MATH364: An Introduction to Number Theory – Congruences, Primes, Cryptography and Primitive Roots

ver. April 2021

A note on the text (April 2021)

In Winter 2021, MATH364 was offered remotely due to COVID-19. Coincidentally, this was the first semester in which *MATH163: Introduction to Mathematical Reasoning* was a pre-requisite for the course. This change in pre-requisite permitted more material to be covered at a greater depth than prior offerings of this course. However, this first offering was hamstrung by the realities of remote delivery.

For students in MATH364 in that semester, the first draft of these notes comprised of their weekly reading for the course. Each section was covered in a single week. Prior to each section of each chapter there was a 5-7 minute video introducing the broad ideas in the readings. Additionally, there were weekly drop-in office hours with the instructor to answer questions about course material. Each of office hours began with a mini-lecture on the previous week's topic. This approach proved successful and the hope is that students taking the course in the classroom can also find use for these notes. Doing so would require further editing to add material and remove commentary related to the specific offering for which these notes were authored.

These notes are written for students who have some familiarity with mathematical formalism from their work in MATH163. Each section of the notes corresponds to roughly one week of material. However, due to the nature of remote learning, some material that could be presented in MATH364 does not appear in these notes. These notes do not include material on solving simultaneous linear congruence. The effect of this missing material ripples into the work on quadratic residues; particularly in counting the number of quadratic residues. Material concerning the existence (and non-existence) of primitive roots modulo n is scant in these notes. As these notes contain only ten complete sections for a thirteen week semester, an classroom version of this course could easily include the missing material.

Each section of the notes begins with a list of learning incomes and outcomes. The former list tells the reader what material they need to be familiar with in order to understand the upcoming material. Readers should return to the latter list once they have finished their work in section to be sure they have attained the learning outcomes.

Throughout the text the reader will find short diversions under the heading **aside**. This fragments of text often present ideas beyond the stated learning outcomes and sometimes require students to have a more mature mathematical background than the stated learning incomes. These parts of the text can be fully ignored without detriment.

These notes are broadly based upon the approach and structure of *Elementary Number Theory* by Jones and Jones, with additional motivation from *Elementary Number Theory: Primes, Congruences and Secrets* by William Stein.

-cd

Contents

---

1 Divisibility Part I

Learning Incomes.

- *Reading mathematics at a level equivalent to student with credit for MATH163. (See Module 1 for a link to the textbook from MATH163)*
- *Reading prose at a level equivalent to at least a SK high-school graduate.*
- *Familiarity with conventions surrounding proof writing.*
- *Familiarity with proof by contradiction.*

Newly Defined Terms and Notation. *divides* ($|$)*, greatest common divisor* ($gcd(a, b)$)*.*

1.1 Divisibility and Greatest Common Divisor

The set $\mathbb{Z}$ of all integers, which this course is all about, consists of all positive and negative integers as well as 0. Thus $\mathbb{Z}$ is the set given by

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

While the set of all positive integers, denoted by $\mathbb{N}$, is defined by

$$\mathbb{N} = \{1, 2, 3, 4, ...\}.$$

(Some authors prefer to include 0 as element of $\mathbb{N}$. This is a perfectly reasonable choice. There is no universal convention. In this course we will not include 0 as an element of $\mathbb{N}$)

Whether or not we have ever noticed, the integers have structure – they are not just a set. We generally think of the integers as being in an order. The integer 8 comes after 7 and before 9. There is also lots of other structure on the integers that we are familiar with. For example, we can easily imagine partitioning the integers into odd integers and even integers. The existence of such structure cannot from the description of the integers as a set. Sets are unordered– we can write down the elements any way we want and still have the same set. So where does this structure come from?

One place that we can think about getting structure from is from addition. We can "add one" to get from one integer to the next. This allows us to put the integers in the order we are familiar with. (... this isn't exactly true, but it isn't worth thinking very hard about it at this point.) In this class most of the structure we are interested in will come from integer multiplication.

Recall the following definition from MATH163.

Definition 1.1. *Let A be a set. A binary operation on A is a function $f : A \times A \to A$*

We don't often think of integer multiplication as a function, but it is. Let $m : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ be given by:

$$m(a, b) = \sum_{i=1}^{a} b$$

For positive integers, a and b the value of $m(a, b)$ is exactly the same as $a \cdot b$. But surely we all agree that writing and manipulating $a \cdot b$ is much easier than working with $\sum_{i=1}^{a} b$.

This definition for positive integer multiplication can be extended to work for all integers. (We won't worry much about how to do this.) The point here is that sets and functions quietly underlie much of what we do in all branches of mathematics.

Our study of the structure of the integers given by multiplication begins with a fairly benign definition–

Definition 1.2. *Let a and b be integers. We say <u>a divides b</u> when there exists an integer k such that $ak = b$. When a divides b we write $a \mid b$. When a does not divide b we write $a \nmid b$. When a divides b we say that <u>b is a multiple of a</u> and <u>a is a divisor of b</u>.*

Example 1.3. *The integer 6 divides 18 as there exists an integer k so that $6k = 18$. However 6 does not divide 20 as for every $k \in \mathbb{Z}$ we have $6k \neq 20$.*

Beware – using "|" to mean "divides" can be quite convenient. But it can lead to confusion when we use | to mean "so that" as part of set-builder notation. For example, this is needless difficult to parse: $\{a \mid a \mid b\}$ (The set of all a such that a divides b). I will generally avoid using this notation, but recognize that it is standard in other places and some of you might already be familiar with it.

Aside 1.4. *A brief note about definitions – we will be defining a lot of terms in this class. This gives us a shared vocabulary and notation with which to express sometimes complicated ideas. Every definition will have underlined the newly defined term. From then on, the use of the term in the notes denotes exactly the meaning defined by the definition of that term.*

There isn't anything particularly interesting or deep about this definition. Our intuition around multiplication can probably tell us what "divides" means without needing a careful definition. However this definition will come to be very important as we prove various facts about integers. For example–

Theorem 1.5. *Let a, b, c be integers. If a divides b and b divides c, then a divides c.*

With a precise and agreed upon definition of the term *divides* our task in proving this theorem is clear: show we can always find $k \in \mathbb{Z}$ so that $ak = c$. The only tool at our disposal is the assumption that a divides b and b divides c. Translating our notation back to a sentence yields – there exists $k_1, k_2 \in \mathbb{Z}$ so that $ak_1 = b$ and $bk_2 = c$. Substituting yields $ak_1k_2 = c$. We can write this as $a(k_1k_2) = c$ to make it more clear that we have found $k \in \mathbb{Z}$ so that $ak = c$. (Let $k = k_1k_2$.)

Proof. Let a, b and c be integers so that a divides b and b divides c. Since a divides b and b divides c there exists $k_1, k_2 \in \mathbb{Z}$ so that $ak_1 = b$ and $bk_2 = c$. Therefore $a(k_1k_2) = c$ and so a divides c. □

Aside 1.6. *Recall the definition of a partial order as a relation that is reflexive, transitive and anti-symmetric. This is a structure that can be imposed on a set. Let R be the relation on $\mathbb{Z} \times \mathbb{Z}$ so that aRb when a divides b. Is R a partial order? What if we change $\mathbb{Z}$ to $\mathbb{N}$?*

Intuitively, when $ak = b$ it seems that a (and k) must be smaller than b. Though this is slightly complicated by allowing b to be negative (which our definition of divides certainly permits), we formalize this idea as follows.

Lemma 1.7. *For $a, b \in \mathbb{Z}$ with $b \neq 0$ so that a divides b, it follows that $|a| \leq |b|$.*

Proof. Consider $a, b \in \mathbb{Z}$ with $b \neq 0$ so that a divides b. Thus there exists $k \in \mathbb{Z}$ so that $ak = b$. If $|a| > |b|$, then $|b| = |ak| = |a||k| > |b||k| > |b|$. This is a contradiction as it cannot be that $|b| > |b|$ □

From this lemma the following theorem immediately follows.

Theorem 1.8. *For $a, b \in \mathbb{Z}$ with $b \neq 0$ so that a divides b, we have $-|b| \leq a \leq |b|$.*

Proof. This follows directly from Lemma 1.7. Since $|a| \leq |b|$ it cannot be that $a > |b|$ or $a < -|b|$. Thus $-|b| \leq a \leq |b|$ □

And from this theorem the following corollary immediately follows.

Corollary 1.9. *For any non-zero integer b, there are finitely many integers a such that a divides b.*

When $b \notin \{-1, 0, 1\}$ we can always find at least four integers that divide a – namely $1, -1, b$ and $-b$. And so in particular we note that for every positive integer b, there is always a positive integer a so that a divides b. (Again, this fact is not particularly deep, but it is worth noting in light of what is to come.)

Aside 1.10. *Why did we disallow $b = 0$ in the statement of Lemma 1.7? For which $a \in \mathbb{Z}$ does a divide 0? Is Corollary 1.9 true when permit $b = 0$?*

Given an integer b, we take for granted that we can find every integer a so that a divides b. These are the numbers that give remainder zero when divided by b (... unless $b = 0$). But how do we know that our intuition around division is actually backed by mathematical fact? We explore this in the following section.

Before we move on to our discussion around division and remainders, we have one last concept to introduce. Rather than come up with a not-very-convincing motivating reason for introducing this idea, I leave it to you own personal curiosity to find the following concept of interest. Be assured, however, that an application for this concept is forthcoming in the course.

Definition 1.11. *Let x and y each be non-zero integers. The greatest common divisor of x and y is the largest integer d such that d divides x and d divides y. We denote the greatest common divisor of x and y as $gcd(x, y)$.*

By our previous remarks, we are certain that x and y have at least one common divisor – namely 1. From Theorem 1.8 it follows that a common divisor can be no larger than the maximum of $|x|$ and $|y|$. And so since x and y have at least one common divisor and there is a bound on the magnitude of any common divisor, we can be sure the greatest common divisor of any two non-zero integers must exist. (We use the definite article "the" with greatest common divisor rather than the indefinite article "a". The greatest common divisor of a and b is unique!)

Aside 1.12. *Recall that an integer $p > 1$ is prime when its only positive divisors are one at itself. What can be said about $gcd(a, p)$ for any integer a?*

1.2 Division of Integers

The following fact seems obvious to us – we learned how to do division in grade school.

Theorem 1.13. *For every pair of integers, a, b, with $b \geq 1$, then there is a unique pair of integers, q and r such that*

$$a = qb + r$$

where $0 \leq r \leq b - 1$.

Here the notations q and r are chosen to remind us of the words quotient and remainder. The word unique is important in this theorem. For fixed integers a and b with $b \geq 1$, each time we divide a by b we can expect to get the same quotient and remainder. (That is, the quotient and remainder are unique.) Despite the fact that we all learned this mathematical fact many years ago, we have probably never taken the time to think about why this is true.

To start to think about justifying why this theorem is true, let us dive in to an example by considering a particular pair a and b. Let $a = 30$ and $b = 7$. There exist many choices of integers q and r so that $30 = 7q + r$. In fact, we can choose any value of q that we want and then find a corresponding r so that $30 = 7q + r$. For example, if we choose $q = 2$, then to have $30 = 7q + r$ we must have $r = 30 - 7(2) = 16$. More precisely, for any $n \in \mathbb{Z}$, there exists $r_n \in \mathbb{Z}$ so that $30 = 7n + r_n$. (Here we use r_n, rather than r to remind ourselves that this value depends on the particular choice of n We have switched from q to n because we want to reserve q to talk about the actual quotient in our division.)

Let S be the set of all possible values of r_n that can arise in this way. That is, let $S = \{30 - 7n \mid n \in \mathbb{Z}\}$. By inspection we see

$$S = \{, \ldots, -12, -5, 2, 9, 16, 23, 30, 37 \ldots\}$$

For example, $16 \in S$ as $16 = 30 - 7(2)$. Using our notation above, when $r_n = 16$ we have $n = 2$. Similarly, $-12 \in S$ as $-12 = 30 - 7(6)$. ($r_n = -12$ for $n = 6$). As we are taking all $n \in \mathbb{Z}$ we can permit n to take a negative value. This is why $37 \in S$ (What is n when $r_n = 37$?).

Thinking back to quotients and remainders, the pair $r_n = 2$ and $n = 4$ seems useful. Indeed when we divide 30 by 7, the quotient is 4 and the remainder is 2. We notice that 2 is the smallest non-negative value of S. If we swap 7 and 30 for other values of a and b, is this is the case. That is, we can find the remainder when a is divided by b by constructing S and finding the smallest non-negative value.

(Take a moment to convince yourself that this is reasonable by choosing a pair of integers a, b with $b > 0$. Find the smallest non-negative value in S and its corresponding n and compare those values with the remainder and quotient you get when dividing a by b.)

We now proceed with a proof of Theorem 1.13. We show that the smallest non-negative value of S and the corresponding value of n satisfy the conditions on r and q in the statement of Theorem 1.8. Once we show that S always has at least one non-negative value, then we can consider the smallest non-negative value in S and then conclude that r and q must exist.

With this intuition in place, we now prove Theorem 1.13.

(There is a video on Canvas that goes along with this proof. I recommend you understand as much of the proof as you can before you watch the video.)

Proof. Let a and b be integers with $b \geq 1$. Let $S = \{a - bn \mid n \in \mathbb{Z}\}$.

We first show that S contains at least one non-negative element. Once we have done this, we can then proceed to consider the smallest non-negative element.

Consider the value of r_n we arrive at when we choose $n = -|a|$. By substitution we have $r_n = a - (-|a|)b = a + |a|b$. Thus $a + |a|b \in S$. We now show $a + |a|b \geq 0$ for every $a, b \in \mathbb{Z}$ with $b \geq 1$.

We proceed based on whether or not a is negative.

If $a \geq 0$, then $a + |a|b = a + ab \geq 0$ as $b \geq 0$.

Otherwise if $a \leq -1$, we notice $|a| = -a$. Therefore $a + |a|b = a - ab = a(1 - b)$. Since b is a positive integer, necessarily $1 - b$ is not positive. Thus the product of a and $(1 - b)$ is not negative. That is, $a(1 - b) \geq 0$. Thus $a + |a|b \geq 0$.

Therefore $a+|a|b \geq 0$ for every $a, b \in \mathbb{Z}$ with $b \geq 1$. Thus S contains at least one non-negative element.

Let r be the smallest non-negative element of S. Since $r \in S$, there exists $q \in \mathbb{Z}$ so that $r = a - qb$. Rearranging we see $a = qb + r$.

To complete the proof we must show $0 \leq r \leq b - 1$ and that q and r are unique.

To show $0 \leq r \leq b-1$ we proceed by contradiction. That is, we assume $r \geq b$. Let $r' = r-b$. Since $r \geq b$, necessarily we have $r' \geq 0$. By construction, we have $r = a - qb$. Substituting this expression in to our expression for r' yields

$$r' = r - b = a - qb - b = a - (q + 1)b$$

Since $q + 1 \in \mathbb{N}$, it then follows that $r' \in S$. However this contradicts that r is the smallest non-negative element of S. Since we chose r to be the smallest non-negative element of S, then r' cannot exist. Thus it must be that $r \leq b - 1$. Therefore $0 \leq r \leq b - 1$.

We now show q and r are unique. Consider $q^\star$ and $r^\star$ so that $a = q^\star b + r^\star$ and $0 \leq r^\star \leq b-1$. To show q and r are unique we must show $q = q^\star$ and $r = r^\star$.

Since $a = q^\star b + r^\star$ and $a = qb + r$ it then follows that $q^\star b + r^\star = qb + r$ and so

$$r - r^\star = (q - q^\star)\, b.$$

We proceed by contradiction. If $q \neq q^\star$, then $|q - q^\star| \geq 1$. And so

$$|q - q^\star| b \geq 1 \cdot b = b.$$

Therefore

$$|r - r^\star| = |\,(q - q^\star)\, b| = |\,(q - q^\star)\,| \cdot |b| \geq 1 \cdot |b| = |b| = b$$

If $r \geq r^\star$, then $|r - r^\star| = r - r^\star$. Recall $0 \leq r, r^\star \leq b - 1$. Thus $r - r^\star \leq b - 1$

Therefore

$$b \leq |r - r^\star| = r - r^\star \leq b - 1.$$

This is a contradiction.

Similarly if $r < r^\star$ we have $|r - r^\star| = -(r - r^\star) \leq b - 1$. Again, a contradiction as $|r - r^\star| \geq b$.

Thus it cannot be that $|r - r^\star| \geq b$. That $|r - r^\star| \geq b$ followed from $q \neq q^\star$. Thus $q = q^\star$. This implies

$$r - r^\star = (q - q^\star)\, b = 0$$

Since $r - r^\star = 0$, it then follows that $r = r^\star$ Therefore q and r are unique, as required. □

Of course division isn't restricted to positive integers. Theorem 1.13 (and its proof) requires $b \geq 1$, but we still have intuition for division with negative integers.

Theorem 1.13 can be extended for when $b \leq 1$. In this case we require $0 \leq r \leq |b| - 1$ (We will see this later in the course.)

Exercises

Selected exercises will be discussed during the weekly synchronous tutorials.

1. Let a, b, c and d be integers. Show that if a divides b and c divides d, then ac divides bd.
2. Let a, b and c be integers so that c divides a and c divides b. Show that c divides $ax + by$ for any integers x and y.
3. Let a and b be integers so that a divides b and b divides a. Is it possible $a \neq b$?

4. Let x and y be non-zero integers. In your own words, explain how you know $gcd(x, y)$ exists? Explain how the statement of Theorem 1.8 imply that there is an upper bound on the size of common divisor of x and y.

5. Do an internet search for `Well-Ordering Principle`. What part of the proof of Theorem 1.13 depends on the Well-Ordered Principle?

Optional things to think about–

- Surprisingly, almost of our work in Module 1 follows from the existence of integer multiplication. Once we write down our definition of integer multiplication, everything else in the module is necessarily true. We were able to access these ideas just by thinking carefully about our definition of divides. (This astounds me! Multiplication seems so simple!) Furthermore, all of these mathematical facts are true regardless of how we choose to express them. Indeed expressing mathematics cannot be extricated from cultural matters. The language and notation we use are mere matters of convenience. They are no more inherently correct than any other system used to express these ideas – the ideas exist independently of their presentation. (This said, it is useful to express our ideas in a way that makes it easy for a reader to understand. This is a reason why we have standard notation for many things.)

---

2 Divisibility Part II

Learning Incomes.

- *Reading mathematics at a level equivalent to student with credit for MATH163. (See Module 1 for a link to the textbook from MATH163)*
- *Recalling the definition of set equality.*
- *Reading prose at a level equivalent to at least a SK high-school graduate.*
- *Familiarity with material from Module 1*

Learning Outcomes.

- *Knowledge of criteria for when a linear diophantine equation has an integer solution*
- *Full understanding of why these criteria are correct and how to apply them*
- *Ability to find the full set of integer solutions to a linear diophantine equation*
- *Improved ability to read written mathematics*

Newly Defined Terms and Notation. *coprime, linear Diophantine equation*

In Module 1 we defined some mostly straight forward terms and notation relating to divisibility. In this module we use these concepts to study linear diophantine equations.

2.1 Linear Diophantine Equations

Among all of the numbers on the number line, the integers seem particularly special. They occur at equal spaces on the number line, adding two integers together always give rise to another integer, multiplying two integers always gives rise to another integer. The positive ones are useful for counting things, which is why they have been so thoroughly studied across so many different cultures across all of human history.

It is so easy to *ask* questions about the integers, but not at all clear which questions have easy answers. For example, consider the following two questions:

1. For which positive integers n does there exist non-zero integers a, b, c so that $a^n + b^n = c^n$.
2. For which integers a, b, c does there exist integers x and y so that $ax + by = c$

The first of these questions is often referred to as *Fermat's Last Theorem*. The difficulty in understanding the solution to Fermat's Last Theorem cannot be understated. Understanding the solution in full would require years of careful study.

Though seemingly not too different from the first question, the solution to the second question can be reached in about four pages worth of lecture notes! We'll take a few diversions along the way, so it will take us ten or so.

Aside 2.1. *Pierre de Fermat was a 17th Century French mathematician. There are lots of things in Number Theory named for French mathematicians of Age of Enlightenment. Most of these things had been considered (and often answered!) long before the any of these mathematicians were even born. Fermat's Last Theorem does not fit this trend. First proposed by Fermat, it was only fully answered twenty-five years ago; such integers a, b, c only exist when $n = 1$ or $n = 2$.*

Let a, b and c be non-zero integers. A linear diophantine equation is one of the form

$$ax + by = c.$$

If we permit $x, y \in \mathbb{R}$, then finding a solution, (x, y) to a linear diophantine equation takes no work at all! The point $(x, y) = (0, c/b)$ will satisfy the equation $ax + by = c$. (In fact we can restrict $x, y \in \mathbb{Q}$ and this solution will still work.) However, if we restrict x and y to be integers, this solution may not work; for almost all choices of c and b, c/b is not an integer.

Diophantus was a Greek mathematician from the third Century. His main contribution, a mathematics text called *Arithmetica*, was a study of equations with unique solutions and those with many solutions. Though his work could probably be easily reproduced by any of us today, one must remember we have the benefit of years of mathematics education that has distilled hard fought results in to easily understood forms. (For example, we expect students to learn in one semester the basics of differential calculus, which required thousands of years of human history before anyone came up with any of the concepts).

Let a, b and c be non-zero integers. Our goal in this module is to understand when linear diophantine equations have integer solutions. That is, we want to understand when there exists a pair of integers (x_0, y_0) so that $ax_0 + by_0 = c$.

We begin with the answer to this question.

Theorem 2.2. *Let a, b and c be non-zero integers. The equation*

$$ax + by = c$$

has an integer solution (x, y) if and only if the greatest common divisor of a and b divides c.

Without thinking about it further, there is no reason to think this theorem ought to be true. (I suppose calling it a theorem suggests that it ought to be true...). There is no reason to think that the greatest common divisor is at all related to linear diophantine equations. As a first step in understanding this theorem, let us consider an example.

Example 2.3. *Determine if there exists integers x and y so that $12x + 9y = 6$.*

By inspection we can find $gcd(12, 9) = 3$. Since 3 divides 6, Theorem 2.2 tells that that there exists an integer solution to $12x + 9y = 6$.

Example 2.4. *Determine if there exists integers x and y so that $12x + 9y = 7$.*

By inspection we can find $gcd(12, 9) = 3$. Since 3 does not divide 7, Theorem 2.2 tells that that there exists no integer solution to $12x + 9y = 7$.

The statement of Theorem 2.2 gives us no idea how to actually find such a solution.

Thinking back to our experiences in other areas of mathematics, we might recognize $12x + 9y = 6$ and $12x + 9y = 7$ as lines in $\mathbb{R}^2$. The points $(x, y) \in \mathbb{R}^2$ that lay on this line are

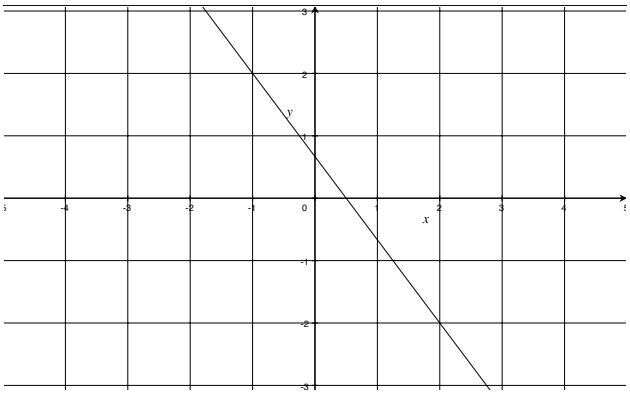


Figure 1: The line $12x + 9y = 6$

exactly those that satisfy the equations of these lines. Thus Theorem 2.2 is telling us which lines in $\mathbb{R}^2$ pass through integer points.

Looking at Figure 2.1 we see that the line $12x + 9y = 6$ passes through the point $(2, -2)$. Thus

$$12(2) + 9(-2) = 6.$$

The point $(2, -2)$ is an integer solution for the linear diophantine equation $12x + 9y = 6$. Looking again we can see another integer solution at $(-1, 2)$. Indeed $12(-1) + 9(2) = 6$. At this point we may wonder if there are more integer solutions lurking out there. Theorem 2.2 isn't helpful – it only tells us about the existence of a single solution. Perhaps we can use one of initial solutions to fine more.

Each of $12, 9$ and 6 are divisible by $gcd(12, 9)$. Notice that integer solutions to $12x + 9y = 6$ are in a one-to-one correspondence with integer solutions to:

$$\frac{12}{3}x + \frac{9}{3}y = \frac{6}{3}$$

Simplifying yields $4x + 3y = 2$. From our figure we see that $(2, -2)$ is an integer solution to $12x + 9y = 6$. It is easy to check that it is also a solution for $4x + 3y = 2$. If we could find another integer solution for $4x + 3y = 2$, say (x', y'), then we would have

$$4(2) + 3(-2) = 2 = 4x' + 3y'$$

Rearranging this equation yields

$$4(2 - x') = -3(-2 - y')$$

Though it is not clear yet why this is helpful, we can see 4 divides $-3(-2 - y')$ (here $k = (2 - x')$). If we knew that 4 divided $(-2 - y')$, then we would be guaranteed the existence of an integer t so that

$$4t = (-2 - y')$$

Since $4(2 - x') = -3(-2 - y')$, we would then have

$$4(2 - x') = -3(4t)$$

Rearranging yields $x' = 2 + 3t$.

Let us try some choices for t what we get. When $t = 0$ we get $x' = 2$ and

$$\begin{aligned} 2 &= 4x' + 3y' \\ 2 &= 4(2 + 3t) + 3y' \\ 2 &= 8 + 3y' \\ -2 &= y' \end{aligned}$$

This is our initial solution, $(2, -2)$. Let us try another value of t. When $t = 1$ we have $x' = 2 + 3 = 5$ and

$$\begin{aligned} 2 &= 4(5) + 3y' \\ -18 &= 3y' \\ -6 &= y' \end{aligned}$$

We can easily check that the solution $(5, -6)$ satisfies the equation $4x+3y = 2$. By our above remarks, this solution satisfies our original equation $12x + 9y = 6$. We have constructed a new solution to our original linear diophantine equation.

Exercise 2.5. *Which choice of t gives the solution $(-1, 2)$?*

There is nothing particularly special about our choice of example here. And so it seems possible that this technique can be used for any linear diophantine equation. Indeed, the existence of a single integer solution to a linear diophantine equation implies the existence of infinitely many integer solutions.

Theorem 2.6. *Let a, b and c be non-zero integers. If (x_0, y_0) is an integer solution to $ax + by = c$, then*

$$x' = x_0 + \frac{b}{gcd(a,b)}t,$$
$$y' = y_0 - \frac{a}{gcd(a,b)}t$$

is an integer solution for each $t \in \mathbb{Z}$. Further, every integer solution to $ax + by = c$ can be expressed in this form.

We prove this theorem on Assignment 1. Let us continue with our example.

Example 2.7. *Describe all solutions to the linear diophantine equation $12x + 9y = 6$.*

We have $gcd(12, 9) = 3$. By inspection we have an initial solution $(x_0, y_0) = (2, -2)$ By Theorem 2.6 the set of all solutions to $12x + 9y = 6$ can be described as (x', y') so that

$$x' = 2 + 3t,$$
$$y' = 2 - 4t$$

for each $t \in \mathbb{Z}$.

Aside 2.8. *Compare this idea with system of linear equations. In a linear algebra course you may have seen that a system of linear equations has either no solution, exactly one solution or infinitely many solutions. We have a similar statement for a linear diophantine equations, except it is impossible for there to be exactly one integer solution. Theorem 2.6 tells us that every linear diophantine equation has either no integer solution or infinitely many integer solutions.*

Applying Theorem 2.6 requires us to have an initial integer solution to our linear diophantine equation. But even knowing that such an initial solution exists requires us to know the greatest common divisor of a and b. Without a method to compute a greatest common divisor, we can't even know if this initial integer solution exists! We explore this problem in the following section.

2.2 Computing The Greatest Common Divisor.

Consider trying to find the greatest common divisor of 104 and 10. We can list out the divisors of each. Since the greatest common divisor must be positive, we can restrict ourselves to just considering the positive divisors of each. The positive divisors of 10 are: $1, 2, 5, 10$. The positive divisors of 104 are $1, 2, 4, 8, 13, 26, 52, 104$. By inspection, $gcd(104, 10) = 2$. These are relatively small numbers and so it is not terribly onerous to take the time and effort to list out their common divisors. This would be much more difficult to do with much larger numbers.

A helpful lemma using Theorem 1.13 provides us with a shortcut.

Lemma 2.9. *Let a, b, q, r be integers so that $b \geq 1$, $a = qb + r$, and $0 \leq r \leq b - 1$. We have $gcd(a, b) = gcd(b, r)$.*

Let us look at example to see how this lemma can be helpful to us.

Example 2.10. *Compute $gcd(9125, 325)$.*

By performing long division we can find

$$9125 = 28(325) + 25.$$

Lemma 2.9 tells us

$$gcd(9125, 325) = gcd(325, 25).$$

Thus to find $gcd(9125, 325)$ we can instead find $gcd(325, 25)$.

We recognize 325 as a multiple of 25 and so we conclude $gcd(325, 25) = 25$. Therefore $gcd(9125, 325) = 25$.

We turn now to providing a proof for Lemma 2.9.

Given a and b as integers so that $a = qb + r$ with $0 \leq r \leq b - 1$, we want to prove $gcd(a, b) = gcd(b, r)$. Let us examine the common divisors of a and b, as well as the common divisors of b and r. Let C_{ab} be the set of common divisors of a and b. Let C_{br} the set of common divisors of b and r.

If we can show $C_{ab} = C_{br}$, then the largest element of the set C_{ab} is the same as the largest element of the set C_{br}. The largest element of C_{ab} is necessarily the greatest common divisor of a and b. Similarly, the largest element of C_{br} is necessarily the greatest common divisor of b and r. Thus, if we can show $C_{ab} = C_{br}$, then necessarily we have $gcd(a, b) = gcd(b, r)$.

(Recall from MATH163, that sets A and B are equal when every element of A is an element of B and every element of B is an element of A.)

Consider $x \in C_{ab}$. Since $x \in C_{ab}$, we have that x divides a and x divides b. Therefore there exists $k_a, k_b \in \mathbb{Z}$ so that $xk_a = a$ and $xk_b = b$. To show $x \in C_{br}$ we must show x divides r. That is, we must find $k \in \mathbb{Z}$ so that $xk = r$.

Recall $a = qb + r$. Thus

$$r = a - qb = xk_a - qxk_b = x(k_a - qk_b)$$

Let $k = k_a - qk_b$. And so we see $r = xk$. Thus x divides r and so $x \in C_{br}$.

To complete our argument we must show that every element of C_{br} is also an element of C_{ab}. This argument has a very similar form as the one above, and so it is left as an exercise.

Thus we are (mostly) convinced that Lemma 2.9 is true.

Aside 2.11. *Asking you to provide a full proof of Lemma 2.9 using the terminology and approach above seems like an excellent midterm question.*

Let us proceed with another example.

Example 2.12. *Find* $gcd(1492, 1066)$.

We apply Lemma 2.9. We find $1492 = 1(1066)+426$. *Therefore* $gcd(1492, 1066) = gcd(1066, 426)$.

We find $1066 = 2(426) + 214$. *Therefore* $gcd(1066, 426) = gcd(426, 214)$.

We find $426 = 1(214) + 212$. *Therefore* $gcd(426, 214) = gcd(214, 212)$.

We find $214 = 1(212) + 2$. *Therefore* $gcd(214, 212) = gcd(212, 2)$.

We find $212 = 2(106) + 0$. *We notice* 212 *is a multiple of* 2 *and so* $gcd(212, 2) = 2$.

Therefore $gcd(1492, 1066) = 2$.

Finding the greatest common divisor using Lemma 2.9 is usually called the Euclidean Algorithm. It is possible to describe it very precisely in a way that occludes the actual application of Lemma 2.9. We will resist doing this as it is not particularly helpful for our purpose.

Aside 2.13. *Euclid's name will come up a lot in this course. His name is pronounced* YOO-KLID. *Euclid was a mathematician in around 300BCE. Mathematicians have a terrible habit of naming things after people. This tradition isn't very helpful for students. Calling this approach the* Euclidean Algorithm *doesn't give us any hints for what it is for. A better name would the* gcd finding algorithm.

Wait– why were we doing this. It seems as if we have come a long way from our original goal of understanding Theorem 2.2. Let us return this with an example.

Example 2.14. *Determine if the linear diophantine equation* $1492x + 1066y = 7$ *has an integer solution.*

By Theorem 2.2, this equation has an integer solution if and only if $gcd(1492, 1066)|7$*. Since* $gcd(1492, 1066) = 2$ *and* $2 \nmid 7$*, it then follows that* $1492x + 1066y = 7$ *has no integer solution.*

Example 2.15. *Determine if the linear diophantine equation* $1492x + 1066y = 6$ *has an integer solution.*

By Theorem 2.2, this equation has an integer solution if and only if $gcd(1492, 1066)|6$*. Since* $gcd(1492, 1066) = 2$ *and* $2 \mid 6$*, it then follows that* $1492x + 1066y = 6$ *has an integer solution.*

In some sense, the statement of Theorem 2.2 is unsatisfying. This theorem gives us a criteria to determine if there is an integer solution, but gives us no information on how to actually find one. All is not lost! The work we have already done is enough to construct an integer solution to $1492x + 1066y = 6$.

Look back at our work in Example 2.12. We have the following equalities:

$$\begin{aligned} 1492 &= 1(1066) + 426 \\ 1066 &= 2(426) + 214 \\ 426 &= 1(214) + 212 \\ 214 &= 1(212) + 2 \end{aligned}$$

Rearranging this last equation yields:

$$2 = 214 - 1(212)$$

Looking at the third equation we can isolate 212 and substitute:

$$\begin{aligned} 2 &= 214 - 1(426 - 214) \\ &= -1(426) + 2(214) \end{aligned}$$

Looking at the second equation we can isolate 214 and substitute:

$$\begin{aligned} 2 &= -1(426) + 2(214) \\ &= -1(426) + 2(1066 - 2(426)) \\ &= 2(1066) - 5(426) \end{aligned}$$

Looking at the first equation we can isolate 426 and substitute:

$$\begin{aligned} 2 &= 2(1066) - 5(426) \\ &= 2(1066) - 5(1492 - 1066) \\ &= 7(1066) - 5(1492) \end{aligned}$$

Thus

$$2 = -5(1492) + 7(1066)2 = 1492(-5) + 1066(7)$$

This isn't exactly what we want, but it is close. We want x and y so that $6 = 1492x + 1066y$. But we have x and y so that $2 = 1492x + 1066y$. However multiplying both sides of the equation

$$2 = 1492(-5) + 1066(7)$$

by 3 yields

$$6 = 1492(-30) + 1066(42)$$

Therefore $(-30, 42)$ is an integer solution to $1492x + 1066y = 6$. Notice that we could replace 6 with any multiple of 2 and this method will still work.

Applying Theorem 2.6 tells us even more! Now that we have found an initial integer solution, we can describe all possible integer solutions to the linear diophantine equation $1492x + 1066y = 6$. By Theorem 2.6, we can substitute in our values for a, b and $gcd(a, b)$ to find that that each solution is of the form:

$$\begin{aligned} x' &= -30 + \frac{1066}{2}t, \\ y' &= 42 - \frac{1492}{2}t \end{aligned}$$

That is we can describe the set of all solutions to the the linear diophantine equation $1492x + 1066y = 6$ as

$$\{(-30 + 533t, 42 - 746t) \mid t \in \mathbb{Z}\}$$

Of course there is nothing particularly special about our choice of $a = 1492$ and $b = 1066$. Given any non-zero integers a and b we can use the above method to find a solution to the linear diophantine equation $ax + by = gcd(a, b)$. If instead we want to find a solution to

$ax + by = c$, where c is a multiple of $gcd(a, b)$ then multiplying our solution for $ax + by = gcd(a, b)$ by $k = c/gcd(a, b)$ gives us a solution to $ax + by = c$. (Perhaps this gives us some insight in to why Theorem 2.2 may be true?).

Aside 2.16. *Theorem 2.2 is traditionally called* Bézout's Identity. *Bézout was a 18th-Century French mathematician and was certainly not the first to write down and prove the statement of Theorem 2.2. For example,* Bézout's Identity *plays a key role in a technique traditionally called the* Chinese Remainder Theorem. *Evidence of this use of this technique goes back as far as the 3rd Century.*

Aside 2.17. *Look at Q7 on Assignment 1. I am intentionally not adding an example that shows, start to finish, how to solve a similar problem. All of the tools you need for this question are here. The reason that I haven't put in an example that puts them altogether is because the challenge of the problem comes in putting together all of the pieces in a clear and well-explained way, and not in actually performing any of the computations. (... you still need to perform the calculations!)*

2.3 Justifying Theorem 2.2

All of our work so far has been predicated on the fact that Theorem 2.2 is actually true. Let us now provide a justification of Theorem 2.2. Theorem 2.2 is an if and only if statement. We want to justify the following two statements:

1. If $gcd(a, b)|c$, then $ax + by = c$ has an integer solution.
2. If $ax + by = c$ has an integer solution, then $gcd(a, b)|c$.

First consider the case where $gcd(a, b)|c$. Let $d = gcd(a, b)$. Since $d|c$ there exists $k \in \mathbb{Z}$ so that $dk = c$. Using the greatest common divisor finding algorithm in Section 2.2 to find the greatest common divisor of two integers we can find integers x and y so that $ax + by = d$. Multiplying both sides of this equation by k yields

$$a(xk) + b(yk) = dk = c$$

Thus (xk, yk) is an integer solution to $ax + by = c$.

Consider now the case where $ax + by = c$ has an integer solution (x_0, y_0). Let $gcd(a, b) = d$ We want to show $d|c$. Since $c = ax_0 + by_0$ it suffices to show $d|ax_0 + by_0$. Since $d = gcd(a, b)$ necessarily there exists $k_a, k_b \in Z$ so that $dk_a = a$ and $dk_b = b$. Therefore

$$c = ax_0 + by_0 = dk_a x_0 + dk_b y_0 = d\,(ax_0 + by_0)$$

Thus $c = dk$ for $k = ax_0 + by_0$. Therefore d divides c.

Aside 2.18. *This isn't really a full proof of Theorem 2.2. We didn't provide a proof that the gcd finding algorithm followed by the substitution will always provide the desired integer solution. Proving that an algorithm works in general can be tricky and isn't always insightful. We will generally avoid such activity in this class.*

2.4 An interesting case: $ax + by = 1$

For a given pair of integers a and b only some values of c give a linear diophantine equation that has an integer solution. Whereas other choices of c give rise to one that does not. For example, when $a = 12$ and $b = 9$ choosing c as a multiple of 3 gives a linear diophantine equation that has an integer solution. Whereas choosing c as not a multiple of 3 gives a linear a diophantine equation that has no integer solution.

Consider $a = 7$ and $b = 4$. For which values of c does the linear diophantine equation $7x + 4y = c$ have an integer solution? Since $gcd(7, 4) = 1$ choosing c as a multiple of 1 yields a linear diophantine equation that has an integer solution. Since every integer is a multiple of 1, the linear diophantine equation $7x + 4y = c$ will have an integer solution no matter which integer value we choose for c. This phenomenon will occur for any pair of integers a and b so that $gcd(a, b) = 1$.

Definition 2.19. *Let a and b be integers. We say that a and b are <u>coprime</u> when $gcd(a, b) = 1$.*

Theorem 2.20. *Let a and b be non-zero integers. The linear diophantine equation $ax+by = 1$ has an integer solution if and only if a and b are coprime.*

Presented in isolation, this theorem would seem quite mysterious! But by applying the result of Theorem 2.2 and the definition of coprime, this theorem seems very reasonable!

Proof. Let a and b be non-zero integers.

Assume the linear diophantine equation $ax + by = 1$ has an integer solution. By Theorem 2.2, we have that $gcd(a, b)$ divides 1. The only divisors of 1 are 1 and -1. Thus $gcd(a, b) \in \{1, -1\}$. Since $gcd(a, b) > 0$ for any pair of integers a and b, we have $gcd(a, b) = 1$. By definition, a and b are coprime.

Assume a and b are coprime. Therefore $gcd(a, b) = 1$. By Theorem 2.2, $ax + by = 1$ necessarily has an integer solution. □

Corollary 2.21. *Let a and b be coprime. For every $c \in \mathbb{Z}$ the linear diophantine equation $ax + by = c$ has an integer solution.*

Theorem 2.20 is an *if and only if statement.* Thus the criteria of *$ax + by = 1$ has an integer solution* is an equivalent to saying *a and b are coprime.* This criteria can be unexpectedly helpful for problems in divisibility.

Example 2.22. *Let a, b and c be integers so that a and b are coprime. Show that if a divides c and b divides c, then ab divides c.*

We want to find $k \in \mathbb{Z}$ so that $abk = c$. Since a divides c and b divides c there exists $k_a, k_b \in \mathbb{Z}$ so that $ak_a = c$ and $bk_b = c$.

Since a and b are coprime there exists integers x and y so that $ax + by = 1$. Multiplying both sides by c yields

$$axc + byc = c.$$

Substituting for c on the left side yields

$$\begin{aligned} ax(bk_b) + by(ak_a) &= c \\ ab(xk_b + yk_a) &= c \end{aligned}$$

Let $k = xk_b + yk_a$ and notice $abk = c$. Thus $ab|c$.

(This example is quite helpful for Q5 on Assignment 1)

2.5 Exercises

Exercises

1. In the statement of Theorem 2.2 there is no restriction that b must be positive. However, in our work we applied the result of Theorem 1.13, which required b to be positive. Explain how you can use solutions for $ax + by = c$ to give solutions for $ax - by = c$.
2. Give a full proof of Lemma 2.9 using the same approach and notation that we used above.
3. We applied the result of Exercise 2 from Module 1 at least twice in our work in Module 2. Can you spot where we could shorten our argument by first stating the result of this exercise as a lemma?

Other things to think about–

- For those with an interest in computer science – the gcd finding algorithm introduced above can be written down as psuedo-code, which can then be converted into an actual running algorithm. Would you expect such an algorithm to be recursive? If so, how many lines of code would you expect the recursive routine to be? Many? Few? The time complexity of this algorithm is related to the Fibonacci numbers (... yes, really!). Do an internet search for `Lame's Theorem` for more information.
- For those with an interest in teaching – I can't imagine it would be terribly difficult for anyone to understand the concept of *greatest common divisor*. Do you think there is value in teaching someone to use our gcd finding algorithm without teaching them to understand why the method works? What is the minimum list of things that someone would need to know to be able to broadly understand why our gcd finding algorithm works?
- For those with an interest in pure mathematics – At the start of the 20th Century, German mathematician David Hilbert proposed a list of 23 unsolved mathematical problems. Many of which would go on to influence the course of 20th Century mathematics.

  Dropping the word linear from linear diophantine equation leads to a much wider class of equations for which finding integer solutions may be quite difficult. Hilbert's tenth problem asked if there was a general method to decide if a diophantine equation admits an integer solution. It was only fully solved in 1970; the answer is no.

---

3 Prime Numbers

Learning Incomes.

- *Concepts from Modules 1 and 2. Specifically familiarity with divides, the method of finding solutions to linear diophantine equations the greatest common divisor of two integers, and coprime integers.*
- *Familiarity with the technique of proof by induction.*

Learning Outcomes.

- *Understanding of the proof of the infinitude of the primes, as well as applications from the proof method.*
- *Understanding of the statement and proof ideas for the fundamental theorem of arithmetic,*
- *Ability to find primes in an interval* $[1, n]$ *by using the Sieve of Eratosthenes*

Newly Defined Terms and Notation. *prime, coprime, prime-power factorization*

Recall from the previous section the following result:

Theorem. *Let* a *and* b *be non-zero integers. The linear diophantine equation* $ax + by = 1$ *has an integer solution if and only if* a *and* b *are coprime.*

In the previous section we saw an algorithm that lets us determine the greatest common divisor of a pair of integers. For example, we find $gcd(53, 17)$ using Lemma 2.9 with the following sequence of equalities:

$$\begin{aligned} 53 &= 3(17) + 10 \\ 17 &= 1(10) + 3 \\ 10 &= 3(3) + 1 \\ 3 &= 3(1) + 0 \end{aligned}$$

From this we can conclude $gcd(53, 17) = 1$.

At the start of Module 2, we briefly saw another method to find the greatest common divisors of a pair of integers: list out all of their positive divisors and look for the largest common one. Let us try the same approach here. The positive divisors of 17 are 1 and 17. The positive divisors of 53 are 1 and 53. Therefore $gcd(53, 17) = 1$.

In this case, finding the greatest common divisor by listing out all divisors seems much easier than applying Lemma 2.9. Each of 17 and 53 only have one divisor other than themselves – namely 1.

Definition 3.1. *Let n be an integer so that $n \geq 2$. We say n is prime when the only positive divisors of n are 1 and n. If n is not prime, we say that n is composite.*

Exercise 3.2. *Show that if n is composite, then it has n divisor in the interval $[2, \lfloor\sqrt{n}\rfloor]$. Proceed by contradiction. Consider the product of two integers in the range $[\sqrt{n}, n-1]$.*

Aside 3.3. *Is 1 prime? According to our definition, 1 is neither prime nor composite. You may wonder why we require $n \geq 2$ in our definition of prime. This is a perfectly reasonable question that does have a reasonable answer. We will return to this question in Section 3.1.*

In this module we study prime numbers. As we start to search for examples of prime numbers, our work begins quite easily; each of $2, 3, 5, 7, 11$ is prime. But as we start to look at larger integers, primes seemingly become more rare. Between 2 and 100 there are 25 primes. Whereas between 900 and 1000 there are only 14. Intuitively, this might make sense: for any integer $n > 2$, all of the numbers from 2 up until $\lfloor n/2 \rfloor$ are possible divisors of n. When n is small there are far fewer numbers in this range as compared to when n is large. And so, intuitively, we expect that the likelihood that a number n is prime decreases and n increases.

Aside 3.4. *A reasonable question to ask at this point is* " what is the probability that a randomly chosen positive integer is prime?" *Intuitively, we may be able to easily make sense of this question, but mathematically it is not so straight forward. We return to this question in Section 3.2.*

Rather than thinking of whether or not an individual integer is prime, let us first consider methods to generate prime numbers.

3.1 Finding Prime Numbers

A very reasonable question at this point is *why might we want to find prime numbers?*. We will see some applications later on the course, but for now our curiousity will have to suffice. If we are satisfied that one can be interested in studying interstellar astronomy, so too must we be satisfied that one can be interested in studying primes numbers. It is in our nature as humans to explore the unknown.

We begin with a method to generate prime numbers based on continuous dividing. Consider the integer 30. This integer is not prime; it is divisible by 3

$$30 = 3 \times 10$$

We notice that 10 is even. And so

$$30 = 3 \times 2 \times 5$$

Each of these factors is prime. We can divide no further.

Consider the integer 6136. This integer is not prime; it is even and so is divisible by 2.

$$6136 = 2 \times 3068$$

Again we see that 3068 is even, and so:

$$6136 = 2^2 \times 1534$$

One more time:

$$6136 = 2^3 \times 767$$

After a bit of calculator work, we can find $767 = 13 \times 59$ and so

$$6136 = 2^3 \times 13 \times 59$$

The integer 13 is prime. To check if 59 is prime we need only check that none of $2, 3, 4, 5, 6, 7$ is a divisor of 59.

Exercise 3.5. *Explain why we need only check integers up to* 7 *to determine if* 59 *is prime?*

None of these integers is a divisor of 59 and so we can conclude 59 is prime.

In this example there is nothing particularly special about 6136. It seems as if this procedure will work for any composite integer. That is, by repeatedly finding factors of a composite integers we eventually return a list of primes that when multiplied together gives our original integer.

Theorem 3.6 (The Fundamental Theorem of Arithmetic)**.** *Every integer $n \geq 2$ can be uniquely expressed as:*

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

where $p_1, p_2, \ldots p_k$ are distinct primes and each of $e_1, e_2, \ldots, e_k$ is a positive integer.

For example, we have

$$6136 = 2^3 \times 13 \times 59$$

And so for 6136 we have $k = 3$, $p_1 = 2, p_2 = 13, p_3 = 59$ and $e_1 = 2, e_2 = e_3 = 1$. We call such a factorization an prime-power factorization.

To begin to think about proving our results, consider the integer 184080. After some calculator work we can find

$$184080 = 30 \times 6136$$

We have already expressed each of 30 and 6136 as the product of primes, and so there is no need to do the work again.

$$\begin{aligned} 184080 &= (2 \times 3 \times 5) \times \left(2^3 \times 13 \times 59\right) \\ &= 3 \times 5 \times 2^4 \times 13 \times 59 \end{aligned}$$

We re-used our work from more than one smaller cases to complete a larger case. Strong induction lurks.

Imagine our theorem is true for all integers from 2 up until some integer ℓ. If $\ell + 1$ is prime, then it has a prime power factorization with $k = 1$, $p_1 = \ell + 1$ and $e_1 = 1$. Otherwise, if $\ell + 1$ is composite, then there exists $a, b \in [2, \ell]$ so that $\ell + 1 = ab$. Since the theorem is true for each of a and b, we can use the prime-power factorization for each of a and b to find the factorization for $\ell + 1$.

Of course, this says nothing of uniqueness. Imagine some integer n had two different prime power factorizations:

$$q_1^{d_1} q_2^{d_2} \cdots q_r^{d_r} = n = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}$$

To complete the argument, we must find a contradiction.

Using these ideas, we will prove return to the proof of this theorem on Assignment 2.

This theorem is often called the *Fundamental Theorem of Arithmetic.* Along with the *Fundamental Theorem of Calculus* and the *Fundamental Theorem of Algebra* it underpins many areas of modern mathematics.

Example 3.7. *At some point we may have encountered the idea that $\sqrt{2}$ is irrational. That is to say, $\sqrt{2}$ cannot be expressed as the ratio of two integers. With the Fundamental Theorem of Arithmetic, we can prove this fact.*

We proceed by contradiction. Assume $\sqrt{2}$ is rational. That is, assume there exist integers a and b so that

$$\sqrt{2} = \frac{a}{b}$$

By the Fundamental Theorem of Arithmetic, each of a and b have a unique prime-power factorization.

$$a = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}$$
$$b = q_1^{d_1} q_2^{d_2} \cdots q_r^{d_r}$$

Rearranging and squaring our expression for $\sqrt{2}$ yields:

$$2\left(q_1^{d_1} q_2^{d_2} \cdots q_r^{d_r}\right)^2 = \left(p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}\right)^2$$
$$2q_1^{2d_1} q_2^{2d_2} \cdots q_r^{2d_r} = p_1^{2e_1} p_2^{2e_2} \cdots p_k^{2e_k}$$

Let $n' = p_1^{2e_1} p_2^{2e_2} \cdots p_k^{2e_k}$. In this prime-power factorization of n', each prime power is raised to an even exponent. Since this prime-power factorization is unique it must be that in the prime-power factorization

$$n' = 2q_1^{2d_1} q_2^{2d_2} \cdots q_r^{2d_r}$$

each prime is raised to an even exponent. Consider the prime 2 in the prime power factorization $n' = 2q_1^{2d_1} q_2^{2d_2} \cdots q_r^{2d_r}$. If $d_1, d_2, \ldots, d_r \neq 2$, then 2 is raised to an odd power, 2^1. This

is a contradiction as all primes in this prime-power factorization must be raised to an even exponent.

Therefore $d_i = 2$ *for some* $1 \leq i \leq t$. *Thus*

$$n' = q_1^{2d_1} q_2^{2d_2} \cdots q_{i-1}^{2d_{i-1}} q_i^{2d_i+1} q_{i+1}^{2d_{i+1}} \cdots q_r^{2d_r}$$

We observe that q_i *is raised to an odd power, a contradiction. And so we conclude that* a *and* b *do not exist. Therefore* $\sqrt{2}$ *is irrational.*

Exercise 3.8. *Let* $n \geq 2$ *be a prime. Show* $\sqrt{p}$ *is irrational. Why does this same argument not work when* n *is composite? Does the argument work if* $n = p_1 p_2$ *where* p_1 *and* p_2 *are distinct primes?*

Aside 3.9. *Back at the start of this section we wondered, briefly, why we might want to exclude* 1 *from being prime. Look again at the statement of the Fundamental Theorem of Arithmetic. If* 1 *is prime, then the uniqueness part of this statement is false; if* 1 *is prime then we can multiply by* 1 *as many times as we wish to get a different prime factorization.*

If 1 *is permitted to be a prime, then our method in Example 3.7 will not work – it relied on the prime factorization being unique. We could permit* 1 *to be prime and then change the statement of the Fundamental Theorem of Arithmetic accordingly to disallow* 1 *as appearing in a prime-power factorization. Consequently the statement of any result that then relied on the Fundamental Theorem of Arithmetic would need to be changed accordingly.*

As a mathematical community, we choose to not allow 1 *to be a prime number for convenience. It is easier for us to disallow* 1 *in the definition of prime number, than to treat* 1 *as a special case in each theorem about prime numbers.*

Definitions in mathematics are a matter of convention and convenience. They are not inherently correct or incorrect; they are just shorthand to let us communicate a complex ideas.

3.1.1 Primality Testing

A common news item in popular science media is an announcement of a new largest prime number. (See sciencedaily.com/releases/2018/01/180104164507.htm for example.). Finding prime numbers requires us to have method to check if a number is prime.

Let $n \geq 2$ be an integer. What strategies can we use to test if n is prime?

Of course, if someone has pre-computed a list of all prime numbers in a given range, our the solution is easy: we check to see if n is contained in the list of prime numbers. But of course this then forces us to consider determining a method to generate a list of all of the prime numbers in a particular interval. We consider a method based upon the following observation:

- if $t \geq 2$ is an integer, then every multiple of t (other than possibly t itself) is composite; and

Consider the following list of positive integers from 2 to 20.

$$2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20$$

We recognize that 2 is prime and we apply our first observation to conclude that none of $4, 6, 8, \ldots, 20$ is prime

$$\textcircled{2}, 3, \not{4}, 5, \not{6}, 7, \not{8}, 9, \not{10}, 11, \not{12}, 13, \not{14}, 15, \not{16}, 17, \not{18}, 19, \not{20}$$

Moving on, we see that 3 must be prime. It has not been crossed out and so it is not a multiple of any smaller prime number. We apply our observation to conclude that none of $6, 9, 12, 15, 18$ is prime.

$$\textcircled{2}, \textcircled{3}, \not{4}, 5, \not{6}, 7, \not{8}, \not{9}, \not{10}, 11, \not{12}, 13, \not{14}, \not{15}, \not{16}, 17, \not{18}, 19, \not{20}$$

Moving on again, we see that 5 must be prime. It is the leftmost number that is not crossed out and not confirmed to be prime. Since it is not crossed out, it is not a multiple of any smaller prime number. And so we conclude it is prime. We apply our observation to cross any multiple 5

$$\textcircled{2}, \textcircled{3}, \not{4}, \textcircled{5}, \not{6}, 7, \not{8}, \not{9}, \not{10}, 11, \not{12}, 13, \not{14}, \not{15}, \not{16}, 17, \not{18}, 19, \not{20}$$

Proceeding in this manner, we arrive at the following list:

$$\textcircled{2}, \textcircled{3}, \not{4}, \textcircled{5}, \not{6}, 7, \not{8}, \not{9}, \not{10}, \textcircled{11}, \not{12}, \textcircled{13}, \not{14}, \not{15}, \not{16}, \textcircled{17}, \not{18}, \textcircled{19}, \not{20}$$

Every number that is circled is necessarily prime. And every number that is crossed out is necessarily composite. Thus we have generated a list of all prime numbers from 2 up to 20.

This process is called the *Sieve of Eratosthenes.* Eratosthenes was a Greek mathematician who was alive approximately 2200 years ago. This method is commonly attributed to Eratosthenes. Undoubtedly it has been discovered and rediscovered by people from cultures all over the world.

Exercise 3.10. *Use the Sieve of Eratosthenes to find all prime numbers in the interval* $[2, 50]$.

Aside 3.11. *The Sieve of Eratosthenes doesn't seem terribly complicated to us. But indeed we stand on the shoulder of giants. For example, place value notation for numbers, which enables our multiplication algorithm, first appeared 1000 years after the death of Eratosthenes. Moreover, Eratosthenes did not even have the notation = to make his computations easier to express.*

As a method for looking for large prime numbers, the Sieve of Eratosthenes does not seem very efficient. It requires us to start by crossing out multiples of 2, and then multiples of 3, and then multiples of 5, etc... This method gives us no real hope in checking if a very large number is prime.

Aside 3.12. *Work on finding algorithms to test whether a particular integer is prime has been the subject of on going research for over 1000 years. In 2002 researchers provided the first deterministic polynomial-time algorithm to decide if an integer is prime. Though the algorithm is not useful in practice (it is $O(n^{12})$), it provides proof that the primality testing is* Polynomial. *This is an interesting result when placed in the broader context of the study of the polynomial hierarchy of computational complexity classes.*

More information on this algorithm can be found by an internet search for `Primes in P`

None of the work we have done in this section implies that there are infinitely many primes. Even very large numbers can have only very small prime factors. For example, $1048576 = 2^{20}$. Perhaps our efforts to find prime numbers will eventually exhaust them all!

3.2 Can we find all of the prime numbers?

No.

Theorem 3.13. *There are infinitely many primes.*

Proof. We proceed by contradiction. If there are not infinitely many primes, then there are exactly k primes for some positive integer k. Let $p_1, p_2, \ldots, p_k$ be these k primes in increasing order of magnitude. That is, we may assume

$$p_1 < p_2 < \cdots < p_k$$

Consider the product

$$p_1 p_2 \cdots p_k + 1$$

As $p_1 p_2 \cdots p_k + 1 > p_k$, it cannot be that $p_1 p_2 \ldots p_k + 1$ is prime; it is larger than the largest of the primes. Therefore $p_1 p_2 \cdots p_k + 1$ is composite. And so by the Fundamental Theorem of Arithmetic, $p_1 p_2 \ldots p_k + 1$ can be expressed as a product of prime powers. Since there are only finitely many primes, there exists $p_i \in \{p_1, p_2, \ldots, p_k\}$ so that p_i divides $p_1 p_2 \cdots p_k + 1$. Since p_i divides $p_1 p_2 \cdots p_k + 1$, there exists an integer t so that

$$p_i t = p_1 p_2 \cdots p_k + 1$$

Therefore

$$p_i(t - p_1 p_2 \cdots p_{i-1} p_{i+1} p_{i+2} \cdots p_k) = 1$$

Thus 1 is a multiple of p_i. The integer 1 is only a multiple of 1 and -1. Therefore $p_i \in \{-1, 1\}$. This is a contradiction as $p_i \geq 2$ □

The proof of Theorem 3.13 proceeds by first constructing the product of the first k primes, adding one, and then showing that none of the first k primes divides this number. This does not imply that $p_1 p_2 \cdots p_k + 1$ is necessarily prime. It is possible that it has a divisor in the range $[p_k + 1, p_1 p_2 \cdots p_k]$. Thus, this proof implies that the $k + 1$st prime must be found in the interval $[p_k + 1, p_1 p_2 \cdots p_k + 1]$. This observation allows us to find an upper bound on the size of the $k + 1st$ prime.

Exercise 3.14. *The proof of Theorem 3.13 does not imply $p_1 p_2 \cdots p_k + 1$ is prime. Testing some small values of k it is easy to mistakenly convince yourself that $p_1 p_2 \cdots p_k + 1$ is always prime. What is the least integer k so that $p_1 p_2 \cdots p_k + 1$ is not prime?*

Corollary 3.15. *The nth prime number, p_n, satisfies:*

$$p_n \leq 2^{2^{n-1}}$$

Before we prove this result, we first build some intuition. When $n = 1$, we are considering the first prime number: 2 (i.e., $p_1 = 2$). We verify

$$2 \leq 2^{2^0}$$

When $n = 2$, we are considering the second prime number: 3 (i.e., $p_2 = 3$) We verify

$$3 \leq 2^{2^1}$$

When $n = 3$, we are considering the third prime number: 5 (i.e., $p_3 = 5$) Rather than computing our bound directly, let us apply the reasoning from Theorem 3.13. We have:

$$p_3 \leq p_1 p_2 + 1 \leq 2^{2^0} \cdot 2^{2^1} + 1 = 2^{2^0+2^1} + 1$$

We recall the following fact:

$$2^0 + 2^1 = 2^2 - 1$$

Thus,

$$p_3 \leq 2^{2^2-1} + 1 = \frac{1}{2}2^{2^2} + 1 \leq 2^{2^2}$$

We consider now $p_4 = 7$

$$p_4 \leq p_1 p_2 p_3 + 1 \leq 2^{2^0} 2^{2^1} 2^{2^2} + 1 = 2^{2^0+2^1+2^2} + 1$$

We recall:

$$2^0 + 2^1 + 2^2 = 2^3 - 1$$

Thus

$$p_4 \leq 2^{2^0+2^1+2^2} + 1 = 2^{2^3-1} + 1 = \frac{1}{2}2^{2^3} + 1 \leq 2^{2^3}$$

We could continue on to consider p_5, but our work would be the same. In this case we would exploit the fact

$$2^0 + 2^1 + 2^2 + 2^3 + 2^4 = 2^5 - 1$$

We can exploit this pattern to construct a proof by strong induction for Corollary 3.15.

Proof. We proceed by strong induction. In this case $P(n)$ is the statement:

$$p_n \leq 2^{2^{n-1}}$$

We begin our induction starting at 1 rather than 0, but the process is the same. We notice that the claim holds when $n = 1$.

Consider the case $n = k + 1$.

Following the reasoning from the proof of Theorem 3.13 we have

$$p_{k+1} \leq p_1 p_2 \cdots p_k + 1$$

We are considering statement 2 in the principle of mathematical induction. And so we assume that $P(k)$ is true. Thus

$$p_i \leq 2^{2^i - 1}$$

for each $1 \leq i \leq k$.

Therefore

$$\begin{aligned} p_{k+1} \leq p_1 p_2 \cdots p_k + 1 &\leq 2^{2^0} 2^{2^1} \cdots 2^{2^{k-1}} + 1 \\ &= 2^{2^0 + 2^1 + \cdots + 2^{k-1}} + 1 \\ &= 2^{2^k} + 1 \\ &= \frac{1}{2} 2^{2^k} + 1 \\ &\leq 2^{2^k} \end{aligned}$$

Thus $p_{k+1} \leq 2^{2^k}$. Our result now follows by induction. □

Aside 3.16. *This bound is not good. Consider the case $t = 5$. Compare p_5 with 2^{2^4}. Though this bound is not good, a bad bound is better than no bound. This bound tells us how far we have to search to find the next prime number. This turns a potentially unbounded search space in to a bounded one (computer scientists rejoice!)*

For any positive integer n, let $\pi(n)$ denote the quantity of primes in the range $[1, n]$. For example, $\pi(10) = 4$, there are four primes between 1 and 10. Using Corollary 3.15 one can deduce a lower bound on $\pi(n)$

$$\pi(n) \geq \lfloor \log_2(\log_2 n) \rfloor + 1$$

As our bound in Corollary 3.15 is very poor, this bound for $\pi(n)$ is very poor.

Aside 3.17. *As there are infinitely many prime numbers, the function $\pi(n)$ is an increasing function (it is not monotonically increasing!). Just over 100 years ago, mathematicians proved a statement about the growth rate of of $\pi(n)$. Considering this theorem carefully requires content beyond the pre-requisite of this course. And so we will not spend any time thinking about the following result:*

Theorem 3.18 (The Prime Number Theorem)**.**

$$\pi(n) \sim \frac{n}{\ln(n)}$$

Here $\sim$ denotes asymptotic equivalence. The statement of the Prime Number Theorem is equivalent to

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

In Aside 3.4 we considered the question " what is the probability that a randomly chosen positive integer is prime?" *To make this precise, we must grapple with the meaning of* probability *and* randomly chosen*.*

As there are infinitely many positive integers, we cannot choose uniformly at random a single one. However, if we restrict our attention a particular range, then this is possible.

For any range $[1, n]$ the probability that a randomly chosen number in that range is prime is $\pi(n)/n$. With this notation we can interpret the question " what is the probability that a randomly chosen positive integer is prime?" *to be the value of the following limit.*

$$\lim_{n\to\infty} \frac{\pi(n)}{n}$$

Using the Prime Number Theorem, one can show this limit equals 0*.*

The proof of Theorem 3.13 provides us both with a bound on the value of of the nth prime, and the number of primes less than a given integer. But this proof is not yet done giving!

Every prime other than 2 is odd. The remainder when an odd prime is divided by 4 must either be 1 or 3. That is to say, each odd prime number can be expressed in the form $4q+1$ or $4q+3$. The proof method for Theorem 3.13 can be adapted to show that there are infinitely many prime numbers that have remainder 3 when divided by 4.

Theorem 3.19. *There are infinitely many prime numbers that can be expressed in the form* $4q + 3$

Proof. We proceed by contradiction. And so we assume there exists an integer k so that there are exactly k prime numbers of the form $4q + 3$. Let $p_1, p_2, \ldots, p_k$ be these prime numbers in increasing order. Since each of these primes numbers is of the form $4q + 3$, there exist integers $q_1, q_2, \ldots, q_k$ so that $p_i = 4q_i + 3$ for each $1 \leq i \leq k$

Consider the integer $n = 4p_1p_2 \cdots p_k - 1$. Notice

$$n = 4(p_1p_2 \cdots p_k - 1) + 3$$

Thus n is of the form $4q + 3$ for $q = p_1p_2 \ldots p_k - 1$. Since $n > p_k$, n is not prime. Therefore n has a prime divisor. Since n is odd, 2 does not divide n.

We claim that n has a prime divisor contained in the set $\{p_1, p_2, \ldots, p_k\}$. If this claim is not true, then every prime divisor of n is of the form $4q + 1$. Notice that the product of a pair of integers of the form $4q + 1$ is again of the form $4q + 1$:

$$(4q + 1)(4q' + 1) = 16qq' + 4q + 4q' + 1 = 4(4qq' + q + q') + 1$$

If every prime divisor of n is of the form $4q + 1$, then necessarily n must be of the form $4q + 1$. This is a contradiction as we showed above that n is of the form $4q + 3$. Thus n has a prime divisor of the form $4q + 3$.

By assumption there are only finitely prime numbers of the form $4q + 3$: $p_1, p_2, \ldots, p_k$. Therefore there exists $p_i \in \{p_1, p_2, \ldots, p_k\}$ so that p_i divides n. Since p_i divides n there exists $t \in \mathbb{Z}$ so that

$$p_it = 4p_1p_2 \cdots p_k - 1.$$

Rearranging and factoring yields

$$p_i(t - 4p_1p_2 \cdots p_{i-1}p_{i+1}p_{i+2} \cdots p_k) = -1$$

Thus -1 is a multiple of p_i. The integer -1 is only a multiple of 1 and -1. Therefore $p_i \in \{-1, 1\}$. This is a contradiction as $p_i \geq 3$ □

Surprisingly, there is no known proof similar to the above that shows there are infinitely many integers of the form $4q + 1$.

3.3 Further Exercises

Exercises

1. Consider the integers 26299, 442 and $gcd(26299, 442)$.
   (a) Is it possible that there is a prime power that divides $gcd(26299, 442)$ that does not divide both of 26299 and 442?
   (b) Is it possible that there is a prime power that divides both of 26299, 442 but does not divide $gcd(26299, 442)$?
   (c) Confirm your answers to the previous parts by finding the prime-power factorizations of 26299, 442 and $gcd(26299, 442)$.
   (d) Use your observations from above to devise a method to find $gcd(a, b)$ and its prime-power factorization given the prime-power factorization of each of a and b.
2. Let $f(x) = x^2 + 18x + 77$
   (a) Compute $f(46)$.
   (b) Factor $f(x)$ into linear factors
   (c) Using your factorization of $f(x)$, explain how you know $f(46)$ is composite.
   (d) Show $f(z)$ is composite for any positive integer z.
   (e) Let $g(x)$ be a polynomial of degree $n \geq 2$ with integer coefficients so that $g(x)$ has n integer roots:
   $$r_1 < r_2 < \cdots < r_n < 0$$
   Show $g(z)$ is not prime for any positive integer z

---

4 Congruences Part I

Learning Incomes.

- *Knowing when $ax + by = c$ has an integer solution.*
- *Definition of* prime

Learning Outcomes.

- *Remembering/understanding the concepts of relation and equivalence relation.*
- *Understanding of notation for $\mathbb{Z}_n$.*
- *Understanding the relationship between multiplicative inverses in $\mathbb{Z}_n$ and greatest common divisor*

Newly Defined Terms and Notation. *relation, equivalence relation, reflexive, symmetric, transitive, equivalence class, $[s]$, $\mathbb{Z}_n$, additive inverse, multiplicative inverse*

Let p be a prime and let a be an integer that is not a multiple of p. Since $gcd(a, p) = 1$, necessarily there exists integers u and v so that

$$au + pv = 1$$

Rearranging yields:

$$au = 1 - pv$$

Let us notice what this is telling us about the multiples of a: necessarily there is a multiple of a that is one more than a multiple of p. Rewriting this equation to remind us of quotients and remainders:

$$au = (-v)p + 1$$

tells us that some multiple of a has remainder 1 when divided by p.

As a and p are coprime, necessarily $ax+py = c$ has a solution for any integer c. Following our line of reasoning above, we can conclude that for any $r \in \{0, 1, 2 \ldots, p-1\}$ there is a multiple of a that has remainder r when divided by p. (To see this, let $c = r$, find a solution to the corresponding linear diophantine equation and then rearrange as appropriate.) Swapping p for an integer n so that $gcd(a, n) = 1$ yields the same behaviour.

Consider now swapping p for some integer $n \geq 2$ so that $gcd(a, n) \neq 1$. Following the same train of thought, we can conclude that there is no multiple of a that has remainder 1 when divided by n.

We have concluded the following: Let a and n be integers. There exists a multiple of a that has remainder 1 when divided by n if and only if a and n are coprime.

In this module we develop some terminology and notation that will help us talk about how the behaviour of remainders changes based upon what number we are dividing by. Though not obvious, this work sets us up to talk about applications of number theory to cryptography.

Our work begins with a review of equivalence relations. We will return to our thoughts about coprimes and remainders at the end of this module.

4.1 Relations and Equivalence Relations – A Refresher (... maybe?)

Aside 4.1. *This is a refresher for students who took MATH163. I don't know how/if this subject is treated in CMPT260. In either case, this good material to review before we move on.*

Sets are unstructured – elements can be written down in any order. Without further context, there is not necessarily any structure within the elements. A relation lets us define a sense of structure for a set. For example, as a set the set $\mathbb{Z}$ has no structure. However, we often think of elements of $\mathbb{Z}$ as being in a particular order.

$$\cdots -3, -2, -1, 0, 1, 2, 3, \ldots$$

Consider the set

$$L = \{(a, b) \mid b - a \in \{1, 2, \ldots\}\}$$

Notice $(a, b) \in L$ if and only if $a < b$. This set of ordered pairs defines the usual ordering of the integers.

Recall the definition of a relation:

Definition 4.2. *Let S be a set, we say R is a relation on S when R is a subset of $S \times S$. When $(s_1, s_2) \in R$ we say s_1 is related to s_2 and we write $s_1 R s_2$.*

Informally we can understand a relation to define structure on a set.

Aside 4.3. *If we label our set L with the symbol $<$ rather than L, then the notation $a < b$ means: a* is related to b with respect to the relation $<$. *This is equivalent to saying a is less than b. Oh look! We have defined the meaning of the symbol $<$!*

Consider the following relation on $\mathbb{R}^2$:

$$C = \{((x_1, y_1), (x_2, y_2)) \mid (x_1, y_1) \text{ is the same distance from the origin as } (x_2, y_2)\}$$

Just as the relation L allows us to put elements of $\mathbb{Z}$ in an ordering, this relation allows us to group elements that are the same distance from the origin. Relations that partition a set into parts are called equivalence relations. By definition, equivalence relations satisfy the following three properties:

Definition 4.4. *Let R be a relation on S.*

1. *We say R is reflexive when $(s, s) \in R$ for each $s \in S$*
2. *We say R is symmetric when $(s, t) \in R$, implies $(t, s) \in R$.*
3. *We say R is transitive when $(r, s), (s, t) \in R$ implies $(r, t) \in R$*

Let S be a set and let R be an equivalence relation on S. We define the following notation:

$$[s] = \{t \mid sRt\}$$

$$S/R = \{[s] \mid s \in S\}$$

For an element $s \in S$, the set $[s]$ is the set of all elements of S that s is related to with respect to R. This set is called the equivalence class of $[s]$. The set S/R is then the set of equivalence classes of elements of S.

Our big idea here is that equivalence relations group together elements of a set that are equal with respect to some property. One can state this formally as follows:

Theorem 4.5. *Let S be a set and let R be an equivalence relation on S. The set*

$$S/R = \{[s] \mid s \in S\}$$

is a partition of S.

Aside 4.6. *Recall the definition of a partition:*

Definition 4.7. *Let X be a set and let $\mathcal{P}$ be a set of subsets of X. We say $\mathcal{P}$ is a partition of X when*

1. $\bigcup_{P \in \mathcal{P}} P = X$*; and*
2. $P_1 \cap P_2 = \emptyset$ *for any pair $P_1, P_2 \in \mathcal{P}$ with $P_1 \neq P_2$*

We understand the two parts of the definition to mean the following

1. *Every element of X is contained in at least one subset contained in $\mathcal{P}$; and*
2. *no element of X is contained in two different subsets that are contained in $\mathcal{P}$*

Let us put this in context by returning to the relation C above. Consider $(0, 1) \in R^2$. The set $[(0, 1)]$ is the set of all elements of $(x, y) \in \mathbb{R}^2$ so that $(0, 1)C(x, y)$. Recalling the criteria for a pair of elements to be in C, it must be that all elements of the set $[(0, 1)]$ are at distance

1 from the origin. Similarly, the set $[(1,0)]$ is the set of all points in $\mathbb{R}^2$ that are at distance 1 from the origin. Thus

$$[(0,1)] = [(1,0)]$$

Generalizing this argument, we see

$$[(x_1,y_1)] = [(x_2,y_2)]$$

whenever (x_1,y_1) and (x_2,y_2) are the same distance from the origin.

Consider $(x,y) \in R^2$ so that (x,y) is at distance r from the origin. Repeating the argument above gives

$$[(x,y)] = [(0,r)]$$

Therefore

$$\mathbb{R}^2/C = \{[(x,y)] \mid (x,y) \in \mathbb{R}^2\} = \{[(0,r)] \mid r \geq 0\}$$

The set $[(0,r')]$ is the set of points in $\mathbb{R}^2$ that are distance r' from the origin. The set of points equidistant from the origin defines a circle centred at the origin. Thus $[(0,r)]$ is the circle centred at the origin with radius r. And so the set $\mathbb{R}^2/C$ defines a partition of $\mathbb{R}^2$ into concentric circles centred at the origin.

4.2 Remainders as an Equivalence Relation

To further our study of prime numbers, we are interested in studying equivalence relations that group together integers that have the same remainder when divided by some fixed integer.

Aside 4.8. *If you have studied any abstract algebra you might recognize that we are defining the ring $\mathbb{Z}_n$ throughout this module.*

Let $n \geq 2$ be an integer. Consider the following relation on $\mathbb{Z} \times \mathbb{Z}$.

$$\equiv_n \ = \{(a, b) \mid n \text{ divides } b - a\}$$

(Don't let this jumble of notation trip you up! The label for the set is $\equiv_n$. There is a good reason for choosing this notation. We'll return to it in the following module.)

One can check that for a fixed integer $n \geq 2$, the relation $\equiv_n$ is an equivalence relation. Just as the set of equivalence classes of C were a partition of $\mathbb{R}^2$ into circles, we wonder how can we describe the equivalence classes of this partition.

Example 4.9. *Let $n = 4$ and $a = 11$. Consider the set*

$$[11] = \{b \mid 11 \equiv_4 b\}$$

and some integer $c \in \mathbb{Z}$.

By definition of $\equiv_4$, we have that $11 \equiv_4 c$ if and only if 4 divides $c - 11$. By Theorem 1.13, there exists integers q and r so that

$$c = 4q + r$$

Therefore

$$c - 11 = (4q + r) - (4(2) + 3) = 4(q - 2) + r - 3$$

Notice $c - 11$ is a multiple of 4 if and only if $r = 3$. Thus $c \in [11]$ if and only if it has remainder 3 when divided by 4. Thus $[11]$ is the set of all integers that have remainder 3 when divided by 4.

$$[11] = \{b \mid b \text{ has remainder } 3 \text{ when divided by } 4\}$$

A similar argument tells us, say,

$$[27] = \{b \mid b \text{ has remainder } 3 \text{ when divided by } 4\}$$

Thus $[11] = [27]$

Consider now $a' \in \mathbb{Z}$ *so that* a' *has remainder* r *when divided by* 4. *An argument similar to the above tells us*

$$[a'] = \{b \mid b \text{ has remainder } r \text{ when divided by } 4\}$$

Since each integer has remainder $0, 1, 2$ *or* 3 *when divided by* 4 *exactly one of the following is true for every* $a \in \mathbb{Z}$

$$[a] = [0]$$
$$[a] = [1]$$
$$[a] = [2]$$
$$[a] = [3]$$

Therefore

$$\mathbb{Z}/\equiv_4 \ = \{[0], [1], [2], [3]\}$$

From this example, we can expect the following theorem to be true.

Theorem 4.10. *Let* a, b *and* n *be integers with* $n \geq 2$. *We have* $a \equiv_n b$ *if and only if* a *and* b *have the same remainder when divided by* n.

Let $n \geq 2$ be an integer. Following from our reasoning above, we have

$$\mathbb{Z}/\equiv_n \ = \{[0], [1], [2], [3], [4], \ldots, [n-1]\}$$

The equivalence classes of this relation correspond to the possible remainders one can obtain when dividing by n.

This piece of notation $\mathbb{Z}/\equiv_n$ is absolutely dreadful! Instead, let us define the following piece of notation for the set $\{[0], [1], [2], [3], [4], \ldots, [n-1]\}$

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

Aside 4.11. *In the study of abstract algebra, this set is sometimes denoted as: $\mathbb{Z}/n\mathbb{Z}$. There are good reasons for opting for $\mathbb{Z}/n\mathbb{Z}$ over $\mathbb{Z}_n$. These reasons are not relevant to this course, and so we stick with $\mathbb{Z}_n$.*

For a fixed $n \geq 2$, $\mathbb{Z}_n$ is a set. Its elements are sets. Each of these sets contain all of the integers that have the same remainder when divided by n.

Notice that $[n]$ does not appear as an element of $\mathbb{Z}_n$. When divided by n, the integer n has remainder 0. Thus

$$[n] = \{b \mid b \text{ has remainder 0 when divided by } n\} = [0]$$

As $[n] = [0]$, we could write

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

This is still the same set. Similarly we could write

$$\mathbb{Z}_n = \{[3], \ldots, [n-1], [n], [2n+1], [7n+2]\}$$

When we consider an element $[a] \in \mathbb{Z}_n$ we can always assume $a \in \{0, 1, \ldots, n-1\}$. If $a \notin \{0, 1, \ldots, n-1\}$, then we can replace a with its remainder when dividing by n.

4.3 Operations in $\mathbb{Z}_n$

Consider the following statements:

$$\begin{aligned} even + even &= even \\ even + odd &= odd \\ odd + odd &= even \end{aligned}$$

Without having to define anything carefully, we likely recognize these statements as telling us the parity of the outcome when we add together odd and even integers.

Consider the relation $\equiv_2$ and the set $\mathbb{Z}_2 = \{[0], [1]\}$. Recalling our notation from the previous section, the set $[0]$ is the set of all integers that have remainder 0 when divided by 2. Thus $[0]$ is the set of even integers. Similarly, $[1]$ is the set of odd integers.

With this in mind, we rewrite our statements above:

$$\begin{aligned} [0] + [0] &= [0] \\ [0] + [1] &= [1] \\ [1] + [1] &= [0] \end{aligned}$$

By analogy, we may also want to write:

$$\begin{aligned} [0] \cdot [0] &= [0] \\ [0] \cdot [1] &= [0] \\ [1] \cdot [1] &= [1] \end{aligned}$$

However these statements should give us pause as mathematicians. The notations $[0]$ and $[1]$ respectively refer to sets. What does it mean to *add together* or *multiply together* two sets? Before we can use an addition sign or multiplication sign for sets we have to define for our reader (and ourselves!) what we mean when we write $[0] + [0]$ or $[0] \cdot [0]$.

Aside 4.12. *Recall the definition of an operation.*

Definition 4.13. *Let A be a set. An operation on A is a function $f : A \times A \to A$.*

In order to make sense of what it means to add together two elements of $\mathbb{Z}_2$ we are defining an operation on $\mathbb{Z}_2$

Let $n \geq 2$ be an integer. We define the operations addition and multiplication on elements of $\mathbb{Z}_n$ as follows. For $[a], [b] \in \mathbb{Z}_n$ let

$$[a] + [b] = [a + b]$$
$$[a] \cdot [b] = [a \cdot b]$$

Aside 4.14. *Look at the statement*

$$[a] + [b] = [a + b]$$

The first appearance of the $+$ is our newly defined $+$ that we can use for elements of $\mathbb{Z}_n$. The second appearance of the $+$ is our usual meaning for addition integers. The $=$ is telling us that when we write $[a] + [b]$ we should take it to have the same meaning as having written $[a + b]$.

Aside 4.15. *We will use, without justification, that $+$ and $\cdot$ are associative and that the distributive laws work as we expect. Asking you to prove these things is an excellent midterm question.*

Looking back to our motivation for defining these operations, we can confirm

$$[0] + [0] = [0 + 0] = [0]$$
$$[0] + [1] = [0 + 1] = [1]$$
$$[1] + [1] = [1 + 1] = [2] = [0]$$

This choice for the definition of $+$ works as intended in $\mathbb{Z}_2$. One can quickly check that the choice for the definition of $\cdot$ works as intended for $\mathbb{Z}_2$.

Let us examine what these operations mean in, say, $\mathbb{Z}_6$. In $\mathbb{Z}_6$ we have the following operation tables.

+	[0]	[1]	[2]	[3]	[4]	[5]
[0]	[0]	[1]	[2]	[3]	[4]	[5]
[1]	[1]	[2]	[3]	[4]	[5]	[0]
[2]	[2]	[3]	[4]	[5]	[0]	[1]
[3]	[3]	[4]	[5]	[0]	[1]	[2]
[4]	[4]	[5]	[0]	[1]	[2]	[3]
[5]	[5]	[0]	[1]	[2]	[3]	[4]

•	[0]	[1]	[2]	[3]	[4]	[5]
[0]	[0]	[0]	[0]	[0]	[0]	[0]
[1]	[0]	[1]	[2]	[3]	[4]	[5]
[2]	[0]	[2]	[4]	[0]	[2]	[4]
[3]	[0]	[3]	[0]	[3]	[0]	[3]
[4]	[0]	[4]	[2]	[0]	[4]	[2]
[5]	[0]	[5]	[4]	[3]	[2]	[1]

Exercise 4.16. *Explain why these operation tables are symmetric.*

Look at the entry for $[4] + [3]$. By definition we have

$$[4] + [3] = [4 + 3] = [7] = [1]$$

As an analogy to addition in $\mathbb{Z}_2$, this seems to be telling us that if we take an integer that has remainder 4 when divided by 6 and add it to an integer that has remainder 3, then the result has remainder 1.

We can use Theorem 1.13 to confirm: Consider $a, b \in \mathbb{Z}$ so that a has remainder 4 when divided by 6 and b has remainder 3 when divided by 6. By Theorem 1.13, there exists unique integers q_a, q_b so that

$$a = 6q_a + 4$$
$$b = 6q_b + 3$$

Therefore

$$a + b = 6(q_a + q_b) + 7 = 6(1 + q_a + q_b) + 1$$

Therefore $a + b$ has remainder 1 when divided by 6. (Here the quotient is $1 + q_a + q_b$.)

There is an issue of representation to consider here. In $\mathbb{Z}_6$ we have $[1] + [1] = [2]$. However, we also have $[1] = [7]$. There is nothing in our definition of addition that guarantees

$$[1] + [1] = [7] + [7]$$

We confirm this with the following theorem.

Theorem 4.17. *Let a, b, a', b' and n be integers so that $n \geq 2$. If $[a] = [a']$ and $[b] = [b']$ in $\mathbb{Z}_n$, then*

1. *$[a] + [b] = [a'] + [b']$ and;*
2. *$[a] \cdot [b] = [a'] \cdot [b']$*

We explore the proof of this theorem on Assignment 3.

Aside 4.18. *Theorem 4.17 shows that our definitions of $+$ and $\cdot$ are* well-defined. *This means that it doesn't matter which representation we use for an element of $\mathbb{Z}_n$ when we perform addition and multiplication.*

4.4 Inverses in $\mathbb{Z}_n$

Now that we have a method to add and multiply in $\mathbb{Z}_n$ let us consider what other features of addition and multiplication we are familiar with also have analogues in $\mathbb{Z}_n$.

Consider the real number 0. Among all of the real numbers, the real number 0 is the only real number z that satisfies the following property for every real number x:

$$z + x = x = x + z$$

Looking at $\mathbb{Z}_n$ we see that $[0]$ plays this same role. Indeed, for any $[a] \in \mathbb{Z}_n$ we have

$$[0] + [a] = [0 + a] = [a] = [a + 0] = [a] + [0]$$

The real number 0 also plays an important role in thinking about negative real numbers. For every real number k_1 there exists a unique real number k_2 so that

$$k_1 + k_2 = k_2 + k_1 = 0$$

When $k_1 + k_2 = 0$ we write $k_2 = -k_1$. By analogy we can meaningfully define $-[a]$.

Definition 4.19. *Let a and n be integers with $n \geq 2$. The additive inverse of $[a]$, denoted $-[a]$, is the unique $[b] \in \mathbb{Z}_n$ so that*

$$[a] + [b] = [b] + [a] = [0]$$

Example 4.20. *Let $n = 8$. Find $-[3]$.*

We want to find $[b] \in \mathbb{Z}_8$ so that

$$[3] + [b] = [b] + [3] = [0]$$

By inspection we can see

$$[3] + [5] = [5] + [3] = [0]$$

Therefore $-[3] = [5]$ in $\mathbb{Z}_8$.

The statement of this definition suggests that when an element of $\mathbb{Z}_n$ has an additive inverse it is unique: However, without a proof this is not certain.

Theorem 4.21. *Let a and n be integers with $n \geq 2$. If there exists $[b], [b'] \in \mathbb{Z}_n$ so that $[a] + [b] = [b] + [a] = [0]$ and $[a] + [b'] = [b'] + [a] = [0]$, then $[b] = [b']$.*

Proof. Let a and n be integers with $n \geq 2$. Consider $[b], [b'] \in \mathbb{Z}_n$ so that

$$[a] + [b] = [b] + [a] = [0]$$

and

$$[a] + [b'] = [b'] + [a] = [0].$$

Without loss of generality, we may assume $a, b, b' \in \{0, 1, 2, \ldots, n-1\}$. Each of $[0], [1], \ldots, [n-1]$ are distinct. Thus to show $[b] = [b']$ it suffices to show $b = b'$.

Since $[a] + [b] = [0]$ we have $[a + b] = [0]$. Therefore $a + b$ has remainder 0 when divided by n. Thus $a + b$ is a multiple of n. Since $a, b \in \{0, 1, 2, \ldots, n - 1\}$ it must be that $a + b = n$. Therefore $b = n - a$.

A similar argument shows $b' = n - a$. Therefore $b = b'$ and so $[b] = [b']$. □

Though our theorem doesn't tell us that every element of $\mathbb{Z}_n$ has an additive inverse, the proof gives an idea to construct an additive inverse for $[a]$ in $\mathbb{Z}_n$. We notice

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

Therefore every element of $\mathbb{Z}_n$ has an additive inverse. Surprisingly, the same statement is not true when we define the notion of multiplicative inverse.

Just as 0 plays a special role with respect to addition of real numbers, so does 1 play the same role with respect to multiplication: Among all of the real numbers, the real number 1 is the only real number n that satisfies the following property for every real number x:

$$n \cdot x = x = x \cdot n$$

Looking at $\mathbb{Z}_n$ we see that $[1]$ plays this same role. Indeed, for any $[a] \in \mathbb{Z}_n$ we have

$$[1] \cdot [a] = [1 \cdot a] = [a] = [a \cdot 1] = [a] \cdot [1]$$

The real number 1 also plays an important role in thinking about reciprocals. For every real number $k_1 \neq 0$ there exists a unique real number k_2 so that

$$k_1 \cdot k_2 = k_2 \cdot k_1 = 1$$

When $k_1 \cdot k_2 = k_2 \cdot k_1 = 1$ we write $k_2 = k_1^{-1}$. By analogy we can meaningfully define $[a]^{-1}$.

Definition 4.22. *Let a and n be integers with $n \geq 2$. The multiplicative inverse of $[a]$, denoted $[a]^{-1}$ is the unique $[b] \in \mathbb{Z}_n$ so that*

$$[a] \cdot [b] = [b] \cdot [a] = [1]$$

Example 4.23. *Find $[4]^{-1}$ in $\mathbb{Z}_7$.*

We want to find $[b] \in \mathbb{Z}_7$ so that

$$[4] \cdot [b] = [b] \cdot [4] = [1]$$

By inspection we can see

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

Therefore $[4]^{-1} = [2]$ in $\mathbb{Z}_7$.

Just as we had for additive inverses, when multiplicative inverses exist they are unique.

Theorem 4.24. *Let a and n be integers with $n \geq 2$. If there exists $[b], [b'] \in \mathbb{Z}_n$ so that $[a] \cdot [b] = [b] \cdot [a] = [1]$ and $[a] \cdot [b'] = [b'] \cdot [a] = [1]$, then $[b] = [b']$.*

The proof of Theorem 4.21 told us that every element of $\mathbb{Z}_n$ has an additive inverse, the same is not true for multiplicative inverses. Recall the multiplication table for $\mathbb{Z}_6$:

•	[0]	[1]	[2]	[3]	[4]	[5]
[0]	[0]	[0]	[0]	[0]	[0]	[0]
[1]	[0]	[1]	[2]	[3]	[4]	[5]
[2]	[0]	[2]	[4]	[0]	[2]	[4]
[3]	[0]	[3]	[0]	[3]	[0]	[3]
[4]	[0]	[4]	[2]	[0]	[4]	[2]
[5]	[0]	[5]	[4]	[3]	[2]	[1]

Looking across the row for [3] we see that [1] does not appear! Thus [3] does not have a multiplicative inverse in $\mathbb{Z}_6$.

A natural thing for us to wonder then is for which pairs a and n, does $[a]$ have a multiplicative inverse in $\mathbb{Z}_n$?

We cast our minds back to the discussion at the start of this module. From our discussion at the start of the module, there existed integers u and v so that

$$au = (-v)n + 1$$

if and only if a and n were coprime

What is this telling us about inverses in $\mathbb{Z}_n$? If $au = (-v)n + 1$, then au has remainder 1 when divided by n. Thus $[au] = [1]$. And so

$$[1] = [a \cdot u] = [a] \cdot u] = [u] \cdot [a]$$

We see that $[u]$ is the multiplicative inverse of $[a]$. That is, $[a]^{-1} = [u]$

Theorem 4.25. *Let a and n be integers with $n \geq 2$. We have that $[a]^{-1}$ exists in $\mathbb{Z}_n$ if and only if $gcd(a, n) = 1$.*

Think about how we may use this fact to detect prime numbers. If every non-zero element of $\mathbb{Z}_n$ has a multiplicative inverse, then $gcd(a, n) = 1$ for each $a \in \{1, 2, 3, \ldots, n - 1\}$, then must n be prime? Is the converse true?

4.5 Further Exercises

1. Find, if it exists, $[6]^{-1}$ in $\mathbb{Z}_{17}$.
   (a) For which n does $[2]^{-1}$ exist in $\mathbb{Z}_n$?
   (b) For which n does $[6]^{-1}$ exist in $\mathbb{Z}_n$?
2. Let $n \geq 2$ be an integer and let $[a], [b] \in \mathbb{Z}_n$.
   (a) Show $[a] + [x] = [b]$ has a solution.
   (b) Show that if $gcd(a, n) = 1$, then $[a][x] = [b]$.

Further Things to Think About

If you have seen any material in linear algebra, how does the definition of A^{-1} compare to the definition of $[a]^{-1}$? Do an internet search for `ring (mathematics)`. For $\mathbb{Z}_n$, $+$ and $\cdot$ are operations that satisfy the *ring axioms*. For $n \times n$ matrices, $+$ and $\cdot$ are operations that satisfy the ring axioms. For $\mathbb{R}$, $+$ and $\cdot$ are operations that satisfy the ring axioms. Instead of studying $\mathbb{R}$ or $\mathbb{Z}_n$ or $n \times n$ matrices in isolation, instead one can all of these structures at once by studying those sets and operations that satisfy the ring axioms.

In this module we showed that the additive inverse is unique. An identical argument shows that the additive inverse is unique for $n \times n$ matrices. Why prove this fact twice? Instead we can prove that additive inverse are unique in any ring.

In semester 2, the Department of Mathematics and Statistics is offering a section of MATH362 – Rings and Fields. This course looks at $\mathbb{R}$, $n \times n$ matrices and $\mathbb{Z}_n$ as examples of rings.

---

5 Congruences Part II

Learning Incomes.

- *Understand the meaning of the notations and operations in* $\mathbb{Z}_n$
- *Understand the connection between multiplicative inverses and greatest common divisor.*

Learning Outcomes.

- *Understand the relationship between* $a \equiv_n b$ *and* $[a] = [b] \in \mathbb{Z}_n$
- *Be able to determine which linear congruence equations have solutions, and how to find their solutions.*
- *Understand the statement and proof of Fermat's Little Theorem*

Newly Defined Terms and Notation.

- *congruent modulo* n*, modulus*

Let $n \geq 2$ be an integer. Recall the following relation on $\mathbb{Z} \times \mathbb{Z}$ from Module 4

$$\equiv_n \ = \{(a, b) \mid n \text{ divides } b - a\}$$

From our work in Module 4 on equivalence relations, we should be comfortable with the following statement:

$$[a] = [b] \text{ if and only if } a \equiv_n b$$

Our work on operations in $\mathbb{Z}_n$ then lets us treat the $\equiv_n$ symbol similar to an equals sign. For example, let $n = 5$. Consider the following sequence of equalities.

$$\begin{aligned} [3] + [2] &= [5] \\ [3] + [2] + -[2] &= [0] + -[2] \\ [3] &= [0] + [3] \\ [3] &= [3] \end{aligned}$$

Using the statement above, we may equivalently write the following.

$$\begin{aligned} 3+2 &\equiv_5 5 \\ 3+2+-2 &\equiv_5 0+-2 \\ 3 &\equiv_5 0+3 \\ 3 &\equiv_5 3 \end{aligned}$$

If we have seen before the notion of *modular arithmetic*, this may all seem very familiar.

Definition 5.1. *Let a, b and n be integers with $n \geq 2$. We say a and b are congruent modulo n when $(a, b) \in \equiv_n$. When a and b are congruent modulo n we write $a \equiv_n b$.*

In this context refer to n as the modulus

Aside 5.2. *We could equivalently use the notation*

$$a \equiv b \mod n$$

to talk about a and b being congruent modulo n. I have chosen $a \equiv_n b$ over $a \equiv b \mod n$ so that we are always reminded of the relation that is providing the structure to the set of integers.

Recall Theorem 4.17 from Module 4

Theorem. *Let a, b, a', b' and n be integers so that $n \geq 2$. If $[a] = [a']$ and $[b] = [b']$ in $\mathbb{Z}_n$, then*

1. *$[a] + [b] = [a'] + [b']$ and;*
2. *$[a] \cdot [b] = [a'] \cdot [b']$*

We can restate this theorem using our notation of congruence.

Theorem 5.3. *Let a, b, a', b' and n be integers so that $n \geq 2$. If $a \equiv_n a'$ and $b \equiv_n b'$, then*

1. *$a + b \equiv_n a' + b'$ and;*
2. *$a \cdot b \equiv_n a' \cdot b'$*

In Module 4 we also saw the following theorem

Theorem. *Let a and n be integers with $n \geq 2$. We have that $[a]^{-1}$ exists in $\mathbb{Z}_n$ if and only if $gcd(a, n) = 1$.*

Just as we did with Theorem 4.17 we can restate this theorem in the language of congruence.

Theorem 5.4. *Let a and n be integers with $n \geq 2$. The congruence equation*

$$ax \equiv_n 1$$

has a solution if and only if $gcd(a, n) = 1$

From our many years of experience in mathematics courses, we are excellent at solving equations in one variable. We know how to manipulate to solve linear equations, we know how to use the atic formula to solve quadratic equations (... and we know when to turn to WolframAlpha to solve more complicated equations).

In this module we consider some techniques to solve congruence equations. However, we will quickly find that even some linear congruence equations our world is more complicated than we are used to.

Aside 5.5. *A note about notation – in the previous module we defined $\cdot$ as our symbol for multiplication in $\mathbb{Z}_n$. Here we switch at times to $\times$ to avoid a notation clash with "...".*

5.1 Linear Equations modulo n

Consider the linear equations

$$\begin{aligned} 3x &= 4 \\ x + 7 &= 9 \end{aligned}$$

Let us take a moment to recall in detail what we do when we solve these equations. For the first of these we multiply both sides by 3^{-1}, as we notice $3^{-1} \cdot 3 = 1$

$$\begin{aligned} 3^{-1} \cdot 3x &= 3^{-1} \cdot 4 \\ 1 \cdot x &= 3^{-1} \cdot 4 \\ x &= \frac{4}{3} \end{aligned}$$

For the second, we add -7 to both sides, as we notice $7 + (-7) = 0$

$$\begin{aligned} x + 7 + (-7) &= 9 + (-7) \\ x + 0 &= 2 \\ x &= 2 \end{aligned}$$

Aside 5.6. *Addition and multiplication are associative, so I have left out any brackets above. This trend will continue throughout the remainder of these notes.*

When we move to consider similar types of equations modulo n, the strategy for the second type of equation seems as if it will always work, regardless of the choice of modulus: additive inverses always exist in $\mathbb{Z}_n$. For example, consider the equation

$$x + 7 \equiv_{15} 2$$

We find that $-[7] = [8]$ in $\mathbb{Z}_{15}$. Thus $7 + 8 \equiv_{15} 0$. And so we find

$$\begin{aligned} x + 7 &\equiv_{15} 2 \\ x + 7 + 8 &\equiv_{15} 2 + 8 \\ x + 15 &\equiv_{15} 10 \\ x &\equiv_{15} 10 \end{aligned}$$

Since $x + 7 \equiv_{15} 2$ holds if and only if $x \equiv_{15} 10$, this congruence equation has infinitely many solutions. Choosing x so that x has remainder 10 when divided by 15 gives a solution to $x + 7 \equiv_{15} 2$. Thus the full set of solutions can be described as $\{15q + 10 \mid q \in \mathbb{Z}\}$

Back to our notation in $\mathbb{Z}_{15}$ this sequence of equalities above is the same as writing

$$\begin{aligned}
[x] + [7] &= [2] \\
[x] + [7] + (-[7]) &= 2 + (-[7]) \\
[x] &= [2] + [8] \\
[x] &= [2 + 8] = [10]
\end{aligned}$$

Again we see that the set of solutions is an entire equivalence class.

Aside 5.7. *It sure it convenient that addition in $\mathbb{Z}_n$ is associative!*

When we try to solve congruence equations of the form $a + x \equiv_n c$ our expected strategy works out as planned. However, when we try to solve congruence equations of the form $ax \equiv_n c$ we may already be able to see that our expected strategy may not work out so nicely. When we solved the equation $3x = 4$ above, we made use of the multiplicative inverse of 3 in $\mathbb{R}$. As we saw at the end of Module 4 the existence of a multiplicative inverse for $[a]$ in $\mathbb{Z}_n$ depends on the greatest common divisor of a and n.

Consider the following congruence equation

$$3x \equiv_7 4.$$

Since $gcd(3, 7) = 1$ we find that $[3]^{-1}$ exists in $\mathbb{Z}_7$.

To find this multiplicative inverse, $[u] = [3]^{-1}$, we must find integers u and v so that

$$3u = (-v)7 + 1$$

Rearranging, this is the same as finding an integer solution to the linear diophantine equation

$$3u + 7v = 1$$

Applying our methods from previous modules, we find a solution with $u = 5$ and $v = -1$. Thus $[3]^{-1} = [5]$ in $\mathbb{Z}_7$. We apply this fact to solve our equation in $\mathbb{Z}_7$

$$\begin{aligned}
[3] \cdot [x] &= [4] \\
[3]^{-1} \cdot [3] \cdot [x] &= [3]^{-1} \cdot [4] \\
[1] \cdot [x] &= [5] \cdot [4] \\
[x] &= [20] \\
[x] &= [6]
\end{aligned}$$

This gives the corresponding sequence of congruences:

$$\begin{aligned} 3 \cdot x &\equiv_7 4 \\ 5 \cdot 3 \cdot x &\equiv_7 5 \cdot 4 \\ 15 \cdot x &\equiv_7 5 \cdot 4 \\ 1 \cdot x &\equiv_7 20 \\ x &\equiv_7 6 \end{aligned}$$

Aside 5.8. *It sure is convenient that multiplication is associative in $\mathbb{Z}_n$!*

However, it is only the existence of the multiplicative inverse in $\mathbb{Z}_7$ that makes this method possible. Consider the following two congruence equations with modulus 8.

$$\begin{aligned} 4x &\equiv_8 3 \\ 2x &\equiv_8 4 \end{aligned}$$

By finding $gcd(2, 8)$ and $gcd(4, 8)$ we see that neither of [4] nor [2] have a multiplicative inverse in $\mathbb{Z}_8$. Thus we cannot use our method above to solve these equations. By checking all possibilities we find that the first equation has no solution and the second equation has a solution $x = 2$.

Following these examples we wonder for which triples a, c and n does $ax \equiv_n c$ have a solution?

Recall the meaning of the notation $ax \equiv_n c$. This means $(ax, c) \in \equiv_n$. Looking back at the definition of $\equiv_n$, this means that n divides $c - ax$. Looking back at the definition of divides, this means that there exists an integer y so that $ny = c - ax$. Rearranging yields

$$ax + ny = c$$

We are back to our linear diophantine equations!

Aside 5.9. *If we had defined the relation $a \equiv_n b$ to mean that a and b have the same remainder when divided by n, this argument above would have taken a few extra steps.*

Theorem 5.10. *Let a, c and n be integers with $n \geq 2$. The congruence equation*

$$ax \equiv_n c$$

has a solution if and only if $gcd(a, n)$ divides c.

5.2 Non-Linear Equations modulo n.

In Section 5.1 we saw that dealing with linear congruence equations can be done using tools we have already developed in previous modules. Let us turn now to considering higher-power equations modulo n. We start by considering such equations in $\mathbb{Z}_n$.

To consider such equations in $\mathbb{Z}_n$ we must first define the meaning of $[a]^k$ in $\mathbb{Z}_n$. Let a, k and n be integers with $k \geq 1$ and $n \geq 2$. We define the notation $[a]^k$ to mean

$$[a]^k = \underbrace{[a] \times [a] \times \cdots \times [a]}_{k \text{ times}}$$

For example $[2]^3 = [2] \times [2] \times [2]$. Notice $[a]^k = [a^k]$.

We start our investigation with quadratic equations by recalling our strategy for quadratic equations in $\mathbb{R}$ – applying the quadratic formula.

The equation $ax^2 + bx + c = 0$ has solutions

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Let us trying to apply this formula in $\mathbb{Z}_n$; perhaps we will find $[a] \cdot [x]^2 + [b] \cdot [x] + [c] = [0]$ has solutions:

$$[x] = \left(-[b] \pm \sqrt{[b]^2 - [4] \cdot [a] \cdot [c]}\right) \cdot ([2] \cdot [a])^{-1}$$

Let us try and see what happens. Consider the equation $[x]^2 + [3] \cdot [x] + [3] = [0]$ in $\mathbb{Z}_9$. Naively applying our equation above we find

$$[x] = \left(-[3] \pm \sqrt{[3]^2 - [4] \cdot [1] \cdot [3]}\right) \cdot ([2] \cdot [1])^{-1}$$

Simplifying in $\mathbb{Z}_9$ yields

$$[x] = \left(-[3] \pm \sqrt{[6]}\right) \cdot [5]$$

What might $\sqrt{[6]}$? mean in this context? By analogy from our study of real numbers, $\sqrt{[6]}$ should be an element $[x] \in \mathbb{Z}_9$ so that $[x]^2 = [6]$. Let us compute some squares in Z_9 to try and find such an element.

$[a]$	$[a]^2$
$[0]$	$[0]$
$[1]$	$[1]$
$[2]$	$[4]$
$[3]$	$[0]$
$[4]$	$[7]$
$[5]$	$[7]$
$[6]$	$[0]$
$[7]$	$[4]$
$[8]$	$[1]$

We see no element of $\mathbb{Z}_9$ so that $[x]^2 = [6]$.

Though every element of $\mathbb{R}$ has an square root, we do not see the same type of behaviour in $\mathbb{Z}_9$. Perhaps this is a matter of 9 being composite. Maybe the situation for square roots is better when n is prime? And so let us consider $n = 5$.

$[a]$	$[a]^2$
$[1]$	$[1]$
$[2]$	$[4]$
$[3]$	$[4]$
$[4]$	$[1]$

Again we can see that square roots don't always exist in $\mathbb{Z}_5$. For example we can find no element $[x] \in \mathbb{Z}_5$ so that $[x]^2 = [3]$.

Our hopes for finding easy conditions for solving quadratic congruence equations doesn't seem too likely at this point. We will return to this topic when we study quadratic residues in Module 12. And so let us turn our attention to some larger powers and leave quadratics behind. Sticking with our example of $n = 5$, let us think about taking powers of elements of $\mathbb{Z}_5$ and see what we find.

$[a]$	$[a]^2$	$[a]^3$	$[a]^4$	$[a]^5$
$[0]$	$[0]$	$[0]$	$[0]$	$[0]$
$[1]$	$[1]$	$[1]$	$[1]$	$[1]$
$[2]$	$[4]$	$[3]$	$[1]$	$[2]$
$[3]$	$[4]$	$[2]$	$[1]$	$[3]$
$[4]$	$[1]$	$[4]$	$[1]$	$[4]$

We notice some curious behaviour: For each $[a] \in \mathbb{Z}_5$ with $[a] \neq [0]$ we have $[a]^4 = [1]$ and $[a]^5 = [a]$.

Let us look at $n = 7$ to see if we see similar behaviour.

$[a]$	$[a]^2$	$[a]^3$	$[a]^4$	$[a]^5$	$[a]^6$	$[a]^7$
$[1]$	$[1]$	$[1]$	$[1]$	$[1]$	$[1]$	$[1]$
$[2]$	$[4]$	$[1]$	$[2]$	$[4]$	$[1]$	$[2]$
$[3]$	$[2]$	$[6]$	$[4]$	$[5]$	$[1]$	$[3]$
$[4]$	$[2]$	$[1]$	$[4]$	$[2]$	$[1]$	$[4]$
$[5]$	$[4]$	$[6]$	$[2]$	$[3]$	$[1]$	$[5]$
$[6]$	$[1]$	$[6]$	$[1]$	$[6]$	$[1]$	$[6]$

For each $[a] \in \mathbb{Z}_7$ with $[a] \neq [0]$ we have $[a]^{7-1} = [1]$ and $[a]^7 = [a]$.

Translating these observations into the language of congruence gives the following famous theorem.

Theorem 5.11 (Fermat's Little Theorem)**.** *If p is prime and a is an integer so that $gcd(a, p) = 1$, then*

$$a^{p-1} \equiv_p 1.$$

Remembering what our notation means, Fermat's Little Theorem is telling us that when a is not a multiple of p, then a^{p-1} has remainder 1 when divided by p.

To prove this statement we consider the corresponding statement in $\mathbb{Z}_n$:

Theorem. *If p is prime and a is an integer so that $gcd(a, p) = 1$, then in $\mathbb{Z}_p$ we have*

$$[a]^{p-1} = [1]$$

Aside 5.12. *There is a video companion for this proof. Go and watch it before you continue reading.*

Let $[p-1]!$ denote the product

$$[p-1]! = [1] \times [2] \times \cdots \times [p-2] \times [p-1]$$

To prove our theorem we will show

$$[p-1]! = [a]^{p-1} \times [p-1]!$$

To see why this is helpful first notice that $[p-1]!$ is equal to some element of $\mathbb{Z}_p$. We have $gcd\,(p, (p-1)!) = 1$ and so we have that $[p-1]!$ has a multiplicative inverse in $\mathbb{Z}_p$. Multiplying on both sides by this inverse yields $[1] = [a]^{p-1}$.

First let us consider expanding out the right side of this equation and reordering the terms.

$$[a]^{p-1} \times [p-1]! = [1] \times [a] \times [2] \times [a] \times \cdots \times [p-2] \times [a] \times [p-1] \times [a]$$

Notice $[k] \times [a] = [ka]$ for each $1 \leq k \leq p-1$. Thus to prove Fermat's Little Theorem, we must prove

$$[1] \times [2] \times \cdots \times [p-2] \times [p-1] = [a] \times [2a] \times \cdots \times [(p-2)a] \times [(p-1)a]$$

The left side of this equation is the product of all of the elements in the set

$$\{[1], [2], \ldots, [p-1]\}$$

The right side of this equation is the product of all of the elements in the set

$$\{[a], [2a], \ldots, [(p-1)a]\}$$

Thus to show

$$[1] \times [2] \times \cdots \times [p-2] \times [p-1] = [a] \times [2a] \times \cdots \times [(p-2)a] \times [(p-1)a]$$

we show

$$\{[1], [2], \ldots, [p-2], [p-1]\} = \{[a], [2a], \ldots, [(p-2)a], [(p-1)a]\}$$

Consider the sequence

$$[a], [2a], \ldots, [(p-2)a], [(p-1)a].$$

This sequence is exactly the sequence of elements of $\mathbb{Z}_p$ that appear in the $[a]$ row of the multiplication table for Z_p.

$\mathbb{Z}_p, \times$	$[1]$	$[2]$	$[3]$	$\ldots$	$[p-2]$	$[p-1]$
$[1]$	$[1]$	$[2]$	$[3]$	$\ldots$	$[p-2]$	$[p-1]$
$[2]$	$[1 \times 2]$	$[2 \times 2]$	$[3 \times 2]$	$\ldots$	$[(p-2) \times 2]$	$[(p-1) \times 2]$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$[a]$	$[1a]$	$[2a]$	$[3a]$	$\ldots$	$[(p-2) \times a]$	$[(p-1) \times a]$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$
$[p-1]$	$[1 \times (p-1)]$	$[2 \times (p-1)]$	$[3 \times (p-1)]$	$\ldots$	$[(p-2) \times (p-1)]$	$[(p-1) \times (p-1)]$

Since each element of $\{[a], [2a], \ldots, [(p-2)a], [(p-1)a]\}$ is a non-zero element of $\mathbb{Z}_p$, to show

$$\{[1], [2], \ldots, [p-2], [p-1]\} = \{[a], [2a], \ldots, [(p-2)a], [(p-1)a]\}$$

it suffices to show

$$\{[a], [2a], \ldots, [(p-2)a], [(p-1)a]\}$$

has exactly $p-1$ elements. We do this by showing that each element of this sequence is distinct.

Aside 5.13. *Notice that this is not the case when n is composite. For example, consider the* [2] *row of the multiplication table in* $\mathbb{Z}_6$*. Not every non-zero row of the multiplication table in* $\mathbb{Z}_6$ *contains every element of* $\mathbb{Z}_6$*.*

We proceed by contradiction to show each element of the sequence is distinct. If each element of the sequence

$$[1a], [2a], \ldots, [(p-2)a], [(p-1)a]$$

is not distinct, then there exists a pair of integers $1 \leq k_1, k_2 \leq p-1$ so that $k_1 \neq k_2$ and $[k_1 a] = [k_2 a]$. Thus we have

$$\begin{aligned} [k_1 a] &= [k_2 a] \\ [k_1] \times [a] &= [k_2] \times [a] \end{aligned}$$

Since $gcd(a, p) = 1$, we have that $[a]^{-1}$ exists in $\mathbb{Z}_p$. Therefore

$$\begin{aligned} [k_1 a] &= [k_2 a] \\ [k_1] \times [a] &= [k_2] \times [a] \\ [k_1] \times [a] \times [a]^{-1} &= [k_2] \times [a] \times [a]^{-1} \\ [k_1] \times [1] &= [k_2] \times [1] \\ [k_1] &= [k_2] \end{aligned}$$

This is a contradiction as $1 \leq k_1, k_2 \leq p-1$ and $k_1 \neq k_2$. Thus there are no repetitions in the sequence

$$[1a], [2a], \ldots, [(p-2)a], [(p-1)a].$$

Therefore

$$\{[1], [2], \ldots, [p-2], [p-1]\} = \{[1a], [2a], \ldots, [(p-2)a], [(p-1)a]\}$$

And so

$$[1] \times [2] \times \cdots \times [p-2] \times [p-1] = [1][a] \times [2][a] \times \cdots \times [p-2][a] \times [p-1][a].$$

Simplifying yields

$$[p-1]! = [a]^{p-1} \times [p-1]!$$

Since $gcd(p, (p-1)!) = 1$, we have that $[p-1]!$ has a multiplicative inverse in $\mathbb{Z}_p$. Therefore

$$[1] = [a]^{p-1}$$

Thus

$$1 \equiv_p a^{p-1}$$

5.3 Further Exercises

Exercises

1. For each of the equations below solve for x or show there is no solution.
   (a) $x + 9 \equiv_{74} 16$
   (b) $3x \equiv_{11} 8$
   (c) $3x + 9 \equiv_{12} 8$
   (d) $[7][x] + [9] = [8]$ in $\mathbb{Z}_{12}$
2. We didn't provide a full proof of Fermat's Little Theorem. In fact there were places where we glossed over some details. In the context of our proof sketch for Fermat's Little Theorem prove each of the following:
   (a) "... each element of $\{[a], [2a], \ldots, [(p-2)a], [(p-1)a]\}$ is a non-zero element of $\mathbb{Z}_p$"
   (b) $gcd(p, (p-1)!) = 1$
3. Let $n \geq 2$ be an odd integer so that $2^{n-1} \not\equiv_n 1$. Show n is composite.

Further Things to Think About

Look back at our proof for Fermat's Little Theorem and notice how rarely we actually cared about divisibility and remainders. The only place that divisibility explicitly came into play was when we needed to show that $[p-1]!$ had a multiplicative inverse in $\mathbb{Z}_p$.

Let p be an integer and let G be a set with p elements so that we can define an associative operation $\cdot$ on G for which each element of G has a multiplicative inverse. Since multiplicative inverses exist, there exists an element $u \in G$ so that $g \cdot u = g \cdot u = g$ for each $g \in G$. (Compare u to [1], 1 and the matrix I_n) Can you use the proof of Fermat's Little Theorem to show

$$g^{p-1} = u$$

for any $g \in G$? If not, what other properties do you need to assume for $\cdot$ so that the proof works out?

---

6 Congruences Part III

Learning Incomes.

- *Understand the statement and proof of Fermat's Little Theorem*
- *Understand the relationship between* $[a] = [b]$ *in* $\mathbb{Z}_n$ *and* $a \equiv_n b$.
- *Understand the the solution to Question 4 on Assignment 2.*

Learning Outcomes.

- *Be able to use the Base-2 Primality Test to show that an integer is composite.*
- *Understand the definition of a pseudo-prime and a Carmichael number.*
- *Understand the proof that there are infinitely many Carmichael numbers.*
- *Understand the definition and how to compute* $\phi(n)$.
- *Be able to determine the elements in* U_n *for a fixed integer* $n \geq 2$.
- *Understand why* $|U_n| = \phi(n)$.
- *Understand the statement and proof of Euler's Theorem.*

Newly Defined Terms and Notation. *base-2 primality test, pseudo-prime, Carmichael number, Euler's phi function,* $\phi(n)$*, group of units of* $\mathbb{Z}_n$*,* U_n

Recall Fermat's Little Theorem from Module 5.

Theorem (Fermat's Little Theorem)**.** *If* p *is prime and* a *is an integer so that* $gcd(a, p) = 1$*, then*

$$a^{p-1} \equiv_p 1.$$

In this module we consider two different topics following from Fermat's Little Theorem.

The first of these topics is Primality Testing. In Module 3 we studied prime numbers, but saw little in the way of tools of testing if a particular number is prime. Fermat's Little Theorem, and some related results, give us some tools to help us determine if a particular positive integer is composite.

The second of these topics is a generalization of Fermat's Little Theorem. Recall the following theorem from Module 4.

Theorem. *Let* a *and* n *be integers with* $n \geq 2$*. We have that* $[a]^{-1}$ *exists in* $\mathbb{Z}_n$ *if and only if* $gcd(a, n) = 1$.

In our proof of Fermat's Little Theorem the need for p to be prime was so that the non-zero elements of $\mathbb{Z}_p$ had a multiplicative inverse. More specifically, we needed $[a]$ and $[p-1]!$ to have a multiplicative inverse for various steps in our proof to work out. For an arbitrary modulus n, it is only those integers a so that $gcd(a, n) = 1$ that have a multiplicative inverse. By restricting our attention to the set of such integers we find a generalization of Fermat's Little Theorem for non-prime moduli.

6.1 Primality Testing with Fermat's Little Theorem

Recall again the statement of Fermat's Little Theorem:

Theorem (Fermat's Little Theorem)**.** *If p is prime and a is an integer so that $gcd(a, p) = 1$, then*

$$a^{p-1} \equiv_n 1$$

Let n be an odd integer. Since $gcd(2, n) = 1$, knowing the remainder when 2^{n-1} is divided by n may be enough to know that n is composite. More specifically, if $2^{n-1} \not\equiv_n 1$, then it must be that n is composite! If n were prime, then we would have $2^{n-1} \equiv_n 1$.

Corollary 6.1. *Let $n \geq 3$ be an odd integer. If*

$$2^{n-1} \not\equiv_n 1,$$

then n is composite.

Example 6.2. *Show $n = 25$ is composite using Corollary 6.1.*

We show $1 \not\equiv_{25} 2^{24}$.

Notice $24 = 6 \cdot 4$ and $2^6 \equiv_{25} 14$. Therefore

$$\begin{aligned} 2^{24} &\equiv_{25} (2^6)^4 \\ &\equiv_{25} 14^4 \end{aligned}$$

We have $14 = 2 \cdot 7$, and so

$$\begin{aligned} 2^{24} &\equiv_{25} 14^4 \\ &\equiv_{25} 7^4 \cdot 2^4 \\ &\equiv_{25} (7^2)^2 \cdot 2^4 \\ &\equiv_{25} (49)^2 \cdot 16 \\ &\equiv_{25} 24^2 \cdot 16 \\ &\equiv_{25} 2^2 \cdot 12^2 \cdot 16 \\ &\equiv_{25} 2^2 \cdot 144 \cdot 16 \\ &\equiv_{25} 4 \cdot 19 \cdot 16 \\ &\equiv_{25} 76 \cdot 16 \\ &\equiv_{25} 1 \cdot 16 \\ &\equiv_{25} 16 \end{aligned}$$

Since $2^{24} \not\equiv_{25} 1$*, it follows that* 25 *is composite.*

Aside 6.3. *Notice that we showed* 25 *was composite without actually finding any factors of* 25*. In terms of a computational process, for large* n*, it is significantly faster to the remainder when* 2^{n-1} *is divided by* n *than it is to factor* n*.*

Corollary 6.1 is called the Base-2 Primality Test. This name is a bit of a misnomer. Using Corollary 6.1 alone, there is no way to determine if an integer is prime. Corollary 6.1 gives us a sufficient condition for an integer to be composite. But this Corollary is not an *if and only if* theorem. We cannot use it to determine if a number is not composite. Consider the following example with $n = 341$.

We find the remainder when 2^{340} is divided by 341

Notice $340 = 10 \times 34$. And so

$$2^{340} \equiv_{341} (2^{10})^{34}$$

We have $2^{10} = 1024$. The integer 1024 has remainder 1 when divided by 341. Therefore $[1024] = [1]$ in $\mathbb{Z}_{341}$. Therefore

$$\begin{aligned} 2^{340} &\equiv_{341} 2^{340} \\ &\equiv_{341} (2^{10})^{34} \\ &\equiv_{341} (1024)^{34} \\ &\equiv_{341} (1)^{34} \\ &\equiv_{341} 1 \end{aligned}$$

From our work we see $2^{340} \equiv_{341} 1$. Does this tell us that 341 is prime? No! Corollary 6.1 tells us nothing about whether or not n is prime in the even that $2^{n-1} \equiv_n 1$. (In fact, $341 = 11 \times 31$)

Definition 6.4. *Let* $n \geq 2$ *be an integer. We say* n *is a pseudo-prime when* n *is composite and* $2^{n-1} \equiv_n 1$*.*

Unfortunately there are infinitely many pseudo-primes.

Theorem 6.5. *Let* $n \geq 3$ *be an odd integer. If* n *is a pseudo-prime, then* $2^n - 1$ *is a pseudo-prime.*

Proof. Let n be a pseudo-prime. By definition n is composite. Thus there exists integers $2 \leq a, b \leq n-1$ so that $n = ab$. By Question 4 on Assignment 2, we have that $2^n - 1$ is composite. Let $n_1 = 2^n - 1$. We show n_1 is a pseudo-prime by showing 2^{n_1-1} has remainder 1 when divided by n_1.

From our work on Question 4 on Assignment 2, recall the factorization of the polynomial $x^t - 1$:

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

To show $n_1 = 2^n - 1$ is a pseudo-prime we must show

$$2^{2^n-2} \equiv_{2^n-1} 1$$

Since n is a pseudo-prime we have

$$2^{n-1} \equiv_n 1$$

Multiplying both sides by 2 yields

$$2^n \equiv_n 2$$

Therefore 2^n has remainder 2 when divided by n. Thus there exists an integer q such that

$$2^n = nq + 2$$

Consider now the factorization of the polynomial above with $x = 2^n$ and $t = q$

$$2^{nq} - 1 = (2^n - 1)(2^{n(q-1)} + 2^{n(q-2)} + \cdots + 2^n + 1)$$

Rearranging yields

$$2^{nq} = (2^n - 1)(2^{n(q-1)} + 2^{n(q-2)} + \cdots + 2^n + 1) + 1$$

Thus 2^{nq} has remainder 1 when divided by divided by $2^n - 1$.

That is,

$$2^{nq} \equiv_{2^n-1} 1$$

Notice $nq = 2^n - 2$. Therefore

$$2^{2^n-2} \equiv_{2^n-1} 1$$

□

Corollary 6.6. *There are infinitely many pseudo-primes.*

Proof. From our work above we have that $n_0 = 341$ is a pseudo prime. By Theorem 6.5 every integer of the sequence

$$n_0, n_1, n_2, \ldots$$

is a pseudo-prime where $n_i = 2^{n_{i-1}} - 1$ for each $i \geq 1$ □

The Base-2 Primality Test is essentially the contrapositive of Fermat's Little Theorem for the case $a = 2$. However, there is nothing particularly special about 2, other than the fact that $gcd(2, n) = 1$ for every odd integer n. And so, if the Base-2 Primality Test fails to conclude that an integer is composite, we can then try another base.

Corollary 6.7. *Let a and n be integers so that $n \geq 2$ be and $gcd(a, n) = 1$. If*

$$a^{n-1} \not\equiv_n 1,$$

then n is composite.

Let us return to our example of $n = 341$. Since $gcd(3, 341) = 1$ we can show 341 is composite by showing

$$3^{341-1} \not\equiv_{341} 1.$$

Using a computer we find

$$3^{340} \equiv_{341} 56.$$

And so by Corollary 6.7 we find that 341 is composite.

Using Corollary 6.7 we propose the following reasonable-seeming procedure for determining if an integer n is composite:

For each $a \in [1, n]$ so that $gcd(a, n) = 1$ find the remainder r_a when a^{n-1} is divided by n.

If we find ever some a so that $r_a \neq 1$, then by Corollary 6.7, then n is necessarily composite.

Consider the following two questions

1. How long will this procedure take?
2. If n is composite will this procedure always find some a so that $a^{n-1} \equiv_n 1$?

To answer the first question we must answer the question:

Given an integer $n \geq 2$, how many integers in the range $[1, n]$ are coprime with n?

We answer this question in the following section. We answer the second question with an unfortunate counterexample, $n = 561$. There are 320 integers a so that $gcd(a, 561) = 1$ and $a \in [1, n]$. For each such a we can compute to find $a^{560} \equiv_{561} 1$. However $561 = 3 \cdot 11 \cdot 17$

Integers that spoil our proposed procedure for testing if an integer is composite are called Carmichael numbers. These numbers are named for Robert Carmichael who is said to have been the first to find an example of such a number in 1910. The existence of these numbers was theorized thirty years earlier by Alwin Korselt. It is only recently that researchers showed there are infinitely many Carmichael numbers.[1]

Aside 6.8. *Most practical algorithms for checking in an integer is prime proceed though a number of sub-algorithms to deal with particular cases. For example, we can begin to determine if an integer n is composite by first performing the Base-2 Test. If the Base-2 Test is inconclusive, then n is either prime or a pseudo-prime. When then proceed with the Base-a Test (i.e., Corollary 6.7) for all a in the range $[1, n]$ with $gcd(a, n) = 1$. If each of these tests is inconclusive them we can conclude that n is either prime or a Carmichael Number. We then proceed with subsequent tests to further narrow the likelihood that n is composite.*

[1] Alford, William R., Andrew Granville, and Carl Pomerance (**1994**) There are infinitely many Carmichael numbers *Annals of Mathematics* 139:3 703-722.

6.2 Euler's Phi Function

We return now to our first question – *given a positive integer n, how many integers in the range $[1, n]$ are coprime with n.* Recall the following theorem from Module 4:

Theorem. *Let a and n be integers with $n \geq 2$. We have that $[a]^{-1}$ exists in $\mathbb{Z}_n$ if and only if $gcd(a, n) = 1$.*

Thus the answer to our question will also answer the question – *How many elements of $\mathbb{Z}_n$ have a multiplicative inverse?*

Definition 6.9. *Let n be a positive integer. We denote by $\phi(n)$ the quantity of integers in the range $[1, n]$ that are coprime with n. We call ϕ the Euler phi function.*

For example, consider $n = 10$. There are 4 integers in the range $[1, 10]$ that are coprime with 10. And so we have $\phi(10) = 4$. To try and find an explicit formula for $\phi(n)$ we begin with integers that are powers of primes.

Consider the integer $5^3 = 125$. The positive divisors of 5^3 are: $1, 5, 5^2$, and 5^3. Consider some integer $a \in [1, 5^3]$. Notice that $gcd(a, 5^3) \in \{5, 5^2, 5^3\}$ if and only if a is a multiple of 5. And so $gcd(a, 5^3) = 1$ if and only if a is not a multiple of 5. There are 5^3 integers in the range $[1, 5^3]$ of which $1/5$ of them are a multiple of 5. (Starting from 1, every fifth number is a multiple of 5.) Therefore there are $5^3\left(1 - \frac{1}{5}\right)$ integers in the range $[1, 5^3]$ that are coprime with 5^3. And so

$$\phi(5^3) = 5^3\left(1 - \frac{1}{5}\right) = 100$$

Aside 6.10. *Why have we written this as: $5^3\left(1 - \frac{1}{5}\right)$ instead of $4 \cdot 5^2$? Expressing it as $5^3\left(1 - \frac{1}{5}\right)$ represents the work we did to get to the result. This will be helpful for us as we generalize.*

In this argument the only thing special about $n = 5^3$ is that 5 is prime. And so we have the following lemma.

Lemma 6.11. *Let p be a prime and let e be a positive integer. We have*

$$\phi(p^e) = p^e\left(1 - \frac{1}{p}\right)$$

Proof. Let p and e be integers so that p is prime and $e \geq 1$ The positive divisors of p^e are: $1, p, p^2, \ldots, p^e$. Consider some integer $a \in [1, p^e]$. Notice that $gcd(a, p^e) \in \{p, p^2, p^3, \ldots, p^e\}$

if and only if a is a multiple of p. And so $gcd(a, p^e) = 1$ if and only if a is not a multiple of p. There are p^e integers in the range $[1, p^e]$ of which $1/p$ of them are a multiple of p. Therefore there are $p^e\left(1 - \frac{1}{p}\right)$ integers in the range $[1, p^e]$ that are coprime with p^e. And so

$$\phi(p^e) = p^e\left(1 - \frac{1}{p}\right)$$

□

Along with Lemma 6.11, we require one more result to aid our computation of $\phi(n)$ for arbitrary n. Let us look at the integer $n = 28$ and arrange all of the integers in the range $[1, 28]$ in a 4×7 grid.

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16
17	18	19	20
21	22	23	24
25	26	27	28

Since $28 = 4 \times 7$, the integers that are not coprime with 28 are those that are a multiple of 2 or of 7. Notice that integers in the same column are congruent modulo 4. And in fact each column corresponds to a different congruence class modulo 4. There are $\phi(4) = 2$ columns that contain integers that are not a multiple of 4. In such row we can notice there are $\phi(7) = 6$ integers that are not a multiple of 7. Thus $\phi(28) = \phi(4) \cdot \phi(7)$.

It takes a little bit more work to show that this observation holds whenever we can express n as the product of 2 coprime factors (here 4 and 7 are coprime.) We will omit this work, as the proof is technical and not particularly insightful.

Lemma 6.12. *Let m_1 and m_2 be positive integers so that $gcd(m_1, m_2) = 1$. We have*

$$\phi(m_1 m_2) = \phi(m_1) \cdot \phi(m_2)$$

Together Lemmas 6.11 and 6.12 are enough to compute $\phi(n)$ for any positive integer n, given that we have its prime power factorization.

Consider the integer $n = 4900 = 2^2 \cdot 5^2 \cdot 7^2$. Notice $gcd(2^2 \cdot 5^2, 7^2) = 1$. By Lemma 6.12 we have

$$\phi(n) = \phi(2^2 \cdot 5^2) \cdot \phi(7^2)$$

We have $gcd(2^2, 5^2) = 1$ and so again by Lemma 6.12 we have

$$\phi(n) = \phi(2^2) \cdot \phi(5^2) \cdot \phi(7^2)$$

By Lemma 6.11 we have

$$\begin{aligned}
\phi(n) &= \phi(2^2) \cdot \phi(5^2) \cdot \phi(7^2) \\
&= 2^2 \left(1 - \frac{1}{2}\right) \cdot 5^2 \left(1 - \frac{1}{5}\right) \cdot 7^2 \left(1 - \frac{1}{7}\right) \\
&= 2^2 \cdot 5^2 \cdot 7^2 \left(1 - \frac{1}{2}\right) \left(1 - \frac{1}{5}\right) \left(1 - \frac{1}{7}\right) \\
&= 1680
\end{aligned}$$

There are 1680 integers in the range $[1, 4900]$ that are coprime with 4900.

Combining Lemmas 6.11 and 6.12 gives us the tools we need to prove the following theorem by induction.

Theorem 6.13. *Let $n \geq 2$ be an integer with prime power factorization:*

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

We have

$$\phi(n) = n \cdot \prod_{i=1}^{k} \left(1 - \frac{1}{p_i}\right)$$

Proof. We proceed by induction on k. We notice that the case $k = 1$ is exactly the statement of Lemma 6.11.

Consider the case now where the prime-power factorization of n has t primes.

$$n = p_1^{e_1} p_2^{e_2} \cdots p_t^{e_t}$$

By inspection we see

$$gcd(p_1^{e_1} p_2^{e_2} \cdots p_{t-1}^{e_{t-1}}, p_t^{e_t}) = 1$$

And so by Lemma 6.12 we have

$$\phi(n) = \phi(p_1^{e_1} p_2^{e_2} \cdots p_{t-1}^{e_{t-1}}) \cdot \phi(p_t^{e_t})$$

By induction we have

$$\phi(p_1^{e_1} p_2^{e_2} \cdots p_{t-1}^{e_{t-1}}) = (p_1^{e_1} p_2^{e_2} \cdots p_{t-1}^{e_{t-1}}) \left(1 - \frac{1}{p_1}\right) \left(1 - \frac{1}{p_2}\right) \cdots \left(1 - \frac{1}{p_{t-1}}\right)$$

By Lemma 6.11 we have $\phi(p_t^{e_t}) = p_t^{e_t}\left(1 - \frac{1}{p_t}\right)$. Therefore

$$\phi(n) = n \cdot \prod_{i=1}^{t}\left(1 - \frac{1}{p_i}\right)$$

□

Aside 6.14. *Without thinking more about it, it is unexpected that* $n \cdot \prod_{i=1}^{k}\left(1 - \frac{1}{p_i}\right)$ *is even an integer, never mind a meaningful integer! Look at all of those fractions in the product!*

Example 6.15. *Compute* $\phi(6136)$*. From Module 3 we had*

$$6136 = 2^3 \times 13 \times 59$$

From Theorem 6.13 and Lemma 6.11 we have

$$\begin{aligned}\phi(6136) &= 6136 \cdot \prod_{i=1}^{3}\left(1 - \frac{1}{p_i}\right)\\ &= 6136 \cdot \left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{13}\right)\left(1 - \frac{1}{59}\right)\\ &= 2784\end{aligned}$$

Thinking back to our motivation for Module 6, we had intended to find a generalization for Fermat's Little Theorem. Let us work our way back that goal. Our proof of Fermat's Little Theorem depends on elements of $\mathbb{Z}_p$ having a multiplicative inverse. And so in our task to generalize Fermat's Little Theorem for arbitrary moduli, let us restrict our attention to those elements of $\mathbb{Z}_n$ that have a multiplicative inverse.

Definition 6.16. *Let* $n \geq 2$ *be an integer. We let* U_n *denote the set of elements of* $\mathbb{Z}_n$ *that have a multiplicative inverse. We call* U_n *the group of units of* $\mathbb{Z}_n$*.*

From our work above we can see $|U_n| = \phi(n)$.

Let us find the group of units of $\mathbb{Z}_{15}$. To do so we find all $a \in [1, 15]$ such that $gcd(a, 15) = 1$. We find

$$U_{15} = \{[1], [2], [4], [7], [8], [11], [13], [14]\}$$

We begin our study of the group of units of $\mathbb{Z}_n$ by considering the multiplication table of U_{15}.

$U_{15}, \cdot$	[1]	[2]	[4]	[7]	[8]	[11]	[13]	[14]
[1]	[1]	[2]	[4]	[7]	[8]	[11]	[13]	[14]
[2]	[2]	[4]	[8]	[14]	[1]	[7]	[1]	[13]
[4]	[4]	[8]	[1]	[13]	[2]	[14]	[7]	[11]
[7]	[7]	[14]	[13]	[4]	[11]	[2]	[1]	[8]
[8]	[8]	[1]	[2]	[11]	[4]	[13]	[14]	[7]
[11]	[11]	[7]	[14]	[2]	[13]	[1]	[8]	[4]
[13]	[13]	[11]	[7]	[1]	[14]	[8]	[4]	[2]
[14]	[14]	[13]	[11]	[8]	[7]	[4]	[2]	[1]

Given that all of the elements of U_{15} are elements of $\mathbb{Z}_{15}$ we shouldn't be surprised that all of the entries in this multiplication table are elements of $\mathbb{Z}_{15}$. However our non-surprise may turn to surprise when we notice that all of the entries of the multiplication table of U_{15} are themselves elements of U_{15}.

Lemma 6.17. *Let $n \geq 2$ be an integer. If $[a] \in U_n$ and $[b] \in U_n$, then $[a] \cdot [b] \in U_n$*

Proof. Let $n \geq 2$ be an integer. Consider $[a], [b] \in U_n$. Recall $[a] \cdot [b] = [ab]$. To show $[a] \cdot [b] \in U_n$ we show $[ab]$ has a multiplicative inverse in $\mathbb{Z}_n$.

Since $[a] \in U_n$ then by definition we have that $[a]^{-1}$ exists in $\mathbb{Z}_n$. Similarly, $[b]^{-1}$ exists in $\mathbb{Z}_n$. We claim $[ab]^{-1} = [b]^{-1} \cdot [a]^{-1}$.

We compute

$$\begin{aligned}
[ab] \cdot [b]^{-1} \cdot [a]^{-1} &= [a] \cdot [b] \cdot [b]^{-1} \cdot [a]^{-1} \\
&= [a] \cdot [1] \cdot [a]^{-1} \\
&= [a] \cdot [a]^{-1} \\
&= [1]
\end{aligned}$$

Thus $[ab]^{-1} = [b]^{-1} \cdot [a]^{-1}$. Since $[ab]^{-1}$ exists in $\mathbb{Z}_n$, by definition we have $[ab] \in U_n$. And so $[a] \cdot [b] \in U_n$. □

Aside 6.18. *For $[a] \cdot [b] \in U_n$ we have $([a] \cdot [b])^{-1} = [b]^{-1} \cdot [a]^{-1}$. Notice how our method of proof here is identical to argument that gives $(AB)^{-1}$ when A and B are invertible matrices.*

Our stage is now set to generalize Fermat's Little Theorem for arbitrary moduli. Recall again the $\mathbb{Z}_p$ version of Fermat's Little Theorem:

Theorem. *If p is prime and a is an integer so that $gcd(a, p) = 1$, then in $\mathbb{Z}_p$ we have*

$$[a]^{p-1} = [1]$$

For a prime p, consider the set U_p. We have $U_p = \{[1], [2], \ldots, [p-1]\}$ and $\phi(p) = p - 1$. Thus we can equivalently state this theorem as

Theorem. *If n is prime, then $[a]^{\phi(n)} = [1]$ for each $[a] \in U_n$.*

Let us see what we find when n is not prime. Consider the case $n = 9$. We have

$$U_9 = \{[1], [2], [4], [5], [7], [8]\}.$$

We compute powers of elements of U_9.

$[a]$	$[a]^2$	$[a]^3$	$[a]^4$	$[a]^5$	$[a]^6$	$[a]^7$
[1]	[1]	[1]	[1]	[1]	[1]	[1]
[2]	[4]	[8]	[7]	[5]	[1]	[2]
[4]	[7]	[1]	[4]	[7]	[1]	[4]
[5]	[7]	[8]	[4]	[2]	[1]	[5]
[7]	[4]	[1]	[7]	[4]	[1]	[7]
[8]	[1]	[8]	[1]	[8]	[1]	[8]

Wow! We have $\phi(9) = 6$ and we see $[a]^{\phi(9)} = [1]$ for each $[a] \in U_9$!

Aside 6.19. *Okay... maybe this shouldn't be surprising given the direction of these notes. But you can imagine how surprising this would be for the first researchers that noticed it.*

Theorem 6.20. *For every integer $n \geq 2$ we have then $[a]^{\phi(n)} = [1]$ for each $[a] \in U_n$.*

To prove this statement let us remember our proof strategy for Fermat's Little Theorem. For every prime p and every $[a] \in \{[1], [2], \ldots, [p-1]\}$ we showed

$$[p-1]! = [a]^{p-1}[p-1]!$$

Recall $[p-1]!$ is the product of all of the elements of the set

$$\{[0], [1], \ldots, [p-1]\}$$

With our new notation, we notice that $[p-1]!$ is the product of all of the elements in the set U_p.

Recall $[a]^{p-1}[p-1]!$ is the product of all of entries in the $[a]$ row of the multiplication table in $\mathbb{Z}_p$. With our new notation, we notice that

$$[1a], [2a], \ldots, [(p-2)a], [(p-1)a].$$

is the sequence of entries in the $[a]$ row of the multiplication table of elements of U_p. With these thoughts in mind we turn to our generalization of Fermat's Little Theorem.

Proof. Let a and n be integers so that $n \geq 2$ and $gcd(a, n) = 1$. Let $U_n = \{[u_1], [u_2], \ldots, [u_{\phi(n)}]\}$. Notice that since $[a] \in U_n$ we have $[a] = [u_i]$ for some $1 \leq i \leq \phi(n)$. Let $[f]$ be the product of the set of all elements in U_n. To show $[a]^{\phi(n)} = [1]$ we first show

$$[f] = [f] \times [a]^{\phi(n)}$$

Expanding and reordering on the right side we have

$$\begin{aligned} [f] \times [a]^{\phi(n)} &= [u_1] \times [a] \times [u_2] \times [a] \times \cdots \times [u_{\phi(n)}] \times [a] \\ &= [u_1 a] \times [u_2 a] \times \cdots \times [u_{\phi(n)} a] \end{aligned}$$

Consider the $[a]$ row of the multiplication table of U_n:

$U_n, \times$	$[u_1]$	$[u_2]$	$[u_3]$	$\ldots$	$[u_{\phi(n)-1}]$	$[u_{\phi(n)}]$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$[a]$	$[u_1 a]$	$[u_2 a]$	$[u_3 a]$	$\ldots$	$[u_{\phi(n)-1} a]$	$[u_{\phi(n)} a]$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$

We then observe that $[f] \times [a]^{\phi(n)}$ is the product of all of the entries in this row. Let A be the set of entries of this row. To show $[f] = [f] \times [a]^{\phi(n)}$ we show $A = U_n$.

By Lemma 6.17 we have that every element of A is an element of U_n. Since U_n has $\phi(n)$ elements, to show $A = U_n$ we show A has $\phi(n)$ elements. To do this, we show that each element in the sequence

$$[(u_1)a], [(u_2)a], \ldots, [(u_{\phi(n)})a]$$

is distinct.

If each of the elements of this sequence is not distinct, then there exists $1 \leq i, j \leq \phi(n)$ so that $u_i \neq u_j$ and $[(u_i)a] = [(u_j)a]$. Since $[a] \in U_n$ we have that $[a]^{-1}$ exists in $\mathbb{Z}_n$. Thus

$$\begin{aligned}
[u_i a] &= [u_j a] \\
[u_i] \times [a] &= [u_j] \times [a] \\
[u_i] \times [a] \times [a]^{-1} &= [u_j] \times [a] \times [a]^{-1} \\
[u_i] \times [1] &= [u_j] \times [1] \\
[u_i] &= [u_j]
\end{aligned}$$

This is a contradiction; we assumed $u_i \neq u_j$. Therefore $A = U_n$. This then implies

$$[f] = [f] \times [a]^{\phi(n)}$$

By Lemma 6.17, we have $[f] \in U_n$. Therefore $[f]^{-1}$ exists in $\mathbb{Z}_n$. And so

$$\begin{aligned}
[f] &= [f] \times [a]^{\phi(n)} \\
[f]^{-1} \cdot [f] &= [f]^{-1} \cdot [f] \times [a]^{\phi(n)} \\
[1] &= [a]^{\phi(n)}
\end{aligned}$$

This completes the proof.

□

Equivalently, we may state Theorem 6.20 as:

Theorem 6.21 (Euler's Theorem)**.** *Let $n \geq 2$ be an integer. For every integer a so that $gcd(a, n) = 1$ we have*

$$a^{\phi(n)} \equiv_n 1$$

In the coming module we look at applications of number theory in cryptography. Euler's Theorem will play a key role.

Aside 6.22. *Euler was a Swiss mathematician in the 18th century. Undoubtedly Euler had no idea that his work would one day be the basis for internet security in the 21st century!*

6.3 Further Exercises

1. Apply the result of Theorem 6.5 to find two pseudo-primes other than $n_0 = 341$. (These numbers are very big!)
2. Compute $\phi(198)$.
3. Give an example of positive integers m_1 and m_2 such that $\phi(m_1 m_2) \neq \phi(m_1) \cdot \phi(m_2)$.
4. For a fixed positive integer n, we can interpret $\phi(n)/n$ to be the probability that a randomly chosen integer in the range $[1, n]$ is coprime with n. Show that for every $\epsilon > 0$ there exists n so that $\phi(n)/n > 1 - \epsilon$.

---

7 An Introduction to Public Key Cryptography

Learning Incomes.

- *Understand the statement of Euler's Theorem*
- *Be able to compute $\phi(n)$ given the prime power factorization of n.*
- *Be able to find the remainder when a very large power is divided by a much smaller integer.*
- *Be comfortable switching between $a \equiv_n b$ and $[a] = [b]$ notations.*

Learning Outcomes.

- *Be able to construct a shared-key using the Diffie-Hellman-Merkle key sharing scheme*
- *Be able to construct public and private keys using the RSA scheme.*
- *Be able to encrypt and decrypt data using the RSA scheme.*

Newly Defined Terms and Notation. *primitive root*

Portions of these notes are adapted from Elementary Number Theory: Primes, Congruences and Secrets *by William Stein*

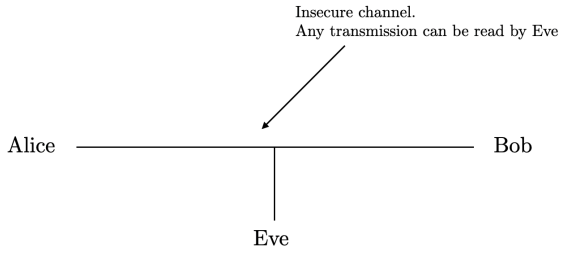
In this module we consider applications of number theory to cryptography – the art of writing and communicating in secret code. The field of cryptography is wide and deep when approached from either a mathematics or computer science viewpoint. Given the scope of this course, we take just a narrow and shallow look. In this module we will focus on the underlying number theory of two common encryption algorithms. Both of these algorithms are used daily by a wide variety of online systems. We begin by setting the stage with some broad generalities about cryptography.

Aside 7.1. *Nothing that follows in this section is objectively false. However some concepts are simplified. This is practical necessity; getting deep into the technical details of various aspects of cryptography is not the goal of this course. If you are taking this course and have already studied cryptography, I apologize for what you are about to read – you may not enjoy it.*

Two people, Alice and Bob, wish to communicate over an insecure channel. That they will be using computers for this communication implies that they have some means to encode their human-readable messages in a form that can be transmitted electronically. We will not concern ourselves with how this is done, we only assume that the messages they wish to send can be represented as positive integers in some fixed range. (In fact, these messages will be

encoded as binary strings. The field of *Coding Theory* looks at the various considerations in play when choosing particular binary strings to use.)

That Alice and Bob are communicating over an insecure channel means that they assume that all of the messages they send will be read by an adversary, Eve. And so Alice and Bob wish to devise systems to encrypt (i.e., scramble) and decrypt (i.e., unscramble) their communications so that Eve cannot understand the messages.



Broadly speaking, an encryption scheme is a publicly-known algorithm to encrypt and decrypt data. As the algorithms are publicly known, Alice and Bob must share some secret information that they can use as part of the input to the algorithm. Such information is, aptly, called a key. Once Alice and Bob have a shared key, they can then use this key to encrypt/decrypt their data. Consider the following example of an encryption scheme.

ROT-s Encryption

- Alice and Bob together choose a fixed integer $s \in \{0, 1, 2, 3 \ldots, 25\}$.
- To encrypt a message Alice takes each letter of the message and converts it to an integer ($A = 0$, $B = 1$, etc...) and then adds s to each integer. Alice then sends this sequence of integers to Bob.
- To decrypt the message, Bob subtracts s from each integer and converts it back to a letter of the alphabet.

An eavesdropping adversary, Eve, only sees the encrypted letters of the message. Without knowing the value of s, Eve cannot decrypt the message. In this case, the integer s is the shared key. Even though Eve may know how the data is encrypted, they cannot directly decrypt the data without knowing s.

Aside 7.2. *ROT-s is a particularly bad example of a* substitution cipher. *Each letter/number in the message is directly substituted for another. Such schemes are vulnerable to* frequency analysis. *Characters that appear often in the encrypted text correspond to characters that appear often in the decrypted text. Do an internet search for* `Ceasar cipher` *for more information.*

Schemes that require a shared key seemingly suffer a serious flaw – if the channel on which Alice and Bob communicate is insecure, how can Alice and Bob choose a shared key in secret? We consider this problem in coming section.

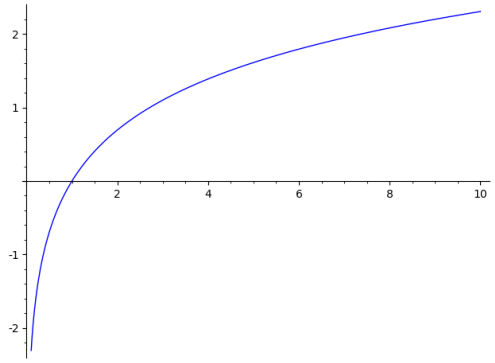
7.1 Diffie-Hellman-Merkle Shared Key Scheme

The Diffie-Hellman-Merkle shared key scheme is a method for Alice and Bob to devise a secret shared key, even when Eve is monitoring all of their communication. The security of the scheme relies on the seemingly difficulty of the following problem:

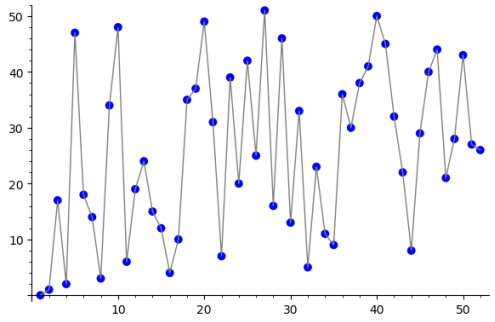
The Discrete-Logarithm Problem
Given an integer $n \geq 2$ and $[g], [h] \in \mathbb{Z}_n$, find an integer e so that $[g]^e = [h]$.

For example, given $n = 7$, find an integer e so that $[3]^e = [5]$. By guessing and checking we can find $e = 4$ satisfies this equation. When n is small, enumerating the powers of $[g]$ is not a computationally difficult task. On the other hand, when n is very large (on the order of, say, 2^{1024}) such a task is computationally infeasible.

The discrete logarithm problem is named because it is exactly the analogue of logarithms for real numbers. (Recall $log_b(x) = e$ when $b^e = x$). For real numbers, logarithms are well understood. Given a pair b and x, there are computational methods to efficiently find $log_b(x)$



On the other hand, the behaviour of the discrete logarithm is seemingly unpredictable. Here the vertical axis is n and the horizontal axis is $[2]^n$ in $\mathbb{Z}_{53}$ for $n \in [1, 52]$. (The points are the pairs $([2]^n, n)$. The lines make the graph easier to parse.)



The discrete-logarithm problem underlies a variety of cryptographic methods. There are no known methods to solve discrete-logarithm problems efficiently.

With our knowledge of discrete logarithms in hand, we proceed with a description of the Diffie-Hellman-Merkle shared key scheme. The output of this scheme is an integer s that is known to both Bob and Alice, but is unknown by Eve. All transmitted information is public.

Diffie-Hellman-Merkle Shared Key Scheme

1. *Alice chooses a prime p, and integers $a, g \in [1, p-1]$. Alice computes g^a and finds the remainder, r_a, when divided by p.*
2. *Alice sends the triple (p, g, r_a) to Bob.*
3. *Bob chooses an integer $b \in [1, p-1]$, computes g^b, finds the remainder, r_b, when divided by p, and sends r_b to Alice.*
4. *Bob computes $(r_a)^b$ and finds the remainder, s, when divided by p.*
5. *Alice computes $(r_b)^a$ and finds the remainder, s, when divided by p.*

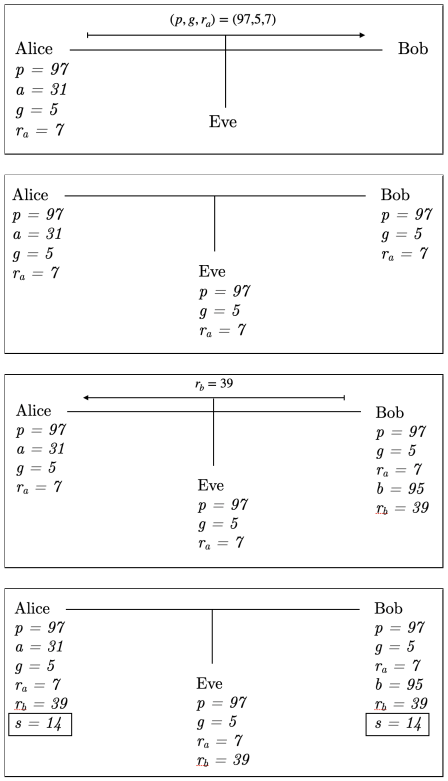
Consider the following example from the point of view of each of the participants.

Alice chooses $p = 97$, $a = 31$ and $g = 5$. She computes $g^a = 5^{31}$ has remainder $r_a = 7$ when divided by $p = 97$. Alice transmits the triple $(p, g, r_a) = (97, 5, 7)$ to Bob. Alice subsequently receives $r_b = 39$ from Bob and computes $(r_b)^a = 39^{31}$. Alice finds that $(r_b)^a = 39^{31}$ has remainder 14 when divided by $p = 97$ and so has $s = 14$.

Bob receives the triple $(p, g, r_a) = (97, 5, 7)$ from Alice. He chooses $b = 95$ and computes $g^b = 5^{95}$ has remainder $r_b = 39$ when divided by $p = 97$. Bob sends $r_b = 39$ to Alice. Bob computes $(r_a)^b = 7^{95}$ has remainder 14 when divided by $p = 97$ and so has $s = 14$.

Eve sees Alice transmit the triple $(p, g, r_a) = (97, 5, 7)$ to Bob. Eve also sees Bob transmit $r_b = 39$. However without knowing a and b, Eve cannot find the remainder, s, when $g^{ab} = 5^{ab}$ is divided by $p = 97$.

The following sequence of images represents the steps of the procedure. Under each name is the information that participant knows.



The description of the scheme tells us that both Alice and Bob have constructed the same number s. Alice constructed this number by taking the remainder when $(r_b)^a$ is divided by p. Bob constructed this number by taking the remainder when $(r_a)^b$ is divided by p. Recall $r_b \equiv_p g^b$ and $r_a \equiv_p g^a$. We observe

$$(r_b)^a \equiv_p (g^b)^a \equiv g^{ab} \equiv_p (g^a)^b \equiv_p (r_a)^b$$

And so both $(r_b)^a$ and $(r_a)^b$ have the same remainder when divided by p.

In our scheme Eve knows p, g, r_a and r_b. Notably, Eve does not know either a or b – the information Eve would need to compute s. Finding a amounts to finding an integer d so that $[g]^d = [r_a]$ in $\mathbb{Z}_p$. Similarly finding b amounts to finding an integer e so that $[g]^e = [r_b]$ in $\mathbb{Z}_p$. Thus to compute s, Eve must solve two separate discrete-logarithm problems.

Eve's task in solving discrete-logarithm problems can be made easier depending on the choice of p and g. For example consider $p = 7$ and the powers of the non-zero elements of $\mathbb{Z}_7$.

$[z]$	$[z]^2$	$[z]^3$	$[z]^4$	$[z]^5$	$[z]^6$
[1]	[1]	[1]	[1]	[1]	[1]
[2]	[4]	[1]	[2]	[4]	[1]
[3]	[2]	[6]	[4]	[5]	[1]
[4]	[2]	[1]	[4]	[2]	[1]
[5]	[4]	[6]	[2]	[3]	[1]
[6]	[1]	[6]	[1]	[6]	[1]

Notice that for $[z] = [3]$ all of the elements of $\mathbb{Z}_7$ can be expressed a power of [3]. On the other hand, for $[z] = [6]$ only [1] and [6] appear as a power of [6]. In the Diffie-Hellman-Merkle shared key scheme with $p = 7$, choosing $g = 3$ gives many more possibilities for g^a and g^b than choosing $g = 6$. Further, choosing $g = 3$ allows s to take any value in $[1, 6]$. In general, we want to choose g so that s can take any value in $[1, p-1]$.

Definition 7.3. *Let p be prime and let $[g] \in \mathbb{Z}_p$. We say $[g]$ is a primitive root modulo p when for every $[a] \in \mathbb{Z}_p$ with $[a] \neq [0]$ there exists $e \in [1, p-1]$ so that $[g]^e = [a]$ in $\mathbb{Z}_p$.*

For $p = 7$ we see that [3] and [5] are primitive roots modulo 7. Every non-zero element of $\mathbb{Z}_7$ can be expressed as a power of both of [3] and [5].

Primitive roots give us a new way to represent elements if $\mathbb{Z}_p$. The $p-1$ powers of a primitive root modulo p are the non-zero elements of $\mathbb{Z}_p$. For $p = 7$, notice

$$\mathbb{Z}_7 = \{[3]^e \mid 1 \leq e \leq p-1\} \cup \{0\}$$

Aside 7.4. *If you are taking MATH361 this term, we have just showed that the multiplicative group $\mathbb{Z}_7 \setminus \{0\}$ has a primitive root. Thus this group is cyclic.*

We return to the study of primitive roots in Modules 8 and 9. We show in Module 8 that every prime modulus admits a primitive root. It is for this reason that Alice chooses a prime Diffie-Hellman-Merkle shared key scheme rather than just an arbitrary integer. Alice can always choose a primitive root as g.

Diffie-Hellman-Merkle shared key scheme allows any pair of people to choose a shared key when their only option for communication is an insecure channel. The Diffie-Hellman-Merkle shared key scheme, however, doesn't allow Alice and Bob to directly exchange encrypted messages. To do this they must use an separate encryption scheme. We turn now to a method of key-sharing that directly permits Alice and Bob to encrypt messages.

Aside 7.5. *The Diffie-Hellman-Merkle shared key scheme was first published in 1976 by Whitfield Diffie and Martin Hellman in a groundbreaking paper* New Directions in Cryptog-

raphy[2]. *Both Diffie and Hellman acknowledge Ralph Merkle as the one to have originally conceived the scheme. This scheme was patented in 1977, with all three researchers named as inventors. Unknown to these researchers, a group of British Intelligence researchers (Ellis, Cocks and Williamson) designed a similar scheme nearly a decade earlier. As these researchers were working in intelligence, their work was classified only becoming public in 1997.*

[2]Diffie, Whitfield, and Martin Hellman. (1976) *New directions in cryptography* IEEE transactions on Information Theory 22(6): 644-654.

7.2 RSA Cryptography

To study the RSA encryption scheme we take another look at Euler's Theorem:

Theorem (Euler's Theorem)**.** *Let $n \geq 2$ be an integer. For every integer a so that $gcd(a, n) = 1$ we have $a^{\phi(n)} \equiv_n 1$.*

Let e and n be integers so that $gcd(e, \phi(n)) = 1$. Since $gcd(e, \phi(n)) = 1$, necessarily $[e]^{-1}$ exists in $\mathbb{Z}_{\phi(n)}$. And so there exists an integer $d \in [1, \phi(n) - 1]$ so that $ed \equiv_{\phi(n)} 1$ (Yes the modulus is $\phi(n)$. This seems weird but it will make sense soon.)

Since ed has remainder 1 when divided by $\phi(n)$, necessarily there exists an integer q so that

$$ed = \phi(n)q + 1$$

For example let $n = 32$. We have $\phi(32) = 32\left(1 - \frac{1}{2}\right) = 16$. Let $e = 5$. We find d so that $5d \equiv_{16} 1$ by solving the linear diophantine equation

$$5d + 16y = 1$$

We find a solution $d = 13$. And so

$$5 \cdot 13 \equiv_{16} 1$$

Consider now the quantity m^{ed} for some integer $m \in [1, n]$ with $gcd(m, n) = 1$. We have

$$m^{ed} = m^{\phi(n)q+1} = (m^{\phi(n)})^q \cdot m$$

The term $m^{\phi(n)}$ is very much suggestive of Euler's Theorem! And so when divided by n, the integer m^{ed} has remainder:

$$m^{ed} \equiv_n (m^{\phi(n)})^q \cdot m \equiv_n (1)^q \cdot m \equiv_n m$$

Continuing with our example, let $m = 5$. We have

$$5^{65} \equiv_{32} 5^{16(4)+1} \equiv_{32} (5^{16})^4 \cdot 5 \equiv_{32} (5^{\phi(32)})^4 \cdot 5 \equiv_{32} (1)^4 \cdot 5 \equiv_{32} 5$$

It is these ideas that form the basis of the RSA encryption scheme. Bob will encrypt his message, m, by computing m^e. Alice will then decrypt by computing $(m^e)^d$

The scheme proceeds in three phases: key generation, encryption and decryption. In this scheme Alice will generate a pair of keys: a private key (d) and a public key (n, e). Anyone can use Alice's public key to encrypt a message for her. Only she can decrypt the message using her private key.

RSA – Key Generation

- *Alice chooses two distinct primes, p and q and computes $n = pq$ and $\phi(n) = \phi(p) \cdot \phi(q)$.*
- *Alice chooses an integer $e \in [1, \phi(n) - 1]$ and finds $d \in [1, \phi(n) - 1]$ so that $ed \equiv_{\phi(n)} 1$.*
- *Alice publishes the pair (n, e) (her public key) and keeps d (her private key) private.*

The pair (n, e) is public. Using this pair, Bob encrypts a message $m \in [1, n - 1]$ with $gcd(m, n) = 1$ as follows:

RSA – Encryption

- *Bob computes m^e, finds the remainder, r_m, when divided by n and sends it to Alice.*

Using d, Alice can decrypt the message as follows. As Alice is the only one who knows d, the message cannot be decrypted by Eve.

RSA – Decryption

- *Alice computes $(r_m)^d$. The remainder when $(r_m)^d$ is divided by n is m as*

$$(r_m)^d \equiv_n (m^e)^d \equiv_n m$$

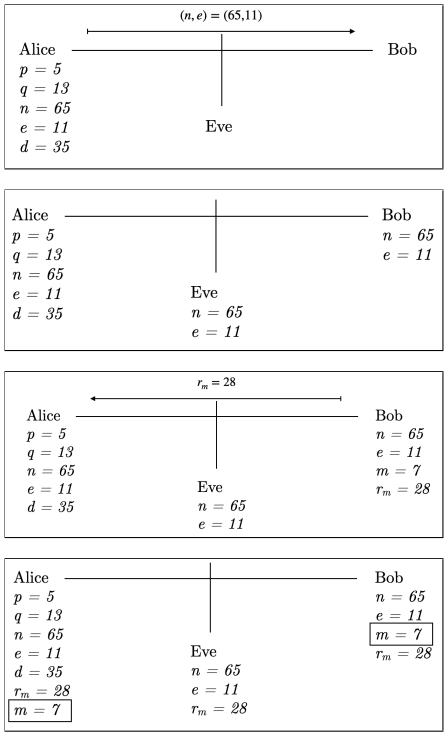
Consider the following example from the point of view of each of the participants.

Alice chooses $p = 5$ and $q = 13$. She computes $n = 5 \cdot 13 = 65$ and $\phi(65) = 48$. She chooses $e = 11$ and computes $d = 35$. She publishes the pair $(n, e) = (65, 11)$. Alice then subsequently receives $r_m = 28$ from Bob. She computes 28^{35} and finds that it has remainder 7 when divided by 65. Thus $m = 7$.

Bob receives the pair $(n, e) = (65, 11)$ from Alice. Bob wishes to encrypt the message $m = 7$. He computes 7^{11} and finds the remainder is 28 when divided by 65. Bob sends $r_m = 28$ to Alice.

Eve witnesses the pair $(n, e) = (65, 11)$ being sent from Alice to Bob. Eve then witnesses $r_m = 28$ being sent from Bob to Alice. Without further information, Eve cannot determine m.

The following sequence of images represents the steps of the procedure. Under each name is the information that participant knows.



Given that n and e are public, a possible line-of-attack for Eve is to try to compute d using this information. As Alice uses $\phi(n)$ to compute d, if Eve can determine $\phi(n)$, then they can also compute d.

Our work in the last chapter for computing $\phi(n)$ in the previous section relied on knowing the prime-power decomposition of n. Thus to compute $\phi(n)$ it suffices for Eve to find the prime-power factorization of n. For the scheme to work, it is not necessary for n to be chosen as the product of two primes. However, choosing $n = pq$ for primes p and q is designed to make this search difficult. It currently unknown if there is a polynomial-time algorithm for integer factorization. This is an important open problem in computer science.

(In our example, Alice's choice of p and q are quite small. It would be no problem for Eve to determine p and q, and thus $\phi(pq)$ for $pq = 65$. In practice, p and q are chosen so that n is on the order of 2^{1024}.)

Aside 7.6. *RSA stands for Rivet-Shamir-Adleman. These three researchers at MIT were granted a patent for this process in the early 1980s, after coming up with the idea in the late 1970s* [3]*. Unknown to them, a researcher working for British Intelligence, Clifford Cocks, had designed the same scheme in the early 1970s. As Cocks was working in intelligence, his*

[3]Rivest, R. L., Shamir, A., and Adleman, L. M. (1983). U.S. Patent No. 4,405,829. Washington, DC: U.S. Patent and Trademark Office.

work was classified only becoming public in 1997. An internet search for `RSA Cryptography` *quickly reveals how integral this scheme is to modern internet commerce.*

7.3 Further Exercises

1. What is Alice and Bob's shared key in the Diffie-Hellman-Merkle scheme with $p = 11$, $g = 2$, $a = 9$ and $b = 4$?
2. Consider RSA where Alice has chosen $p = 11$ and $q = 13$ and $e = 7$.
   (a) Compute n and $\phi(n)$.
   (b) Compute d.
   (c) If Alice receives $r_m = 48$, what is m?
3. In the description of RSA encryption above, Bob chooses a message that is co-prime with n. What happens if Bob chooses a message $m \in [1, n-1]$ that isn't coprime with n?

---

8 Primitive Roots Part I

Learning Incomes.

- *Understand the statement of Fermat's Little Theorem.*
- *Understand the relationship between $a \equiv_p b$ and $[a] = [b]$ in $\mathbb{Z}_p$.*
- *Be able to manipulate equations in $\mathbb{Z}_p$*

Learning Outcomes.

- *Understand the broad steps to construct a primitive root modulo p.*
- *Understand the broad steps to proving that a primitive root exists modulo p.*

Newly Defined Terms and Notation. *primitive root modulo p, order*

A Note About Module 8. *In Module 8 we prove that a primitive modulo p exists for every prime p. Along the way we prove a variety of smaller results. We combine these results together to prove our main result. Your primary goal in this module is understanding how the smaller results combine to give the main result. A secondary goal should be understanding the proofs of all of the smaller results.*

Recall from Module 7 the definition of a primitive root.

Definition 8.1. *Let p be prime and let $[g] \in \mathbb{Z}_p$. We say $[g]$ is a primitive root modulo p when for every $[a] \in \mathbb{Z}_p$ with $[a] \neq [0]$ there exists $e \in [1, p-1]$ so that $[g]^e = [a]$ in $\mathbb{Z}_p$.*

Informally, a primitive root $[g]$ is an element of $\mathbb{Z}_p$ so that every non-zero element of $\mathbb{Z}_p$ can be expressed as a power of $[g]$.

Recall the exponentiation table for $\mathbb{Z}_7$:

$[z]$	$[z]^2$	$[z]^3$	$[z]^4$	$[z]^5$	$[z]^6$
[1]	[1]	[1]	[1]	[1]	[1]
[2]	[4]	[1]	[2]	[4]	[1]
[3]	[2]	[6]	[4]	[5]	[1]
[4]	[2]	[1]	[4]	[2]	[1]
[5]	[4]	[6]	[2]	[3]	[1]
[6]	[1]	[6]	[1]	[6]	[1]

Every non-zero element of $\mathbb{Z}_7$ appears in the row corresponding to [3]. Thus

$$\mathbb{Z}_7 = \{[3]^e \mid 1 \leq e \leq 6\} \cup \{[0]\}$$

Fermat's Little Theorem tells us that for every non-zero $[a] \in \mathbb{Z}_p$, there exists an integer e so that $[a]^e = [1]$. (We may take always $e = p - 1$.) Looking at the table above, we see that there are non-zero elements of $\mathbb{Z}_7$ for which $[a]^e = [1]$ and $e < 7 - 1 = 6$. Primitive roots are exactly those non-zero elements of $\mathbb{Z}_p$ for which $e = p - 1$ is the least integer so that $[a]^e = [1]$.

Let us take a moment to justify this last sentence. Assume e is the least integer so that$[a]^e = [1]$ If $[a]^e = [1]$, then $[a]^{e+1} = [a]$. If $e < p - 1$, then there is no way that the sequence

$$[a], [a]^2, [a]^3, \ldots, [a]^{p-1}$$

can contain all of the non-zero elements of $\mathbb{Z}_p$ as $[a]$ appears twice. And so $[a]$ is not a primitive root modulo p.

On the other hand, if $e = p-1$ is the least integer so that $[a]^e = [1]$, but $[a]$ is not a primitive root, then the sequence

$$[a], [a]^2, [a]^3, \ldots, [a]^{p-1}$$

does not contain all non-zero elements of $\mathbb{Z}_p$. As there are $p-1$ non-zero elements of $\mathbb{Z}_p$ and this sequence has $p - 1$ entries, if this sequence does not contain every non-zero element of $\mathbb{Z}_p$, then some element must appear twice. Let i be the smallest integer so that $[a]^i = [b]$ is repeated in the sequence. Therefore there exists an integer $i < j \leq p - 1$ so that $[a]^j = [b]$. If $i \neq 1$, then

$$\begin{aligned}
[a]^i[a]^{-1} &= [a]^j[a]^{-1} \\
\left([a]^{i-1}\right)[a][a]^{-1} &= \left([a]^{j-1}\right)[a][a]^{-1} \\
[a]^{i-1} &= [a]^{j-1}
\end{aligned}$$

This equality contradicts that $[a]^i$ was the first element to be repeated in the sequence. Therefore $i = 1$. But this then implies $[a]^{j-1} = [1]$ as

$$\begin{aligned}
[a][a]^{-1} &= [a]^j[a]^{-1} \\
[1] &= \left([a]^{j-1}\right)[a][a]^{-1} \\
[1] &= [a]^{j-1}
\end{aligned}$$

This now contradicts that e is the least integer so that $[a]^e = [1]$. Therefore i does not exist. Which in turn implies that every element of the sequence is unique. And so each non-zero element appears in the sequence. Which then implies $[a]$ is a primitive root modulo p.

Our work in Module 8 heads us towards a proof that $\mathbb{Z}_p$ has a primitive root for every prime p. In particular we prove the following theorem:

Theorem 8.2. *For every prime p, $\mathbb{Z}_p$ has at least one primitive root.*

In Sections 8.1 and 8.2 we put together the tools we need for our proof. A broad outline of how the proof proceeds appears at the end of Section 8.1. As a first step, we formalize our discussion of this integer e mentioned above.

8.1 Order

Definition 8.3. *Let p be a prime and let $[a]$ be a non-zero element of $\mathbb{Z}_p$. The order of $[a]$ is the least integer e so that $[a]^e = [1]$.*

Fermat's Little Theorem implies that every non-zero element of $\mathbb{Z}_p$ has an order. Recall again the exponentiation table for $\mathbb{Z}_7$.

$[z]$	$[z]^2$	$[z]^3$	$[z]^4$	$[z]^5$	$[z]^6$
[1]	[1]	[1]	[1]	[1]	[1]
[2]	[4]	[1]	[2]	[4]	[1]
[3]	[2]	[6]	[4]	[5]	[1]
[4]	[2]	[1]	[4]	[2]	[1]
[5]	[4]	[6]	[2]	[3]	[1]
[6]	[1]	[6]	[1]	[6]	[1]

The following table gives the order of each non-zero element of $\mathbb{Z}_7$:

$[z]$	e
[1]	1
[2]	3
[3]	6
[4]	3
[5]	6
[6]	2

And so, $\mathbb{Z}_7 \setminus \{0\}$ contains one element of order 1, one element of order 2, two elements of order 3 and two elements of order 6.

Let us consider another example: $\mathbb{Z}_{11}$. Using a computer one can find that the non-zero elements of $\mathbb{Z}_{11}$ have the following orders:

$[z]$	e
[1]	1
[2]	10
[3]	5
[4]	5
[5]	5
[6]	10
[7]	10
[8]	10
[9]	5
[10]	2

And so $\mathbb{Z}_{11} \setminus \{0\}$ contains one element of order 1, one element of order 2, four elements of order 5 and four elements of order 10.

Aside 8.4. *Are you convinced that* $[2]$ *is a primitive root modulo* 11*?*

There is something curious to notice here. We have that $\mathbb{Z}_7 \setminus \{0\}$ has 6 elements and the possible orders of elements in $\mathbb{Z}_7 \setminus \{0\}$ are exactly the positive divisors of 6. Similarly, we have that $\mathbb{Z}_{11} \setminus \{0\}$ has 10 elements and the possible orders of elements in $\mathbb{Z}_{11} \setminus \{0\}$ are exactly the positive divisors of 10.

Aside 8.5. *There is another curiosity here that is a little harder to notice. Count the number of elements of order* e *and compare it to* $\phi(e)$*. This observation, if true, would directly imply that* $\mathbb{Z}_p$ *has a primitive root (consider* $e = p - 1$*). We do not take this approach in our proof of Theorem 8.2*

If we knew in advance that the order of an non-zero element of $\mathbb{Z}_p$ was necessarily a divisor of $p - 1$, then our work in computing the order of an element becomes much easier! In trying to find the order of $[a] \in \mathbb{Z}_p$. We need only compute $[a]^d$ for each d that divides $p - 1$.

To see why our observation is true, let us take a look at the powers of $[3]$ in $\mathbb{Z}_{11}$

$[z]$	$[z]^2$	$[z]^3$	$[z]^4$	$[z]^5$	$[z]^6$	$[z]^7$	$[z]^8$	$[z]^9$	$[z]^{10}$	$[z]^{11}$	$[z]^{12}$	$[z]^{13}$	$\ldots$
$[3]$	$[9]$	$[5]$	$[4]$	$[1]$	$[3]$	$[9]$	$[5]$	$[4]$	$[1]$	$[3]$	$[9]$	$[5]$	$\ldots$

Notice that the values repeat with a period of $e = 5$. In other words–

Theorem 8.6. *Let* p *be a prime, and let* $[a]$ *be a non-zero element of* $\mathbb{Z}_p$ *with order* e*. For any integers* $i, j \geq 1$ *we have* $[a]^i = [a]^j$ *if and only if* i *and* j *are congruent modulo* e*.*

Proof. Let p be prime and let $[a]$ be a non-zero element of $\mathbb{Z}_p$ with order e.

We first show that if i and j are congruent modulo e, then $[a]^i = [a]^j$. Notice that for any multiple of e we have

$$[a]^{ke} = ([a]^e)^k = [1]^k = [1]$$

Consider now $i \in [1, p - 1]$. We show $[a]^i = [a]^{qe+i}$ for every $q \in \mathbb{N}$.

$$[a]^{qe+i} = [a]^i \cdot ([a]^e)^q = [a]^i \cdot [1]^q = [a]^i$$

To complete the proof it suffices to show that each element of the sequence

$$[a], [a]^2, [a]^3, \ldots, [a]^e$$

is distinct.

Assume otherwise. Let $i \geq 1$ be the smallest integer so that $[a]^i = [b]$ is repeated in the sequence. Therefore there exists an integer $i < j \leq e$ so that $[a]^j = [b]$. If $i \neq 1$, then

$$\begin{aligned} [a]^i[a]^{-1} &= [a]^j[a]^{-1} \\ \left([a]^{i-1}\right)[a][a]^{-1} &= \left([a]^{j-1}\right)[a][a]^{-1} \\ [a]^{i-1} &= [a]^{j-1} \end{aligned}$$

This contradicts that i was the first integer so that an element of the sequence repeated. Therefore $i = 1$. But this then implies $[a]^{j-1} = [1]$ as

$$\begin{aligned} [a][a]^{-1} &= [a]^j[a]^{-1} \\ [1] &= \left([a]^{j-1}\right)[a][a]^{-1} \\ [1] &= [a]^{j-1} \end{aligned}$$

Since $j - 1 < e$, this contradicts that $[a]$ has order e.

Therefore each element of the sequence

$$[a], [a]^2, [a]^3, \ldots, [a]^e$$

is distinct. □

We now prove that the order of a non-zero element in $\mathbb{Z}_p$ divides $p - 1$.

Corollary 8.7. *Let p be a prime. For any non-zero $[a] \in \mathbb{Z}_p$ the order of $[a]$ divides $p - 1$.*

Proof. Let e be the order of $[a]$. By definition and Fermat's Little Theorem we have $[a]^e = [1] = [a]^{p-1}$. By Theorem 8.6 we have $e \equiv_e p - 1$. Since $e \equiv_e 0$ we have $p - 1 \equiv_e 0$. Therefore $p - 1$ is a multiple of e. And so, e divides $p - 1$. □

Returning to our motivation for studying the order of elements, we now have the terminology to state the result we justified at the end of the previous section.

Theorem 8.8. *Let p be a prime and let $[a]$ be a non-zero element of $\mathbb{Z}_p$. We have that $[a]$ is a primitive root modulo p if and only if $[a]$ has order $p - 1$.*

With this theorem, our work in proving $\mathbb{Z}_p$ always has a primitive root is reduced to proving $\mathbb{Z}_p$ necessarily has an element of order $p - 1$.

A key ingredient in this proof will be a technique to find the order of a product of two non-zero elements of $\mathbb{Z}_p$ whose orders are coprime. To see this in action, let us return to $\mathbb{Z}_{11}$. From our table above, we see that [4] has order 5 and [10] and has order 2. The product of these elements, $[4][10] = [40] = [7]$, has order $5 \cdot 2 = 10$

We state this observation in general:

Theorem 8.9. *Let p be a prime. Let $[a]$ and $[b]$ be non-zero elements of $\mathbb{Z}_p$. Let e and f respectively be the orders of $[a]$ and $[b]$. If e and f are coprime, then $[a][b]$ has order ef.*

Our proof of this lemma requires nothing more than some clever algebraic manipulations in $\mathbb{Z}_p$. We omit this proof as it doesn't contribute to our broader understanding of primitive roots.

Before we move on, let us take a moment to think about how this lemma is helpful for us in our quest to construct a primitive root modulo p, for some prime p.

To find an element of order $p-1$ in $\mathbb{Z}_p$ we can instead find a pair of elements of orders e and f where e and f are coprime and $ef = p-1$. We can inductively extend this idea to a factorization of $p-1$ into a product of coprime factors. What better coprime factors are there than prime powers!

Consider $p = 241$. We have

$$p - 1 = 240 = 2^4 \cdot 3 \cdot 5$$

From our thoughts above, if we can find an element $[a_1]$ of order 2^4, and element $[a_2]$ of order 3 and an element $[a_3]$ of order 5, then by two applications of Theorem 8.9 we have that the element $[a_1][a_2][a_3] \in \mathbb{Z}_{241}$ has order $2^4 \cdot 3 \cdot 5 = 240$. By Theorem 8.8, such an element is a primitive root modulo 241.

Let p be a prime. Since $p-1$ is a positive integer, it has a prime power factorization

$$p - 1 = q_1^{f_1} q_2^{f_2} \cdots q_k^{f_k}$$

If we can find an element of order q_i^i for each $1 \leq i \leq k$, then necessarily the product of these elements will have order $q_1^{f_1} q_2^{f_2} \cdots q_k^{f_k} = p - 1$. By Theorem 8.8, such a product necessarily is a primitive root modulo p.

Corollary 8.10. *Let p be prime, let*

$$p - 1 = q_1^{f_1} q_2^{f_2} \cdots q_k^{f_k}$$

be the unique prime-power factorization of $p-1$. For all $1 \leq i \leq k$, if $[a_i] \in \mathbb{Z}_p$ has order $q_i^{f_i}$, then

$$[g] = \prod_{i=1}^{k} [a_i]$$

has order $p-1$.

The proof proceeds by induction but isn't terribly interesting. And so we omit this proof as it doesn't contribute to our broader understanding of primitive roots.

Before you go any further, make sure that you understand the meaning of this corollary. Look back above to the example $p = 241$. The corollary gives us a road map to prove that there is a primitive root modulo p for every prime p: we show $\mathbb{Z}_p$ has an element of order q_i^i for each $1 \leq i \leq k$ where

$$p - 1 = q_1^{f_1} q_2^{f_2} \cdots q_k^{f_k}$$

The product of these elements in $\mathbb{Z}_p$ is necessarily a primitive root modulo p. We shelf this idea for now and develop some other necessary tools in the following section.

8.2 Polynomials Modulo p

We turn our minds back to our study of polynomials in $\mathbb{Z}_p$. For clarity, we use to our notation based on the equivalence relation $\equiv_p$.

Let be a p be a prime. We may equivalently define the order of an element $a \in [1, p-1]$ as the least integer e so that $a^e \equiv_p 1$. Alternatively, we may express this congruence as $a^e - 1 \equiv_p 0$. And so we notice that if a has order e, then $x = a$ is a solution to the congruence

$$x^e - 1 \equiv_p 0$$

(This is not an if and only if statement. By Fermat's Little Theorem, every element $a \in [1, p-1]$ satisfies $a^{p-1} \equiv_p 1$, but not every element has order equal to $p-1$.)

And so to find a primitive root modulo p we must at the very least find a solution to

$$x^{p-1} - 1 \equiv_p 0$$

that is not a solution to any polynomial congruence

$$x^d - 1 \equiv_p 0$$

where $d < p-1$ and d is a divisor of $p-1$. We begin our work by thinking about the number of possible solutions to polynomial congruences. Recall the Fundamental Theorem of Algebra.

Theorem (Fundamental Theorem of Algebra)**.** *A non-zero polynomial of degree n with complex coefficients has, counted with multiplicity, exactly n complex roots.*

When we restrict our coefficients and our roots to $\mathbb{R}$ we arrive at the following corollary.

Corollary 8.11. *A non-zero polynomial of degree n with real coefficients has at most n distinct real roots.*

Though Corollary 8.11 follows almost directly from the Fundamental Theorem of Algebra, let us take a moment to examine a proof of this corollary that does not invoke the Fundamental Theorem of Algebra.

Proof. We proceed by induction on n.

Every non-zero polynomial of degree $n = 0$ is of the form $f(x) = c$ for some $c \neq 0$. Since $c \neq 0$, such a polynomial does not cross the x-axis and so has no roots.

Let $g(x)$ be a polynomial of degree $n = k + 1$.

$$g(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

If $g(x)$ has at least one root $x = r$, then we have

$$\begin{aligned} g(x) &= g(x) - 0 \\ &= g(x) - g(r) \\ &= a_n(x^n - r^n) + a_{n-1}(x^{n-1} - r^{n-1}) + \cdots + a_1(x - r) + a_0(1 - 1) \end{aligned}$$

From each of these terms (except the last!) we may factor out $x - r$. That is

$$a_i(x^i - r^i) = (x - r)a_i \sum_{j=0}^{i} x^{i-j} r^{j-1}$$

for each $1 \leq i \leq n$. And so we may express $g(x)$ as a product

$$g(x) = (x - r)h(x)$$

where $h(x)$ is a polynomial of degree $n - 1$.

We claim that every root of $g(x)$, except for possibly $x = r$ is a root of $h(x)$. If this holds, then the result follows by induction as $h(x)$ has degree $n-1 = k$ and so has at most $n-1 = k$ distinct roots.

Consider $r' \neq r$ so that $g(r') = 0$. We have

$$0 = g(r') = (r - r')h(r')$$

Since $r \neq r'$, we have $r - r' \neq 0$. Since $(r - r')h(r') = 0$ and $r - r' \neq 0$, it must be that $h(r') = 0$. And so r' is a root of $h(x)$.

By induction $h(x)$ has at most $n - 1 = k$ distinct roots. And so it follows that $g(x)$ has at most $n = k + 1$ distinct roots. The result now follows by induction. □

In thinking about transplanting this proof to work modulo p, almost nothing needs to change! The only cause for concern is the part where we confirm r' is a root of $h(x)$. In $\mathbb{R}$, the product of non-zero elements is necessarily non-zero. And so we conclude that if a product is zero, then one of terms must be zero. This statement is not true in general in $\mathbb{Z}_m$. (For example, $[2][2] = [0]$ in $\mathbb{Z}_4$) Fortunately, this statement is true in $\mathbb{Z}_p$ (See Midterm Q5). And so we arrive at the following theorem.

Theorem 8.12. *Let p be a prime. The polynomial congruence*

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \equiv_p 0$$

has at most n distinct solutions in the range $[0, p - 1]$.

Recall from our work at the end of the last section, we are interested in finding elements of order q^f where q^f appears as part of the prime-power factorization of $p-1$. And so we are interested in solutions for $x^{q^f} - 1 \equiv_p 0$. From Theorem 8.12, this polynomial congruence has at most q^f solutions. We show this polynomial congruence has exactly q^f solutions.

Lemma 8.13. *Let p be a prime and let d be a divisor of $p-1$. The polynomial congruence $x^d - 1 \equiv_p 0$ has exactly d distinct solutions in the range $[0, p-1]$.*

Proof. Since d divides $p-1$, there exists an integer t so that $dt = p-1$. We have

$$\begin{aligned} x^{p-1} - 1 &= \left(x^d\right)^t - 1 \\ &= (x^d - 1)\left(\left(x^d\right)^{t-1} + \left(x^d\right)^{t-2} + \cdots + 1\right) \end{aligned}$$

Let $g(x) = \left(x^d\right)^{t-1} + \left(x^d\right)^{t-2} + \cdots + 1$. By Fermat's Little Theorem, $(x^d - 1)g(x)$ has exactly $p-1$ distinct roots in the range $[0, p-1]$. By Theorem 8.12, $g(x)$ has at most $d(t-1) = dt - d = p - 1 - d$ distinct roots in the range $[0, p-1]$. Therefore $x^d - 1$ has at least $(p-1) - (p-1-d) = d$ distinct roots in the range $[0, p-1]$. Theorem 8.12 implies $x^d - 1$ has at most d such roots. Therefore $x^d - 1 \equiv_p 0$ has exactly d distinct solutions in the range $[0, p-1]$. □

Let us return to our example of $p = 241$. Recall our interest in finding an element of order 2^4. Consider the polynomials

$$\begin{aligned} f_0(x) &= x^{2^0} - 1 \\ f_1(x) &= x^{2^1} - 1 \\ f_2(x) &= x^{2^2} - 1 \\ f_3(x) &= x^{2^3} - 1 \\ f_4(x) &= x^{2^4} - 1 \end{aligned}$$

Every element of order 2^4 in $\mathbb{Z}_{241}$ is a solution to the polynomial congruence

$$f_4(x) \equiv_{241} 0$$

However, not every solution corresponds to an element of order 2^4 in $\mathbb{Z}_{241}$. For example,

$$1^{2^4} - 1 \equiv_{241}$$

but [1] has order 1.

For $i \in \{0,1,2,3,4\}$, let $S_i \subseteq [1,240]$ be the set of solutions to $f_i(x) \equiv_{241} 0$. By Lemma 8.13, since $f_i(x)$ is degree 2^i we expect $|S_i| = 2^i$. Using a computer we find

$$\begin{aligned}
S_0 &= \{1\} \\
S_1 &= \{1, 240\} \\
S_2 &= \{1, 64, 177, 240\} \\
S_3 &= \{1, 8, 30, 64, 177, 211, 233, 240\} \\
S_4 &= \{1, 8, 30, 44, 64, 76, 111, 115, 126, 130, 165, 177, 197, 211, 233, 240\}
\end{aligned}$$

An element of order 2^4 must correspond to an element of S_4 that does not appear in any of S_0, S_1, S_2, S_3. In looking for such an element we notice

$$S_0 \subset S_1 \subset S_2 \subset S_3 \subset S_4$$

Let us think for a moment about why this should be true. Consider $a \in S_3$. Since $a \in S_3$ we have $a^{2^3} - 1 \equiv_{241} 0$. In other words, $a^{2^3} \equiv_{241} 1$. Since $S_3 \subset S_4$ we should be able to conclude $a^{2^4} - 1 \equiv_{241} 0$. Indeed

$$a^{2^4} - 1 \equiv_{241} a^{2^{3+1}} - 1 \equiv_{241} a^{2^3 \cdot 2} - 1 \equiv_{241} (a^{2^3})^2 - 1 \equiv_{241} (1)^2 - 1 \equiv_{241} 0$$

In this argument, there is nothing at all special about $p = 241$, S_3 and S_4. And so in general we have:

Lemma 8.14. *Let p be prime and let $t \geq 1$ be an integer. If $a \in [1, p-1]$ is a solution of $x^{q^t} - 1 \equiv_p 0$, then a is a solution of $x^{q^{t+1}} - 1 \equiv_p 0$*

Proof. If $a \in [1, p-1]$ is a solution for $x^{q^t} - 1 \equiv_p 0$, then $a^{q^t} \equiv_p 1$. Therefore

$$a^{q^{t+1}} - 1 \equiv_p a^{q^t \cdot q} - 1 \equiv_p \left(a^{q^t}\right)^q - 1 \equiv_p (1)^q - 1 \equiv_p 0$$

Therefore a is a solution for $x^{q^{t+1}} - 1 \equiv_p 0$. □

And so we arrive at our main result connecting solutions to polynomial congruences and the order of an element.

Theorem 8.15. *Let p be prime and let q^f appear as a prime-power as part of the prime power factorization of $p-1$. We have that $[a]$ has order q^f in $\mathbb{Z}_p$ if and only if a is a solution to $x^{q^f} - 1 \equiv_p 0$, but not a solution to $x^{q^{f-1}} - 1 \equiv_p 0$.*

Proof. Let p be prime, let q^f appear as a prime-power as part of the prime power factorization of $p-1$ and let $[a]$ be an element of $\mathbb{Z}_p$.

If $[a]$ has order q^f in $\mathbb{Z}_p$, then q^f is the least integer e such that $a^e \equiv_p 1$. Therefore a is a solution to $x^{q^f} - 1 \equiv_p 0$. Since e is the least such integer, for any $1 \leq e' < e$ we have $a^{e'} - 1 \not\equiv_p 0$. In particular, $a^{q^{f-1}} - 1 \not\equiv_p 0$.

Assume now $x = a$ is a solution to $x^{q^f} - 1 \equiv_p 0$, but not a solution to $x^{q^{f-1}} - 1 \equiv_p 0$. Let e be the order of $[a]$ in $\mathbb{Z}_p$ we claim $e = q^f$.

By Theorem 8.6, since $[a]^e = [1]$ and $[a]^{q^f} = [1]$ we have that e and q^f are congruent modulo e. Therefore q^f is a multiple of e. Therefore e divides q^f. The only divisors of q^f are $1, q, q^2, \ldots q^f$. Therefore $e \in \{1, q, q^2, \ldots, q^f\}$.

Since e is the order of $[a]$ we have that a is a solution for $x^e - 1 \equiv_p 0$. Since a is not a solution to $x^{q^{f-1}} - 1 \equiv_p 0$ it follows $e \neq q^{f-1}$. By Lemma 8.14 it follows that a is not a solution to $x^r - 1 \equiv_p 0$ for any $r \in \{1, q, q^2, \ldots q^{f-1}\}$. Therefore $e = q^f$. □

Let p be a prime and let q^f appear as part of the prime-power factorization of p. Theorem 8.15 fully characterizes those elements of $\mathbb{Z}_p$ that have order q^f. They correspond exactly to those $a \in [1, p-1]$ that are a solution to $x^{q^f} - 1 \equiv_p 0$, but not a solution to $x^{q^{f-1}} - 1$.

8.3 Constructing Primitive Roots Modulo p.

With all of these pieces, we are now ready to prove our theorem.

Before we do so, let us return to our example $p = 241$. For $i \in \{1, 2, 3, 4\}$ recall that S_i is the set of solutions (in $[0, 240]$) to the polynomial congruence $x^{2^i} - 1 \equiv_{241} 0$. By Theorem 8.15 the set of elements of order 2^4 in $\mathbb{Z}_{241}$ corresponds the set $S_4 \setminus S_3$. And so the set of elements of order 2^4 in $\mathbb{Z}_{241}$ is

$$\{[44], [76], [111], [115], [126], [130], [165], [197]\}$$

Recalling our strategy for finding a primitive root modulo $p = 241$ we have $p - 1 = 240 = 2^4 \cdot 3 \cdot 5$. We have found an element of order 2^4 (in fact we have found many!). And so now we proceed to find an element of order 3 and an element of order 5.

To find an element of order 3 we apply the result of Theorem 8.15 with $q^f = 3^1$. And so we find a solution to $x^{3^1} - 1 \equiv_{241} 0$ that is not a solution to $x^{3^{1-1}} - 1 \equiv_{241} 0$.

Using a computer we find the solutions to $x^3 - 1 \equiv_{241} 0$ are $\{1, 15, 225\}$. By inspection we see that 1 is a solution to $x - 1 \equiv_{241} 0$. Since this polynomial is of degree 1, Theorem 8.13 tells us that this is the only solution. Therefore the set of elements of order 3 in $\mathbb{Z}_{241}$ is $\{[15], [225]\}$

To find an element of order 5 we apply the result of Theorem 8.15 with $q^f = 5^1$. And so we find a solution to $x^{5^1} - 1 \equiv_{241} 0$ that is not a solution to $x^{5^{1-1}} - 1 \equiv_{241} 0$.

Using a computer we find that the solutions to $x^5 - 1 \equiv_{241} 0$ are $\{1, 87, 91, 98, 205\}$. By inspection we see that 1 is a solution to $x - 1 \equiv_{241} 0$. Since this polynomial is of degree 1, Theorem 8.13 tells us that this is the only solution. Therefore the set of elements of order 5 in $\mathbb{Z}_{241}$ is $\{[87], [91], [98], [205]\}$.

By Theorem 8.9, an element of order 240 in $\mathbb{Z}_{241}$ arises as a product of elements respectively of orders 2^4, 3 and 5. We compute $[44] \cdot [15] \cdot [87] = [57420] = [12]$. Therefore $[12]$ has order 240 in $\mathbb{Z}_{241}$. And so we conclude $[12]$ is a primitive root modulo 241.

By replacing $[44]$ in this product by another element of $\mathbb{Z}_{241}$ of order 2^4 we may construct other primitive roots modulo 241. Similarly we can construct other primitive roots by replacing $[15]$ and $[87]$.

We generalize. Let p be an odd prime. By Theorem 8.9 an element of order $p - 1$ in $\mathbb{Z}_p$ arises as a product of elements of order $q_i^{f_i}$ where

$$p - 1 = q_1^{f_1} q_2^{f_2} \dots q_k^{f_k}$$

and $1 \leq i \leq k$. By Lemma 8.14, to find an element of order $q_i^{f_i}$ it suffices to find a solution to the polynomial congruence $x^{q_i^{f_i}} - 1 \equiv_p 1$ that is not a solution to the polynomial congruence $x^{q_i^{f_i - 1}} - 1 \equiv_p 1$. By Lemma 8.13, there are $q_i^{f_i}$ solutions to $x^{q_i^{f_i}} - 1 \equiv_p 1$ and $q_i^{f_i - 1}$ solutions to $x^{q_i^{f_i - 1}} - 1 \equiv_p 1$. Thus there are $q_i^{f_i} - q_i^{f_i - 1} \geq 1$ elements of order $q_i^{f_i}$ in $\mathbb{Z}_p$.

Theorem. *For every prime p, $\mathbb{Z}_p$ has at least one primitive root.*

Proof. When $p = 2$, we have $\mathbb{Z}_2 = \{[0], [1]\}$. Notice $[1]$ has order $p - 1 = 2 - 1 = 1$. By Theorem 8.8, $[1]$ is a primitive root modulo 2.

Assume now p is an odd prime. By the Fundamental Theorem of Arithmetic we may express $p - 1$ as a unique product of prime powers.

$$p - 1 = q_1^{f_1} q_2^{f_2} \dots q_k^{f_k}$$

We first show that for every $1 \leq i \leq k$ there exists an element of order $q_i^{f_i}$.

By Theorem 8.15, to find an element of order $q_i^{f_i}$ it suffices to find a solution to the polynomial congruence $x^{q_i^{f_i}} - 1 \equiv_p 0$ that is not a solution to the polynomial congruence $x^{q_i^{f_i-1}} - 1 \equiv_p 0$. By Lemma 8.13, there are $q_i^{f_i}$ solutions to $x^{q_i^{f_i}} - 1 \equiv_p 0$ and $q_i^{f_i-1}$ solutions to $x^{q_i^{f_i-1}} - 1 \equiv_p 0$. Thus there are $q_i^{f_i} - q_i^{f_i-1} \geq 1$ elements of order $q_i^{f_i}$ in $\mathbb{Z}_p$.

For each $1 \leq i \leq k$ let $[a_i]$ be an element of order $q_i^{f_i}$. Let

$$[g] = \prod_{i=1}^{k} [a_i]$$

By Corollary 8.10 it follows that $[g]$ has order

$$q_1^{f_1} q_2^{f_2} \dots q_k^{f_k} = p - 1$$

Since $[g]$ has order $p - 1$, by Theorem 8.8 it follows that $[g]$ is a primitive root modulo $p - 1$. □

Consider the construction of our primitive root $[g]$ in the proof above. We have

$$[g] = \prod_{i=1}^{k} [a_i]$$

where each $[a_i]$ is an element of order $q_i^{f_i}$.

In our construction of a primitive root in $\mathbb{Z}_{241}$ we chose $[44]$ as our element of order 2^4. However, we could have chosen any element of S_4 that was not an element of S_3. Each of these choices would give rise to a different primitive root modulo 241.

In general we may choose any element of order $q_i^{f_i}$ for each $1 \leq i \leq k$ in our construction of $[g]$. By Theorem 8.15 these are the solutions to $x^{q_i^{f_i}} - 1 \equiv_p 0$ that are not solutions to $x^{q_i^{f_i-1}} - 1 \equiv_p 0$. By Theorem 8.13 there are exactly $q_i^{f_i} - q_i^{f_i-1}$ such solutions. Therefore there

are $q_i^{f_i} - q_i^{f_i-1}$ elements of order $q_i^{f_i}$. And so in general, when constructing our primitive root for each $[a_i]$ we have

$$q_i^{f_i} - q_i^{f_i-1} = q_i^{f_i}\left(1 - \frac{1}{q_i}\right)$$

choices.

This parameter looks very familiar! Recall from Module 5 we have

$$\phi(q_i^{f_i}) = q_i^{f_i}\left(1 - \frac{1}{q_i}\right)$$

Therefore there are $\phi(q_i^{f_i})$ elements of order $q_i^{f_i}$ for each $1 \leq i \leq k$. Thus when constructing $[g]$ we have $\phi(q_i^{f_i})$ choices for each $[a_i]$. And so we can construct the product $\Pi_{i=1}^{k}[a_i]$ in

$$\prod_{i=1}^{k} \phi(q_i^{f_i}) = \prod_{i=1}^{k} q_i^{f_i}\left(1 - \frac{1}{q_i}\right) = (p-1)\prod_{i=1}^{k}\left(1 - \frac{1}{q_i}\right) = \phi(p-1)$$

ways.

It turns out that each of the ways to construct this product gives rise a unique primitive root modulo p.

Corollary 8.16. *For every prime p, $\mathbb{Z}_p$ has $\phi(p-1)$ primitive roots modulo p.*

Aside 8.17. *Wow!*

Using tools from group theory, we give a proof of this remarkable fact in the following optional section.

Aside 8.18. *On the face of it, the proof of this theorem is constructive. Given solutions to the polynomial congruences of the form $x^t - 1$ for particularly useful values of t, we can construct a primitive root modulo n. Even before computing the solutions for some polynomial congruences of the form $x^t - 1$ we would need to know for which values of t we would need solutions. To find these values of t, we would need to know the prime-power factorization of $p-1$. As we saw in our discussion on cryptographic schemes, finding prime power factorizations is no easy task.*

In practice, primitive roots modulo p are found with a method that amounts to guess and check*. There is no publicly known method to efficiently find a primitive root modulo p. However, though this problem seems difficult, the problem is not known to be NP-hard.*

8.4 $\mathbb{Z}_p$ is a group with respect to multiplication (Optional)

In the last section we ended by stating there that for each prime p there are $\phi(p-1)$ primitive roots modulo p. However in our work we did not manage to fully justify this fact. Using some tools from group theory, we can verify this fact.

Let $\mathbb{Z}_p^\star$ be the set of non-zero elements of $\mathbb{Z}_p$. We denote the canonical multiplicative group on this set as $\left(\mathbb{Z}_p^\star, \cdot\right)$. (In some contexts this group is denoted as $\mathbb{Z}/p\mathbb{Z}$)

Recall Lagrange's Theorem

Theorem 8.19 (Lagrange's Theorem)**.** *Let G be a finite group and let H be a subgroup of G. The order of H divides the order of G.*

Aside 8.20. *It is annoying that order has two slightly different meanings in group theory. I can't imagine the countless hours of confusion that have arisen from using order to refer to the number of elements of a group and order to refer to the least integer e so that $g^e = 1$. I suppose the small upside getting to say* an element of order e generates a cyclic subgroup of order e.

Since $\left(\mathbb{Z}_p^\star, \cdot\right)$ is a finite group, the order of any subgroup of $\left(\mathbb{Z}_p^\star, \cdot\right)$ divides $p-1$, the number of elements in this multiplicative group. For any element $[a] \in \mathbb{Z}_p^\star$, the cyclic subgroup generated by $[a]$ is the elements of the set

$$\langle [a] \rangle = \{[a]^i \mid i \geq 1\}$$

If $[a]$ has order e, then necessarily

$$\langle [a] \rangle = \{[a]^i \mid i \geq 1\} = \{[a]^i \mid 1 \leq i \leq e\}$$

Since $[a]$ has order e, for $1 \leq i \leq e$ each element of the form $[a]^i$ is distinct. And so $\langle [a] \rangle$ has e elements. Since $\langle [a] \rangle$ is a subgroup of $\left(\mathbb{Z}_p^\star, \cdot\right)$, then, by Lagrange's Theorem, it follows that e is a divisor of $p-1$.

Since $\left(\mathbb{Z}_p^\star, \cdot\right)$ has a primitive root, it follows that $\left(\mathbb{Z}_p^\star, \cdot\right)$ is a cyclic group of order $p-1$. (A primitive root is a generator.) A cyclic group of order n is isomorphic to $(\mathbb{Z}_n, +)$. In particular, $\left(\mathbb{Z}_p^\star, \cdot\right)$ is isomorphic to $(\mathbb{Z}_{p-1}, +)$ by way of the following isomorphism.

Let $[g]$ be a generator (i.e., primitive root) of $\left(\mathbb{Z}_p^\star, \cdot\right)$. And so

$$\mathbb{Z}_p^\star = \langle [g] \rangle = \{[g]^i \mid 1 \leq i \leq p-1\}$$

Consider the function $\beta : \mathbb{Z}_p^\star \to \mathbb{Z}_{p-1}$ given by

$$\beta([g]^i) = [i]$$

Example 8.21. *Look at the exponentiation table for $\mathbb{Z}_7$ above. We observe* $[3]$ *is a primitive root modulo* 7*. Using ϕ above we have*

$$\begin{aligned}
\beta([3]) &= \beta([3]^1) = [1] \\
\beta([2]) &= \beta([3]^2) = [2] \\
\beta([6]) &= \beta([3]^3) = [3] \\
\beta([4]) &= \beta([3]^4) = [4] \\
\beta([5]) &= \beta([3]^5) = [5] \\
\beta([1]) &= \beta([3]^6) = [6] = [0]
\end{aligned}$$

Here the values on the right side of the equals sign are elements of $\mathbb{Z}_6$ as $\beta : \mathbb{Z}_7^\star \to \mathbb{Z}_6$

We claim this function is a group isomorphism. To show this, we check

$$\beta\left([a] \cdot [b]\right) = \beta\left([a]\right) + \beta\left([b]\right)$$

for every $[a], [b] \in \mathbb{Z}_p^\star$.

Since $[a], [b] \in \mathbb{Z}_p^\star$ and $\mathbb{Z}_p^\star$ is generated by $[g]$ it follows that there exists $1, \leq i, j \leq p-1$ so that $[g]^i = [a]$ and $[g]^j = [b]$. Therefore

$$\beta\left([a] \cdot [b]\right) = \beta\left([g]^i \cdot [g]^j\right) = \beta\left([g]^{i+j}\right) = i + j = \beta([g]^i) + \beta([g]^j) = \beta\left([a]\right) + \beta\left([b]\right)$$

Therefore β is an isomorphism.

Since β is an isomorphism, the order of the element $[a] \in \mathbb{Z}_p^\star$ is equal to the order of the element $\beta([a]) \in \mathbb{Z}_{p-1}$. And so the image (with respect to β) of a generator in $\left(\mathbb{Z}_p^\star, \cdot\right)$ must generate $(\mathbb{Z}_{p-1}, +)$ (and vice versa). And so to count the number of primitive roots modulo p it suffices to count the number of generators in $(\mathbb{Z}_{p-1}, +)$.

The order of an element $[a] \in \mathbb{Z}_n$ in $(\mathbb{Z}_n, +)$ is least integer e such that $e[a] = [0]$. (This is an additive group, not multiplicative group, and so our definition of order changes accordingly.) That is, it is the least integer e so that ea is a multiple of n. That is, it is the least integer e so that $lcm(a, n) = ea$. Recall

$$gcd(a, n) \cdot lcm(a, n) = an$$

Therefore

$$e = \frac{n}{gcd(a, n)}$$

We have $[a]$ generates $(\mathbb{Z}_n, +)$ if and only if has order n. Therefore $[a]$ generates $(\mathbb{Z}_n, +)$ if and only if $gcd(a, n) = 1$. And so $(\mathbb{Z}_n, +)$ has $\phi(n)$ generators.

Since

$$\left(\mathbb{Z}_p^\star, \cdot\right) \cong (\mathbb{Z}_{p-1}, +)$$

it then follows that there are $\phi(p-1)$ primitive roots modulo p for every prime p. Further, we have

Theorem 8.22. *Let p be prime and let $[g]$ be a primitive root modulo p. An element $[g'] \in \mathbb{Z}_p$ is a primitive root modulo p if and only if*

$$[g'] \in \{[g]^i \mid gcd(i, p-1) = 1\}$$

8.5 Further Exercises

1. Let p be a prime. Show that $[p-1]$ is only element of its order in $\mathbb{Z}_p$.
2. Find the order of $[2]$ in $\mathbb{Z}_{17}$.
3. Let p be a prime and let $[a] \in \mathbb{Z}_p$ have order e. Show $[a]^{e/2} = [p-1]$
4. Find a second primitive root modulo 241.

---

9 Primitive Roots Part II

Learning Incomes.

- *Understand the definition of primitive root modulo p*
- *Recall the meaning of the notation $[n-1]!$ and recognize*

Learning Outcomes.

- *Be able to, given a primitive root modulo p, find a solution (if one exists) to the polynomial congruence of the form $x^t \equiv_p a$*
- *Understand the relationship between the primality of n and the product $[n-1]!$ in $\mathbb{Z}_n$.*
- *Be able to determine, given n, if there exists a primitive root modulo n.*

Newly Defined Terms and Notation. *primitive root modulo n, order*

Recall the definition of a primitive root modulo p:

Definition 9.1. *Let p be prime and let $[g] \in \mathbb{Z}_p$. We say $[g]$ is a primitive root modulo p when for every $[a] \in \mathbb{Z}_p$ with $[a] \neq [0]$ there exists $e \in [1, p-1]$ so that $[g]^e = [a]$ in $\mathbb{Z}_p$.*

Informally, a primitive root $[g]$ is an element of $\mathbb{Z}_p$ so that every non-zero element of $\mathbb{Z}_p$ can be expressed as a power of $[g]$. In the previous module we proved that primitive roots were exactly the elements of order $p-1$ in $\mathbb{Z}_p$.

In this module we first look how primitive roots are a useful tool in studying polynomial congruences modulo p. We then return to our proof Fermat's Theorem and devise a primality test using the product $[n-1]!$. Finally, we conclude our study of primitive roots by considering primitive roots for non-prime moduli.

9.1 Solving Polynomial Equations Using Primitive Roots

We find an application for primitive roots in the search for solutions to polynomial congruences modulo p. Consider the polynomial congruence

$$x^7 \equiv_{11} 9$$

In the previous module we saw that [6] is a primitive root modulo 11. We compute the powers of [6] in $\mathbb{Z}_{11}$.

$[z]$	$[z]^2$	$[z]^3$	$[z]^4$	$[z]^5$	$[z]^6$	$[z]^7$	$[z]^8$	$[z]^9$	$[z]^{10}$
[6]	[3]	[7]	[9]	[10]	[5]	[8]	[4]	[2]	[1]

In $\mathbb{Z}_{11}$ every non-zero element can be expressed as $[6]^z$ for some $1 \leq z \leq 10$. Therefore in $\mathbb{Z}_{11}$ we may express the equation $[x]^7 = [9]$ as

$$\begin{aligned} ([6]^z)^7 &= [6]^4 \\ [6]^{7z} &= [6]^4 \end{aligned}$$

By Theorem 8.6, we have $[6]^{7z} = [6]^4$ if and only if $7z$ and 4 are congruent modulo 10. Recall we write $7z \equiv_{10} 4$, to mean 10 divides $4 - 7z$. That is, there exists an integer y so that $10y = 4 - 7z$. Rearranging yields

$$7z + 10y = 4$$

An integer solution (z_0, y_0) to this linear diophantine equation yields z_0 so that $([6]^{z_0})^7 = [6]^4$. In other words, it yields a solution $[x] = [6]^{z_0}$ for the equation $[x]^7 = [9]$.

Since $gcd(7, 10) = 1$, this linear diophantine equation has a solution. Using our work from previous modules we find a solution $(12, -8)$. Since 12 is congruent to 2 modulo 10 we find

$$[6]^{12} = [6]^2 = [3]$$

Therefore $[3]^7 = [9]$. And so $x = 3$ is a solution to $x^7 \equiv_{11} 9$.

As our work in linear diophantine equations tells us when such an equation does not have an integer solution, we can show that some polynomial congruences have no solution with this same technique. For example, consider the polynomial congruence

$$x^5 \equiv_{11} 4$$

We have $[3]^8 = [4]$ and so we may express this congruence as

$$[3]^{5z} = [3]^8$$

Following our reasoning above, to find z it suffices to solve the linear diophantine equation

$$5z + 10y = 8$$

However since $gcd(5, 10) = 5$ and 5 does not divide 8, this linear diophantine equation has no solution. And so $x^5 \equiv_{11} 4$ has no solution.

Polynomial congruences were first introduced in Module 6. After only a quick look at quadratic congruences we concluded that our standard tools for thinking about quadratic things (i.e., the quadratic equation) would not suffice as square roots didn't always exist. That is, for any $a \in [1, p-1]$ there was no guarantee that there existed $x \in [1, p-1]$ so that $x^2 \equiv_p a$. Our work above seemingly gives us a route forward in this task. Using a primitive root modulo p we may be able to determine for which $x \in [1, p-1]$ the polynomial $x^2 \equiv_p a$ has a solution. We return to this thought in Module 10 when we revisit quadratic equations in $\mathbb{Z}_p$.

9.2 Primality Testing With Primitive Roots

Let p be an odd prime. In our proof of Fermat's Little Theorem we encountered the product.

$$[p-1]! = [1] \times [2] \times \cdots \times [p-1]$$

As p is prime, there exists at least one primitive root modulo p. Let $[g]$ be a primitive root modulo p. Since $[g]$ is a primitive root we have

$$\{[g]^i \mid 1 \leq i \leq p-1\} = \{[1], [2], \ldots, [p-1]\}$$

Thus we may express $[p-1]!$ as

$$\begin{aligned}
[p-1]! &= [1] \times [2] \times \cdots \times [p-1] && (1)\\
&= [g] \times [g]^2 \times \cdots \times [g]^{p-2} \times [g]^{p-1} && (2)\\
&= [g] \times [g]^2 \times \cdots \times [g]^{p-2} \times [1] && (3)\\
&= [g] \times [g]^2 \times \cdots \times [g]^{p-2} && (4)
\end{aligned}$$

(The terms in this product aren't in the same order on line (1) and line (2). In the first line we are taking the product of the elements $\{[1], [2], \ldots, [p-1]\}$. In the second line we are taking the product of the elements of $\{[g]^i \mid 1 \leq i \leq p-1\}$. Since these two sets are the same, it follows that the two products are the same.)

Simplifying yields

$$\begin{aligned}
[p-1]! &= [g] \times [g]^2 \times \cdots \times [g]^{p-2}\\
&= [g]^{1+2+3+\cdots+(p-2)}\\
&= [g]^{\frac{(p-2)(p-1)}{2}}\\
&= \left([g]^{\frac{p-1}{2}}\right)^{p-2}
\end{aligned}$$

Let $[a] = [g]^{\frac{p-1}{2}}$. Notice

$$[a]^2 = \left([g]^{\frac{p-1}{2}}\right)^2 = [g]^{p-1} = [1]$$

Therefore a is a solution to

$$x^2 - 1 \equiv_p 0$$

This polynomial congruence is of degree 2 and is of the form $x^k - 1 \equiv_p 0$. And so by Lemma 8.13 it has exactly two solutions. By inspection we find 1 and -1. Since $-1 \equiv_p p-1$ it must be $[a] = [1]$ or $[a] = [p-1]$.

Recall $[g]$ has order $p-1$. (It is a primitive root modulo p.) Therefore $[g]^t \neq [1]$ for all $1 \leq t < p-1$. As

$$1 \leq \frac{p-1}{2} < p-1$$

it cannot be that $[a] = [g]^{\frac{p-1}{2}} = [1]$. Therefore $[a] = [p-1]$.

Recall that we chose p to be an odd prime. Therefore $p-2$ is an odd integer. And so

$$\begin{aligned}[p-1]! &= \left([g]^{\frac{p-1}{2}}\right)^{p-2} \\ &= [p-1]^{p-2} \\ &= [-1]^{p-2} \\ &= [-1] \\ &= [p-1]\end{aligned}$$

Putting this altogether, we see that when p is an odd prime we have $[p-1]! = [p-1]$. When $p = 2$, similarly we have $[2-1]! = [1] = [2-1]$. And so

if p is a prime, then $(p-1)! \equiv_p -1$.

Much as considering the contrapositive of Fermat's Little Theorem gives rise to a primality test, so too does the contrapositive of the previous sentence. After some massaging, we can express the contrapositive as:

Let $n > 1$ be an integer. If the remainder when $(n-1)!$ is divided by n is not $n-1$, then n is not prime.

When we looked as primality testing with Fermat's Little Theorem we ran into the problem of pseudo-primes: those composite numbers n for which $2^{n-1} \equiv_n 1$. The existence of these pseudo-primes confirmed to us that we could not use Fermat's Little Theorem to determine if an integer was prime. We wonder if the same phenomenon occurs here. That is, we wonder if is possible to have $(n-1)! \equiv_n -1$ for a composite integer n.

If n is composite, then there exists $a, b \in [2, n-1]$ so that $n = ab$. If a and b are distinct, then each of them will appear in the product $(n-1)!$. In this case we see that ab divides $(n-1)!$ and so $(n-1)! \equiv_n 0$.

Otherwise if n can only be expressed as a product of a pair of identical factors, it follows that $n = p^2$ for some prime p. (This fact is not immediate, but admits a short proof.) In this case, however, we notice that as long as we have $n > 4$, then p and $2p$ will appear as terms in the product $(n-1)!$. (Again, not immediate, but admits a short proof.) And so again we see $(n-1)! \equiv_n 0$.

Putting this all together yields the following theorem.

Theorem 9.2. *An integer $n \geq 5$ is prime if and only if*

$$(n-1)! \equiv_n -1$$

Aside 9.3. *This number theoretic fact astounds me any time I see it. Once primitive roots are in place, the proof is not overly complicated. But yet, the conclusion is entirely unexpected.*

This theorem is usually called Wilson's Theorem*, named for 18th Century English mathematician John Wilson. In the grand tradition of mathematics, Wilson did not not actually prove this theorem.*

In the other grand tradition of (western) mathematics, though this theorem is named for Wilson, these ideas appeared nearly 1000 years earlier in work by an Indian mathematician Ibn al-Haytham. The breadth of Ibn al-Haytham's contributions to science and mathematics are overwhelming, including work on optics, the movement of celestial bodies, number theory and non-euclidean geometry.

Though the statement of theorem is an nice if and only if *classification of prime numbers, it does not lead to an efficient (i.e., polynomial) algorithm for testing if a number is prime. The product* $(n-1)!$ *is exponentially large (as a function of* n*) (Stirling's Approximation:* $k! \sim \sqrt{2\pi k}\left(\frac{k}{e}\right)^k$*)*

9.3 Primitive Roots Modulo n.

Continuing our trend in this course, after finding an answer when n is prime, we turn to considering the general case. Almost immediately we find that when n is composite, we have no hope of expressing every non-zero element of $\mathbb{Z}_n$ as a power of some $[g] \in \mathbb{Z}_n$.

Recall the following definition.

Definition 9.4. *Let $n > 1$ be an integer. The group of units modulo n is the subset of $\mathbb{Z}_n$ containing all of the elements of $\mathbb{Z}_n$ that have a multiplicative inverse. We denote this set as U_n.*

Further recall that $|U_n| = \phi(n)$.

For a composite integer n we can partition the non-zero elements of $\mathbb{Z}_n$ into those that have a multiplicative inverse (i.e., the elements of U_n) and those that do not. As n is composite we have, $U_n \neq \mathbb{Z}_n \setminus \{0\}$. As we showed in Module 6, the product of any pair of elements of U_n is again an element of U_n. And so for $[a] \in U_n$ we have $[a]^k \in U_n$ for every $k \geq 1$. Therefore there are elements of $\mathbb{Z}_n$ that cannot be expressed as a power of $[a]$ – namely those $[d] \in \mathbb{Z}_n$ so that $gcd(d, n) \neq 1$

Consider now a non-zero $[b] \in \mathbb{Z}_n$ so that $[b] \notin U_n$. Since $[b] \notin U_n$ necessarily $gcd(b, n) \neq 1$. Therefore b and n have at least one common prime factor. Therefore b^k and n have at least one common prime factor for every $k \geq 1$. Therefore $gcd(b^k, n) \neq 1$. And so $[b^k] = [b]^k \notin U_n$. Therefore there are elements of $\mathbb{Z}_n$ that cannot be expressed as power of $[b]$ – namely those $[d] \in \mathbb{Z}_n$ so that $gcd(d, n) = 1$.

And so we see that for every non-zero $[c] \in \mathbb{Z}_n$ there exist $[d] \in \mathbb{Z}_n$ so that $[d]$ cannot be expressed as a power of $[c]$.

We manoeuvre around this problem by restricting our attention to U_n. Looking back at our work at the end of Section 8.0, we see we have made liberal use of multiplicative inverses. And so when thinking about primitive roots modulo n when n is composite it make sense to restrict our attention to those elements of $\mathbb{Z}_n$ that have a multiplicative inverse.

Definition 9.5. *Let $n > 1$ be an integer and let $[g] \in U_n$. We say $[g]$ is a primitive root modulo n when for every $[a] \in U_n$ there exists $e \in [1, \phi(n)]$ so that $[g]^e = [a]$ in U_n.*

Consider $n = 9$. We find

$$U_n = \{[1], [2], [4], [5], [7], [8]\}$$

We compute the powers of these elements in $\mathbb{Z}_9$.

$[z]$	$[z]^2$	$[z]^3$	$[z]^4$	$[z]^5$	$[z]^6$
[1]	[1]	[1]	[1]	[1]	[1]
[2]	[4]	[8]	[7]	[5]	[1]
[4]	[7]	[1]	[4]	[7]	[1]
[5]	[7]	[8]	[4]	[2]	[1]
[7]	[4]	[1]	[7]	[4]	[1]
[8]	[1]	[8]	[1]	[8]	[1]

We notice $U_9 = \{[2]^i \mid 1 \leq i \leq \phi(9)\}$ and so [2] is a primitive root modulo 9.

Just as we did when we studied primitive roots modulo p, we analogously define order for elements of U_n.

Definition 9.6. *Let $n > 1$ be an integer and let $[a]$ be an element of U_n. The order of $[a]$ is the least integer e so that $[a]^e = [1]$.*

Just as Fermat's Little Theorem implied that every non-zero element of $\mathbb{Z}_p$ has an order, so too does Euler's Theorem imply that every element of U_n has an order. And so we may generalize our results in Section 8.1

Theorem 9.7. *Let $n > 1$ be an integer, and let $[a]$ be an element of U_n with order e. For any integers $i, j > 1$ we have $[a]^i = [a]^j$ if and only if i and j are congruent modulo e.*

Corollary 9.8. *Let $n > 1$ be an integer. For any non-zero $[a] \in U_n$ the order of $[a]$ divides $\phi(n)$.*

Theorem 9.9. *Let $n > 1$ be an integer and let $[a]$ be an element of U_n. We have that $[a]$ is a primitive root modulo n if and only if $[a]$ has order $\phi(n)$*

Theorem 9.10. *Let $n > 1$ be an integer. Let $[a]$ and $[b]$ be non-zero elements of U_n. Let e and f respectively be the orders of $[a]$ and $[b]$. If e and f are coprime, then $[a][b]$ has order ef.*

Unfortunately, this is where our progress stops. Theorem 8.12 for counting the number of solutions to polynomial congruences modulo p does not extend modulo n. And so our strategy for constructing primitive roots modulo p does not generalize to constructing primitive roots modulo n.

Aside 9.11. *Thinking about why the proof of Theorem 8.12 fails in U_n is an excellent final exam question.*

In attempt to make some progress for composite modulus we consider the case where $n = 2^k$ for $k \geq 1$. For any integer $a \geq 1$ we have $gcd(a, 2^k) \in \{1, 2, 2^2, \ldots, 2^k\}$. And so $gcd(a, 2^k) = 1$ if and only if a is odd. Therefore

$$U_{2^k} = \{[1], [3], [5], \ldots, [2^k - 1]\}$$

When $n = 2^2$ we have

$$U_4 = \{[1], [3]\}$$

Since $[3]^2 = [1]$ in $\mathbb{Z}_4$ we have that $[3]$ is a primitive root modulo 4.

Consider the exponentiation table for U_8:

$[z]$	$[z]^2$	$[z]^3$	$[z]^4$
$[1]$	$[1]$	$[1]$	$[1]$
$[3]$	$[1]$	$[3]$	$[1]$
$[5]$	$[1]$	$[5]$	$[1]$
$[7]$	$[1]$	$[7]$	$[1]$

By observation we see that no element of U_8 has order $\phi(8) = 4$ and so there is no primitive root modulo 8.

By Theorem 9.9, a primitive root modulo 2^k is an element of U_{2^k} of order $\phi(2^k) = 2^{k-1}$. By our definition of order, if $[g] \in U_{2^k}$ is a primitive root modulo 2^k, then 2^{k-1} is the least integer e so that $g^e \equiv_{2^k} 1$. We show in fact that that for all integers $k \geq 3$ and all odd integers $a \in [1, 2^k - 1]$ we have $a^{2^{k-2}} \equiv_{2^k} 1$. This then implies that each element of U_{2^k} has order no more than 2^{k-2}. Which in turn implies that there are no primitive root modulo 2^k.

Theorem 9.12. *For all integers $k \geq 3$ and all odd integers $a \in [1, 2^k - 1]$ we have $a^{2^{k-2}} \equiv_{2^k} 1$.*

Proof. We proceed by induction on k. When $k = 3$ we have $2^{k-2} = 2^1 = 2$ and so we compute $[a]^2$ for each $[a] \in U_8$

$$1^2 \equiv_8 1$$
$$3^2 \equiv_8 9 \equiv_8 1$$
$$5^2 \equiv_8 25 \equiv_8 1$$
$$7^2 \equiv_8 49 \equiv_8 1$$

Assume now the claim holds when $k = t$. That is, assume for every odd integer $a \in [1, 2^k - 1]$ we have

$$a^{2^{t-2}} \equiv_{2^t} 1$$

Since $a^{2^{t-2}} \equiv_{2^t} 1$, there exists an integer q so that $a^{2^{t-2}} = q2^t + 1$. And so

$$a^{2^{(t+1)-2}} \equiv_{2^{t+1}} a^{2^{(t-2)+1}} \equiv_{2^{t+1}} a^{2^{(t-2)} \cdot 2} \equiv_{2^{t+1}} (q2^t + 1)^2 \equiv_{2^{t+1}} q^2 2^{t+1} + q2^{t+1} + 1 \equiv_{2^{t+1}} 1$$

Therefore $a^{2^{(t+1)-2}} \equiv_{2^{t+1}} 1$. The result follows by induction. $\square$

Theorem 9.12 gives an upper bound on the order of an element in U_{2^k} when $k \geq 3$. Since $a^{2^{k-2}} \equiv_{2^k} 1$, it follows that the order of a is bounded above by 2^{k-2}. Since $2^{k-2} < \phi(2^k)$, no element has order $\phi(2^k)$ in U_{2^k}.

Corollary 9.13. *There exists a primitive root modulo 2^k if and only if $k = 1, 2$.*

Turning to powers of odd primes, we find the opposite behaviour.

Theorem 9.14. *For every odd prime p and every integer $p > 1$ there exists a primitive root modulo p^k.*

We omit the proof of this fact, as time in this course is beginning to run, short and the proof require some tools we did not have the opportunity to develop this semester. In this spirit, we also state the main result of this section without proof or intuition. (Sorry!)

Theorem 9.15. *Let $n > 1$ be an integer. The group of units, U_n has a primitive root if and only if*

1. *$n = 2^k$ for $k = 1, 2$,*
2. *$n = p^k$ for $k \geq 1$ where p is an odd prime, or*
3. *$n = 2p^k$ for $k \geq 1$ where p is an odd prime.*

Much like the proof of our main result in the previous module, the proof of this theorem is constructive-ish. In proving this result one proceeds by first finding [g] a primitive root modulo p and then showing that at least one of $[g]$ or $[g+p]$ is a primitive root modulo p^2. From there, one shows that any primitive root modulo p^2 is also a primitive root modulo p^k for any $k \geq 3$.

9.4 $\mathbb{U}_n$ is a group with respect to multiplication (Optional)

In our optional section in Module 8 we saw how we can use the tools from group theory to talk about primitive roots modulo p. In this work we realize that primitive roots modulo p corresponded to a generator in a cyclic group. Using the isomorphism $\beta : \mathbb{Z}_p^\star \to \mathbb{Z}_{p-1}$ given by $\beta([g]^i) = [i]$ we proved that that for every prime there are $\phi(p-1)$ primitive roots modulo p. (Here $\mathbb{Z}_{p-1}$ is the canonical additive group on this set.) Since $\phi(p) = p-1$, we may write:

For every prime p there are $\phi(\phi(p))$ primitive roots modulo p.

Replacing $\mathbb{Z}_p^\star$ with U_n and $p-1$ with $\phi(n)$ yields an identical result by the same method. In particular, we observe that when U_n has a primitive root, we have $U_n \cong \mathbb{Z}_{\phi(n)}$ where $\mathbb{Z}_{\phi(n)}$ denotes the canonical additive group on this set. Such an isomorphism maps generators (i.e., primitive roots) in the multiplicative group U_n to generators in the additive group $\mathbb{Z}_{\phi(n)}$.

Following our argument from the previous section, the generators of the additive group $\mathbb{Z}_{\phi(n)}$ are exactly those $[a] \in \mathbb{Z}_{\phi(n)}$ so that $gcd(a, \phi(n)) = 1$. By definition there are exactly $\phi(\phi(n))$ such elements of $\mathbb{Z}_{\phi(n)}$. And so we have:

Theorem 9.16. *For every $n \geq 2$, if there exists a primitive root modulo n, then there are exactly $\phi(\phi(n))$ primitive roots modulo n.*

Aside 9.17. *Wow!*

---

10 Quadratic Residues Modulo p

Learning Incomes.

- *Recall the meaning of primitive root modulo p.*

Learning Outcomes.

- *Understand notation related to quadratic residues*
- *Understand which powers of a primitive root are a quadratic residue.*
- *Be able to use Euler's Criterion to determine if an element of $\mathbb{Z}_p$ is a quadratic residue*

Newly Defined Terms and Notation. *quadratic residue modulo p., Legendre Symbol,* $\left(\frac{a}{p}\right)$

Back in Module 5 we looked at techniques to solve non-linear congruences modulo n. We were stymied early on as we realize that the quadratic formula didn't nicely generalize to $\mathbb{Z}_n$; there was no guarantee that $\sqrt{[a]}$ existed. We saw no method to be able to discern for which $[r]$ the congruence $x^2 \equiv_n r$ had a solution.

To consider this problem more carefully, we define the following terminology.

Definition 10.1. *Let p be prime and let $[r]$ be a non-zero element of $\mathbb{Z}_p$ We say $[r]$ is a quadratic residue modulo p when there exists $[a] \in \mathbb{Z}_p$ so that $[a]^2 = [r]$. We denote the set of quadratic residues as QR_p.*

To start our exploration let us first compute QR_p for some small values of p. By definition we have

$$QR_p = \{[a]^2 \mid [a] \neq [0]\}$$

And so to find the set of quadratic residues modulo p it suffices to compute $[a]^2$ for each $[a] \neq [0]$.

$\mathbb{Z}_3$

$[a]$	$[1]$	$[2]$
$[a]^2$	$[1]$	$[1]$

$QR_3 = \{[1]\}$

$\mathbb{Z}_5$

$[a]$	$[1]$	$[2]$	$[3]$	$[4]$
$[a]^2$	$[1]$	$[4]$	$[4]$	$[1]$

$QR_5 = \{[1], [4]\}$

$\mathbb{Z}_7$

$[a]$	$[1]$	$[2]$	$[3]$	$[4]$	$[5]$	$[6]$
$[a]^2$	$[1]$	$[4]$	$[2]$	$[2]$	$[4]$	$[1]$

$QR_7 = \{[1], [2], [4]\}$

$\mathbb{Z}_{11}$

$[a]$	$[1]$	$[2]$	$[3]$	$[4]$	$[5]$	$[6]$	$[7]$	$[8]$	$[9]$	$[10]$
$[a]^2$	$[1]$	$[4]$	$[9]$	$[5]$	$[3]$	$[3]$	$[5]$	$[9]$	$[4]$	$[1]$

$QR_{11} = \{[1], [3], [4], [5], [9]\}$

$\mathbb{Z}_{13}$

$[a]$	$[1]$	$[2]$	$[3]$	$[4]$	$[5]$	$[6]$	$[7]$	$[8]$	$[9]$	$[10]$	$[11]$	$[12]$
$[a]^2$	$[1]$	$[4]$	$[9]$	$[3]$	$[12]$	$[10]$	$[10]$	$[12]$	$[3]$	$[9]$	$[4]$	$[1]$

$$QR_{13} = \{[1], [3], [4], [9], [10], [12]\}$$

$\mathbb{Z}_{17}$

$[a]$	$[1]$	$[2]$	$[3]$	$[4]$	$[5]$	$[6]$	$[7]$	$[8]$	$[9]$	$[10]$	$[11]$	$[12]$	$[13]$	$[14]$	$[15]$	$[16]$
$[a]^2$	$[1]$	$[4]$	$[9]$	$[16]$	$[8]$	$[2]$	$[15]$	$[13]$	$[13]$	$[15]$	$[2]$	$[8]$	$[16]$	$[9]$	$[4]$	$[1]$

$$QR_{17} = \{[1], [2], [4], [8], [9], [13], [15], [16]\}$$

Some interesting patterns emerge:

(1) $[a]^2 = [p-a]^2$.

(2) For fixed p, each quadratic residue occurs twice when we compute $[a]^2$ over all $[a] \neq [0]$.

(3) For fixed p, the total number of quadratic residues is $(p-1)/2$.

(4) $[p-1]$ appears as a quadratic residue each time $p \equiv_4 1$

To verify (1) we notice

$$\begin{aligned}
[p-a]^2 &= [(p-a)^2] \\
&= [p^2 - 2pa + a^2] \\
&= [p]^2 + [-2pa] + [a]^2 \\
&= [0] + [0] + [a]^2 \\
&= [a]^2
\end{aligned}$$

Let $p \geq 3$ be prime and let $[r] \in QR_p$. Consider the congruence

$$x^2 \equiv_p r$$

Since $[r] \in QR_p$ there exists $[a] \in \mathbb{Z}_p$ so that $[a]^2 = [r]$. From our work above, if $x = a$ is a solution to this congruence, then so is $x = p - a$. Thus the congruence $x^2 - r \equiv_p 0$ has at least two solutions.

Recall the following theorem from Module 8.

Theorem. *Let p be a prime. The polynomial congruence*

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \equiv_p 0$$

has at most n distinct solutions in the range $[0, p-1]$.

This theorem implies $x^2 - r \equiv_p 0$ has at most 2 solutions. Therefore $x^2 - r \equiv_p 0$ has exactly 2 solutions. And so $[r]$ appears twice as a square of a non-zero element of $\mathbb{Z}_p$. This verifies observation (2) above.

The verify (3) we recall the following fact from Assignment 4:

Theorem 10.2. *Let $p \geq 3$ be a prime and let $[g]$ be a primitive root modulo p. The congruence $x^2 \equiv_p a$ has a solution if and only if $[a]$ can be expressed as an even power of $[g]$ in $\mathbb{Z}_p$.*

In effect this theorem fully classifies the quadratic residues modulo p.

Theorem 10.3. *Let $p \geq 3$ be a prime and let $[g]$ be a primitive root modulo p. We have*

$$QR_p = \{[g]^2, [g]^4, [g]^6 \ldots, [g]^{p-3}, [g]^{p-1}\}$$

When we list out the powers of a primitive root in order, the list alternates between quadratic residues and quadratic non-residues.

$$[g], [g]^2, [g]^3 \ldots, [g]^{p-1}$$

Every second element of this sequence is a quadratic residue and so $|QR_p| = \frac{p-1}{2}$. This verifies observation (3).

Using Theorem 10.3 we verify observation (4). In Assignment 4 you showed that for all primes $p \geq 3$ that $[p-1] = -[1]$ is the unique element of $\mathbb{Z}_p$ of order 2.

Let us consider where in the sequence

$$[g], [g]^2, \ldots, [g]^{p-1}$$

the element $-[1]$ can appear. Since $-[1]$ has order 2 we want to find $1 \leq i < p-1$ so that $([g]^i)^2 = [g]^{2i} = [1]$. Notice that for $i = \frac{p-1}{2}$ we have

$$[g]^{2 \cdot \frac{p-1}{2}} = [g]^{p-1} = [1]$$

And so we conclude $[g]^{\frac{p-1}{2}} = -[1]$.

By Theorem 10.3, we have that $-[1] \in QR_p$ if and only if $\frac{p-1}{2}$ is even. Observe that for $p \geq 3$ we have that $\frac{p-1}{2}$ is even if and only if $p-1$ is divisible by 4. In other words, for $p \geq 3$, $[p-1]$ appears as a quadratic residue if and only if $p \equiv_1 4$. This verifies observation (4).

Using Theorem 10.3, for a prime p, it is straightforward to compute the set of quadratic residues once we have a primitive root. However, as we discussed in Module 9, our method

for finding a primitive root modulo p requires us to have a prime-power factorization of $p-1$. As finding such a factorization is computationally infeasible, we consider some other methods of determining when an element of $\mathbb{Z}_p$ is a quadratic residue modulo p. For this end we define the following notation.

Definition 10.4. *Let p be an odd prime and let $a \in \mathbb{Z}$. The Legendre symbol of a is*

$$\left(\frac{a}{p}\right) = \begin{cases} 0 & gcd(a,p) \neq 1 \\ 1 & [a] \in QR_p \\ -1 & otherwise \end{cases}$$

For example, we have $\left(\frac{2}{7}\right) = 1$ since $[2] \in QR_7$ and $\left(\frac{10}{7}\right) = -1$ as $[10] = [3]$ in $\mathbb{Z}_7$ and $[3] \notin QR_7$.

Let us combine the Legendre symbol with the statement of Theorem 10.3. In the sequence

$$[g], [g]^2, [g]^3, \ldots, [g]^{p-1}$$

every second element is a quadratic residue. And so when we consider the corresponding sequence of Legendre symbols we find

$$\left(\frac{g}{p}\right), \left(\frac{g^2}{p}\right), \left(\frac{g^3}{p}\right), \left(\frac{g^4}{p}\right), \ldots, \left(\frac{g^{p-2}}{p}\right), \left(\frac{g^{p-1}}{p}\right) = -1, 1, -1, 1 \ldots, -1, 1$$

From this observation arises the following lemma.

Lemma 10.5. *If p is an odd prime and $[g]$ is a primitive root modulo p, then for all $i \geq 1$ we have*

$$\left(\frac{g^i}{p}\right) = (-1)^i.$$

We omit the proof of this result because I need to find something substantive from this module to ask you about on the final exam.

Though this lemma is statement about primitive roots, it gives rise to a test for quadratic residues that does not require us to know a primitive root. For any non-zero $[a] \in \mathbb{Z}_p$ we may express $[a] = [g]^i$ for a primitive root $[g]$ and some $1 \leq i \leq p-1$. Consider the element $[a]^{\frac{p-1}{2}}$.

$$[a]^{\frac{p-1}{2}} = \left([g]^i\right)^{\frac{p-1}{2}} = \left([g]^{\frac{p-1}{2}}\right)^i = [-1]^i$$

(Recall from above we have $[g]^{\frac{p-1}{2}} = [-1]$)

When i is even, we have $[a] \in QR_p$, $\left(\frac{a}{p}\right) = 1$, and $[a]^{\frac{p-1}{2}} = [-1]^i = [1]$. Similarly when i odd, we have $[a] \notin QR_p$, $\left(\frac{a}{p}\right) = -1$, and $[a]^{\frac{p-1}{2}} = [-1]^i = [-1]$. And so:

Theorem 10.6 (Euler's Criterion)**.** *For an odd prime p we have*

$$\left(\frac{a}{p}\right) \equiv_p a^{\frac{p-1}{2}}$$

Proof. When $gcd(a, p) \neq 1$ we have $\left(\frac{a}{p}\right) = 0$ and $a^{\frac{p-1}{2}} \equiv_p 0^{\frac{p-1}{2}} \equiv_p 0$. Therefore $\left(\frac{a}{p}\right) \equiv_p a^{\frac{p-1}{2}}$.

Otherwise, assume $gcd(a, p) = 1$. Without loss of generality, assume $1 \leq a \leq p-1$. Let $[g]$ be a primitive root modulo p. Therefore there exists $1 \leq i \leq p-1$ so that $[a] = [g]^i$. In other words we have $a \equiv_p g^i$. And so

$$a^{\frac{p-1}{2}} \equiv_p \left(g^i\right)^{\frac{p-1}{2}} = \left(g^{\frac{p-1}{2}}\right)^i$$

From our work above, we have that $[g]^{\frac{p-1}{2}} = -[1]$. And so

$$a^{\frac{p-1}{2}} \equiv_p (-1)^i$$

By Lemma 10.5 we have $(-1)^i = \left(\frac{g^i}{p}\right)$. Therefore

$$\left(\frac{a}{p}\right) \equiv_p a^{\frac{p-1}{2}}$$

□

Using Euler's Criterion we can test if $[a]$ is a quadratic residue modulo p by computing $[a]^{\frac{p-1}{2}}$. For example, consider $[2] \in \mathbb{Z}_{17}$.

$$\begin{aligned} 2^{\frac{16}{2}} &= 2^8 \\ &\equiv_{17} 2^8 \\ &\equiv_{17} (2^4)^2 \\ &\equiv_{17} (16)^2 \\ &\equiv_{17} (-1)^2 \\ &\equiv_{17} 1 \end{aligned}$$

By Euler's Criterion we have $\left(\frac{2}{17}\right) = 1$. Therefore $[2]$ is a quadratic residue modulo 17.

Aside 10.7. *Notice that Euler's Criterion is not constructive. It tells us that $x^2 - 2 \equiv 0$ has a solution but does not tell us how to find it.*

We close our discussion on quadratic residues with a brief look at one of the highlights in the history of number theory – the Law of Quadratic Reciprocity. Let p and q be odd primes. Thinking about the Legendre symbols $\left(\frac{p}{q}\right)$ and $\left(\frac{q}{p}\right)$ we see no reason why their values should be related. Knowing information about elements of $\mathbb{Z}_q$ seemingly tells us nothing about elements of $\mathbb{Z}_p$. And yet,

Theorem 10.8 (Law of Quadratic Reciprocity)**.** *If p and q are distinct odd primes, then*

$$\left(\frac{q}{p}\right) = \left(\frac{p}{q}\right)$$

except when $p \equiv_4 q \equiv_4 3$. In which case,

$$\left(\frac{q}{p}\right) = -\left(\frac{p}{q}\right)$$

For example, consider $p = 23$ and $q = 5$. Since $5 \equiv_4 1$ we have

$$\left(\frac{5}{23}\right) = \left(\frac{23}{5}\right)$$

Since $23 \equiv_5 3$ we have $\left(\frac{23}{5}\right) = \left(\frac{3}{5}\right)$. From our work above, $[3] \notin QR_5$. Therefore $-1 = \left(\frac{3}{5}\right) = \left(\frac{5}{23}\right)$. And so we conclude $[5]$ is not a quadratic residue modulo 23.

The Legendre symbol was first introduced by the 18th Century French mathematician Adrien-Marie Legendre in his unsuccessful attempts to prove the Law of Quadratic Reciprocity. Despite this, his symbol lives on as a means to confuse undergraduate students who miss where it is defined and mistakenly interpret as a fraction.

Though a proof of the Law of Quadratic Reciprocity does not come easily enough for us to consider it in this last week of term, it is not so difficult that it is beyond what we would be capable of understanding as undergraduate students. In fact, more than two hundred proofs of this theorem appear in the literature. Some of these proofs, notably one provided by Gauss, admit a geometric interpretation as points in a lattice. Proofs of this theorem are so ubiquitous that there is an entire Wikipedia page[4] dedicated to them

On this note we end our work in quadratic residues and close the book on MATH 364 for the semester. Take a moment to appreciate the work you have done this semester. I suspect this is the first time you have read nearly 150 pages of mathematics on a single subject. In itself, this is an impressive feat!

Good luck with your final exams!

[4] https://en.wikipedia.org/wiki/Proofs_of_quadratic_reciprocity

Appendices

A Mathematical Induction – A refresher

Mathematical induction is a powerful proof technique. Mathematical induction appears both in MATH163 and CMPT260. And so I will generally assume that you have seen this technique somewhere before. (If you have not, then please be sure to come chat with me to follow up on what is below.)

We begin with an example.

Let n be a non-negative integer. Consider the following sum:

$$s(n) = \sum_{i=0}^{n} 2^i = 2^0 + 2^1 + 2^2 + \cdots + 2^n$$

When n is small it is easy to compute values of $s(n)$:

$$\begin{aligned} s(0) &= 1 \\ s(1) &= 3 \\ s(2) &= 7 \\ s(3) &= 15 \\ s(4) &= 31 \end{aligned}$$

You may notice each of these values is one fewer than a power of 2.

$$\begin{aligned} s(0) &= 2^1 - 1 \\ s(1) &= 2^2 - 1 \\ s(2) &= 2^3 - 1 \\ s(3) &= 2^4 - 1 \\ s(4) &= 2^5 - 1 \end{aligned}$$

We wonder does this pattern hold in general. That is, does $s(n) = 2^{n+1} - 1$ for each non-negative integer n? Let us consider the case $n = 5$, but rather than substitute directly, let us try a different approach:

$$s(5) = 2^0 + 2^1 + 2^2 + 2^3 + 2^4 + 2^5$$

Notice $2^0 + 2^1 + 2^2 + 2^3 + 2^4 = s(4)$. We have already confirmed $s(4) = 2^5 - 1$. Thus

$$s(5) = \left(2^0 + 2^1 + 2^2 + 2^3 + 2^4\right) + 2^5 = \left(2^5 - 1\right) + 2^5$$

Simplifying, we notice $2^5 + 2^5 = 2(2^5) = 2^6$. Therefore

$$s(5) = 2^6 - 1$$

Using this same technique, we can verify $s(6) = 2^7 - 1$. And using $s(6) = 2^7 - 1$ we can use the same technique to verify $s(7) = 2^8 - 1$.

(If this is not clear to you, take a moment to do the calculation. Compute $s(6)$ by noticing that the sum of the first 6 terms in $s(6)$ is equal to $s(5)$)

For this example, once we have verified that our formula holds for $s(k)$, we can then verify that our formula holds for $s(k+1)$. This idea is the central component of mathematical induction. Let us formalize.

Let $P(n)$ be a statement that is either true or false for each integer $n \geq 0$. For example, if $P(n)$ is the statement *"n is odd"* then $P(0)$ is false and $P(1)$ is true. If $P(n)$ is the statement

$$\sum_{i=0}^{n} 2^i = 2^0 + 2^1 + 2^2 + \cdots + 2^n = 2^{n+1} - 1$$

then we have confirmed that each of $P(0), P(1), P(2), P(3), P(4)$ and $P(5)$ is true. In asking if

$$\sum_{i=0}^{n} 2^i = 2^{n+1} - 1$$

holds for each integer $n \geq 0$, we are asking if $P(n)$ is true for each $n \geq 0$.

Imagine you are provided with the following information about $P(n)$:

1. $P(0)$ is true; and
2. for each $k \geq 0$, if $P(k)$ is true, then $P(k+1)$ is true

Can we conclude $P(1)$ is true?

Statement 2 tells us that knowing $P(0)$ is true tells that that $P(1)$ is true. Thus, to know if $P(1)$ is true, it is enough to know that $P(0)$ is true. Statement 1 tells us that $P(0)$ is true, thus we can conclude $P(1)$ is true.

Can we conclude $P(2)$ is true?

Statement 2 tells us that knowing $P(1)$ is true tells that that $P(2)$ is true. Thus, to know if $P(2)$ is true, it is enough to know that $P(1)$ is true. We have already concluded that $P(1)$ is true, thus we can conclude $P(2)$ is true.

A similar line of reasoning tells that that $P(3)$ is true. Once we have concluded $P(3)$ is true, then we can conclude $P(4)$ is true, and so on. Thus, knowing that statements 1 and 2 hold is enough to convince us that $P(n)$ is true for each $n \geq 0$.

We express this idea with the following theorem.

Theorem (The Principle of Mathematical Induction)**.** *Let $P(n)$ be a statement that is either true or false for each integer $n \geq 0$. If the following two statements hold, then $P(n)$ is true for each integer $n \geq 0$.*

1. *$P(0)$ is true; and*
2. *for each $k \geq 0$, if $P(k)$ is true, then $P(k+1)$ is true*

Let us consider applying The Principle of Mathematical Induction to study $s(n)$ above. To prove a result using The Principle of Mathematical Induction one must show that hypothesis of the theorem above holds. In this case the hypothesis is: *statements 1 and 2 are true.*

$P(n)$ is the statement

$$\sum_{i=0}^{n} 2^i = 2^{n+1} - 1$$

We can use the Theorem of Mathematical Induction, to prove $P(n)$ is true for each integer $n \geq 0$ by verifying that 1 and 2 hold for this $P(n)$.

1. $P(0)$ is true

$P(0)$ is the statement

$$\sum_{i=0}^{0} 2^i = 2^1 - 1$$

We observe $\sum_{i=0}^{0} 2^i = 2^0 = 1$ and $2^1 - 1 = 1$. Therefore $P(0)$ is true. And so statement 1 holds in the Principle of Mathematical Induction

2. *for each $k \geq 0$, if $P(k)$ is true, then $P(k+1)$ is true*
To show this statement is true, we assume the hypothesis is true and we use that to show that the conclusion is true. Let $k \geq 0$ be an integer so that $P(k)$ is true. Thus the following statement is true:

$$\sum_{i=0}^{k} 2^i = 2^{k+1} - 1$$

We want to show $P(k+1)$ is true. That is, we want to show

$$\sum_{i=0}^{k+1} 2^i = 2^{k+2} - 1$$

is true.

Consider the sum

$$\sum_{i=0}^{k+1} 2^i$$

The last term of this sum is 2^{k+1}. Therefore

$$\sum_{i=0}^{k+1} 2^i = \left(\sum_{i=0}^{k} 2^i \right) + 2^{k+1}$$

By hypothesis $P(k)$ is true. And so $\sum_{i=0}^{k} 2^i = 2^{k+1} - 1$. Thus

$$\sum_{i=0}^{k+1} 2^i = \left(\sum_{i=0}^{k} 2^i \right) + 2^{k+1} = \left(2^{k+1} - 1\right) + 2^{k+1} = 2^{k+2} - 1$$

Therefore $P(k+1)$ is true. And so we see statement 2 holds the Principle of Mathematical Induction.

Since statement 1 and statement 2 hold in the Principle of Mathematical Induction, so must the conclusion. Therefore $P(n)$ is true for each $n \geq 0$. That is,

$$\sum_{i=0}^{n} 2^i = 2^{n+1} - 1$$

for each $n \geq 0$.

Aside A.1. *If you have seen induction before you might find notable the lack of appearance of the phrases* base case*,* induction hypothesis *and* induction step*. As long as it is clear that the author is showing that the two statements hold, these phrases need not be used.* Base Case *refers to statement 1.* Induction hypothesis *refers to the hypothesis of statement 2.* Induction step *refers to proving that the conclusion of statement 2 holds, given that the hypothesis is true. This is likely the only time these phrases will appear in these notes.*

The application of the Theorem of Mathematical Induction is so ubiquitous in mathematics, that one need only mention the word *induction* in the proof for the reader to know what to expect.

For example, if the author knows that their reader is well-experienced with induction the following proof would suffice.

Theorem. *If n is a non-negative integer, then*

$$\sum_{i=0}^{n} 2^i = 2^{n+1} - 1$$

Proof. We proceed by induction on n, noting $2^0 = 1$.

Consider $n = k + 1$. By induction, $\sum_{i=0}^{k} 2^i = 2^{k+1} - 1$.

Thus,

$$\sum_{i=0}^{k+1} 2^i = \left(\sum_{i=0}^{k} 2^i\right) + 2^{k+1} = \left(2^{k+1} - 1\right) + 2^{k+1} = 2^{k+2} - 1.$$

□

Aside A.2. *The Principle of Mathematical Induction is a theorem. This suggests that one can write down a proof of the Principle of Mathematical Induction. This is true, but not immediately relevant to us. We will not concern ourselves with the proof of the Principle of Mathematical Induction.*

Let us turn now to a related technique: the Principle of Strong Mathematical Induction.

Our modern base-ten place-value system of notation is incredibly useful. It is so engrained in us, that most of the time we don't even notice. Recall that the notation

$$1342$$

Refers to the number equal to

$$1 \times 10^4 + 3 \times 10^2 + 4 \times 10^1 + 2 \times 2^0$$

This system is based upon powers of 10 and uses the digits $0 - 9$. Such a system is likely based upon us humans having ten fingers.

Imagine what our number system would be like if we only had thumbs. We may have developed a place-value system based upon powers of 2 using digits 0 and 1. For example we would write:

$$1101$$

To refer to the number equal to

$$1 \times 2^3 + 1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0.$$

We would express this number as 13 in our base-ten notation.

Is it clear that every number can be expressed as a sum of powers of 2?

Numbers on the left are in base ten. The equivalent representation in base two is on the right. The equals sign means that the two pieces of notation represent the same integer.

$$\begin{aligned} 0 &= 0 \\ 1 &= 1 \\ 2 &= 10 \\ 3 &= 11 \\ 4 &= 100 \\ 5 &= 101 \\ 6 &= 110 \end{aligned}$$

So far so good, now what about 7? Rather than continue with the pattern, let us try and be slightly clever. We notice that $7 = 4 + 3$. The integer 4 can be expressed as 100 in base two. The integer 3 can be expressed as 011 in base two. Thus

$$\begin{aligned} 4 &= 1 \times 2^2 + 0 \times 2^1 + 0 \times 2^0 \\ 3 &= 1 \times 2^1 + 1 \times 2^0 \end{aligned}$$

Adding these together yields

$$7 = 1 \times 2^2 + 1 \times 2^1 + 1 \times 2^0$$

Thus 7 can be expressed as 111 in base-two.

Aside A.3. *Notice how this differs from our previous example. In the previous example we only needed to remember information from the previous case. In this case we need to remember information more than one previous case.*

Just as we did for induction, let us generalize.

Let $P(n)$ be a statement that is either true or false for each integer $n \geq 0$. Imagine you are provided with the following information:

1. $P(0)$ is true; and
2. for each $k \geq 0$, if $P(k')$ is true for each $0 \leq k' \leq k$, then $P(k+1)$ is true

As compared to our first example, the second statement here is stronger. Instead of considering $k \geq 0$ so that $P(k)$ is true, we are considering $k \geq 0$ so that each of $P(0), P(1), P(2), \ldots, P(k)$ is true. Regardless, our intuition here should be the same as in the Principle of Mathematical Induction. If we know that both of statement 1 and statement 2 hold, then necessarily $P(n)$ is true for each $n \geq 0$.

Theorem (The Principle of Strong Mathematical Induction)**.** *Let $P(n)$ be a statement that is either true or false for each integer $n \geq 0$. If the following two statements hold, then $P(n)$ is true for each integer $n \geq 0$.*

1. *$P(0)$ is true; and*
2. *for each $k \geq 0$, if $P(k')$ is true for each $0 \leq k' \leq k$, then $P(k+1)$ is true.*

Let $P(n)$ be the statement "*The integer n can be represented in base two*". We want to show $P(n)$ is true for each $n \geq 0$. To apply the Principle of Strong Mathematical Induction we must show that statements 1 and 2 both hold.

1. $P(0)$ is true

$P(0)$ is the statement *"the integer 0 can be represented in base two"*. We have $0 = 0 \times 2^0$. Thus the integer 0 is expressed as 0 in base two.

2. for each $k \geq 0$, if $P(k')$ is true for each $0 \leq k' \leq k$, then $P(k+1)$ is true

To show this statement holds, we assume the hypothesis is true and then show that the conclusion follows. Here our hypothesis is: $P(k')$ is true for each $0 \leq k' \leq k$.

Consider some $k \geq 0$ so that each of $P(0), P(1), \ldots, P(k)$ is true. This is equivalent to saying that all non-negative integers up to k can be expressed in base-two notation.

We want to show that $P(k+1)$ is true. That is, we want to show that the integer $k+1$ can be represented in base two.

If $k+1$ is a power of two, then $k+1 = 2^\ell$ for some integer ℓ. Thus

$$k+1 = 1 \times 2^{\ell} + 0 \times 2^{\ell-1} + \cdots + 0 \times 2^{1} + 0 \times 2^{0}$$

And so the representation of $k+1$ in base two is a 1 followed by ℓ zeroes.

Otherwise, assume $k+1$ is not a power of 2. Let 2^{ℓ} be the largest power of 2 that is smaller than $k+1$. Thus

$$k+1 = 2^{\ell} + t$$

where t is an integer strictly less than 2^{ℓ}. By hypothesis, $P(2^{\ell})$ and $P(t)$ are both true.

And so we can express 2^{ℓ} and t as follows:

$$\begin{aligned} 2^{\ell} &= 1 \times 2^{\ell} + 0 \times 2^{\ell-1} + \cdots + 0 \times 2^{1} + 0 \times 2^{0} \\ t &= 0 \times 2^{\ell} + e_{\ell-1} \times 2^{\ell-1} + e_{\ell-2} \times 2^{\ell-2} + \cdots + e_1 \times 2^{1} + e_0 \times 2^{0} \end{aligned}$$

where $e_i \in \{0, 1\}$ for each $0 \leq e_i \leq \ell - 1$.

Therefore

$$k+1 = 1 \times 2^{\ell} + e_{\ell-1} \times 2^{\ell-1} + e_{\ell-2} \times 2^{\ell-2} + \cdots + e_1 \times 2^{1} + e_0 \times 2^{0}$$

This implies that $k+1$ can be expressed as $1e_{\ell-1}e_{\ell-2}\cdots e_1e_0$ in base two. Therefore $P(k+1)$ is true.

Since both statements of the Theorem of Strong Mathematical Induction hold, necessarily the conclusion holds. Therefore $P(n)$ is true for each $n \geq 0$. In other words, the integer n can be represented in base two for every integer $n \geq 0$.

In this course you will generally not be left to your own devices when it comes to proving things by induction. If a question calls for us to use induction, then it will be indicated. For these questions I will generally provide steps, rather than leaving you to do the entire thing from scratch.