Chapter 3

Modular Arithmetic

In this chapter, we'll look at some applications of modular arithmetic, i.e., applications of the rings $\mathbb{Z}/n\mathbb{Z}$ to number theory. In particular, we'll get applications to divisibility tests, necessary conditions for solutions of various Diophantine equations (including non-solvability results), as well as an application to modern cryptography. For some of these applications, we will need a deeper understanding of the arithmetic structure of $\mathbb{Z}/n\mathbb{Z}$, such as knowing which elements of $\mathbb{Z}/n\mathbb{Z}$ have a multiplicative inverse and when $\mathbb{Z}/n\mathbb{Z}$ is a field. For this, will take another little detour into abstract algebra with the notion of groups. (Thus we will have hit the 3 main types of algebraic structures covered in an abstract algebra course: groups, rings, and fields—albeit mainly restricted to the commutative setting.)

3.1 Divisibility criteria

One of the most basic applications of modular arithmetic is to obtaining the classic divisibility tests based on the decimal (base 10) representation of n.

Proposition 3.1.1. *Let $n \in \mathbb{N}$. Then n is divisible by 2, 5 or 10 if and only if its last digit is. Similarly, n is divisible by 4, 25 or 100 if and only if the integer consisting of its last two digits is.*

If $n < 10$, we interpret the last two digits to just mean n (i.e., write n in decimal with a preceding 0).

Proof. Write $a_d a_{d-1} \cdots a_1 a_0$ as the base 10 representation of n, i.e., $0 \leq a_i \leq 9$ and

$$n = 10^d a_d + 10^{d-1} a_{d-1} + \cdots + 10^1 a_1 + 10^0 a_0.$$

If $m = 2$, 5 or 10 then $m|10$ so $n \equiv a_0 \bmod m$. Hence $m|n$ (i.e., $n \equiv 0 \bmod m$) if and only if $m|a_0$.

If $m = 4$, 25 or 100, then, then $10^j \equiv 0 \bmod m$ for $j \geq 2$, so $n \equiv 10a_1 + a_0 \bmod m$. So again, $m|n$ if and only if $m|(10a_1 + a_0)$. □

The above argument can be written easily enough without modular arithmetic, but the the standard divisibility tests for 3 and 9 are really much more transparent with modular arithmetic.

Proposition 3.1.2. *Let $n \in \mathbb{N}$. Then n is divisible by 3 or 9 if and only if the sum of its digits is.*

Proof. Let $a_d a_{d-1} \cdots a_1 a_0$ be the base 10 representation of n, i.e., $0 \le a_i \le 9$ and

$$n = 10^d a_d + 10^{d-1} a_{d-1} + \cdots + 10^1 a_1 + 10^0 a_0.$$

Let $m = 3$ or 9. Since $10 \equiv 1 \bmod m$, we have $10^j \equiv 1^i \equiv 1 \bmod m$ for any i. Hence

$$n \equiv a_d + a_{d-1} + \cdots + a_1 + a_0 \bmod m.$$

Again, this means $m|n$ if and only if $m| \sum a_i$. □

> **Exercise 3.1.1.** Let $n \in \mathbb{N}$. Show n is divisible by 11 if and only if the alternating sum of its digits is. (E.g., by the alternating sum of the digits of the number 12345, we mean $1 - 2 + 3 - 4 + 5$.)

We can use the same idea to give divisibility criteria in terms of representations of numbers in other bases. Here is a simple example which is similar to the last problem.

> **Exercise 3.1.2.** Consider the binary expansion of $n \in \mathbb{N}$, which consists of a string of *bits* ("binary digits"). Show that n is divisible by 3 if and only if the alternating sum of its bits is.

From above we have tests for divisibility of n in terms of its digits for dividing by any number up to 10, except for 7 and 8. We didn't state one explicitly for divisibility by 6, but clearly you can just test for divisibility by 2 and by 3 thanks to unique factorization, or more directly the prime divisor property. (Think about why the prime divisor property is relevant.) You can also use a simple test for 8, generalizing the ones for 2 and 4, which is a special case of the following:

> **Exercise 3.1.3.** Let $k, n \in \mathbb{N}$. Show n is divisible by 2^k if and only if the number consisting of just the last k digits of n is. Moreover, show that looking at the last $k-1$ digits does not suffice to determine divisibility by 2^k.

Probably you knew about most of these divisibility tests already (though maybe you didn't know how to prove some of them). On the other hand, you probably don't know a divisibility test for 7, and that's because such a test is more complicated, though you can still write one down:

> **Exercise 3.1.4.** Let $n \in \mathbb{N}$. Devise a test to determine if n is divisible by 7 or not, based on looking at certain combinations of digits.

3.2 Applications to Diophantine equations

Recall from the introduction that one of standard descriptions of number theory is the study of Diophantine equations. To be formal, here is a proper definition:

Definition 3.2.1. *A* **Diophantine equation** *is an equation of the form $f(x_1, \dots, x_n) = g(x_1, \dots, x_n)$, where $x_1, \dots, x_n$ are variables in $\mathbb{Z}$ and f, g are polynomials with coefficients in $\mathbb{Z}$.*

Note such an equation is equivalent to the equation $F(x_1, \dots, x_n) = 0$ where F is the polynomial $f - g$, so when we discuss Diophantine equations it suffices to assume the equation is in the form $F(x_1, \dots, x_n) = 0$.

Since we take $x_1, \dots, x_n$ to be variables in $\mathbb{Z}$, by a solution to a Diophantine equation $F(x_1, \dots, x_n) = 0$ we mean a solution with each $x_i \in \mathbb{Z}$, which we call a **solution over** $\mathbb{Z}$.[1] Thus solving Diophantine equations is equivalent to finding integer roots of polynomials with integer coefficients.

To remind you where we're going, the following families of Diophantine equations—all of which were discussed in the introduction—are the main Diophantine equations we are focused on in this course.

(1) $x^2 + y^2 = n$ (which numbers are sums of 2 squares?)

(2) $x^2 + y^2 + z^2 + w^2 = n$ (which numbers are sums of 4 squares?)

(3) $x^2 - dy^2 = 1$ (Pell's equation, related to finding rational approximations for $\sqrt{d}$)

(4) $x^3 + y^3 = z^3$ (we'll also say a bit more generally about $x^n + y^n = z^n$, the subject of Fermat's Last Theorem)

Here we regard n and d as constants in these equations. The goal is to determine when these equations have solutions and, if possible, describe all solutions or explain how to find all solutions. We already treated the simple case of linear Diophantine equations $ax + by = c$ in 2 variables in Proposition 2.3.1, where a, b, c are constants.

For instance, for the first family of equations above, $x^2 + y^2 = n$, we mainly want to do two things: (i) for n such that a solution exists prove one exists, and (ii) for n such that no solution exists prove there is no solution. In this case, for given n, it is not hard to determine solutions algorithmically—one can simply check values of $x^2 + y^2$ for $0 \le x, y \le \sqrt{n}$ similar to the proof of part (1) of Proposition 1.5.5. There are of course more efficient algorithms, but we will not focus on algorithmic aspects too much in this course. While there's no simple formula in general (in terms of n) for solutions to $x^2 + y^2 = n$, another thing one can do is count the number of solutions, which is a refinement of just determining whether solutions exist or not (i.e., determine when the count is positive versus zero). We won't focus too much on actually counting the number of solutions in this course, but we'll say a little about

[1] Technically, the phrasing "a solution *in* $\mathbb{Z}$" would mean that the solution to the equation is a single integer in $\mathbb{Z}$, rather than a tuple of integers, so I will try to say a solution *over* $\mathbb{Z}$ when there is more than one variable, but forgive me if I make a *faux pas*. On the other hand, I may say "integer solution" or "integral solution" for a solution over $\mathbb{Z}$ which is not a single integer in $\mathbb{Z}$ but a tuple in $\mathbb{Z}^n$. (This can be grammatically justified by calling $\mathbb{Z}^n$ the set of integer or integral points in $\mathbb{R}^n$ or $\mathbb{C}^n$.)

this also. (At a crude level, we've already noted that the number of solutions to (1), and similarly (2), must be finite, whereas the number of solutions to (3) can be infinite.)

The easiest way to show that a Diophantine equation has a solution is exhibit a solution. Recall, for $ax + by = c$, we didn't give a formula for solutions x, y but rather an algorithm for finding solutions x, y when they exist, which the most practical thing one can hope for as there are typically no simple formulas for solutions to Diophantine equations. For the above equations, one needs to work harder to show solutions exist.

On the other hand, much of the time there is an easy way to show solutions don't exist. That comes via modular arithmetic.

Proposition 3.2.2. *Let $F(x_1, \ldots, x_n) = 0$ be a Diophantine equation. If this equation has a solution, then*

$$F(x_1, \ldots, x_n) \equiv 0 \bmod m, \tag{3.2.1}$$

has a solution for all $m \in \mathbb{N}$.

The equation (3.2.1) is called the **reduction mod** m of $F(x_1, \ldots, x_n) = 0$, and we may view it as an equation in n variables in $\mathbb{Z}/m\mathbb{Z}$.

Proof. Suppose $x_1, \ldots, x_n \in \mathbb{Z}$ such that $F(x_1, \ldots, x_n) = 0$. Then $m | F(x_1, \ldots, x_n)$ for all $m \in \mathbb{N}$ (in fact for all $m \in \mathbb{Z}$ if one wants). □

The point is that it is often easy to show an equation mod m doesn't have any solutions. Algorithmically, certainly it's very simple: there are only m possibilities for $x_1, \ldots, x_n$ regarded as elements of $\mathbb{Z}/m\mathbb{Z}$, so at most we need to compute $F(x_1, \ldots, x_n)$ mod m for a total of m^n possible inputs.

Remark 3.2.3. It is *not* true that the converse of the proposition holds. Namely, there are Diophantine equations which have solutions mod m for all m, but do not have solutions over $\mathbb{Z}$. A couple of famous examples are $x^2 + y^2 + z^2 + w^2 = -1$ and $3x^3 + 4y^3 + 5z^3 = 0$. The problem in some sense is that while these have solutions mod m for all m, you can't choose the solutions in a compatible way to "lift" them to solutions over $\mathbb{Z}$. One of the major themes in modern number theory is to study to what extent you can lift solutions mod m to solutions over $\mathbb{Z}$. To read more about this, look up the *local-global principle.* One particularly fascinating situation is the family of equations of the form $x^2 + dy^2 = n$ (here $d > 0$). It turns out that the problem of lifting solutions mod m to solutions over $\mathbb{Z}$ is related to the failure of unique factorization in $\mathbb{Z}[\sqrt{-d}]$. In particular, if one has unique factorization in $\mathbb{Z}[\sqrt{-d}]$ (or if unique factorization doesn't fail "too badly") then $x^2 + dy^2 = n$ has a solution over $\mathbb{Z}$ if and only if it does mod m for all m and $n \geq 0$. On the other hand, this is not true for $d = 23$, where unique factorization fails "sufficiently badly." In particular, $x^2 + 23y^2 = 41$ has a solution mod m for all m but no integer solution.

> **Example 3.2.1.** Let $n \in \mathbb{Z}$ and $f(x) = x^2 + x$. If n is odd, then $f(x) = n$ has no solution. To see this, look at the equation mod 2, which is simply $x^2 + x \equiv n$ mod 2 Now either $x \equiv 0$ mod 2 or $x \equiv 1$ mod 2. In either case, we see $x^2 + x \equiv x(x + 1) \equiv 0$ mod 2, whence n must be even to get a solution.
>
> Of course we could just make this argument in terms of even and odd numbers, but the benefit of this language of modular arithmetic is that it greatly generalizes what you can

> easily do just by thinking in terms of evens and odds. For instance, consider x^2+x mod 3. This is 0 when $x \equiv 0, 2$ mod 3 and 2 when $x \equiv 1$ mod 3, so $x^2+x \equiv 1$ mod 3 has no solutions. Thus we can conclude that any integer n of the form x^2+x ($x \in \mathbb{Z}$) must be even and not $\equiv 1$ mod 3, i.e., $6|n$ or $6|(n-2)$.

> **Exercise 3.2.1.** Determine the possibilities for x^2+x mod 5 and x^2+x mod 7. Using this, and the previous example, completely determine which $0 \le n \le 20$ are of the form x^2+x.

Since our next example is important in determining which numbers are sums of two squares, one of the main goals of the course, we elevate its status to a proposition.

Proposition 3.2.4. *Let $n \in \mathbb{N}$. If $n \equiv 3$ mod 4, then n is not a sum of 2 (integer) squares, i.e., $x^2+y^2=n$ has no solution over $\mathbb{Z}$.*

Note this criterion provides a great speed-up to the algorithmic approach to looking for solutions to $x^2+y^2=n$. We can just first check n mod 4 (for which it suffices to check the last 2 digits), and if we get 3 mod 4 stop. Of course if n is not 3 mod 4, we still need to look for solutions.

The proof requires the notion of squares mod n. We also say an integer $a \in \mathbb{Z}$ is a **square mod** n if $a+n\mathbb{Z}$ is a square in $\mathbb{Z}/n\mathbb{Z}$, i.e., $a \equiv x^2$ mod n for some $x \in \mathbb{Z}$. Otherwise, we say a is a **nonsquare mod** n. Since being a square (or nonsquare) mod n only depends upon the equivalence class mod n, we will sometimes think of the squares (or nonsquares) mod n as elements of $\mathbb{Z}/n\mathbb{Z}$.

> **Example 3.2.2.** Let $n \ge 2$. Then $0^2 \equiv 0$ mod n and $1^2 \equiv 1$ mod n, so there are always at least 2 squares mod n (thought of as elements of $\mathbb{Z}/n\mathbb{Z}$). On the other hand there are at most n, as there are n elements of $\mathbb{Z}/n\mathbb{Z}$. In particular, all numbers are squares mod 2.

> **Example 3.2.3.** Note that $0^2 \equiv 2^2 \equiv 0$ mod 4 and $1^2 \equiv 3^2 \equiv 1$ mod 4. Put another way, the square of an even number is 0 mod 4 and the square of an odd number is 1 mod 4. Hence the squares mod 4 are simply 0 and 1 (thought of as elements of $\mathbb{Z}/4\mathbb{Z}$), and 2 and 3 (as elements of $\mathbb{Z}/4\mathbb{Z}$, i.e., technically $2+4\mathbb{Z}$ and $3+4\mathbb{Z}$) are nonsquares mod 4.

> **Example 3.2.4.** Note $2^2 \equiv (-1)^2 \equiv 1$ mod 3, so the elements 0 and 1 of $\mathbb{Z}/3\mathbb{Z}$ are squares and $-1 \equiv 2$ mod 3 is a nonsquare.

Proof of proposition. Since the squares mod 4 are 0 and 1, we have one of the following possibilities for $x, y \in \mathbb{Z}$:

$$x^2+y^2 \equiv \begin{cases} 0+0 \equiv 0 \bmod 4 \\ 1+0 \equiv 1 \bmod 4 \\ 0+1 \equiv 1 \bmod 4 \\ 1+1 \equiv 2 \bmod 4. \end{cases}$$

Thus the sum of 2 squares is never 3 mod 4. □

We remark one can also formulate the proposition as a divisibility statement: the: for any x, y, $x^2 + y^2 - i$ is divisible by 4 for some $i = 0, 1, 2$, i.e., $f(x, y) = (x^2 + y^2)(x^2 + y^2 - 1)(x^2 + y^2 - 2)$ is always divisible by 4. Here are some similar, rather well known, examples.

Exercise 3.2.2. Show $x^2 + 2y^2 = n$ has no solution over $\mathbb{Z}$ if $n \equiv 5, 7 \bmod 8$.

Exercise 3.2.3. Show $x^2 + 3y^2 = n$ has no solution over $\mathbb{Z}$ if $n \equiv 2 \bmod 3$.

Exercise 3.2.4. Show that $n \in \mathbb{N}$ is not a sum of 3 (integer) squares if $n \equiv 7 \bmod 8$.

Exercise 3.2.5. Show that $n \in \mathbb{N}$ is not a sum of two (integer) cubes if $n \equiv 4, 5 \bmod 9$.

More generally than just getting non-existence of solutions to certain Diophantine equations, we can also obtain necessary conditions for solutions to Diophantine equations. This is useful for (i) helping to find solutions when they exist, and (ii) as an intermediary step for proving the non-existence of solutions when they don't exist. Here's a simple example of this technique.

Proposition 3.2.5. *Suppose $x, y, z, w \in \mathbb{Z}$ such that $x^2 + y^2 + z^2 = w^2$. If w is odd, then exactly one of x, y, z is odd. If w is even, then all of x, y, z are even.*

Proof. Recall that the squares mod 4 are 0 and 1. Note that if w is odd, then an odd number of x, y, and z are odd, i.e., an odd number of x^2, y^2 and z^2 are 1 mod 4. If all three are, then $x^2 + y^2 + z^2 \equiv 3 \bmod 4$, but $w^2 \equiv 1 \bmod 4$. Hence exactly one of x, y and z is odd.

The argument for w even is similar, and we leave it to the reader. □

Note even though the original statement is only about the parity of solutions, looking at things mod 2 is not sufficient to prove this statement, as all numbers are squares mod 2. For instance, when w is odd, then looking at parities only tells you that an odd number of x, y and z must be odd.

Example 3.2.5. Now let's use the above proposition to determine all solutions to $x^2 + y^2 + z^2 = 9$ with $x, y, z \in \mathbb{N}$. We know exactly one of x, y and z must be odd. So two of them must be at least 2, which forces the other to be 1. Consequently all solutions over $\mathbb{N}$ are $(2, 2, 1)$, $(2, 1, 2)$ and $(1, 2, 2)$.

Exercise 3.2.6. Determine if $x^2 + y^2 + z^2 = 25$ has any solutions with $x, y, z \in \mathbb{N}$. If so, find all solutions.

Exercise 3.2.7. Determine if $x^2 + y^2 + z^2 = 64$ has any solutions with $x, y, z \in \mathbb{N}$. If so, find all solutions.

In Chapter 6, we'll see how we can use this technique to make a little progress on Fermat's last theorem, however the only known ways to prove Fermat's last theorem use much more advanced machinery than just simple considerations mod m.

3.3 Groups and invertibility mod n

To go a bit further with applications of modular arithmetic, we need to understand some things about the multiplicative structure of $\mathbb{Z}/n\mathbb{Z}$. In this section, except where noted otherwise, we assume $n > 1$.

Definition 3.3.1. *We say $a \in \mathbb{Z}$ is* **invertible** *mod n if $a + n\mathbb{Z}$ has a multiplicative inverse in $\mathbb{Z}/n\mathbb{Z}$, i.e., if there exists $b \in \mathbb{Z}$ such that $ab \equiv 1 \bmod n$. In this case b is called a* **(multiplicative) inverse** *of a mod n.*

Note that this only depends on the congruence class, i.e., if $a \equiv a' \bmod n$, then a is invertible mod n if and only if a' is, and the inverse only depends on the congruence class as well. As with the notion of squares mod n, we sometimes think of inverses mod n as integers, and sometimes as elements of $\mathbb{Z}/n\mathbb{Z}$, depending on which is more convenient.

The notion of invertibility can also be phrased in terms of Diophantine equations mod n: a is invertible mod n if and only if $ax \equiv 1 \bmod n$ has a solution in $\mathbb{Z}/n\mathbb{Z}$.

The invertible elements of $\mathbb{Z}/n\mathbb{Z}$ (or more generally a ring) will give us an algebraic structure known as a group.

Definition 3.3.2. *Let G be a set with a binary operation $\cdot$. We say $(G, \cdot)$ (or just G if the operation is understood) is a* **group**, *if the following three properties hold:*

(1) $\cdot$ is associative: $(g \cdot h) \cdot k = g \cdot (h \cdot k)$ for all $g, h, k \in G$;

(2) there is an identity $1 \in G$ such that $1 \cdot g = g \cdot 1 = g$ for all $g \in G$;

(3) every $g \in G$ has an **inverse** *g^{-1} such that $g^{-1} \cdot g = g \cdot g^{-1} = 1$;*

If G is a group which also satisfies

(4) $\cdot$ is commutative: $g \cdot h = h \cdot g$ for all $g, h \in G$,

then we say $(G, \cdot)$ (or just G) is an **abelian group**.

When the operation is understood, we typically write gh for $g \cdot h$, and this notation is called **multiplicative notation**. However, for some abelian groups, the operation $\cdot$ will be written as $+$, which is called **additive notation**. In the case of additive notation, we denote the identity by 0 instead of 1, and the inverse of g by $-g$ instead of g^{-1}. Accordingly, an **additive group** will mean an abelian group in additive notation, and a **multiplicative group**.

The reason for these conventions should be clear from following simple examples (the proofs are easy, and you may fill them in for yourself).

> **Example 3.3.1.** $(\mathbb{Z}, +)$ is an additive (abelian) group. So is $(\mathbb{Q}, +)$ and $(\mathbb{R}, +)$, or more generally $(R, +)$ where R is any ring. In all cases, the identity of the group is the zero element 0 of the ring, and the inverse of any a in the ring is $-a$. (Our notation for 0, $+$ and $-$ in a ring R is consistent with the additive notation for the group $(R, +)$.) On the other hand, $(\mathbb{N}, +)$ is not a group as it does not have the identity or (additive) inverses.

> **Example 3.3.2.** $(\mathbb{Q}^\times, \times) = \mathbb{Q} - \{0\}$ is an infinite abelian multiplicative group. So is $\mathbb{R}^\times$ and $\mathbb{C}^\times$. We will generalize these examples (with proof) to an arbitrary ring below. Similarly, the positive rational (or reals) also are. In all cases, the identity is the integer 1 and the inverse of any element x is $x^{-1} = \frac{1}{x}$.

Lemma 3.3.3. *Let G be a group. Then there is a unique identity, and each $g \in G$ has a unique inverse.*

Proof. You already proved that any binary operation has at most 1 identity (Exercise 1.2.4), so the identity of G is unique. Now let $g \in G$ and suppose $h, h' \in G$ such that h and h' are inverses of g. Then on one hand $hgh' = (hg)h' = 1 \cdot h' = h'$, but also $hgh' = h(gh') = h \cdot 1 = h$, whence $h = h'$. □

We say a group $(G, \cdot)$ is **finite** if the set G is finite. The finite abelian groups are in some sense the simplest class of groups and have a simple characterization. If G is a finite group with n elements, we say it has **order** n, and write $|G| = n$.

Here are some more examples, mostly without proofs.

> **Example 3.3.3.** $(\mathbb{Z}/n\mathbb{Z}, +)$ is an finite abelian group of order n.

> **Example 3.3.4.** (n-th roots of unity) Recall the n-th roots of unity $\mu_n = \left\{e^{2\pi i k/n} : 0 \le k < n\right\}$. Then, with the standard multiplication, μ_n is a finite abelian group of order n (see exercise below).

We remark that the group μ_n has the same structure (is "isomorphic" to) $(\mathbb{Z}/n\mathbb{Z}, +)$, the only difference being one group is written with multiplicative notation and one with additive notation. (Recall the pictures of $\mathbb{Z}/n\mathbb{Z}$ and μ_n as n points around a circle.) Precisely, if we write down the operation table for $(\mathbb{Z}/n\mathbb{Z}, +)$, with elements represented as $0, 1, \dots, n-1$ in the obvious way, and change each element label i to ζ_n^i and relabel our operation $+$ for $\mathbb{Z}/n\mathbb{Z}$ to $\cdot$, we get exactly the multiplication table for μ_n.[2]

[2] Determining when two groups have the same structure is one of the basic problems in group theory. We remark it is a hard (as in research level) problem do determine exactly the number of different kinds of (the number of "isomorphism classes") of groups of a fixed order n. No exact formula is known (except for n of special type) and the number of distinct groups (up to isomorphism) of order n grows very quickly as the number of factors of n grows. (There is only type of group of order n when $n = p$ is prime, which is the isomorphism class of $\mathbb{Z}/p\mathbb{Z}$.)

Exercise 3.3.1. Prove μ_n is a group under multiplication. Write down the multiplication table when $n = 3$ and check it looks the same as the addition table for $\mathbb{Z}/3\mathbb{Z}$.

Example 3.3.5. (dihedral groups) Fix $n > 2$. Let P be a regular polygon with n vertices. The the set of automorphisms of P, namely the rotations and reflections which map P to itself, form a finite *non-abelian* group of order $2n$ called the **dihedral group** D_{2n}, where the operation is composition.

Example 3.3.6. (symmetric groups) Let S_n be the set of permutations of $\{1, 2, \dots, n\}$. Then S_n is a finite group of order $n!$ with the composition operation, called the **symmetric group** on n elements. It is non-abelian for $n > 2$.

Note that in the groups in the previous two examples can be naturally thought of as the symmetries of some object—D_{2n} is the set of geometric symmetries of a regular n-gon in a plane, and S_n is the set of "combinatorial" symmetries of a set of size n (though one can also realize S_n geometrically, e.g., as the symmetries of the standard basis of $\mathbb{R}^n$). The standard way of thinking about what the notion of a group represents is the notion of the symmetries of some object: given two symmetries g and h one can compose them to get a new symmetry $g \cdot h$; this composition is associative, the "do nothing" symmetry is the identity, and each symmetry can be applied in reverse giving an inverse. (Historically, group theory was developed to study permutations of roots of polynomials by Galois and others. The term "abelian" is in honor of Niels Abel, who proved that the "Galois group" of a polynomial being commutative means the roots of that polynomial can be found with radicals).

Example 3.3.7. Let $\mathrm{GL}_n(\mathbb{R})$ denote the set of $n \times n$ invertible matrices with real entries. From linear algebra, being invertible simply means the determinant is nonzero. Then $\mathrm{GL}_n(\mathbb{R})$ forms a group with respect to matrix multiplication. (In linear algebra, probably you essentially proved this was a group without using the word group.)

Example 3.3.8. Let $\mathrm{SL}(2, \mathbb{Z}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, b, c, d \in \mathbb{Z},\ ad - bc = 1 \right\}$. This is an infinite non-abelian group with usual matrix multiplication, and is an important group in number theory. To prove it is a group, the main point is to show that the matrix inverse of an element in $\mathrm{SL}(2, \mathbb{Z})$ is again in $\mathrm{SL}(2, \mathbb{Z})$. (Here it does not suffice to look at matrices with integer entries whose determinant is nonzero—you need that the determinant is a unit in $\mathbb{Z}$—e.g., $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ has integer entries and nonzero determinant, but its inverse has fractional entries.)

Okay, so those were some examples. Basically, a group (in multiplicative notation) is a collection of objects that you can multiply and divide, and has "1." Recall we are interested in the invertible elements mod n, or equivalently, the invertible elements of $\mathbb{Z}/n\mathbb{Z}$:

$$(\mathbb{Z}/n\mathbb{Z})^\times = \{n\mathbb{Z} + a \in \mathbb{Z}/n\mathbb{Z} : a \text{ invertible mod } n\}.$$

More generally, for a ring R, we denote the set of **invertible** elements of R by $R^\times$, i.e.,

$$R^\times = \{a \in R : ab = 1 \text{ for some } b \in R\}.$$

Proposition 3.3.4. *Let R be a (commutative) ring. Then $R^\times$ is an abelian group. In particular, $(\mathbb{Z}/n\mathbb{Z})^\times$ is an abelian group for any n.*

Recall that for $a \in R$, a^{-1} denotes an inverse when it exists. Furthermore, by the same argument as for Exercise 1.2.4, an inverse is unique when it exists.

This result generalizes the earlier examples of $\mathbb{Q}^\times$, $\mathbb{R}^\times$ and $\mathbb{C}^\times$. Similarly, there is an analogue for non-commutative rings which generalizes the example of $\mathrm{GL}_n(\mathbb{R}) = M_n(\mathbb{R})^\times$.

Proof. Consider $a, b \in R^\times$, which have inverses a^{-1}, b^{-1}. Then $(ab)(b^{-1}a^{-1}) = a(bb^{-1})a^{-1} = aa^{-1} = 1$, so ab is also invertible, and thus $ab \in R^\times$. This means multiplication defines a binary operation on $R^\times$. Further, it is associative since multiplication on R is.

First note that $1 \in R^\times$, so $R^\times$ is non-empty and has a multiplicative identity. Next, if $a \in R^\times$, then there exists $a^{-1} \in R$ such that $aa^{-1} = a^{-1}a = 1$, so also $a^{-1} \in R^\times$, and thus $R^\times$ (essentially by definition) contains inverses. □

Proposition 3.3.5. *We have*

$$(\mathbb{Z}/n\mathbb{Z})^\times = \{a + n\mathbb{Z} \in \mathbb{Z}/n\mathbb{Z} : \gcd(a, n) = 1\}.$$

Hence $|(\mathbb{Z}/n\mathbb{Z})|^\times$ is the number of integers $1 \le a < n$ with $\gcd(a, n) = 1$.

Proof. Let $a \in \mathbb{Z}$. Note a is invertible mod n if and only if

$$ax + ny = 1 \tag{3.3.1}$$

has a solution for some $x, y \in \mathbb{Z}$. By the Euclidean algorithm (see Proposition 2.3.1), this happens if and only if $\gcd(a, n)$. This proves the first statement, and the second statement follows immediately. □

Definition 3.3.6. *The function $\phi : \mathbb{N} \to \mathbb{N}$ given by $\phi(n) = |(\mathbb{Z}/n\mathbb{Z})^\times|$ (where we interpret $\phi(1) = 1$) is called the* **Euler phi** *or* **Euler totient** *function.*

> **Example 3.3.9.** When $n = 2$, we have $(\mathbb{Z}/2\mathbb{Z})^\times$ consists of 1 element, $1 + 2\mathbb{Z}$. It is its own inverse. Thus $\phi(2) = 1$.

> **Example 3.3.10.** When $n = 3$, we have $(\mathbb{Z}/3\mathbb{Z})^\times$ consists of 2 elements, $1+3\mathbb{Z}$ and $2+3\mathbb{Z}$. Since $1 \cdot 1 \equiv 1 \bmod 3$ and $2 \cdot 2 \equiv 1 \bmod 3$, we see they are each their own inverse. Thus $\phi(3) = 2$.

Recall that, to avoid the cumbersome notation $a + n\mathbb{Z}$, we often denote the elements of $\mathbb{Z}/n\mathbb{Z}$ using a set of representatives $\{0, 1, 2 \ldots, n-1\}$ from $\mathbb{Z}$, e.g., we will often write 2 instead of $2 + n\mathbb{Z}$. We hope this will not cause any confusion.

Example 3.3.11. For $n = 4$, a set of representatives for $(\mathbb{Z}/4\mathbb{Z})^\times$ is $\{1, 3\}$. Again, each element is its own inverse, and we see $\phi(4) = 2$.

Example 3.3.12. For $n = 5$, a set of representatives for $(\mathbb{Z}/5\mathbb{Z})^\times$ is $\{1, 2, 3, 4\}$, so $\phi(5) = 4$. We see $2 \cdot 3 \equiv 1 \bmod 5$ and $4^2 \equiv (-1)^2 \equiv 1 \bmod 5$, so 1 and 4 are their own inverses, while 2 and 3 are inverses of each other.

Exercise 3.3.2. For $6 \leq n \leq 10$, write down a set of representatives for $(\mathbb{Z}/n\mathbb{Z})^\times$, determine the inverse of each representative, and compute $\phi(n)$.

If (x, y) is a solution to (3.3.1), then x is an inverse to a mod n. Hence we can compute inverses of a mod n using the extended Euclidean algorithm/tableau method. This will be useful when n is very large, and is an important step in the RSA cryptosystem below.

Exercise 3.3.3. Use the extended Euclidean algorithm to find by hand an inverse to 37 mod 100. Check that your solution is indeed an inverse.

The above proposition readily gives:

Corollary 3.3.7. *For $n \geq 2$, we have $\phi(n) \leq n - 1$, with equality if and only if n is prime. Hence $\mathbb{Z}/n\mathbb{Z}$ is a field if and only if n is prime.*

Proof. Since there are only n elements of $\mathbb{Z}/n\mathbb{Z}$ and 0 is never invertible if $n \geq 2$, we immediately get $\phi(n) \leq n - 1$. If n is prime, then each $1 \leq a < n$ has $\gcd(a, n) = 1$, so $\phi(n) = n - 1$. If n is not prime, it has a nontrivial divisor $1 < m < n$. Then m is not invertible mod n by the above proposition, so $\phi(n) < n - 1$. This proves the first statement.

For the second, recall that $\mathbb{Z}/n\mathbb{Z}$ is a field if and only if each nonzero element is invertible, i.e., if and only if $\phi(n) = n - 1$. □

We've seen $\phi(n)$ is easy to compute when n is prime, and you might wonder about other values. Indeed, $\phi(n)$ is a basic function in number theory, and many elementary number theory courses derive a precise formula in terms of the prime factorization of n. We will just do a couple of special cases now, but see how to say something more general later.

Proposition 3.3.8. *For any prime p, $\phi(p^2) = p(p-1)$.*

Proof. We just need to count the numbers between 1 and $p^2 - 1$ which are relatively prime to p^2, i.e., relatively prime to p. Since p is prime, these are just the multiples of p up to $p^2 - 1$:

$$p, 2p, 3p, \ldots, (p-1)p,$$

of which there are $p - 1$. So we have $(p^2 - 1) - (p - 1) = p^2 - p = p(p-1)$ numbers up to $p^2 - 1$ which are relatively prime to p. □

Exercise 3.3.4. Determine $\phi(p^n)$. Test your formula on small powers of 2 and 3.

The following situation will come up in RSA below:

Exercise 3.3.5. Prove $\phi(pq) = (p-1)(q-1)$ when p and q are distinct primes.

Exercise 3.3.6. Determine $\phi(60)$.

Exercise 3.3.7. Write $n = p_1^{e_1} \cdots p_r^{e_r}$ (the prime-power factorization). Conjecture a formula for $\phi(n)$ in terms of p_i's and e_i's, and provide some evidence for your conjecture.

3.4 Cosets and Lagrange's theorem

Definition 3.4.1. *Let $(G, \cdot)$ be a group. Let H be a subset of G. If $(H, \cdot)$ is also a group then H is called a* **subgroup** *of G. The* **(left) cosets** *of H in G are the subsets of G of the form*

$$g \cdot H = \{g \cdot h : h \in H\} \; g \in G.$$

Just like the subring test from Lemma 1.2.6, we have the following subgroup test.

Lemma 3.4.2. *If G is a group and H is a nonempty subset of G, then H is a subgroup of G if and only if it is closed under multiplication ($h_1h_2 \in H$ for $h_1, h_2 \in H$) and inversion ($h^{-1} \in H$ for $h \in H$).*

Proof. ($\Leftarrow$) Suppose H is closed under multiplication and inversion. Being closed under multiplication implies that the multiplication on G restricts to a well defined binary operation on H. Associativity holds because it does in G. If H is closed under inversion, then pick any $h \in H$ so $h^{-1} \in H$. (Here is where we need H nonempty.) By closure under multiplication $hh^{-1} = 1 \in H$. This takes care of all 3 properties required to be a group.

($\Rightarrow$) If H is a group, it is closed under multiplication and inverses by definition. □

Example 3.4.1. The set $n\mathbb{Z} \subseteq \mathbb{Z}$ consisting of multiples of n is a subgroup of $\mathbb{Z}$. To check this, obvserve $0 \in n\mathbb{Z}$ (so $n\mathbb{Z}$ is nonempty), the sum of two multiples of n is a multiple of n, and for any $kn \in n\mathbb{Z}$, $-kn \in n\mathbb{Z}$. Then the cosets of $n\mathbb{Z}$ in $\mathbb{Z}$ are the subsets of G of the form $a + n\mathbb{Z}$ for $a \in \mathbb{Z}$. In other words, the cosets of $n\mathbb{Z}$ in $\mathbb{Z}$ are precisely the congruences classes mod n.

Just as congruences mod n partition $\mathbb{Z}$ into n different classes, we will see in the proof of Lagrange's theorem below that the cosets partition G into a certain number of different classes (which we now call cosets). (In fact, in general we can view cosets as equivalence classes with the equivalence relation—see Exercise 3.4.4—but we will not emphasize this point of view in this course.) Let us first look at a few more examples.

Example 3.4.2. Let G be any group, and 1 the identity. Then it is easy to see $H = \{1\} \subseteq G$ is a subgroup, called the **trivial (sub)group**. For any $g \in G$, $g \cdot H = \{g\}$. Hence there are $|G|$ cosets of H in G, each consisting of a single element. This corresponds to the unique partition of G into $|G|$ singleton sets.

Example 3.4.3. Clearly $H = G$ is a subgroup of G. Then for any $g \in G$, $gH = H = G$ (a proof is contained in the proof of Lagrange's theorem below), so there is only one coset, $H = G$ itself. This corresponds to the "trivial partition" of G into one set, G.

Example 3.4.4. $\mu_2 = \{\pm 1\}$ is a subgroup of $\mu_4 = \{\pm 1, \pm i\}$. Note $1 \cdot \mu_2 = -1 \cdot \mu_2 = \mu_2$, and $i \cdot \mu_2 = -i \cdot \mu_2 = \{\pm i\}$. So there are two cosets of μ_2 in μ_4, and they give the following partition of μ_4:

$$\mu_4 = \mu_2 \sqcup i\mu_2 = \{1, -1\} \sqcup \{i, -i\}.$$

Exercise 3.4.1. Show that the only subgroups of μ_4 are $\mu_1 = \{1\}$, μ_2 and μ_4.

Example 3.4.5. Both μ_2, μ_3 are subgroups of $G = \mu_6 = \left\{\zeta_6^i : 0 \le i \le 5\right\} = \left\{\pm 1, \pm\zeta_6, \pm\zeta_6^2\right\}$.

First consider $H = \mu_2 = \{\pm 1\}$. Then the cosets are

$$1 \cdot \mu_2 = -1 \cdot \mu_2 = \{\pm 1\}, \quad \zeta_6 \cdot \mu_2 = \zeta_6^4 \cdot \mu_2 = \{\pm\zeta_6\}, \quad \zeta_6^2 \cdot \mu_2 = -\zeta_6^2 \cdot \mu_2 = \left\{\pm\zeta_6^2\right\}.$$

For $H = \mu_3 = \left\{1, \zeta_3, \zeta_3^2\right\} = \left\{1, \zeta_6^2, \zeta_6^4\right\}$, the cosets are

$$1 \cdot \mu_3 = \zeta_6^2 \cdot \mu_3 = \zeta_6^4 \cdot \mu_3 = \left\{1, \zeta_6^2, \zeta_6^4\right\}, \quad \zeta_6 \cdot \mu_3 = \zeta_6^3 \cdot \mu_3 = \zeta_6^5 \cdot \mu_3 = \left\{\zeta_6, \zeta_6^3, \zeta_6^5\right\}.$$

Exercise 3.4.2. Show that the only subgroups of μ_6 are μ_1, μ_2, μ_3 and μ_6.

You might have noticed that in the examples above that all cosets of H in G have the same size, and if $\{g_1, \ldots, g_r\}$ is a coset, we can represent it as $g_i \cdot H$. All of this will fall out of the proof of our next result.

Proposition 3.4.3. (Lagrange's theorem) *Suppose H is a subgroup of a finite group G. Then there are $|G|/|H|$ distinct cosets of H in G, each of size $|H|$. In particular, $|H|$ divides $|G|$.*

Proof. First note that any the size of any coset gH is $|H|$: if $h, h' \in H$, then

$$gh = gh' \implies g^{-1}gh = g^{-1}gh' \implies h = h',$$

hence for a fixed g, all the products gh are distinct.

Now we claim that any two distinct g_1H and g_2H are disjoint. For if they intersect, then for some $h_1, h_2 \in H$, we have $g_1h_1 = g_2h_2$. We can write any $h \in H$ as $h_1h_1^{-1}h$, so

$$g_1h_1 = g_2h_2 \implies g_1h = g_1h_1h_1^{-1}h = g_2(h_2h_1^{-1}h) \in g_2H,$$

i.e., any element of g_1H must be inside g_2H also. But since the have the same size ($|H|$), we must have $g_1H = g_2H$, proving the claim.

Hence the cosets $\{gH\}$ partition G into disjoint subsets, all of size $|H|$. In particular there must be $|G|/|H|$ cosets, which proves Lagrange's theorem. □

Exercise 3.4.3. Let H be a subgroup of a finite group G, and $C = \{g_1, \ldots, g_r\}$ a coset of H in G. Prove that, for $g \in G$, $g \cdot H = C$ if and only if $g \in C$.

Exercise 3.4.4. Let H be a subgroup of a finite group G. Define $g_1 \equiv g_2 \bmod H$ if $g_2^{-1}g_1 \in H$.

(i) Show $g_1 \equiv g_2 \bmod H$ if and only if $g_1H = g_2H$.

(ii) Prove this defines an equivalence relation on G, and that the equivalence classes are simply the cosets of H in G.

Exercise 3.4.5. Let $G = (\mathbb{Z}/8\mathbb{Z})^\times$, which we represent as $\{1, 3, 5, 7\}$.

(i) Write down the multiplication table for G.

(ii) Let $H = \{1, 7\}$. Show H is a subgroup of G.

(iii) Determine the cosets of H in G.

Exercise 3.4.6. Let $G = (\mathbb{Z}/7\mathbb{Z})^\times$. We represent the elements of G by $1, 2, \ldots, 6$.

(i) Write down the multiplication table for G.

(ii) Let $H = \{1, 6\}$. Show H is a subgroup of G.

(iii) Determine the cosets of H in G.

(iv) Repeat (ii) and (iii) for the set $H = \{1, 2, 4\}$.

Exercise 3.4.7. Let $n > 2$. Recall D_{2n} is the symmetries of a regular n-gon P.

(i) Label the vertices of P by $1, 2, \ldots n$. Use this to realize D_{2n} as a subgroup of the symmetric group S_n.

(ii) Show $D_6 = S_3$.

(iii) Determine the cosets of D_8 in S_4.

Lemma 3.4.4. *Let G be a finite group and $a \in G$.*

(i) There is some $n \in \mathbb{N}$ such that $a^n = 1$.

(ii) Take the smallest such n, called the **order** *of a. Then $C = \{a, a^2, a^3, \cdots, a^n\}$ is a subgroup of G of order n.*

Proof. (i) Since G is finite, and $a^k \in G$ for all $k \in \mathbb{N}$ there must be some $j, k \in \mathbb{N}$ with $j \neq k$ such that $a^j = a^k$. Assume $j < k$ and let $n = k - j$. Then $a^j a^n = a^j a^{k-j} = a^k = a^j$. Multiplying by $(a^j)^{-1}$, we see $a^n = 1$.

(ii) Let n be the order of a, i.e., $n \in \mathbb{N}$ is the smallest such that $a^n = 1$. Then the argument in (i) shows we can't have $a^j = a^k$ for $1 \le j < k \le n$—otherwise $a^{k-j} = 1$ but $k - j < n$. Hence C has precisely n elements.

By the lemma above, to check it is a subgroup it suffices to check closure under multiplication and inverses. Take any a^j and a^k in C (with $1 \le j, k \le n$). If $j + k \le n$, $a^j a^k = a^{j+k} \in C$ trivially; if $j + k > n$, we see $a^j a^k = a^{j+k} = a^{j+k-n} a^n = a^{j+k-n} \in C$ since $1 \le j + k - n \le n$. Hence C is closed under multiplication.

Note since $a^n = 1$, $(a^n)^{-1} = 1 = a^n \in C$. For any $1 \le j < n$, we have $1 \le n - j < n$. Then since $a^j a^{n-j} = a^n$ □

The group C in this lemma is called the **cyclic subgroup generated by** a because it consists only of elements that are powers of a single element a. (It is called cyclic because these powers cyclically repeat: $a^n = 1, a^{n+1} = a^n a = a, a^{n+2} = a^n a^2 = a^2, \ldots$.)

> **Exercise 3.4.8.** Check that the powers of a cyclically repeat in this example.
>
> (i) In $(\mathbb{Z}/7\mathbb{Z})^\times$), compute 3^k for $1 \le k \le 10$.
>
> (ii) What is the cyclic subgroup of $(\mathbb{Z}/7\mathbb{Z})^\times$ generated by 3? What about generated by 2?

Proposition 3.4.5. *Let G be a finite group of order n. Then, for any $a \in G$, $a^n = 1$.*

Proof. Say m is the order of a in G. Then a generates a cyclic subgroup H of G of order m, by the previous lemma. Now by Lagrange's theorem, $m|n$, say $n = km$. Then

$$a^n = a^{km} = (a^m)^k = 1^k = 1.$$

□

The proof is essentially summarized in the following phrase: the order of any element of G divides the order of G.

Corollary 3.4.6. (Fermat's little theorem) *If p is prime and $a \not\equiv 0 \bmod p$, then $a^{p-1} \equiv 1 \bmod p$.*

Proof. Apply the previous proposition to $G = (\mathbb{Z}/p\mathbb{Z})^\times$, which has order $p - 1$. □

Corollary 3.4.7. (Formula for inverses mod p) *Suppose* $\gcd(a, p) = 1$. *Then the inverse a^{-1} of a mod p is given by $a^{-1} \equiv a^{p-2} \bmod p$.*

Proof. Note $a^{-1} a \equiv a^{p-2} a \equiv a^{p-1} \equiv 1 \bmod p$. □

This gives a quick way to compute inverses mod p, using what is what is called the method of **repeated squaring**. We just illustrate this procedure with an example.

Example 3.4.6. Let's compute the inverse of 3 mod 19. By the above corollary, we have $3^{-1} \equiv 3^{17} \bmod 19$. We first use repeated squaring to compute the $3^{2^j} \bmod 19$ for $2^j \leq 17$:

$$\begin{aligned}
3^2 &\equiv 9 \bmod 19 \\
3^4 &\equiv (3^2)^2 \equiv 9^2 \equiv 81 \equiv 5 \bmod 19 \\
3^8 &\equiv (3^4)^2 \equiv 5^2 \equiv 25 \equiv 6 \bmod 19 \\
3^{16} &\equiv (3^8)^2 \equiv 6^2 \equiv 36 \equiv 17 \bmod 19.
\end{aligned}$$

Then we write 17 as a sum of powers of 2: $17 = 16 + 1$, and use this to compute

$$3^{-1} \equiv 3^{17} \equiv 3^{16+1} \equiv 3^{16} \cdot 3^1 \equiv 17 \cdot 3 \equiv (-2) \cdot 3 \equiv -6 \equiv 13 \bmod 19.$$

We can check this is indeed the inverse: $3 \cdot 13 \equiv 39 \equiv 1 \bmod 19$.

We also noted we can compute inverses mod p (in fact, mod n for any n) via the extended Euclidean algorithm in the last section. While that method is quite fast (and often faster than the above method), it is often useful in theory to have a *formula* rather than just an *algorithm*. On the other hand, it is sometimes more useful to have an algorithm than a formula, and we will see both Euler's theorem and the extended Euclidean algorithm being used in (different parts of) RSA in the next section.

Exercise 3.4.9. Use the formula $a^{-1} \equiv a^{p-2} \bmod p$ with repeated squaring to compute by hand the inverse of 5 mod 23. Check your answer is correct.

In fact, we will want to use the following generalization of Fermat's little theorem.

Corollary 3.4.8. (Euler's theorem) *For any invertible a mod n, we have $a^{\phi(n)} \equiv 1 \bmod n$.*

Proof. Apply the above proposition to $G = (\mathbb{Z}/n\mathbb{Z})^\times$, which, by definition of the totient function, has order $\phi(n)$. □

Exercise 3.4.10. Use Euler's theorem and repeated squaring to compute by hand $3^{-1} \bmod 14$.

As an addendum, we say a little more about cyclicity and orders. We say a finite group G is **cyclic** if there exists $g \in G$ such that the order of g is the order of G. Note for such a g, then the cyclic subgroup of G generated by g has order $|G|$, and so must be all of G. Note $(\mathbb{Z}/n\mathbb{Z}, +)$ and μ_n are cyclic groups of order n, and we can take for generators 1 and ζ_n, respectively. (By a **generator** of a cyclic group G, we mean any element of g which cyclically generates G, i.e., any element of order $|G|$.)

Exercise 3.4.11. For $2 \leq n \leq 10$, determine if $(\mathbb{Z}/n\mathbb{Z})^\times$ is cyclic or not. When it is cyclic, list all of the generators.

3.5 RSA

Beyond the uses of very elementary arithmetic, number theory has long been regarded largely as a purely theoretical study, with little practical applications, particularly when compared with fields like calculus and differential equations, which have many more obvious connections with real world applications. Things have greatly changed in our modern information age, and now aspects of number theory and algebra that were long considered purely theoretical have found many applications to computer science and information theory (so by extension are of interest to computer and electrical engineers as well). Two of the main sources of applications are cryptography and error-correcting codes, which can be bundled together under the name of coding theory.

Cryptography, the more famous of the two, is about keeping information secret from intruders, whereas error-correcting codes are about the opposite situation: how to send and receive information across a noisy channel (e.g., communicating with satellites). Both of these subjects are now a fundamental part of modern life, with most people not realizing what they are doing for us "under the hood." Essentially, any time you use a modern electronic device, you're relying on coding theory from things to making secure purchases online (or credit card purchases in store) and keeping other people from logging into your accounts (both cryptography), to having any sort of network reliability on a mobile device and not losing information on your hard drive anytime a butterfly flaps its wings (both error-correcting codes).

In this section, we'll just explain one beautiful and very practical application of number theory to cryptography: the RSA cryptosystem (the main idea, without all of the implementation details). To put this into context, let us first just very briefly discuss some general cryptography. Here is the basic problem in cryptography, which involves 3 characters:

- Alice, our protagonist. She want to send Bob a secret message.
- Bob, his name is Bob.
- Eve, the specific antagonist, and general ne'er-do-well. She eavesdrops on communications between Alice and Bob.

Problem: How can Alice send Bob a message in such a way that only Bob will be able to read it?

Private-key cryptography

The classical approach to this is using what is known as **private-key cryptography**. In this, Alice and Bob agree upon a secret code, or cipher, in advance. This involves 3 things: (i) an encryption algorithm, (ii) a decryption algorithm, and (iii) a secret key. One of the simplest and oldest ciphers is the **(Caesar) shift cipher**. Let's assume messages just consist of letters A, B, ..., Z. Our encryption will be to just to cyclically shift the letters by k "to the right". E.g., if $k = 1$, A will get encrypted as B, B as C, and so on, until Z which gets encrypted to A. If $k = 2$, A gets encrypted to C, B to D, and so on until Y to A, and Z to B. To decrypt, you simply cyclically shift the letters "to the left" by k. Both encryption and decryption, besides requiring that we know we are using a Caesar shift, require knowing the number $0 \leq k \leq 25$ (note: $k = 0$ is not so great!), which we call the key for this cryptosystem.

For instance, if Alice wants to send Bob the message EVESUCKS, they might agree on a shift cipher with $k = 3$ in advance[3], so Alice would encrypt letter by letter, send Bob the message

HYHVXFNV

which Bob easily decrypts, and if Eve sees the message along the way, it looks like nonsense to her. However, if she knows or guesses that they are using a shift cipher, and really wants to decode it, she can try all possible values for k, and notice that decrypting with $k = 3$ gives the only message that makes sense, and figure out the message. Of course, since the time of Caesar, ciphers have gotten incredibly more complex, and for good cryptosystems are nigh impossible to crack even if you know the algorithm (but not the key) being used.[4]

Public-key cryptography

The main problem with private-key cryptography, is that both Alice and Bob need to know the key without Eve knowing. This is fine if Alice and Bob can meet in advance in private to decide upon a key, but if they can't, or if they need to choose a new key, this is going to be quite difficult. In the 1970's, cryptographers were thinking about a way around this issue of making Alice and Bob agree on a key in advance, which led to what is now called **public-key cryptography**. The basic idea is the following: Bob makes generates a two keys: a **public key** e for encryption and a **private key** d for decryption. The public key e he announces to the world, and Alice can use e to encrypt her message, and send it to Bob. Then Bob, and only Bob, can decrypt the message using d, as only he knows d.

This idea requires two things. First, a cryptosystem where the encryption and decryption keys are different. E.g., if d and e are inverses in a ring R, encryption of a ring element $m \in R$ could be multiplication by e, giving the encrypted message $x = em$, and decryption could be multiplication by d: $dx = (de)m = m$. Second, since everyone knows the public key e, it should be hard to determine the decryption key d from just only e, but it should be easy for Bob to generate a pair of keys (e, d). Note that in our toy example of multiplication by e and d in R, at least in the case $R = \mathbb{Z}/n\mathbb{Z}$, it is easy to compute d from e via Euler's theorem or the extended Euclidean algorithm, so this would not make a good public-key cryptosystem. Note it's not at all obvious that a cryptosystem is possible, and for awhile cryptographers weren't sure if it was.

In 1978, Rivest, Shamir and Adleman published a paper with such a cryptosystem, now known as **RSA**, whose security is based upon the belief that factoring integers is hard. RSA is now widely used, and you are probably using RSA anytime you do something securely online. For instance, anytime you use an https website (e.g., any secure login webpage, credit card payment page, etc), both your web browser and the server are using RSA. When

[3] Julius Caesar reportedly used this shift with $k = 3$ to communicate with his generals.

[4] Typically, the weakest link in computer security is not the cryptosystem. Usually in hacking/data breach scandals, hackers are exploiting people not following good security protocols, rather than "cracking cryptosystems." E.g., people have easily guessable password, sensitive information is stored unencrypted, account number printouts are just found in a bank dumpster, a company allows someone to reset your password without really proving they are you, you download a virus that logs all your keystrokes, etc.

the server sends you information, they encrypt it with RSA using your browser's public key, and when you send information back you encrypt it with the server's public key.

Here is the basic algorithm, with explanations to follow:

(1) Bob chooses 2 large primes p and q, sets $n = pq$, and chooses some $1 < d, e < \phi(n)$ such that $de \equiv 1 \bmod \phi(n)$. Then Bob posts (e, n) as the public key, keeping, p, q and d secret, with d being the private key.

(2) Alice has a message m, which she represents as an integer $< n$. (If the message m is too long, she can break it up into pieces and encrypt each piece separately.) She encrypts it with Bob's public key as $x \equiv m^e \bmod n$, and sends the cipher text x to Bob.

(3) Bob decrypts the message x by computing $x^d \equiv m \bmod n$.

The first step, called key generation, only needs to be done once to initialize the process, and then Alice can send Bob as many encrypted messages as she wants with using Bob's fixed public key. The Prime Number Theorem, about distribution of primes, says that it's not too hard to find big primes. Basically, just choose a really big random number (say around 1,000 bits, or around 200 digits) and test nearby numbers to see if they are prime. What's important here are two things (i) there are primality tests which are fast (they don't rely on factoring integers)[5], and (ii) the Prime Number Theorem says you only need to try around $\log m$ numbers to find a prime near some big number m. Do this twice, once to find p, and once to find q. Then, calculation of $n = pq$ is not hard. Then Bob can just randomly choose $1 < d < \phi(n)$, and with probability near 1. Since we know $n = pq$, we know $\phi(n) = (p-1)(q-1)$ by Exercise 3.3.5, so we can quickly compute $\phi(n)$ also. Then we can quickly invert $d \bmod \phi(n)$ with the extended Euclidean algorithm to get $1 < e < \phi(n)$ such that $de \equiv 1 \bmod \phi(n)$.[6] These are all the calculations needed for Step 1.

> **Example 3.5.1.** To work with a small example, say Bob wants to take p, q around 3 digits. (I did the following calculations in the Sage mathematical software package.) We generate a couple numbers between 100 and 1000. I got 582 and 959. Starting with 582, I test successive numbers for primality, and I get $p = 587$ is prime, and similarly $q = 967$ is prime. So Bob computes $n = pq = 567629$ and $\phi(n) = (p-1)(q-1) = 566076$. Now we randomly take a number between 1 and 566076, say 154951. Using the Euclidean algorithm, we find 154951 is not invertible mod $\phi(n)$—their gcd is 23. Testing the next

[5] Here a fast probabilistic primality test to see if for some integer m is most likely prime: First, you can use divisibility tests to quickly check for divisibility of m by small primes. If m is divisible by some small prime, we know m is not prime. This rules out most numbers quickly: e.g., 1/2 of numbers are divisible by 2, $1/2 + 1/3 - 1/6 = 2/3$ of numbers are divisible by 2 or 3, and so on. If m is not divisible by a small prime, we can take a random number a less than m, and compute by repeated squaring $a^{m-1} \bmod m$—if m is prime, this is $\equiv a \bmod m$ by Fermat's little theorem. But if m is not prime, it turns out it's very unlikely that $a^{m-1} \equiv a \bmod m$ (though it happens occasionally). So if $a^{m-1} \equiv a \bmod m$, we conclude m is probably prime (and we can try this for a few values of a if we like). Otherwise, we showed m is not prime.

[6] Here it's *much* better to use the extended Euclidean algorithm as opposed to Euler's theorem to compute e—if we tried to use Euler's theorem, we'd need to compute $\phi(\phi(n))$, which essentially requires factoring $\phi(n)$, which may be infeasible as we're working with very big numbers in practice.

few numbers, we see $d = 154955$ is invertible mod $\phi(n)$, and its inverse is $e = 402575$. So Bob publishes e and n as his public key.

Exercise 3.5.1. Say Bob takes $p = 7$, $q = 11$ and $d = 17$. Determine Bob's public key.

Alice's part in this is easy. She has some message m, which she can realize as a number in some standard way. In practice this is a file, which is written in binary, that you can break up into bite-size blocks and encrypt separately. What's important is that $m < n$. Then, since she (and Eve and everyone else) knows n and e, she can compute $x \equiv m^e \bmod n$ (interpreted as a number between 0 and $n-1$) quickly using repeated squaring, as discussed in the previous section.

Example 3.5.2. Continue with the set up from the previous example.

Let's say Alice, again wants to send the message EVESUCKS to Bob. We can convert this to a number as follows: realize each letter as a number between 0 and 25 in the obvious way (A=0, B=1, ..., Z=25). (If you wanted to, you could include punctuation, and what not into your conversion scheme, say using the ASCII code which represents each character as a number between 0 and 255. Or just view a computer file, which is stored as a string of 1's and 0's, as a number in binary.) So we can think of EVESUCKS as representing a base 26 number of length 8. Since $26^3 < n$, we can break this up into blocks of length 3 as EVE, SUC, and KSZ. (Here I needed to pad the last block with some symbol such as Z. In practice, you can use a special character just for padding.) Let's just do the first block, EVE. Since E corresponds to 4 and V corresponds 21, EVE represents the base 26 form of the number $m = 4 * 26^2 + 21 * 26 + 4 * 1 = 3254$ in decimal. Then Alice computes the cipher text (encrypted message) using repeated squaring as

$$x = (m^e \bmod n) = 391820.$$

Exercise 3.5.2. Using Bob's public key from the previous exercise, encrypt the message $m = 15$.

Just like the previous step, Bob's decryption is easy. Since $de = k\phi(n) + 1$ for some k, we have

$$x^d \equiv (m^e)^d \equiv m^{k\phi(n)+1} \equiv (m^\phi(n))^k m \equiv m \bmod n,$$

using Euler's theorem at the last step. (The original paper of RSA gave a slightly different proof that $x^d \equiv m \bmod n$ using Fermat's little theorem.) So when Bob computes $x^d \bmod n$ (again by repeated squaring), he recovers m.

Example 3.5.3. Continue with the setup from the previous two examples.

Bob receives the message $x = 320576$ as the first encrypted block of the message. He computes

$$(x^d \bmod n) = 3254 = m.$$

Now, Bob can convert this back to the first block of the plain text (unencrypted) message

> EVE.

We remark that, in practice, there are various standard protocols to automate the way to messages (or files) are converted into blocks of numbers, but our goal is not to get into the technical aspects of implementation on a computer, just the main idea of RSA.

> **Exercise 3.5.3.** Perform Bob's decryption of the Alice's message from the previous exercise.

Now, why is this algorithm (believed to be) secure? Well, let's suppose Eve intercepts the cipher text x. To decrypt, she needs to know what to exponentiate it by (mod n) to get back to m. That is, she needs some d' such that $m^{d'e} \equiv m \bmod n$. Euler's theorem says this will be true if d' is an inverse of $e \bmod \phi(n)$. While there are a few choices of messages where other d''s can work (e.g., if $m = 1$, then $m^{d'e} \equiv m$ for any d'), for almost all messages m you really need d' to be an inverse of $e \bmod \phi(n)$. By the extended Euclidean algorithm, Eve can compute the inverse of $e \bmod \phi(n)$ to get Bob's decryption key d if she knows $\phi(n)$. But the point is that there are no known fast ways to compute $\phi(n) = (p-1)(q-1)$ without knowing p and q, and there are no known fast ways to factor n to get p and q.

We remark the actual implementation of RSA involves a little more to avoid encountering special situations where the message can be easy to decrypt (e.g., if $m = 1$, or d or e is too small). Also, since the encryption and decryption process in RSA is slower than many private-key methods such as AES (the current government standard), sometimes RSA is used to exchange a private-key when sending large amounts of encrypted data.

Moreover, RSA can be used for **message authentication**. What's to prevent Eve from intercepting x and sending Bob a different fake message m' (which she can also encrypt with Bob's public key)? Well, if Alice wants to authenticate her message, she can add a **digital signature**. Basically the idea is to run RSA in reverse. First, she generates her own public key (n_A, e_A) and private key d_A. She can encrypt her message m using her *private* key d_A as $s \equiv m^{d_A} \bmod n_A$. This is her signature, and she can send both m and s. Then anyone in the world can check that s decrypts to m using Alices public key: $s^{e_A} \equiv m^{d_A e_A} \equiv m \bmod n_A$. Since no one else could generate s from m without knowing d_A, this proves m is from Alice. So Alice could instead of just sending Bob m, she could send him the pair (m, s) to prove a message wants to come from her, and if she doesn't want Eve to be able to read the message, she can first encrypt both m and s using Bob's public key. (She needs to encrypt s also, otherwise Eve could decode s to get m from Alice's public key.)

If you're interested in learning cryptography, there are many good references out there. One possibility is William Stein's *Elementary Number Theory* book mentioned in the introduction. We also have a course here in the math department, *Applied Modern Algebra*, whose actual content varies according to the instructor, but it is usually largely cryptography. (The last time I taught it it was 75% cryptography and 25% error-correcting codes, but another faculty who teaches it often makes it 100% cryptography.)