\left(\frac{A}{I}\right) \to \left(\frac{AE_1}{E_1}\right) \to \cdots \to \left(\frac{AE_1E_2\cdots E_k}{E_1E_2\cdots E_k}\right) = \left(\frac{AP}{P}\right).$$

The idea of Hermite reduction is to perform both of these processes **simultaneously**. To begin we start with a matrix of the form

$$\left(\begin{array}{c|c} A & I \\ \hline I & \end{array}\right).$$

It doesn't matter what we put in the bottom right because we are never going to touch it. If we perform a column operation on the first n columns (suppose A is an $n \times n$ matrix) then this has the effect of multiplying on the **right** by an elementary matrix

$$\left(\begin{array}{c|c} A & I \\ \hline I & \end{array}\right) \to \left(\begin{array}{c|c} AE & I \\ \hline E & \end{array}\right).$$

Note that the upper-right corner is left untouched by this process. Now we perform "the same" row operation on the on the first n rows. This has the effect of multiplying on the **left** by the **transposed** elementary matrix:

$$\left(\begin{array}{c|c} A & I \\ \hline I & \end{array}\right) \to \left(\begin{array}{c|c} AE & I \\ \hline E & \end{array}\right) \to \left(\begin{array}{c|c} E^TAE & E^T \\ \hline E & \end{array}\right).$$

It doesn't matter in which order we perform the two operations because of the associative property of matrix multiplication: $(E^TA)E = E^T(AE)$. After performing a sequence of simultaneous row/column operations then we obtain

$$\left(\begin{array}{c|c} A & I \\ \hline I & \end{array}\right) \to \left(\begin{array}{c|c} E_1^TAE_1 & E_1^T \\ \hline E_1 & \end{array}\right) \to \cdots \to \left(\begin{array}{c|c} E_k^T\cdots E_1^TAE_1\cdots E_k & E_k^T\cdots E_1^T \\ \hline E_1\cdots E_k & \end{array}\right).$$

The end result is a matrix of the form

$$\left(\begin{array}{c|c} A & I \\ \hline I & \end{array}\right) \to \left(\begin{array}{c|c} P^TAP & P^T \\ \hline P & \end{array}\right)$$

where $P = E_1E_2\cdots E_k$ is the product of the elementary matrices. The goal now is to choose the simultaneous row/column operations so that we can reduce A to a diagonal matrix.

It turns out that if $A^T = A$ has rational entries then this is always possible.[26] Let's see how the Hermite reduction algorithm works on our favorite 2×2 matrix

$$A = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}.$$

[26]More generally, it is always possible for a symmetric matrix with entries in a given *field*. It is not always possible over the integers.

First let's assume that $a \neq 0$. Then we can subtract $b/2a$ times the first row from the second row and subtract $b/2a$ times the first column from the second column to obtain

$$\left(\begin{array}{cc|cc} a & b/2 & 1 & 0 \\ b/2 & c & 0 & 1 \\ \hline 1 & 0 & & \\ 0 & 1 & & \end{array}\right) \to \left(\begin{array}{cc|cc} a & b/2 & 1 & 0 \\ 0 & c-b^2/4a & -b/2a & 1 \\ \hline 1 & 0 & & \\ 0 & 1 & & \end{array}\right)$$

$$\to \left(\begin{array}{cc|cc} a & 0 & 1 & 0 \\ 0 & c-b^2/4a & -b/2a & 1 \\ \hline 1 & -b/2a & & \\ 0 & 1 & & \end{array}\right).$$

From the remarks above, it follows that by choosing

$$P = \begin{pmatrix} 1 & -b/2a \\ 0 & 1 \end{pmatrix}$$

we obtain the diagonalization

$$P^T AP = \begin{pmatrix} 1 & 0 \\ -b/2a & 1 \end{pmatrix} \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \begin{pmatrix} 1 & -b/2a \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & c-b^2/4a \end{pmatrix}.$$

Note that the matrix P is certainly not a rotation matrix, but it is still invertible with

$$P^{-1} = \begin{pmatrix} 1 & +b/2a \\ 0 & 1 \end{pmatrix}.$$

The main feature of the matrix P is that it has **rational entries** so that it preserves the rationality of solutions of the Diophantine equation.

The case $c \neq 0$ is similar. Here we subtract $b/2c$ times the **second** row/column from the **first** row/column to obtain

$$\left(\begin{array}{cc|cc} a & b/2 & 1 & 0 \\ b/2 & c & 0 & 1 \\ \hline 1 & 0 & & \\ 0 & 1 & & \end{array}\right) \to \left(\begin{array}{cc|cc} a-b^2/4c & 0 & 1 & -b/2c \\ b/2 & c & 0 & 1 \\ \hline 1 & 0 & & \\ 0 & 1 & & \end{array}\right)$$

$$\to \left(\begin{array}{cc|cc} a-b^2/4c & 0 & 1 & -b/2c \\ 0 & c & 0 & 1 \\ \hline 1 & 0 & & \\ -b/2c & 1 & & \end{array}\right).$$

Thus by choosing

$$P = \begin{pmatrix} 1 & 0 \\ -b/2c & 1 \end{pmatrix}$$

we obtain the rational diagonalization

$$P^T AP = \begin{pmatrix} 1 & 0 \\ -b/2c & 1 \end{pmatrix} \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \begin{pmatrix} 1 & -b/2c \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a-b^2/4c & 0 \\ 0 & c \end{pmatrix}.$$

If $a = b = c = 0$ then there is nothing to do, thus the last case we must consider is when $a = c = 0$ and $b \neq 0$. This one is a bit harder. First we have to add the **second** row/column to the **first** row/column to get a nonzero entry on the diagonal. Then we subtract $1/2$ of the **first** row/column from the **second** row/column to eliminate the off-diagonal entries. At this point the matrix is diagonalized, but we can clean it up a bit more by multiplying the **second** row/column by 2:

$$\left(\begin{array}{cc|cc} 0 & b/2 & 1 & 0 \\ b/2 & 0 & 0 & 1 \\ \hline 1 & 0 & & \\ 0 & 1 & & \end{array}\right) \to \left(\begin{array}{cc|cc} b & b/2 & 1 & 1 \\ b/2 & 0 & 0 & 1 \\ \hline 1 & 0 & & \\ 1 & 1 & & \end{array}\right)$$

$$\to \left(\begin{array}{cc|cc} b & 0 & 1 & 1 \\ 0 & -b/4 & -1/2 & 1/2 \\ \hline 1 & -1/2 & & \\ 1 & 1/2 & & \end{array}\right)$$

$$\to \left(\begin{array}{cc|cc} b & 0 & 1 & 1 \\ 0 & -b & -1 & 1 \\ \hline 1 & -1 & & \\ 1 & 1 & & \end{array}\right).$$

In summary, by choosing

$$P = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

we obtain the diagonalization

$$P^T A P = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 0 & b/2 \\ b/2 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} b & 0 \\ 0 & -b \end{pmatrix}.$$

Maybe we could have come up with all of these tricks through cleverness, but I prefer the systematic way.

The rest of the reduction is identical to the real case, since all of the translation vectors $\mathbf{u}$ we chose in that case had entries that were rational expressions of the previous entries. In summary, by an invertible rational affine transformation we can reduce (QDE) to one of the following three forms:

- If $A = 0$ then the original equation (QDE) had the form

  $$a'x + b'y + c' = 0$$

  for some integers $a', b', c' \in \mathbb{Z}$.

- If $\Delta = b^2 - 4ac = 0$ and $A \neq 0$ then the equation (QDE) can be reduced to the form

  $$a'x^2 + b'y + c' = 0 \quad \text{or} \quad a'y^2 + b'x + c' = 0$$

  for some integers $a', b', c' \in \mathbb{Z}$ with $a' \neq 0$.

- If $\Delta = b^2 - 4ac \neq 0$ then the equation (QDE) can be reduced to the form

$$a'x^2 + b'y^2 + c' = 0$$

for some integers $a', b', c' \in \mathbb{Z}$ with $a'b' \neq 0$.

This was all pure algebra. In the next chapter we will return to the problem of finding the complete rational solution $(x, y) \in \mathbb{Q}^2$ to each of these equations. But first:

4.3 Why Do We Call Them Conic Sections?

To end the chapter I'll show you something that I wish someone had shown to me a long time ago. We have seen that the equation

$$\text{(QDE)} \qquad ax^2 + bxy + cy^2 + dx + ey + f = 0$$

defines either a point, a line, two lines, a parabola, an ellipse or a hyperbola in the real x, y-plane. It turns out that each of these shapes can be described as the intersection of a plane with a cone[27] in three dimensional space, thus these shapes are often referred to as *conic sections*.

But **why** does this happen? In other words:

why do solutions of quadratic equations look like conic sections?

To understand this we must return to the homogenized version of (QDE):

$$\text{(HQDE')} \qquad ax^2 + bxy + cy^2 + dxz + eyz + fz^2 = 0.$$

If we can find the geometric shape of the solution to (HQDE') in cartesian x, y, z-space, then we will obtain the geometric shape of the solution to (QDE) by intersecting this shape with the plane defined by $z = 1$. Thus our goal is to show that the solutions of (HQDE') look like a cone (in the generic case).

To do this, we first express the equation in terms of matrix arithmetic as follows: we have

$$\begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} a & b/2 & d/2 \\ b/2 & c & e/2 \\ d/2 & e/2 & f \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0$$

$$\mathbf{x}^T H \mathbf{x} = 0,$$

where the 3×1 column vector $\mathbf{x}$ and the 3×3 symmetric matrix $H^T = H$ are defined by

$$\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \qquad \text{and} \qquad H = \begin{pmatrix} a & b/2 & d/2 \\ b/2 & c & e/2 \\ d/2 & e/2 & f \end{pmatrix}.$$

[27] Except for the degenerate case of two parallel lines, which we will see is the intersection of a plane with two other planes.

Now we will use the Principal Axes Theorem which was stated in the previous section (but we still won't prove it). Since H is a symmetric real matrix we know that there exists an 3×3 orthogonal matrix $Q^T = Q^{-1}$ such that $Q^T H Q$ is diagonal, i.e., such that

$$H' = Q^T H Q = \begin{pmatrix} a' & 0 & 0 \\ 0 & c' & 0 \\ 0 & 0 & f' \end{pmatrix}.$$

Then after performing the change of variables $\mathbf{x} = Q\mathbf{x}'$ we see that (HQDE') is equivalent to the transformed equation

$$\begin{aligned} \mathbf{x}^T H \mathbf{x} &= 0 \\ (Q\mathbf{x})^T H \left(Q\mathbf{x}'\right) &= 0 \\ (\mathbf{x}')^T \left(Q^T H Q\right) \mathbf{x}' &= 0 \\ (\mathbf{x}')^T H' \mathbf{x}' &= 0 \\ (*) \qquad a'(x')^2 + c'(y')^2 + f'(z')^2 &= 0. \end{aligned}$$

Note that $(x', y', z') = (0, 0, 0)$ is always a solution. If $a' = c' = f' = 0$ then every triple $(x', y', z') \in \mathbb{R}^3$ is a solution to $(*)$ and there is nothing to show. If exactly one of a', c', f' equals zero then the solution of $(*)$ is either a line or a pair of planes through the origin in x', y', z'-space. Finally, if $a'c'f' \neq 0$ then the solution of $(*)$ is either the single point $(0, 0, 0)$ or an **elliptic cone** (a cone whose perpendicular cross sections are ellipses) in x', y', z'-space.

Suppose that we are in the most interesting case when $(*)$ defines an elliptic cone in x', y', z'-space. Then since (HQDE') is equivalent to $(*)$ under an orthogonal (i.e., distance-preserving) change of coordinates we see that the equation (HQDE') defines an elliptic cone in x, y, z-space. Finally, the solution to the original (QDE) is the intersection of this elliptic cone with the plane $z = 1$ in x, y, z-space. This explains why we call the general equation (QDE) a "conic section".

Remark: I believe that it is possible in all cases to realize the solution of (QDE) as the intersection of the plane $z = 1$ with a **right circular cone**[28] in x, y, z-space (i.e., not just an elliptic cone). Unfortunately I do not know a similarly slick (or really any) proof of this fact. If you know one please tell me.

5 Rational Points on Conics

The topic of this chapter is to find the complete **rational solution** $(x, y) \in \mathbb{Q}^2$ to the quadratic equation

$$\text{(QDE)} \qquad ax^2 + bxy + cy^2 + dx + ey + f = 0$$

[28] or a circular cylinder, which we think of as a cone with its apex "at infinity".

for general integers $a, b, c, d, e, f \in \mathbb{Z}$. The steps involved in the solution are the same as the steps involved when we solved the linear equation in Chapter 2. Here are the steps in increasing order of difficulty:

- Reduce the general equation to a standard form (primitive, homogeneous, etc.)
- Determine whether a solution exists.
- Find one specific solution.
- Find the complete solution.

In the previous chapter we saw that the general equation (QDE) can be reduced by an invertible affine transformation with rational coefficients to one of three standard forms. To be specific, if the equation (QDE) is not of the form $0 = 0$ or $0 = 1$ then there exist integers $a, b, c \in \mathbb{Z}$ with $\gcd(a, b, c) = 1$ and $a \neq 0$ such that (QDE) is rationally equivalent to one of the following:

(linear)	$ax + by + c = 0$
(parabolic)	$ax^2 + by + c = 0$
(elliptic/hyperbolic)	$ax^2 + by^2 + c = 0$

Furthermore, we have seen that the problem of finding all rational solutions $(x, y) \in \mathbb{Q}^2$ is equivalent to the problem of finding all integer solutions $(x, y, z) \in \mathbb{Z}^3$ with $\gcd(x, y, z) = 1$ and $z \geqslant 1$ to the associated homogenized equation:

(linear)	$ax + by + cz = 0$
(parabolic)	$ax^2 + byz + cz^2 = 0$
(elliptic/hyperbolic)	$ax^2 + by^2 + cz^2 = 0$

We already understand the linear case so let's quickly dispense with it. Suppose that we have integers $a, b, c, x, y, z \in \mathbb{Z}$ of the stated form satisfying $ax + by + cz = 0$. If $d = \gcd(a, b)$ then the equality tells us that $d|cz$. Then since $\gcd(a, b, c) = \gcd(d, c) = 1$ Euclid's Lemma tells us that $d|z$, say $z = dk$. Since $z \geqslant 1$ we must have $k \geqslant 1$. Finally, for any fixed value of $k \geqslant 1$ we can find the complete integer solution $(x, y) \in \mathbb{Z}^2$ to the equation $ax + by + cdk = 0$ exactly as in Chapter 2.

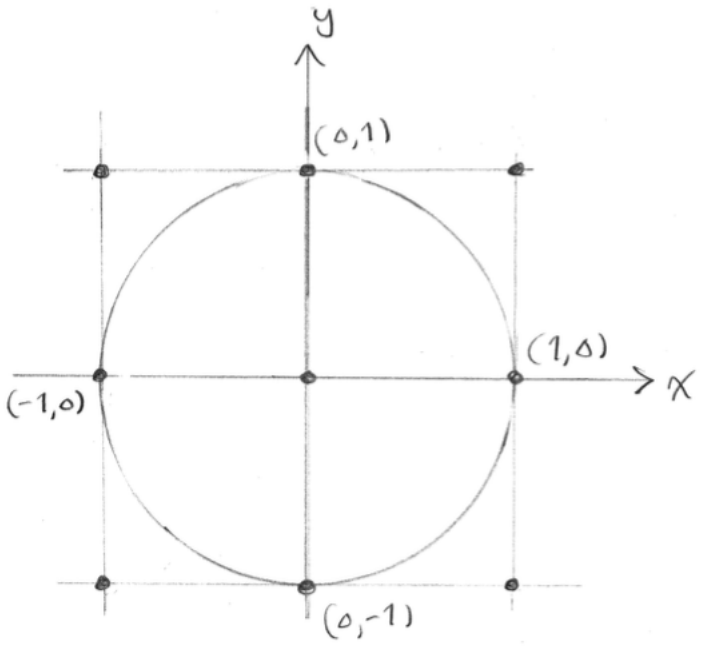
The rest of this chapter will deal with the parabolic and the elliptic/hyperbolic cases.

5.1 Pythagorean Triples

Just as with the case of linear equations, it turns out that if we can find **one particular solution** to a quadratic Diophane equation (that is, one particular rational solution or one particular integer solution of the homogenized equation) then the complete solution is easy to obtain. I will illustrate the general method as it applies to the following specific equation:

(UC) $$x^2 + y^2 = 1.$$

If we temporarily allow x and y to be real numbers, then we can think of (UC) as the equation of a circle of radius 1 centered at the origin of the x, y-plane:



Note that the equation (UC) has exactly four integer solutions:

$$(x, y) \in \{(1, 0), (0, 1), (-1, 0), (0, -1)\}.$$

There is nothing more to say about this, so we move on to rational solutions of (UC).[29] Suppose that we have rational solution $(x, y) \in \mathbb{Q}^2$. By finding a common denominator we can write $x = a/c$ and $y = b/c$ and by canceling the greatest common divisor we can assume that $\gcd(a, b, c) = 1$ with $c \geqslant 1$. Then equation (UC) becomes

$$\begin{aligned} (a/c)^2 + (b/c)^2 &= 1 \\ a^2/c^2 + b^2/c^2 &= 1 \\ a^2 + b^2 &= c^2. \end{aligned} \tag{PT}$$

Integer solutions to the equation (PT) are called *Pythagorean triples* and solutions with $\gcd(a, b, c) = 1$ and $c \geqslant 1$ are called *primitive* Pythagorean triples. You are probably familiar with the primitive Pythagorean triple $3^2 + 4^2 = 5^2$ and the fact that for any integer $\lambda \in \mathbb{Z}$ we have $(3\lambda)^2 + (4\lambda)^2 = (5\lambda)^2$. The results of Section 4.1 immediately imply the following general theorem.

[29] In general, elliptic Diophantine equations have only finitely many integer solutions. Integer solutions of hyperbolic equations are more interesting. However, for rational solutions there is no difference.

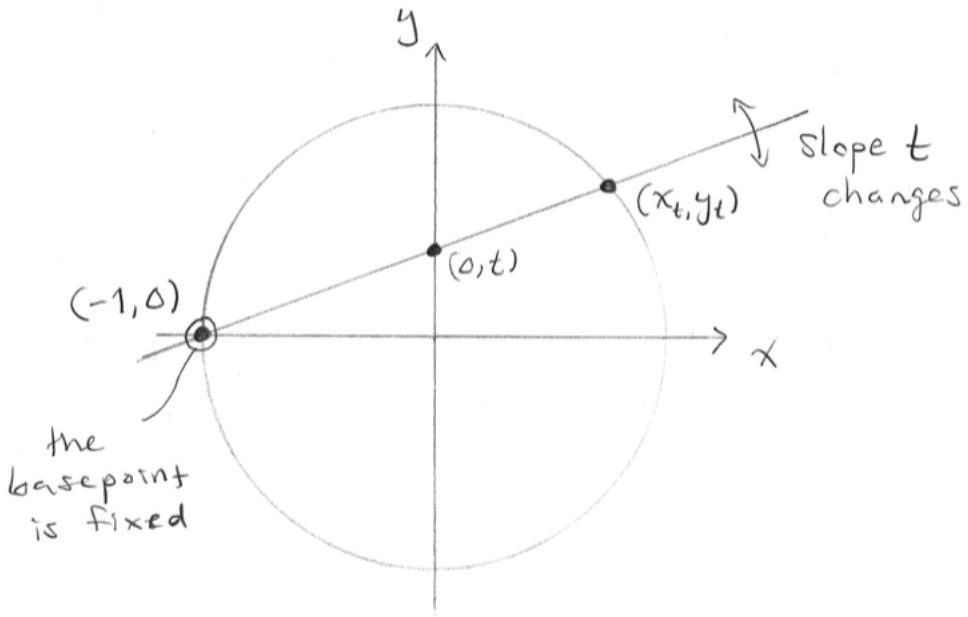
Theorem (Reduction to Primitive Pythagorean Triples). Each Pythagorean triple $(a, b, c) \in \mathbb{Z}^3$ has a unique expression of the form

$$(a, b, c) = \lambda \cdot (a', b', c')$$

where $\lambda \in \mathbb{Z}$ and where $(a', b', c') \in \mathbb{Z}^3$ is a primitive Pythagorean triple. //

Thus the problem of classifying Pythagorean triples is reduced to the problem of classifying **primitive** Pythagorean triples. And the classification of primitive triples is reduced to the problem of classifying rational points on the unit circle. It turns out that there is a beautiful geometric trick for finding these points. The method was hinted at in the *Arithmetica* by Diophantus of Alexandria (c. 200–284), althouth he didn't describe it in geometric terms.

The idea is to choose **one specific rational point** and to consider the line of slope t passing through this point. In this case we'll choose the point $(-1, 0)$:



For any finite value of t this will intersect the circle in exactly one other point, which we call (x_t, y_t). To compute the coordinates of this point, first note that the equation of the line is

$$\begin{aligned} t &= (\text{rise})/(\text{run}) \\ t &= (y - 0)/(x - (-1)) \\ t(x + 1) &= y. \end{aligned}$$

We substitute this into the equation of the circle to obtain

$$\begin{aligned} 1 &= x^2 + y^2 \\ 1 &= x^2 + t^2(x+1)^2 \\ 0 &= x^2(t^2+1) + x(2t^2) + (t^2-1). \end{aligned}$$

Now we can use the quadratic formula to solve for x. Note that a very lucky cancellation happens under the square-root sign:

$$\begin{aligned} x &= \frac{-2t^2 \pm \sqrt{(2t^2)^2 - 4(t^2+1)(t^2-1)}}{2(t^2+1)} \\ &= \frac{-2t^2 \pm \sqrt{4t^4 - 4(t^4-1))}}{2(t^2+1)} \\ &= \frac{-2t^2 \pm \sqrt{4t^4 - 4(t^4-1))}}{2(t^2+1)} \\ &= \frac{-2t^2 \pm \sqrt{4}}{2(t^2+1)} \\ &= \frac{-2t^2 \pm 2}{2(t^2+1)} \\ &= \frac{-t^2 \pm 1}{t^2+1} \\ &= \frac{-t^2-1}{t^2+1} \quad \text{or} \quad \frac{-t^2+1}{t^2+1} \\ &= -1 \quad \text{or} \quad \frac{1-t^2}{1+t^2} \end{aligned}$$

The solution $x = -1$ corresponds to the point $(x, y) = (-1, 0)$ and thus we have $x_t = (1 - t^2)/(1 + t^2)$. Finally, we substitute this formula for x_t into the equation of the line to obtain

$$y_t = t(x_t + 1) = t\left(\frac{1-t^2}{1+t^2} + 1\right) = t\left(\frac{1-t^2}{1-t^2} + \frac{1-t^2}{1+t^2}\right) = \frac{2t}{1+t^2}.$$

In summary, we have the following two equations relating the slope t to the coordinates of the point (x_t, y_t). These equations hold for any real number t:

$$t = \frac{y_t}{x_t + 1} \tag{1}$$

$$(x_t, y_t) = \left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right) \tag{2}$$

But recall that we are only interested in the **rational** points on the circle. Here is the key fact.

Theorem (Diophantus' Chord Method for the Circle). The mapping $t \mapsto (x_t, y_t)$ defines a one-to-one correspondence between rational numbers $t \in \mathbb{Q}$ and the rational points on the unit circle, excluding $(-1, 0)$. //

Proof. Every **real** point on the unit circle except for $(-1, 0)$ has the form (x_t, y_t) for some unique **real** number t. Furthermore, from equation (1) above we see that

$$(x_t, y_t) \in \mathbb{Q}^2 \quad \Longrightarrow \quad t \in \mathbb{Q}$$

and from equation (2) above we see that

$$t \in \mathbb{Q} \quad \Longrightarrow \quad (x_t, y_t) \in \mathbb{Q}^2.$$

In other words, we have a one-to-one correspondence between rational values of t and rational points on the circle except for $(-1, 0)$. □

It turns out that exactly the same trick works for any quadratic Diophantine equation as long as we are able to find **one specific rational point** to begin with. Finding a rational point on the unit circle was easy, but unfortunately this will not always be the case.

Next let's investigate the rational points (x_t, y_t) in detail so we can extract a classification of Pythagorean triples.

The general rational point on the unit circle has the form $(x_t, y_t) \in \mathbb{Q}^2$ for some rational number $t \in \mathbb{Q}$. Let us write t in lowest terms so that $t = u/v$ for some unique integers $u, v \in \mathbb{Z}$ with $\gcd(u, v) = 1$ and $v \geqslant 1$. Then we have

$$\begin{aligned}
(x_t, y_t) &= \left(\frac{1 - t^2}{1 + t^2}, \frac{2t}{1 + t^2} \right) \\
&= \left(\frac{1 - (u/v)^2}{1 + (u/v)^2}, \frac{2(u/v)}{1 + (u/v)^2} \right) \\
&= \left(\frac{v^2}{v^2} \cdot \frac{1 - (u/v)^2}{1 + (u/v)^2}, \frac{v^2}{v^2} \cdot \frac{2(u/v)}{1 + (u/v)^2} \right) \\
&= \left(\frac{v^2 - u^2}{v^2 + u^2}, \frac{2uv}{v^2 + u^2} \right).
\end{aligned}$$

From the above remarks we also know that

$$(*) \qquad \left(\frac{a}{c}, \frac{b}{c} \right) = (x_t, y_t) = \left(\frac{v^2 - u^2}{v^2 + u^2}, \frac{2uv}{v^2 + u^2} \right)$$

for some unique integers $a, b, c \in \mathbb{Z}$ with $\gcd(a, b, c) = 1$ and $c \geqslant 1$. To determine the relationship between the unique integers a, b, c and the unique integers u, v it only remains to determine the greatest common divisor of the integers $v^2 - u^2, 2uv, v^2 + u^2$.

Lemma. Consider integers $u, v \in \mathbb{Z}$ with $\gcd(u, v) = 1$. Then we have

$$\gcd(v^2 - u^2, 2uv, v^2 + u^2) = 1 \text{ or } 2.$$

//

Proof. Let $d := \gcd(v^2 - u^2, 2uv, v^2 + u^2)$. Now let p be any **odd prime divisor** of d. Since p divides each of $v^2 - u^2$ and $v^2 + u^2$ it must also divide

$$\big[(v^2 - u^2) + (v^2 + u^2)\big] = 2v^2.$$

Then since $p \nmid 2$, Euclid's Lemma tells us that $p|v^2$ and hence $p|v$. Similarly we see that p divides

$$\big[(v^2 + u^2) - (v^2 - u^2)\big] = 2u^2$$

and it follows from Euclid's Lemma that $p|u$. But this contradicts the fact that $\gcd(u, v) = 1$ so we conclude that d has no odd prime divisors and it follows that d is a power of 2. I claim that $d = 2^k$ for $k = 0$ or $k = 1$.

To see this, assume for contradiction that d is divisible by $4 = 2^2$. Then by the same argument as above we see that 4 divides each of $2v^2$ and $2u^2$. Since $4|2v^2$ there exists an integer $\ell \in \mathbb{Z}$ such that $2v^2 = 4\ell = 2(2\ell)$ and canceling ℓ gives $v^2 = 2\ell$. Then since 2 is prime, Euclid's Lemma tells us that $2|v$. Then a similar argument gives $2|u$, which again contradicts the fact that $\gcd(u, v) = 1$. This completes the proof. □

In the case $\gcd(v^2 - u^2, 2uv, v^2 + u^2) = 1$ we conclude from equation $(*)$ that

$$(a, b, c) = (v^2 - u^2, 2uv, v^2 + u^2).$$

In the case $\gcd(v^2 - u^2, 2uv, v^2 + u^2) = 2$ we can divide through by 2 to obtain

$$\gcd\left(\frac{v^2 - u^2}{2}, uv, \frac{v^2 + u^2}{2}\right) = 1,$$

and then it follows from equation $(*)$ that

$$(a, b, c) = \left(\frac{v^2 - u^2}{2}, uv, \frac{v^2 + u^2}{2}\right).$$

This completes the classification of Pythagorean triples, but I don't like the look of the fractions in the previous equation. Maybe we can get rid of them?

Observe that since $a = (v^2 - u^2)/2$ is an integer, it must be the case that v^2 and u^2 have the same parity, and it follows from this that u and v also have the same parity. Thus we can define new integers $u', v' \in \mathbb{Z}$ with the following change of variables:

$$\begin{cases} u' = (u - v)/2 \\ v' = (u + v)/2 \end{cases} \qquad \begin{cases} u = v' - u' \\ v = v' + u' \end{cases}$$

From the system of equations on the right we see that the common divisors of u and v are the same as the common divisors of u' and v' and hence $\gcd(u', v') = \gcd(u, v) = 1$. Finally, we have the miraculous simplification

$$\left(\frac{v^2 - u^2}{2}, uv, \frac{v^2 + u^2}{2}\right) = (2u'v', (v')^2 - (u')^2, (v')^2 + (u')^2).$$

In summary, we have the following theorem.

Theorem (Classification of Pythagorean Triples). Consider a nonzero integer vector $(0, 0, 0) \neq (a, b, c) \in \mathbb{Z}^3$ such that $a^2 + b^2 = c^2$. Then *exactly one of the following applies*:

- There exist unique integers $\lambda, u, v \in \mathbb{Z}$ with $\gcd(u, v) = 1$ and $v \geqslant 1$ such that

$$(a, b, c) = \lambda \cdot \left(v^2 - u^2, 2uv, v^2 + u^2\right).$$

- There exist unique integers $\lambda, u, v \in \mathbb{Z}$ with $\gcd(u, v) = 1$ and $v \geqslant 1$ such that

$$(a, b, c) = \lambda \cdot \left(2uv, v^2 - u^2, v^2 + u^2\right).$$

In either case we have $\gcd(v^2 - u^2, 2uv, u^2 + v^2) = 1$. //

This theorem was by no means trivial to prove. The algebraic step of parametrizing the rational points in terms of $t \in \mathbb{Q}$ was straightforward; however, the process of finding a unique representation for the integer Pythagorean triples involved some tricky number-theoretic arguments. And there are still some mysteries hiding in the final answer. For example, here is a puzzle:

> It follows from the previous theorm that if (a, b, c) is a Pythagorean triple then a and b cannot both be odd. But this fact never showed up explicitly in the proof. Why is it true?

Believe it or not, the easeist way to "explain" this phenomenon is by thinking about the "square elements" in the ring $\mathbb{Z}/4\mathbb{Z}$!

Here is the relevant definition, which will play a central role later in this chapter.

Definition of Quadratic Residue. Consider integers $a, n \in \mathbb{Z}$ with $n > 0$. We say that a is a *quadratic residue* mod n if there exists an integer $x \in \mathbb{Z}$ such that

$$[a]_n = [x^2]_n = ([x]_n)^2.$$

Equivalently, the element $[a]_n \in \mathbb{Z}/n\mathbb{Z}$ has some square root $[x]_n$ in the ring $\mathbb{Z}/n\mathbb{Z}$.

//

The following exercise explains the puzzle.

Exercise.

(a) Show that $[0]_4$ and $[1]_4$ are the only square elements in the ring $\mathbb{Z}/4\mathbb{Z}$.

(b) Consider any integers $a, b, c \in \mathbb{Z}$ such that $a^2 + b^2 = c^2$. Reduce this equation mod 4 to obtain

$$[a^2]_4 + [b^2]_4 = [c^2]_4.$$

Now apply part (a) to show that a and b cannot both be odd.

5.2 Diophantus' Chord Method in General

5.3 Legendre's Theorem

In the previous section we showed that if we can find **one specific rational solution** to a parabolic or hyperbolic/elliptic Diophantine equation then we can find the complete rational solution by using Diophantus' chord method. Thus the problem of finding the complete rational solution of a quaderatic Diophantine equations has been reduced to the question of

existence of solutions.

The following example shows that rational solutions do not necessarily exist.

Example: There are no rational points on the circle $x^2 + y^2 = 3$.

Remark: We could prove this by reducing the equation $a^2 + b^2 = 3c^2$ mod 4 as we did at the end of section 4.1. However, that was a bit of a lucky trick. Now I want to follow a method of proof that will extend to the general hyperbolic/elliptic case.

Proof. Assume for contradiction that there exist rational numbers $(x, y) \in \mathbb{Q}^2$ such that $x^2 + y^2 = 3$. By finding a common denominator we can write $(x, y) = (a/c, b/c)$ for some integers $a, b, c \in \mathbb{Z}$ with $c \geqslant 1$ and by canceling common factors as in Section 4.1 we can assume that $\gcd(a, b, c) = 1$. We obtain the equation

$$\begin{aligned} x^2 + y^2 &= 3 \\ (a/c)^2 + (b/c)^2 &= 3 \\ a^2 + b^2 &= 3c^2. \end{aligned}$$

Now I claim that 3 is not a common divisor of a and b. Indeed, if we had $a = 3a'$ and $b = 3b'$ for some integers $a', b' \in \mathbb{Z}$ then we would obtain

$$\begin{aligned} a^2 + b^2 &= 3c^2 \\ (3a')^2 + (3b')^2 &= 3c^2 \\ 3^2\left[(a')^2 + (b')^2\right] &= 3c^2 \\ 3\left[(a')^2 + (b')^2\right] &= c^2. \end{aligned}$$

Now since $3|c^2$, Euclid's Lemma tells us that $3|c$ and we conclude that 3 is a common factor of a, b, c. This contradicts the fact that $\gcd(a, b, c) = 1$.

Finally, let us reduce the equation $a^2 + b^2 = 3c^2 \bmod 3$ to obtain

$$\begin{aligned} [a^2 + b^2]_3 &= [3c^2]_3 \\ [a^2]_3 + [b^2]_3 &= [0]_3. \end{aligned}$$

Since 3 is not a common multiple of a and b we can assume without loss of generality that $3 \nmid a$ so that the elements $[a]_3$ and $[a^2]_3$ are invertible, and we obtain

$$\begin{aligned} [a^2]_3 + [b^2]_3 &= [0]_3 \\ [a^2]_3 &= [-b^2]_3 \\ [a^2]_3 \cdot [a^{-2}]_3 &= [-1]_3 \cdot [b^2]_3 \cdot [a^{-2}]_3 \\ [1]_3 &= [2]_3 \cdot [b^2]_3 \cdot [a^{-2}]_3 \\ [2]_3 \cdot [1]_3 &= [2]_3 \cdot [2]_3 \cdot [b^2]_3 \cdot [a^{-2}]_3 \\ [2]_3 &= \left([2]_3 \cdot [b]_3 \cdot [a^{-1}]_3\right)^2. \end{aligned}$$

This last equation says that the element $[2]_3 \in \mathbb{Z}/3\mathbb{Z}$ is a perfect square. But we can see that this is a contradiction by squaring all three elements of $\mathbb{Z}/3\mathbb{Z}$:

$$[0^2]_3 = [0]_3 \neq [2]_3, \quad [1^2]_3 = [1]_3 \neq [2]_3 \quad \text{and} \quad [2^2]_3 = [1]_3 \neq [2]_3.$$

□

In summary, we find that the equation $x^2 + y^2 = 3$ has no rational solution because the element $[2]_3$ has no square root in the ring $\mathbb{Z}/3\mathbb{Z}$. In this section we will prove a celebrated theorem of Legendre (1785) which says that the existence of rational solutions to a general quadratic Diophantine equation is controlled by the existence of certain modular square roots.

Legendre's Theorem. Fix integers $(a, b, c) \in \mathbb{Z}^3$ such that:

- a, b, c are squarefree (i.e., have no repeated prime factors),
- a, b, c are pairwise coprime (i.e., $\gcd(a, b) = \gcd(a, c) = \gcd(b, c) = 1$),

- $abc \neq 0$, not all of the same sign.

Then the equation

$$ax^2 + by^2 + cz^2 = 0$$

has a solution $(0,0,0) \neq (x,y,z) \in \mathbb{Z}^3$ if and only if the following elements have square roots:

$$[-ab]_c \in \mathbb{Z}/c\mathbb{Z}, \quad [-ac]_b \in \mathbb{Z}/b\mathbb{Z} \quad \text{and} \quad [-bc]_a \in \mathbb{Z}/a\mathbb{Z}.$$

Moreover, in this case we will show that there exists a solution satisfying

$$0 < |a|x^2 + |b|y^2 + |c|z^2 < 8|abc|.$$

//

[Remark: The bound $< 8|abc|$ is not the best possible[30] but we include it because it follows from our proof of existence with no extra work. The bound shows that we can find a solution (or prove that none exists) by a **finite computation**.]

Before proving the theorem let us oberve why it completely solves the problem of the existence of rational points on conic sections. There are two cases to consider:

Existence of Rational Points on a Parabola: Given integers $(a,b,c) \in \mathbb{Z}^3$ with $\gcd(a,b,c) = 1$ and $a \neq 0$, we want to determine if there exist rational numbers $(x,y) \in \mathbb{Q}^2$ such that

$$ax^2 + by + c = 0.$$

If $b \neq 0$ then we note that $(x,y) = (0,-c/b)$ is a solution, so let us assume that $b = 0$. Then the equation becomes

$$\begin{aligned} ax^2 + c &= 0 \\ x^2 &= -c/a. \end{aligned}$$

In other words, we need to determine whether the rational number $-c/a$ has a rational square root. If we write $d = \gcd(a,c)$ with $a = da'$ and $b = dc'$ then we observe that $-c/a = -c'/a'$ has a rational square root if and only if

- a' and c' have opposite signs,
- $|a'|$ and $|c'|$ are perfect squares.

Existence of Rational Points on a Hyperbola or Ellipse: Given integers $(a,b,c) \in \mathbb{Z}^3$ we want to determine if there exist rational numbers $(x,y) \in \mathbb{Q}^2$ such that

$$ax^2 + by^2 + c = 0.$$

[30] The best possible bound is $\leqslant 2|abc|$. See the paper *Small Solutions of the Legendre Equation* (1998) by Cochrane and Mitchell.

If $a = 0$ or $b = 0$ then this was already solved in the parabolic case above so we will assume that $ab \neq 0$. If $c = 0$ then the equation becomes $(x/y)^2 = -b/a$, which was also solved above, so we can assume that $abc \neq 0$. If a, b, c all have the same sign (say positive) then for any $(x, y) \in \mathbb{Q}^2$ we obtain $ax^2 + by^2 \geqslant 0 > -c$, so the equation has no solution. Furthermore, by dividing the equation by the greatest common denominator of a, b, c we can assume that $\gcd(a, b, c) = 1$.

We have reduced the problem of existence to the case when $\gcd(a, b, c) = 1$ and $abc \neq 0$, not all of the same sign. By finding a common denominator we also see that the existence of a rational solution $(x, y) \in \mathbb{Q}^2$ is equivalent to the existence of integers $(x, y, z) \in \mathbb{Z}^3$ with $z \neq 0$ such that

$$ax^2 + by^2 + cz^2 = 0,$$

and since this equation is symmetric in a, b, c we might as well look for nontrivial integer solutions $(0, 0, 0) \neq (x, y, z) \in \mathbb{Z}^3$.

Next we will reduce to the "squarefree" case, but first we need a definition.

Definition/Theorem. Consider an integer $0 \neq a \in \mathbb{Z}$ with unique prime factorization

$$a = \pm p_1^{e_1} p_2^{e_2} p_3^{e_3} \cdots .$$

By reducing each exponent e_i mod 2 we obtain unique quotients and remainders $e_i = 2q_i + r_i$ with $r_i \in \{0, 1\}$. Then we can write $a = \pm \bar{a}\alpha^2$ where

$$\bar{a} := p_1^{r_1} p_2^{r_2} p_3^{r_3} \cdots$$
$$\alpha := p_1^{q_1} p_2^{q_2} p_3^{q_3} \cdots .$$

The unique integers $\bar{a}$ and α^2 are called the *squarefree part* and the *square part of* a, respectively. We say that $0 \neq a \in \mathbb{Z}$ is *squarefree* if and only if $\alpha = 1$. //

So consider integers $(a, b, c) \in \mathbb{Z}^3$ with $\gcd(a, b, c) = 1$ and $abc \neq 0$, not all of the same sign, and consider the unique square/squarefree decompositions: $a = \bar{a}\alpha^2$, $b = \bar{b}\beta^2$, $c = \bar{c}\gamma^2$. Note that we have $\gcd(\bar{a}, \bar{b}, \bar{c}) = 1$ because any common divisor of $\bar{a}, \bar{b}, \bar{c}$ is also a common divisor of a, b, c. Now consider the following equations:

$$ax^2 + by^2 + cz^2 = 0 \tag{1}$$
$$\bar{a}x^2 + \bar{b}y^2 + \bar{c}z^2 = 0. \tag{2}$$

I claim that (1) has a nontrivial integer solution if and only if (2) does. To see this let $(0, 0, 0) \neq (x, y, z) \in \mathbb{Z}^3$ be a solution to (1). Then we have

$$\begin{aligned} ax^2 + by^2 + cz^2 &= 0 \\ (\bar{a}\alpha^2)x^2 + (\bar{b}\beta^2)y^2 + (\bar{c}\gamma^2)z^2 &= 0 \\ \bar{a}(\alpha x)^2 + \bar{b}(\beta y)^2 + \bar{c}(\gamma z)^2 &= 0 \end{aligned}$$

and it follows that the equation (2) has a solution $(0,0,0) \neq (\alpha x, \beta y, \gamma z) \in \mathbb{Z}^3$. Conversely, if (2) has a solution $(0,0,0) \neq (x,y,z) \in \mathbb{Z}^3$ then we have

$$\begin{aligned} \bar{a}x^2 + \bar{b}y^2 + \bar{c}z^2 &= 0 \\ \frac{a}{\alpha^2}x^2 + \frac{b}{\beta^2}y^2 + \frac{c}{\gamma^2}z^2 &= 0 \\ ax^2\beta^2\gamma^2 + by^2\alpha^2\gamma^2 + cz^2\alpha^2\beta^2 &= 0 \\ a(x\beta\gamma)^2 + b(y\alpha\gamma)^2 + c(z\alpha\beta)^2 &= 0 \end{aligned}$$

and it follows that equation (1) has a solution $(0,0,0) \neq (x\beta\gamma, y\alpha\gamma, z\alpha\beta) \in \mathbb{Z}^3$.

Thus we have reduced our problem to the following. Given **squarefree** integers $(a,b,c) \in \mathbb{Z}^3$ such that $\gcd(a,b,c) = 1$ and $abc \neq 0$, not all of the same sign, determine whether there exist integers $(0,0,0) \neq (x,y,z) \in \mathbb{Z}^3$ such that

$$ax^2 + by^2 + cz^2 = 0.$$

To complete the reduction, I claim that we can also assume that $\gcd(a,b) = \gcd(a,c) = \gcd(b,c) = 1$. To see this, suppose that a,b,c are **not pairwise-coprime**. By the symmetry of a,b,c we can assume without loss of generality that $d := \gcd(a,b) > 1$. Now let $a = da'$ and $b = db'$ and consider the following two equations:

$$ax^2 + by^2 + cz^2 = 0, \tag{1}$$

$$a'x^2 + b'y^2 + cdz^2 = 0. \tag{2}$$

I claim that (1) has a nontrivial integer solution if and only if (2) does. Indeed, suppose that (1) has a solution $(0,0,0) \neq (x,y,z) \in \mathbb{Z}^3$ so that $ax^2 + by^2 = -cz^2$. Since d is a common divisor of a and b this implies that d divides cz^2. But we also know that

$$1 = \gcd(a,b,c) = \gcd(\gcd(a,b),c) = \gcd(d,c),$$

so Euclid's Lemma tells us that $d|z^2$. Finally, since a and b are squarefree, d is also squarefree so that $d|z^2$ implies $d|z$, say $z = dz'$. It follows that

$$\begin{aligned} ax^2 + by^2 + cz^2 &= 0 \\ (da')x^2 + (db')y^2 + c(dz')^2 &= \\ d(a'x^2 + b'y^2 + cd(z')^2) &= 0 \\ a'x^2 + b'y^2 + cd(z')^2 &= 0 \end{aligned}$$

and hence (2) has a nontrivial solution $(0,0,0) \neq (x,y,z') \in \mathbb{Z}^3$. Conversely, suppose that (2) has a solution $(0,0,0) \neq (x,y,z) \in \mathbb{Z}^3$. Then it follows that

$$\begin{aligned} a'x^2 + b'y^2 + cdz^2 &= 0 \\ d(a'x^2 + b'y^2 + cdz^2) &= 0 \\ (da')x^2 + (db')y^2 + cd^2z^2 &= 0 \\ ax^2 + by^2 + c(dz)^2 &= 0 \end{aligned}$$

and hence (1) has a nontrivial solution $(0,0,0) \neq (x,y,dz) \in \mathbb{Z}^3$.

We have shown that equations (1) and (2) are equivalent. Now observe the following:

- Since $\gcd(d,c) = 1$ the integers $(a',b',cd) \in \mathbb{Z}^3$ are squarefree.
- Since $\gcd(a',b') = 1$ we have $\gcd(a',b',cd) = 1$.
- Since $1 < d$ we have

$$0 < |a'b'(cd)| < d \cdot |a'b'(cd)| = |(a'd)(b'd)c| = |abc|.$$

If the coefficients of (2) are still not pairwise-coprime then the first two observations above tell us that we can repeat the argument, and the third observation tells us that the process will eventually stop. In the end we will arrive at an equation that is equivalent to (1) in which the coefficients are pairwise-coprime. As an example of this reduction procedure consider three distinct (positive or negative) primes p,q,r. Then the solvability of the following equations are equivalent:

$$\begin{aligned} pqx^2 + pry^2 + qrz^2 &= 0 \\ qx^2 + ry^2 + pqrz^2 &= 0 \\ x^2 + qry^2 + prz^2 &= 0 \\ rx^2 + qy^2 + pz^2 &= 0. \end{aligned}$$

In summary, we have reduced the problem of the existence of rational points on a conic section to the case of Legendre's Theorem. //

Proof of Legendre's Theorem. So consider any squarefree integers $(a,b,c) \in \mathbb{Z}^3$ with $\gcd(a,b) = \gcd(a,c) = \gcd(b,c) = 1$ and $abc \neq 0$, not all of the same sign. We want to show that the equation

$$ax^2 + by^2 + cz^2 = 0$$

has a nontrivial solution $(0,0,0) \neq (x,y,z) \in \mathbb{Z}^3$ if and only if the elements $[-ab]_c$, $[-ac]_b$ and $[-bc]_a$ have square roots.

First the easy direction. Assume that a nontrivial solution exists. We will show that $[-ab]_c$ has a square root and then the other cases will follow from symmetry. Since the equation

$$ax^2 + by^2 + cz^2 = 0$$

is homogeneous in x,y,z we can assume that $\gcd(x,y,z) = 1$. Then I claim that $\gcd(x,c) = 1$. To see this, **assume for contradiction** that c and x have a common prime divisor p, so that p divides $ax^2 + cz^2 = -by^2$. But we know that $p \nmid b$ because $\gcd(b,c) = 1$ so Euclid's Lemma tells us that $p|y$. Then since $p|x$ and $p|y$ we see that p^2 divides $ax^2 + by^2 = -cz^2$. But we already know that $p|c$, and since c is squarefree this implies that $p^2|cz^2 \Rightarrow p|z$. We have shown that p is a common divisor of x,y,z which contradicts the fact that $\gcd(x,y,z) = 1$.

In summary we conclude that $\gcd(x,c) = 1$ and hence the element $[x]_c \in \mathbb{Z}/c\mathbb{Z}$ is invertible. Finally, we reduce the equation $ax^2 + by^2 + cz^2 = 0 \bmod c$ to obtain

$$\begin{aligned}
[ax^2 + by^2 + cz^2]_c &= [0]_c \\
[ax^2 + by^2]_c &= [-cz^2]_c \\
[ax^2 + by^2]_c &= [0]_c \\
[ax^2]_c &= [-by^2]_c \\
[a]_c &= [-by^2]_c \cdot [x^{-2}]_c \\
[-b]_c \cdot [a]_c &= [-b]_c \cdot [-by]_c \cdot [x^{-2}]_c \\
[-ab]_c &= \left([-b]_c \cdot [y]_c \cdot [x^{-1}]_c\right)^2,
\end{aligned}$$

and hence $[-ab]_c$ is square.

Now the hard direction. Assume that each of the elements $[-ab]_c$, $[-ac]_b$ and $[-bc]_a$ has a square root. Our goal is to prove that there exist integers $(0,0,0) \neq (x,y,z) \in \mathbb{Z}^3$ such that

$$ax^2 + by^2 + cz^2 = 0.$$

There are two steps:

(1) Since we have $abc \neq 0$, not all of the same sign, we can assume without loss of generality that exactly two of a, b, c are negative and hence $abc \geqslant 1$. We will prove that there exist integers $A, B, C, D, E, F \in \mathbb{Z}$ such that for any integers $(x,y,z) \in \mathbb{Z}^3$ we have

$$[ax^2 + by^2 + cz^2]_{abc} = [Ax + By + Cz]_{abc} \cdot [Dx + Ey + Fz]_{abc}.$$

To do this we consider any integers $(x,y,z) \in \mathbb{Z}^3$ and then we reduce the integer $ax^2+by^2+cz^2 \in \mathbb{Z} \bmod a$. Since $\gcd(a,b) = 1$ we know that there exists $b^* \in \mathbb{Z}$ with $[bb^*]_a = [1]_a$ and since $[-bc]_a$ is square we have $[-ba]_a = [k^2]_a$ for some $k \in \mathbb{Z}$. Then we obtain

$$\begin{aligned}
[ax^2 + by^2 + cz^2]_a &= [by^2 + cy^2]_a \\
&= [b]_a \cdot [y^2 + b^*cz^2]_a \\
&= [b]_a \cdot [y^2 - (b^*)^2(-bc)z^2]_a \\
&= [b]_a \cdot [y^2 - (b^*)^2k^2z^2]_a \\
&= [b]_a \cdot [y^2 - (b^*kz)^2]_a \\
&= [b]_a \cdot [(y - b^*kz)(y + b^*kz)]_a \\
&= [0x + by - kz]_a \cdot [0x + y + b^*kz]_a.
\end{aligned}$$

In other words, there exist integers $A_1, B_1, C_1, D_1, E_1, F_1 \in \mathbb{Z}$ such that

$$[ax^2 + by^2 + cz^2]_a = [A_1x + B_1y + C_1z]_a \cdot [D_1x + E_1y + F_1z]_a$$

and similar arguments show that there exist integers $A_2, \ldots, F_3 \in \mathbb{Z}$ such that

$$\begin{aligned}
[ax^2 + by^2 + cz^2]_b &= [A_2x + B_2y + C_2z]_b \cdot [D_2x + E_2y + F_2]_b \\
[ax^2 + by^2 + cz^2]_c &= [A_3x + B_3y + C_3z]_c \cdot [D_3x + E_3y + F_3]_c.
\end{aligned}$$

Now since a, b, c are pairwise coprime, the Chinese Remainder Theorem from Section 3.5 tells us that there exists an integer $A \in \mathbb{Z}$ satisfying

$$\begin{aligned}[A]_a &= [A_1]_a \\ [A]_b &= [A_2]_b \\ [A]_c &= [A_3]_c,\end{aligned}$$

and similarly we have integers $B, C, D, E, F \in \mathbb{Z}$ so that

$$\begin{aligned}[ax^2 + by^2 + cz^2]_a &= [Ax + By + Cz]_a \cdot [Dx + Ey + Fz]_a \\ [ax^2 + by^2 + cz^2]_b &= [Ax + By + Cz]_b \cdot [Dx + Ey + Fz]_b \\ [ax^2 + by^2 + cz^2]_c &= [Ax + By + Cz]_c \cdot [Dx + Ey + Fz]_c.\end{aligned}$$

Finally, recall from HW3.4 that if we have $[d]_a = [e]_a$ and $[d]_b = [e]_b$ for some integers a, b, d, e with $\gcd(a, b) = 1$ then it follows that $[d]_{ab} = [e]_{ab}$. Since a, b, c are pairwise coprime we can apply this argument twice to the above system of three congruences to obtain

$$(*) \qquad [ax^2 + by^2 + cz^2]_{abc} = [Ax + By + Cz]_{abc} \cdot [Dx + Ey + Fz]_{abc}$$

as desired.

(2) The congruence $(*)$ suggests a strategy to find integers $(0, 0, 0) \neq (x, y, z) \in \mathbb{Z}^3$ such that $ax^2 + by^2 + cz^2 = 0$. Observe that the expression

$$[Ax + By + Cz]_{abc} \in \mathbb{Z}/abc\mathbb{Z}$$

can take on at most abc distinct values. Since there exist **infinitely many** integer triples $(x, y, z) \in \mathbb{Z}^3$ there must be two distinct triples $(x_1, y_1, z_1) \neq (x_2, y_2, z_2)$ with the property

$$\begin{aligned}[Ax_1 + By_1 + Cz_1]_{abc} &= [Ax_2 + By_2 + Cz_2]_{abc} \\ [(Ax_1 + By_1 + Cz_1) - (Ax_2 + By_2 + Cz_2)]_{abc} &= [0]_{abc} \\ [A(x_1 - x_2) + B(y_1 - y_2) + C(z_1 - z_2)]_{abc} &= [0]_{abc}.\end{aligned}$$

In other words, we have found integers $(x, y, z) := (x_1 - x_2, y_1 - y_2, z_1 - z_2) \neq (0, 0, 0)$ such that $[Ax + By + Cz]_{abc} = [0]_{abc}$ and then from the congruence $(*)$ we obtain

$$(**) \qquad [ax^2 + by^2 + cz^2]_{abc} = [0]_{abc}.$$

This doesn't necessarily mean that $ax^2 + by^2 + cz^2 = 0$, but it does mean that $ax^2 + by^2 + cz^2 = abck$ for some $k \in \mathbb{Z}$. Our goal is to choose the points (x_1, y_1, z_1) and (x_2, y_2, z_2) **sufficiently close together** so that $k = 0$. To do this we consider the following rectangular box of integer points:

$$\text{Box} := \{(x, y, z) \in \mathbb{Z}^3 : 0 \leqslant x \leqslant \sqrt{|bc|}, 0 \leqslant y \leqslant \sqrt{|ac|}, 0 \leqslant z \leqslant \sqrt{|ab|}\}$$

Observe that the number of $x \in \mathbb{Z}$ satisfying $0 \leqslant x \leqslant \sqrt{|bc|}$ is strictly greater than $\sqrt{|bc|}$ and a similar observation holds for y and z. Thus the number of points in the box satisfies

$$\#\text{Box} > \sqrt{|bc|}\sqrt{|ac|}\sqrt{|ab|} = \sqrt{(abc)^2} = |abc|.$$

Since there are more than $|abc|$ points in the box it follows that we can choose the two points $(x_1, y_1, z_1) \neq (x_2, y_2, z_2)$ from inside the box. Then since a, b, c are pairwise coprime and $abc \neq 0$ is squarefree we see that none of $\sqrt{|bc|}$, $\sqrt{|ac|}$ and $\sqrt{|ab|}$ is an integer. Thus the point $(x, y, z) = (x_1 - x_2, y_1 - y_2, z_1 - z_2) \neq (0, 0, 0)$ satisfies

$$|x| < \sqrt{|bc|}, \quad |y| < \sqrt{|ac|} \quad \text{and} \quad |z| < \sqrt{|ab|},$$

and hence also

$$0 < |a|x^2 + |b|y^2 + |c|z^2 < 3|abc|.$$

Finally, since a, b, c don't all have the same sign **we can assume without loss of generality that a is positive and b, c are negative**. In particular, this implies that $|abc| = abc > 0$. Then we have

$$ax^2 + by^2 + cz^2 < ax^2 < abc$$

and

$$ax^2 + by^2 + cz^2 \geqslant by^2 + cz^2 > b(-ac) + c(-ab) = -2abc.$$

And combining these inequalities with the congruence $(**)$ gives

$$ax^2 + by^2 + cz^2 \in \{0, -abc\}.$$

If $ax^2 + by^2 + cz^2 = 0$ then we are **done** so let us assume that $ax^2 + by^2 + cz^2 = -abc$. Then we can make the clever[31] change of variables

$$(x', y', z') := (xz - by, yz + ax, z^2 + ab) \neq (0, 0, 0)$$

to obtain

$$\begin{aligned} a(x')^2 + b(y'^2) + c(z')^2 &= a(xz - by)^2 + b(yz + ax)^2 + c(z^2 + ab)^2 \\ &= (ax^2 + by^2 + cz^2)z^2 + 2abcz^2 + ab^2y^2 + a^2bx^2 + a^2b^2c \\ &= (-abc)z^2 + 2abcz^2 + ab^2y^2 + a^2bx^2 + a^2b^2c \\ &= abcz^2 + ab^2y^2 + a^2bx^2 + a^2b^2c \\ &= ab(ax^2 + by^2 + cz^2) + a^2b^2c \\ &= ab(-abc) + a^2b^2c \\ &= 0. \end{aligned}$$

Thus we have found the desired solution $(0, 0, 0) \neq (x', y', z') \in \mathbb{Z}^3$. To complete the proof of the bound, one can show by an easy and tedious computation that

$$0 < |a|(x')^2 + |b|(y')^2 + |c|(z')^2 = a(x')^2 - b(y')^2 - c(z')^2 < 8|abc|.$$

Since the bound is symmetric in a, b, c we observe that it is independent of our assumption that $a > 0 > b, c$. This completes the proof of Legendre's Theorem. □

[31] too clever

5.4 Primitive Roots and Euler's Criterion

In the last section we considered the *Legendre equation*

$$ax^2 + by^2 + cz^2 = 0$$

with integer coefficients $(a, b, c) \in \mathbb{Z}^3$ satisfying $abc \neq 0$. We proved Legendre's Theorem which says that a nontrivial integer solution $(0, 0, 0) \neq (x, y, z) \in \mathbb{Z}^3$ exists if and only if certain elements have modular square roots. As a free corollary of the existence proof we also obtained a bound on the size of the smallest nontrivial solution, however our bound was not optimal. The sharpest possible bound was obtained by Holzer (1950) for the case when a, b, c are squarefree and pairwise relatively prime. Mordell (1951) gave an elementary proof of Holzer's result and then Williams (1988) generalized the result to arbitrary a, b, c. I will state their result without proof.

Theorem (Smallest Solution to Legendre's Equation). Consider any integers $(a, b, c) \in \mathbb{Z}^3$ with $abc \neq 0$ and $d = \gcd(a, b, c)$. If the Legendre equation

$$ax^2 + by^2 + cz^2 = 0$$

has a nontrivial integer solution $(0, 0, 0) \neq (x, y, z) \in \mathbb{Z}^3$ then it has a solution satisfying

$$|x| \leqslant \frac{\sqrt{|bc|}}{d}, \quad |y| \leqslant \frac{\sqrt{|ac|}}{d} \quad \text{and} \quad |z| \leqslant \frac{\sqrt{|ab|}}{d}.$$

//

Thus by testing every integer point in the box

$$\left\{(x, y, z) \in \mathbb{Z}^3 : 0 \leqslant x \leqslant \sqrt{|bc|}/d, 0 \leqslant y \leqslant \sqrt{|ac|}/d, 0 \leqslant z \leqslant \sqrt{|ab|}/d\right\}$$

we obtain an algorithm of complexity $|abc|/d^3$ that either finds a nontrivial solution to Legendre's Equation or proves that no such solution exists. For the purpose of computing a solution there is probably no faster method.

However, this algorithm is in some sense unsatisfying because it ignores the criterion from Legendre's Theorem on the existence of certain square roots. In this section and the next we will pursue a deeper study of square roots in order to understand the nature of the solutions. At the end of the chapter we will obtain a much faster algorithm that determines whether a solution exists without actually finding a solution. This discussion will lead us naturally into some deeper concepts of number theory.

Our general goal is to determine when a given integer is a quadratic residue (i.e., has a square root) mod n. To begin the study we assume that $n = p$ is **prime**.

Definition of the Legendre Symbol. For any integers $a, p \in \mathbb{Z}$ with p prime we define the *Legendre symbol* as follows:

$$\left(\frac{a}{p}\right)_2 = \begin{cases} 1 & \text{if } \exists\, x \in \mathbb{Z} \text{ with } [a]_p = [x^2]_p \\ -1 & \text{if } \nexists\, x \in \mathbb{Z} \text{ with } [a]_p = [x^2]_p \\ 0 & \text{if } [a]_p = [0]_p \end{cases}$$

Warning: Most authors omit the subscript "2" from the Legendre symbol, which I think results in one of the most confusing notations in any branch of mathematics. //

When $p = 2$ we observe that $(a/2)_2 = 0$ for a even and $(a/2)_2 = 1$ for a odd, and there is nothing else to say, so let us assume that p is an odd prime. The main theorem of this section is an explicit and easily computable[32] **formula** for the Legendre symbol.

Theorem (Euler's Criterion). Let p be an odd prime. Then for any integer p we have

$$\left[\left(\frac{a}{p}\right)_2\right]_p = \left[a^{(p-1)/2}\right]_p.$$

//

As an immediate corollary we obtain the following important fact.

Corollary of Euler's Criterion. Let p be an odd prime. Then for all $a, b \in \mathbb{Z}$ we have

$$\left(\frac{a}{p}\right)_2 \cdot \left(\frac{b}{p}\right)_2 = \left(\frac{ab}{p}\right)_2.$$

//

In fancier terms, we can say that the function $(\mathbb{Z}/p\mathbb{Z})^\times \to \{\pm 1\}$ defined by $[a]_p \mapsto (a/p)_2$ is a "homomorphism" of multiplicative groups. In even fancier terms, the function $[a]_p \mapsto (a/p)_2$ is called a "character" of the group $(\mathbb{Z}/p\mathbb{Z})^\times$, and for this reason the Legendre symbol $(a/p)_2$ is also called the *quadratic character* of a mod p.

Proof of the Corollary. We don't really **need** Euler's Criterion to prove this, but with Euler's Criterion the proof becomse trivial:

$$\begin{aligned} \left[\left(\frac{a}{p}\right)_2\right]_p \cdot \left[\left(\frac{b}{p}\right)_2\right]_p &= [a^{(p-1)/2}]_p \cdot [b^{(p-1)/2}]_p \\ &= [a^{(p-1)/2} \cdot b^{(p-1)/2}]_p \\ &= [(ab)^{(p-1)/2}]_p \\ &= \left[\left(\frac{ab}{p}\right)_2\right]_p. \end{aligned}$$

[32] because modular exponentiation is easy

Then the result follows because $p > 2$. $\square$

Instead of presenting the **quickest** proof of Euler's Criterion I will present the **best** proof, and this will also give me an excuse to introduce some ideas that should be part of every undergraduate number theory course. I will state these ideas as three separate lemmas.

Lemma 1 (Counting Reduced Fractions). For all $n \in \mathbb{Z}$ with $n \geqslant 1$ we have

$$n = \sum_{1 \leqslant d | n} \varphi(d),$$

where the sum runs over all positive divisors of n. //

Proof. For each integer $n \geqslant 1$ we define the following two sets of fractions:

$$F_n := \left\{ \frac{1}{n}, \frac{2}{n}, \cdots, \frac{n}{n} \right\},$$
$$F'_n := \left\{ \frac{k}{n} : 1 \leqslant k \leqslant n \wedge \gcd(k, n) = 1 \right\}.$$

Note that by definition we have $\#F_n = n$ and $\#F'_n = \varphi(n)$. By reducing each fraction in F_n to lowest terms, I will show that F_n decomposes as the disjoint union

$$F_n = \coprod_{1 \leqslant d | n} F'_d,$$

and then the theorem will follow by taking the cardinality of each side. There are three things to show:

(1) $F_n \subseteq \cup_{1 \leqslant d | n} F'_d$**:** Consider any fraction $k/n \in F_n$, i.e., with $1 \leqslant k \leqslant n$, and suppose that we have $\lambda = \gcd(k, n)$ with $k = \lambda k'$ and $n = \lambda n'$. By by dividing the numberator and denominator by their greatest common divisor we obtain

$$\frac{k}{n} = \frac{\lambda k'}{\lambda n'} = \frac{k'}{n'}.$$

with $\gcd(k', n') = 1$ and $1 \leqslant k' \leqslant n'$. [Why?] It follows that $k/n \in F'_{n'}$ and then since n' is a positive divisor of n we obtain $k/n \in \cup_{1 \leqslant d | n} F'_d$ as desired.

(2) $\cup_{1 \leqslant d | n} F'_d \subseteq F_n$**:** Suppose that the fraction $\alpha \in \mathbb{Q}$ is an element of the union $\cup_{1 \leqslant d | n} F'_d$. Then by definition we must have $\alpha \in F_d$ for some positive divisor $d | n$, i.e., we must have $\alpha = k/d$ with $\gcd(k, d) = 1$ and $1 \leqslant k \leqslant d$. But then since d is a divisor of n we have $n = \lambda d$ for some $1 \leqslant \lambda \in \mathbb{Z}$ and it follows that

$$\frac{k}{d} = \frac{\lambda k}{\lambda d} = \frac{\lambda k}{n} \in F_n$$

as desired. [Why is $1 \leqslant \lambda k \leqslant n$?]

(3) $\cup_{1\leqslant d|n} F'_d = \coprod_{1\leqslant d|n} F'_d$ **:** To show that the union is disjoint, assume for contradiction that there exists a fraction $\alpha \in \mathbb{Q}$ such that we have $\alpha \in F'_d \cap F'_e$ for distinct positive integers $d \neq e$. Since $\alpha \in F'_d$ we must have $\alpha = k/d$ for some $1 \leqslant k \leqslant d$ with $\gcd(k,d) = 1$ and since $\alpha \in F'_e$ we must have $\alpha = k'/e$ for some $1 \leqslant k' \leqslant e$ with $\gcd(k',e) = 1$. Since $k/d = \alpha = k'/e$ we see that

$$ke = k'd.$$

But now since $e|k'd$ with $\gcd(k',e) = 1$, Euclid's Lemma says that $e|d$, and a similar argument shows that $d|e$. Finally, since d and e are both positive we must have $d = e$, which is the desired contradiction. □

For example, note that the positive divisors of $n = 15 = 3 \cdot 5$ are 1, 3, 5 and 15. Then note that we have

$$\varphi(1) + \varphi(3) + \varphi(5) + \varphi(15) = 1 + (3-1) + (5-1) + (3-1)(5-1) = 1 + 2 + 4 + 8 = 15$$

as expected. The totient function value $\varphi(1)$ is not really defined but we will adopt the convention $\varphi(1) := 1$ precisely so this formula works out.

The next lemma has to do with counting solutions of polynomial equations in the rings $\mathbb{Z}/n\mathbb{Z}$ for various n. As an extreme case, one can check by hand that the equation

$$[x^2 - 1]_8 = [0]_8$$

is true for **every** element of the ring $[x]_8 \in \mathbb{Z}/8\mathbb{Z}$. However, it turns out that something very special happens in the rings $\mathbb{Z}/p\mathbb{Z}$ for prime p.

Lemma 2 (Lagrange's Polynomial Congruence Theorem). Let $p \in \mathbb{Z}$ be prime and consider a polynomial of degree d with integer coefficients:

$$f(x) = a_d x^d + \cdots + a_1 x + a_0$$

with $a_0, a_1, \ldots, a_d \in \mathbb{Z}$ and $a_d \neq 0$. If $[a_d]_p \neq [0]_p$ then I claim that the equation

$$[f(x)]_p = [0]_p$$

has at most d distinct solutions $[x]_p \in \mathbb{Z}/p\mathbb{Z}$. //

Remark: This follows from a well-known theorem in abstract algebra. That is, if K is any field and if $f(x) \in K[x]$ is a polynomial of degree d then there exist at most d distinct solutions $x \in K$ of the equation $f(x) = 0$. The lemma then follows because $\mathbb{Z}/p\mathbb{Z}$ is a field. However, I prefer to present the proof in language that Lagrange would recognize.

Proof. We will use induction on the degree d. If $d = 1$ then the polynomial has the form

$$f(x) = a_1 x + a_0$$

with $a_1, a_0 \in \mathbb{Z}$ and $[a_1]_p \neq [0]_p$. Then since $\gcd(a_1, p) = 1$ we see that the element $[a_1]_p$ is invertible and we obtain a unique solution:

$$\begin{aligned} [a_1 x + a_0]_p &= [0]_p \\ [a_1]_p \cdot [x]_p &= [-a_0]_p \\ [x]_p &= [a_1^{-1}]_p \cdot [-a_0]_p. \end{aligned}$$

Now let us fix $d \geqslant 2$ and **assume for induction** that the theorem is true for polynomials of degree $d-1$. If the equation $[f(x)]_p = [0]_p$ has no solutions then we are done, so suppose that there exists $b \in \mathbb{Z}$ with $[f(b)]_p = [0]_p$. If this is the only solution then we are still done because $d \geqslant 2$. So suppose that we have another solution $[f(c)]_p = [0]_p$ with $[b]_p \neq [c]_p$. If we can show that there are at most $d-1$ such solutions $[c]_p$ with $[c]_p \neq [b]_p$ then it will follow that $[f(x)]_p = [0]_p$ has at most d solutions as desired.

To prove this we will use the fact that for all integers $x, n \in \mathbb{Z}$ with $n \geqslant 1$ we have

$$(x^n - b^n) = (x-b)(x^{n-1} + x^{n-2}b + \cdots + xb^{n-2} + b^{n-1}).$$

Then since $[f(b)]_p = [0]_p$ we compute that

$$\begin{aligned} [f(x)]_p &= [f(x)]_p - [f(b)]_p \\ &= [f(x) - f(b)]_p \\ &= [a_d(x^d - b^d) + \cdots + a_1(x-b) + 0]_p \\ &= [(x-b)(a_d x^{d-1} + \text{ lower terms in } x\ )]_p \\ &= [x-b]_p \cdot [g(x)]_p \end{aligned}$$

for some polynomial $g(x) \in \mathbb{Z}[x]$ of degree $d-1$ whose leading coefficient a_d is not divisible by p. Now if $[c]_p$ is any solution of $[f(c)]_p = [0]_p$ with $[c]_p \neq [b]_p$, then since the element $[c-b]_p \neq [0]_p$ is **invertible** we obtain

$$\begin{aligned} [0]_p &= [f(c)]_p \\ [0]_p &= [c-b]_p \cdot [g(c)]_p \\ [0]_p &= [g(c)]_p. \end{aligned}$$

It follows from the induction hypothesis that there exist at most $d-1$ distinct such $[c]_p$, which completes the proof. □

To set up the last of the three lemmas, recall from Chapter 3 that for all integers $a, n \in \mathbb{Z}$ with $\gcd(a, n) = 1$ there exists a positive integer $\mathrm{ord}_n(a) \geqslant 1$ with the following properties:

- $[a^{\mathrm{ord}_n(a)}]_n = [1]_n$,
- $[a^k]_n \neq [1]_n$ for all $0 < k < \mathrm{ord}_n(a)$.

The integer $\text{ord}_n(a)$ is called the *multiplicative order of a mod n* and it follows from Euler's Totient Theorem that $\text{ord}_n(a)|\varphi(n-1)$. However, the precise value of $\text{ord}_n(a)$ is very difficult to predict in general. Suppose for the sake of argument that there exists an element $[g]_n \in (\mathbb{Z}/n\mathbb{Z})^\times$ with the property that $\text{ord}_n(g) = \varphi(n)$. Then it follows that every element of $(\mathbb{Z}/n\mathbb{Z})^\times$ can be expressed as a power of this element:

$$(\mathbb{Z}/n\mathbb{Z})^\times = \left\{ [g]_n, [g^2]_n, \ldots, [g^{\varphi(n)}]_n = [1]_n \right\}.$$

In this case we say that the group $(\mathbb{Z}/n\mathbb{Z})^\times$ is *cyclic* and we say that $[g]_n \in (\mathbb{Z}/n\mathbb{Z})^\times$ is a *genertor* (hence the letter "g"). An alternative and older notation would call g a *primitive root mod n*.

[Warning: The **additive** group $(\mathbb{Z}/n\mathbb{Z}, +, [0]_n)$ is always cyclic with generator $[1]_n$. Here we are asking whether the **multiplicative** group $((\mathbb{Z}/n\mathbb{Z})^\times, \times, [1]_n)$ is cyclic, which is a separate issue.]

A primitive root is a nice thing to have because then we can phrase all properties of $(\mathbb{Z}/n\mathbb{Z})^\times$ in terms of powers of g. In particular, we would see that an element $[a]_n \in (\mathbb{Z}/n\mathbb{Z})^\times$ has a square root if and only if it is an **even** power of $[g]_n$. Unfortunately, primitive roots don't always exist. For example, recall that we have $\varphi(8) = 4$ with

$$(\mathbb{Z}/8\mathbb{Z})^\times = \{[1]_8, [3]_8, [5]_8, [7]_8\}.$$

The following table lists the multiplicative order of each element of the group:

a	1	3	5	7
$\text{ord}_8(a)$	1	2	2	2

Observe that we always have $\text{ord}_8(a)|\varphi(8) = 4$ as required by Euler's Totient Theorem, but it is **never** the case that $\text{ord}_8(a) = \varphi(8)$. This is related to the fact that the equation $[x^2 - 1]_8 = [0]_8$ has too many solutions (i.e., more than 2) in the ring $\mathbb{Z}/8\mathbb{Z}$.

The celebrated "Primitive Root Theorem" says that primitive roots always exist in the group $(\mathbb{Z}/p\mathbb{Z})^\times$ when p is prime.[33] Unfortunately, the proof is non-constructive, i.e., it does not tell us how to actually **find** a primitive root. However, we will see that there exist exactly $\varphi(\varphi(p)) = \varphi(p-1)$ primitive roots mod p so at least we know how long it will take us to find one via random search.

Lemma 3 (The Primitive Root Theorem). For any prime $p \in \mathbb{Z}$ the group of units $(\mathbb{Z}/p\mathbb{Z})^\times$ is cyclic. That is, there exists an element $[g]_p \in (\mathbb{Z}/p\mathbb{Z})^\times$ with multiplicative order $\text{ord}_p(g) = \varphi(p) = p-1$. More precisely, we will show that there are exactly $\varphi(\varphi(p)) = \varphi(p-1)$ such "primitive roots". //

[33] The general theorem says that the group $(\mathbb{Z}/n\mathbb{Z})^\times$ is cyclic if and only if $n = 2p^k$ for some odd prime p, but we won't prove this.

Sadly, there is no really short proof of this fact. We will have to get our hands dirty.

Proof. Recall from Euler's Totient Theorem that each element of $(\mathbb{Z}/p\mathbb{Z})^\times$ has order d for some positive divisor $1 \leqslant d|\varphi(p)$. We will prove that the number of elements of order d is exactly $\varphi(d)$. Then the result follows by putting $d = \varphi(p)$.

So fix any positive divisor $1 \leqslant d|\varphi(p)$ and suppose that there exists an element $[a]_p \in (\mathbb{Z}/p\mathbb{Z})^\times$ of order d. By definition this means that the elements

$$(*) \qquad [a]_p, [a^2]_p, \ldots, [a^{d-1}]_p, [a^d]_p = [1]_p$$

are all distinct, since otherwise we would have $[a^k]_p = [a^\ell]_p$ for some $1 \leqslant k < \ell \leqslant d$ and it would follow that $[a^{\ell-k}]_p = [1]_p$ for some $1 \leqslant \ell - k < d$, contradicting the fact that d is the smallest positive integer such that $[a^d]_p = [1]_p$. Furthermore, since $[a^d]_p = [1]_p$ we see that

$$[(a^k)^d]_p = [(a^d)^k]_p = \Big([a^d]_p\Big)^k = ([1]_p)^k = [1]_p$$

for all integers $k \in \mathbb{Z}$. It follows that the d distinct elements $(*)$ are all solutions to the equation $[x^d - 1]_p = [0]_p$. But **Lemma 2** says that this equation has **at most** d solutions, so $(*)$ is the complete solution.

We have seen that every element $[x]_p \in (\mathbb{Z}/p\mathbb{Z})^\times$ of order d is in the list $(*)$, and hence the number of such elements is $\leqslant d$. But this is not a sharp estimate because some of the elements $(*)$ have order **less than** d. To be precise, consider a fixed element $[a^k]_p$ and let $\lambda := \gcd(k, d)$ with $k = \lambda k'$ and $d = \lambda d'$. Then I claim that the order of the element $[a^k]_p$ is precisely $d' = d/\gcd(k, d)$. To see this, first note that

$$[(a^k)^{d'}]_p = [a^{\lambda k' d'}]_p = [(a^{\lambda d'})^{k'}]_p = [(a^d)^{k'}]_p = \Big([a^d]_p\Big)^{k'} = ([1]_p)^{k'} = [1]_p.$$

Now assume that we have $[(a^k)^n]_p = [a^{kn}]_p = [1]_p$ for some positive integer $1 \leqslant n \in \mathbb{Z}$. Since d is the order of $[a]_p$ this implies that d divides kn. [Remind yourself why this is true. Hint: Divide d by kn and show that the remainder must be zero.] Thus we have $kn = d\ell$ for some $\ell \in \mathbb{Z}$. Then since $\lambda \neq 0$ we have

$$\begin{aligned} kn &= d\ell \\ (\lambda k')n &= (\lambda d')\ell \\ \lambda(k'n) &= \lambda(d'\ell) \\ k'n &= d'\ell. \end{aligned}$$

Finally, since $d'|k'n$ and $\gcd(k', d') = 1$, Euclid's Lemma says that $d'|n$ and since n is positive this implies that $d' \leqslant n$. This completes the proof that $\mathrm{ord}_p(a^k) = d/\gcd(k, d)$. We conclude that **if** the group $(\mathbb{Z}/p\mathbb{Z})^\times$ contains an element $[a]_p$ of order d, **then** the complete set of elements of order d is

$$\Big\{[a^k]_p : 1 \leqslant k \leqslant d \wedge d/\gcd(k, d) = d\Big\} = \Big\{[a^k]_p : 1 \leqslant k \leqslant d \wedge \gcd(k, d) = 1\Big\},$$

and there are precisely $\varphi(d)$ of these elements. In summary, for any positive divisor $1 \leqslant d|\varphi(p)$, the group $(\mathbb{Z}/p\mathbb{Z})^\times$ contains either 0 or $\varphi(d)$ elements of order d.

It only remains to count them up and see what we have. So let ν_d denote the number of elements of order d and recall that $\nu_d \in \{0, \varphi(d)\}$. Then on the one hand we have

$$(*) \qquad p-1 = \varphi(p) = \#(\mathbb{Z}/p\mathbb{Z})^\times = \sum_{1\leqslant d|n} \nu_d \leqslant \sum_{1\leqslant d|n} \varphi(d)$$

with equality if and only if $\nu_d = \varphi(d)$ for all $1 \leqslant d|\varphi(p)$. On the other hand, we know from **Lemma 1** that

$$\sum_{1\leqslant d|\varphi(p)} \varphi(d) = \varphi(p) = p-1.$$

Thus the inequality $(*)$ is really an **equality**, and it follows that $\nu_d = \varphi(d)$ for all $1 \leqslant d|\varphi(p)$ as desired. □

That was a lot of work, but we learned some wholesome things that will help us later.

Proof of Euler's Criterion. Let p be an odd prime so that $(p-1)/2 \in \mathbb{Z}$ and consider any integer $a \in \mathbb{Z}$. We want to prove that $[(a/p)_2]_p = [a^{(p-1)/2}]_p$.

If $[a]_p = [0]_p$ then by definition we have $(a/p)_2 = 0$ and hence

$$[a^{(p-1)/2}]_p = ([a]_p)^{(p-1)/2} = ([0]_p)^{(p-1)/2} = [0]_p = \left[\left(\frac{a}{p}\right)_2\right]_p.$$

So let us assume that $[a]_p \neq [0]_p$, i.e., $[a]_p \in (\mathbb{Z}/p\mathbb{Z})^\times$. Then Euler's Totient Theorem gives

$$\left[\left(a^{(p-1)/2}\right)^2 - 1\right]_p = [a^{p-1}]_p - [1]_p = [a^{\varphi(p)}]_p - [1]_p = [1]_p - [1]_p = [0]_p.$$

But Lagrange's Congruence Theorem says that the equation $[x^2-1]_p = [0]_p$ has at most two solutions $[x]_p \in \mathbb{Z}/p\mathbb{Z}$. Since $[x]_p = [1]_p$ and $[x]_p = [-1]_p$ are solutions we conclude that

$$[a^{(p-1)/2}]_p = [1]_p \text{ or } [-1]_p.$$

Similarly we have by definition that

$$\left(\frac{a}{p}\right)_2 = 1 \text{ or } -1$$

and it only remains to show that the functions $[a^{(p-1)/2}]_p$ and $[(a/p)_2]_p$ are equal to $[1]_p$ for the same values of a.

To show this we will use the **Primitive Root Theorem**, which tells us that there exists a generator $[g]_p$ such that $(\mathbb{Z}/p\mathbb{Z})^\times = \{[g]_p, [g^2]_p, \ldots, [g^{p-1}]_p = [1]_p\}$. In particular, we have $[a]_p = [g^k]_p$ for some $1 \leqslant k \leqslant p-1$. There are two things to show:

(1) We have $[a^{(p-1)/2}]_p = [1]_p$ if and only if k is even. Suppose that k is even with $k = 2k'$. Then by Euler's Totient Theorem we have

$$[a^{(p-1)/2}]_p = \left[\left(g^{2k'}\right)^{(p-1)/2}\right]_p = \left[(g^{k'})^{p-1}\right]_p = [1]_p$$

and hence $[a^{(p-1)/2} - 1]_p = [0]_p$. We have seen that the equation $[x^{(p-1)/2} - 1]_p = [1]_p$ holds for the $(p-1)/2$ distinct elements

$$[x]_p \in \left\{[g^2]_p, [g^4]_p, \ldots, [g^{p-1}]_p = [1]_p\right\}.$$

But then since the polynomial $x^{(p-1)/2} - 1 \in \mathbb{Z}[x]$ has degree $(p-1)/2$, Lagrange's Congruence Theorm tells us that this is the full solution.

(2) We have $(a/p)_2 = 1$ if and only if k is even. If k is even (say $k = 2k'$) then we see that $[a]_p = [g^{2k'}]_p = \left([g^{k'}]_p\right)^2$ is square and hence $(a/p)_2 = 1$. Conversely, let k be odd and assume for contradiction that $(a/p)_2 = 1$, i.e., that we have $[g^k]_p = [x^2]_p$ for some $x \in \mathbb{Z}$. Since $[g]_p$ is a generator we have $[x]_p = [g^\ell]_p$ for some $\ell \in \mathbb{Z}$ and then

$$[g^k]_p = [x^2]_p \quad \Longrightarrow \quad [g^k]_p = [(g^\ell)^2]_p \quad \Longrightarrow \quad [g^{k-2\ell}]_p = [1]_p.$$

But since $\varphi(p) = p - 1$ is the order of $[g]_p$ this implies that $(p-1)|(k-2\ell)$. Finally, since $p-1$ is even this implies that $2|(k-2\ell)$ which contradicts the fact that k is even.

This completes the proof of Euler's Criterion. □

Special cases $a = -1$ and $a = 2$. For odd primes p we have

$$\left(\frac{-1}{p}\right)_2 = (-1)^{(p-1)/2} = \begin{cases} 1 & p = +1 \pmod 4 \\ -1 & p = -1 \pmod 4 \end{cases}$$

$$\left(\frac{2}{p}\right)_2 = (-1)^{(p^2-1)/8} = \begin{cases} 1 & p = \pm 1 \pmod 8 \\ -1 & p = \pm 3 \pmod 8 \end{cases}$$

Application: Infinitely many primes $= 1, 3, 5, 7 \bmod 8$.

5.5 Quadratic Reciprocity

A bit of group theory.

Zolotarev Reciprocity and dealing cards.

6 Integer Points on Conics

The equations $x^2 + y^2 = p$ and $x^2 - 2y^2 = p$.

Unique factorization in $\mathbb{Z}[\sqrt{D}]$.

Integer solutions of $x^2 - Dy^2 = k$.