Notes on Number Theory, Fall 2022

Randall Holmes

October 18, 2023

10/23/2022 some extra remarks in Homework 11 instructions

10/20/2022 finishing up this unit, adding proofs to the last section.

On 10/13/22 updating with notes on the Wednesday 10/12 lecture and including the new homework.

10/12 On page 6 you will find promised theorems about divisibility. I am inserting a section on modular arithmetic (the 10/10 lecture topic) as well.

Updated 10/8. 7 pm fixed a "typo" in the extended Euclidean algorithm theorem: wrote all the quotients backward!

Im posting these on Monday 10/3; I still need to add more to cover the lecture the previous Monday. The notes on that lecture should be complete sometime today, to be expanded to cover what I lecture today (10/3). The homework assigned for 10/3 will also appear as an update to this file.

Contents

1 Basic concepts and axioms

The mathematical system we are interested in is officially the set of integers, whose official name is $\mathbb{Z}$ (this stands for **Zahlen**, numbers, in German).

I'm going to state a definition of the set of integers as a subset of the real numbers $\mathbb{R}$, just because I would like you to know in the back of your mind that things like this can be done. This definition is not examinable, but it might in some ways be useful to understand it.

Definition (integer-closed set of reals): A set A of real numbers is said to be "integer-closed" if and only if $0 \in A$ and for every real number x, if $x \in A$ then $x+1$ and $x-1$ are both in A.

There are lots of integer-closed subsets of $\mathbb{R}$: $\mathbb{R}$ itself is integer-closed; the set of rationals is integer-closed; the set of all fractions with denominator 2 is integer closed (it contains all integers n and also contains $n + \frac{1}{2}$ for each integer n).

Definition of $\mathbb{Z}$: $\mathbb{Z}$ is defined as $\{x \in \mathbb{R} :$ for every set $A \in \mathcal{P}(\mathbb{R})$, if A is integer closed then $x \in A\}$. It really works...think about it.

There is something odd about this definition, because it presupposes that we know what the reals are. But if you get to an advanced course where a formal definition of the natural numbers, integers, and reals is actually given, you might get an idea of what I'm up to. Really, I just want to be able to say something more accurate than $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2 \ldots\}$.

The mathematical system of the integers does not have the set of integers as its only component. It also has operations and relations for which we will state axioms. These operations and relations are familiar. The constants 0 and 1, the operations of addition, multiplication, and additive inverse, and the relation "less than" are the basics in our presentation.

First set of axioms (basic algebra): commutative laws: For any integers a, b, $a + b = b + a$ and $a \cdot b = b \cdot a$.

associative laws: For any integers a, b, c, $(a + b) + c = a + (b + c)$ and $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.

distributive law: For any integers a, b, c, $a(b + c) = a \cdot b + a \cdot c$.

distributive laws? (comment): We could give another distributive law $(a + b)c = ac + bc$ too, but we do not need to, since $(a + b)c = c(a+b) = ca+cb = ac+bc$ justifies the alternative distributive law in terms of the one we give and commutativity of multiplication.

It is useful to notice that matrix algebra is a familiar system (to some of you, at least) in which both forms of the distributive law have to be given, because matrix multiplication is not commutative.

identity laws: For any integer a $a + 0 = a$ and $a \cdot 1 = a$; $0 \neq 1$

additive inverse law: For any integer a, $a + (-a) = 0$.

multiplicative cancellation: For any integers a, b, c, If $c \neq 0$ and $ac = bc$, then $a = b$.

Normally, we would prove multiplicative cancellation (in the reals or rationals) using the existence of a multiplicative inverse (reciprocal) a^{-1} of each nonzero real number a such that $a \cdot a^{-1} = 1$. But this property doesn't hold in the integers. Notice that we can prove additive cancellation using the additive inverse property (if $a + c = b + c$ then $a = b$: try to demonstrate this).

Observation: All of these axioms hold in the systems of reals or rationals familiar to you. We will also study systems of modular arithmetic which look very different from the reals or rationals (they are finite systems!) in which all of the axioms of the first set hold.

A second set of axioms rules out the modular arithmetic systems.

Second set of axioms (properties of order): $<$ is transitive: For all integers a, b, c, if $a < b$ and $b < c$, then $a < c$.

trichotomy: For any integers a, b, exactly one of the following is true: $a < b, a = b, b < a$.

additive monotonicity: For any integers a, b, c, if $a < b$ then $a + c < b + c$ (it is actually if and only if, because you can add the additive inverse to go the other way).

multiplicative monotonicity: For any integers a, b, c, if $c > 0$ and $a < b$, then $ac < bc$.

The other version, if $c < 0$ and $a < b$, then $bc < ac$, is provable from the axioms we already have.

Definitions: $a > b$ means $b < a$. $a \leq b$ means $a < b$ or $a = b$. $a \geq b$ means $b \leq a$. These relations all have similar properties to $<$, which you are already informally familiar with.

Observation: These axioms all hold in the familiar systems of reals and rationals. The axiom which separates the integers from the reals or rationals is something we have already discussed.

Axiom of Mathematical Induction: Let $P(n)$ be a statement about an integer variable n and let b be an integer. If $P(b)$ is true and for every integer $k \geq b$ such that $P(k)$ is true, it follows that $P(k + 1)$ is true, we can conclude that $P(n)$ is true for every integer $n \geq b$.

Proof structure: A proof of "for all integers $n \geq b$, $P(n)$ by math induction falls into two parts:

the basis steps shows $P(b)$

the induction step introduces an arbitrarily chosen integer k and the assumption that $P(k)$ is true [called the inductive hypothesis] and shows that $P(k + 1)$ follows.

A variation is proof by strong induction, in which the inductive hypothesis is strengthened to "for all integers m with $b \leq m \leq k$, $P(m)$.

The axiom of mathematical induction drives a wedge between the integers and the reals or rationals. For example, we can prove that there is no multiplicative inverse of 2.

Theorem: There is no integer x such that $0 < x < 1$.

Proof: We prove by mathematical induction that for every integer $x \geq 0$, either $x = 0$ or $x \geq 1$.

Basis: $0 = 0$ or $0 \geq 1$ is true because $0 = 0$.

Induction: Suppose $k = 0$ or $k \geq 1$. In either case we have $k \geq 0$. $k+1 \geq 1$ follows by additive monotonicity, so we have either $k+1 = 0$ or $k + 1 \geq 1$ because we have the latter.

Theorem: There is no integer x such that $2x = 1$.

Proof: Suppose there is such an x. We have to establish $x > 0$: clearly x is not 0 ($2 \cdot 0 = 0 \neq 1$). If $x < 0$ we would have $2x < 0$ by multiplicative monotonocity, so $1 < 0$, which is absurd. So $x > 0$. It follows by additive monotonicity that $x + x > x$, so $1 > x$. But then $0 < x < 1$, which we have just shown is not possible.

Sanity check ($1 > 0$): We have in our axioms that $1 \neq 0$. Thus we can only have $1 > 0$ (which we believe) or $1 < 0$. Suppose that $1 < 0$. Add -1 to both sides to get $0 < -1$, so $-1 > 0$. Then we can use $-1 > 0$ and multiplicative monotonicity: it follows that $(-1)(-1) > 0$, $1 > 0$, contradicting the assumption that $1 < 0$ (by an appeal to trichotomy). So we have ruled out $1 < 0$ and $1 > 0$ is the only possibility.

Again, Im appealing to common sense in the equation $(-1)(-1) = 1$. We may look into how to prove that.

This is an example of a general point. You have seen something like these axioms before, and you have been told that all of your ninth-grade algebra knowledge (say) follows from these principles. But you weren't really shown this, and proofs of "easy and obvious" things from basic sets of axioms in any area of math may turn out to be tricky.

Challenge Problem: Prove from the axioms that for any integer a, $a \cdot 0 = 0$. It is rather tricky! (Notice that we used this fact freely in our discussion above.)

2 Divisibility and a little about prime numbers

Definition (divisibility): Let a and b be integers.

We define $a|b$ (a goes into b, or equivalently b is divisible by a) as meaning "There is an integer k such that $ak = b$".

We also say "a is a factor of b", or "a is a divisor of b". We usually restrict our attention to positive divisors, but we will try always to say this explicitly.

Important Observation: Please notice that $a|b$ is not a fraction or any kind of expression: it is a sentence. And notice that writing $\frac{b}{a}$ is not a way of saying b is divisible by a: I saw this on many test papers.

What is (almost) true is that $a|b$ is equivalent to "$\frac{b}{a}$ is an integer". The exception is when a and b are both 0: $0|0$ is true but $\frac{0}{0}$ is undefined.

We will do some work on basic theorems about divisiblity, notes on which will be inserted at this point. The facts about divisibility which we will prove are fairly obvious, but writing the proofs will be useful for general thinking about how proofs are to be written.

Theorem: If a, b, c are integers and $a|b$, then $a|(bc)$

Proof: Because $a|b$, we can choose an integer k such that $ak = b$.

To show $a|(bc)$, we need to find an integer m such that $am = bc$.

We know that $bc = (ak)c = a(kc)$, so $m := kc$ does the trick: we have $a|(bc)$ because there is an integer $m[= kc]$ such that $mm = bc$.

The point here is not that it is hard to see that it is true, but to see what a formal proof of such a statement looks like.

Theorem: If a, b, c are integers, and $c|a$ and $c|b$, then $c|(a + b)$.

Proof: Because $c|a$, we can choose an integer x such that $cx = a$.

Because $c|b$, we can choose an integer y such that $cy = b$.

We need an integer m such that $cm = a + b$ to establish the theorem.

Now $a + b = cx + cy = c(x + y)$, so if we choose $m = x + y$, we have shown that there is an integer m such that $cm = a+b$, that is, $c|(a+b)$.

We define a basic notion already familiar to you.

Definition (prime numbers): An integer n is prime iff $n > 1$ and the only positive divisors of n are 1 and n.

Note that 1 is not a prime.

Theorem: Each integer $n \geq 2$ is a prime or a finite product of primes.

Proof: We prove this by strong induction:

Basis ($n = 2$): 2 is a prime.

Induction step: Let k be an arbitrarily chosen integer. Suppose (ind hyp) that every m with $2 \leq m \leq k$ is a prime or a finite product of primes.

Our goal is then to show that $k + 1$ is a prime or finite product of primes.

If $k + 1$ is prime, we are done.

If $k + 1$ is not prime, then $k + 1 = LM$ for some $2 < L \leq M < k + 1$ (it has positive divisors other than 1 and itself). But by inductive hypothesis each of L, M is either a prime or a finite product of primes, so $LM = k + 1$ is a finite product of primes, and we are done.

There is a bigger result which we will prove soon: each integer ≥ 2 factors into primes in exactly one way.

The theorem which follows is ancient, and it has been seriously argued that any educated person should know it. Certainly its proof is examinable in this class.

Theorem (Euclid): There are infinitely many prime numbers.

Proof: We have at least one prime, 2. Suppose we have a list of n prime numbers $p_1, p_2 \ldots, p_n$ which contains all of the prime numbers (if there were finitely many primes we could make such a list).

Consider $P = 1 + \prod_{i=1}^{n} p_i$, one plus the product of all of the primes on our list.

By the preceding theorem, P (which is ≥ 2 because the list has 2 in it) has a prime divisor q, because it is either a prime or a product of primes.

Iff $q = p_m$ then q goes into $P = 1 + \prod_{i=0}^{n} p_i$, and $q|(P-1) = \prod_{i=0}^{n} p_i$ because it is one of the factors in the product, so $q|P - (P-1) = 1$, which is absurd. So q cannot be one of the primes we listed, so there cannot be a finite list of all the primes, which is what we wanted to prove.

3 The Division Algorithm: integer division and remainder operations

Theorem (division algorithm): For any integer a and any integer $b > 0$, there are uniquely determined integers q and r such that $a = bq + r$ and $0 \leq r < 1$.

integer division and remainder operations: Because q and r are uniquely determined, we can define a `div` b as q and a `mod` b as r. The div operation is integer division and the mod operation is remainder. These operations should actually be familiar from elementary school.

Proof of the Theorem: This falls into three sections:

induction proof of existence of q and r (not uniqueness) for $a \geq 0$:

We prove the theorem (actually, just part of it to begin with) by induction on a.

We let b be a fixed positive integer.

We prove by induction that for each $n \geq 0$ there are integers q and r such that $n = bq + r$ and $0 \leq r < 1$.

We use this result to prove the rest of the full theorem afterward.

Basis ($n = 0$): $0 = b0 + 0$ and $0 \leq 0 < b$. $q = r = 0$ works.

Induction step: Let $k \geq 0$ be chosen arbitrarily and suppose there are q_1, r_1 such that $k = bq_1 + r_1$ and $0 \leq r_1 < b$ (this is the ind hyp).

Our aim is to find q and r such that $k+1 = bq+r$ and $0 \leq r < 1$.

If $r_1 < b - 1$ then $k + 1 = bq_1 + (r_1 + 1)$ and $0 \leq r_1 + 1 < b$.

In this case let $q = q_1$ and $r = r_1 + 1$.

If $r_1 < b - 1$ is false, then $r_1 = b - 1$ (because there is no integer between 0 and 1, so there is no integer between $b - 1$ and 1). In this case, $k + 1 = q_1 b + r_1 + 1 = q_1 b + b = (q_1 + 1)b + 0$

and we can let $q = q_1 + 1$, $r = 0$.

So we have shown by math induction that for each $n \geq 0$ there are integers q and r such that $n = bq + r$ and $0 \leq r < 1$.

existence of q and r when $a < 0$: We need to deal with the case $a < 0$ in the main theorem: if $a < 0$ then $-a > 0$ and we have shown that there are q_1, r_1 such that $-a = bq_1 + r_1$ and $0 \leq r_1 < b$.

If $r_1 = 0$, then let $q = -q_1$, $r = 0$ and we have $a = -(-a) = -(bq_1) = b(-q_1) + r$, and of course $0 \leq 0 = r < b$.

If $r_1 > 0$, then let $q = -(q_1 + 1), r = b - r_1$. $bq + r = b(-q_1 + 1) + b - r_1 = -(bq_1 + r_1) = -(-a) = a$ and $0 < b - r_1 < b$ follows because $0 < r_1 < b$.

proof of uniqueness of q and r: Now we need to show that in all these cases, there can be only one such q and r.

Suppose $a = bq_1 + r_1$ and $a = bq_2 + r_2$, with $b > 0$.

We can suppose further that $r_1 \geq r_2$ (trichotomy, and pick the smaller one to be r_2).

Subtract to get $b(q_2 - q_1) = r_1 - r_2$. We have $r_1 - r_2$ nonnegative and strictly less than b.

From $0 \leq b(q_2 - q_1) < b$ we can conclude $0 \leq q_2 - q_1 < 1$ so $q_2 - q_1 = 0; q_1 = q_2$, because there is no integer strictly between 0 and 1.

So we have shown that if $a = bq_1 + r_1$ and $a = bq_2 + r_2$, with $b > 0$, it follows that $q_1 = q_2$.

Now if $bq + r_1 = bq + r_2$, it follows that $r_1 = r_2$ by adding $-bq$ to both sides of the equation.

So we are done.

4 Greatest common divisor (gcd) and the Euclidean algorithm

Definition (common divisor): Let a, b be integers. We say that d is a common divisor of a and b just in case $d|a$ and $d|b$.

Observations: If $a = b = 0$ then every integer is a common divisor of a and b, because every integer is a divisor of 0.

If $d|a$ and $a \neq 0$ then $a = kd$ for some integer k and so $|a| = k'd$ for $k' = \pm k$. Now if $d < 0$ we certainly have $d \leq |a|$, and if $d > 0$ we clearly have $d \leq k'd = |a|$. That is, $|a|$ is the largest divisor of a if $a \neq 0$.

This further implies that if a, b are not both 0 and d is a common divisor of a and b then $d \leq \texttt{max}(|a|, |b|)$.

Finally, a fact which we will prove later: any set of integers which has an upper bound has a largest element. So there is a greatest common divisor of a, b if a and b are not both zero.

Another way to see that the greatest common divisor exists is to note that an integer other than zero has only finitely many divisors, so the set of common divisors of two numbers which are not both zero is finite, and a finite set has a largest element. These statements actually require some analysis too!

Definition (greatest common divisor, gcd): For any pair of integers a, b which are not both zero, we define $\texttt{gcd}(a, b)$, the greatest common divisor of a and b, as the largest element of the set of common divisors of a and b, which we have shown above exists.

This concept is familiar: The greatest common divisor is again a concept familiar from elementary school. When you simplify a fraction $\frac{a}{b}$, the common factor by which you divide the numerator and denominator is the gcd of a and b.

You learned a method for finding these common factors using prime factorizations. We will teach a much better method (at least, better for large numbers) which is ancient: it was known to Euclid).

Theorem (facts about the gcd):

1. $\texttt{gcd}(a, b) = \texttt{gcd}(b, a)$ Obvious by symmetry of the definition. Because of this, we can assume $a \geq b$.

2. $\texttt{gcd}(a, b) = \texttt{gcd}(|a|, |b|)$

   This is obvious: the divisors of a are the same as the divisors of $|a|$ and the same is true of b, so the common divisors of a, b make up the same set as the common divisors of $|a|, |b|$ and of course these sets have the same largest element.

   Because of this, we can assume a, b are nonnegative in gcd calculations (with possible fixes later if we have to think about the case where one of them might be negative). So our default assumption is that a is positive and $a \geq b \geq 0$.

3. $\texttt{gcd}(a, 0) = |a|$

   Common divisors of a and 0 are exactly the divisors of a, of which the largest is $|a|$. If a is positive of course this simplifies to $\texttt{gcd}(a, 0) = a$.

4. If $b > 0$, $\texttt{gcd}(a, b) = \texttt{gcd}(b, a \texttt{ mod } b)$.

   This takes a little more work. Notice that if $q = a \,\texttt{div}\, b$ and $r = a \,\texttt{mod}\, b$, we have $a = bq + r$ and $r = a - bq$. We will use q, r with these meanings.

   Now we argue that the set of common divisors of a and b is the same set as the set of common divisors of b and $r = a \,\texttt{mod}\, b$.

   To show this, we show that any element of the first set belongs to the second and any element of the second set belongs to the first.

   If x belongs to the set of common divisors of a and b, then $x|a$ and $x|b$. If $x|b$ then certainly $x|qb$. If $x|a$ and $x|qb$ then $x|(a - qb) = r$. So x belongs to the set of common divisors of b and $r = a \,\texttt{mod}\, b$.

   If x belongs to the set of common divisors of b and $r = a \,\texttt{mod}\, b$, then $x|b$ and $x|r$. If $x|b$ then $x|bq$. If $x|bq$ and $x|r$, it follows that $x|bq + r = a$. So x also belongs to the set of common divisors of a and b.

   Since these two sets are the same, they have the same largest element, so $\texttt{gcd}(a, b) = \texttt{gcd}(b, a \texttt{ mod } b)$.

   Notice that our default assumption that the first argument is positive and the second is nonnegative and smaller holds automatically for $\texttt{gcd}(b, a \texttt{ mod } b)$.

Theorem (Euclidean algorithm): Let a, b be integers with $b > 0$.

Define a sequence D by $D_0 = a$, $D_1 = b$ and $D_{n+2} = D_n \,\texttt{mod}\, D_{n+1}$. Notice that D_{n+2} is only defined if $D_{n+1} \neq 0$.

We argue that for every a, b there is an n such that $D_{n+1} = 0$, D_{n+2} is undefined, and $D_n = \texttt{gcd}(a, b)$.

First we prove by induction that for every integer n, $\texttt{gcd}(D_n, D_{n+1}) = \texttt{gcd}(a, b)$ if D_{n+1} is defined (and so in particular if $D_{n+1} = 0$ then $D_n = \texttt{gcd}(a, b)$).

The basis is the assertion $\texttt{gcd}(D_0, D_1) = \texttt{gcd}(a, b)$ which is true by definition of D.

The induction step: If we assume $\texttt{gcd}(D_n, D_{n+1}) = \texttt{gcd}(a, b)$, then by the previous theorem $\texttt{gcd}(D_n, D_{n+1}) = \texttt{gcd}(D_{n+1}, D_n \,\texttt{mod}\, D_{n+1}) = \texttt{gcd}(D_{n+1}, D_{n+2})$ if D_{n+2} is defined.

We aren't done! We have to prove that for every a, b, there is n such that D_n is undefined (that the process stops). We prove this by induction (there is another way to prove it which we will discuss shortly). The trick, as is often the case, is to see which variable to apply induction to. We choose to do strong induction on b.

For $b = 1$, the result is clearly true: $D_0 = a, D_1 = 1, D_2 = 0$, and D_3 is undefined.

Suppose that for every a and for every b greater than one and less than k a D sequence must terminate.

The sequence which starts $D_0 = a, D_1 = k + 1, D_2 = a \,\texttt{mod}\, (k + 1) \ldots$ terminates if $a \,\texttt{mod}\, (k+1)$ is zero. Otherwise, it can be seen to terminate because the sequence beginning $D_1 = k+1, D_2 = a \,\texttt{mod}\, (k+1) \ldots$ must terminate by ind hyp, because $a \,\texttt{mod}\, (k + 1) \leq k$.

And this completes the argument. I'll expand these notes with sample calculations after Friday's lecture, when I will have more computational techiques for you to try out. The next homework assignment will appear as a section in these notes after the lecture on Friday.

We comment that not only does this method of computation of gcd's always terminate, but it actually terminates fairly fast: the number of steps until the D sequence terminates is proportional to the logarithm of a.

We give a computational example.

Example: Compute the greatest common denominator of 925 and 851.

$\texttt{gcd}(925, 851)$

$= \texttt{gcd}(851, 925 \,\texttt{mod}\, 851)$

$= \texttt{gcd}(851, 925 - (1)851)$

$= \texttt{gcd}(851, 74)$

$= \texttt{gcd}(74, 851 \,\texttt{mod}\, 74)$

$= \texttt{gcd}(74, 851 - (11)(74))$

$= \texttt{gcd}(74, 37) = 37$ (37 goes evenly into 74).

Now we present a more impressive result, probably entirely unexpected to you.

Definition: An integer c is said to be a linear combination of integers a and b iff there are integers x and y such that $c = ax + by$.

Theorem (extended Euclidean algorithm): For any integers a, b which are not both zero, $\texttt{gcd}(a, b)$ is a linear combination of a and b, that is, there are integers x and y such that $ax + by = \texttt{gcd}(a, b)$.

Proof: We prove this by a method similar to our verification of the Euclidean algorithm, by defining additional sequences which will give us our x and y.

We assume that a, b are integers, not both zero, with $b > 0$ (notice that the result for negative b follows very easily from the result for positive b).

Define D_n as above.

Define additional sequences X and Y: $X_0 = 1; X_1 = 0; Y_0 = 0; Y_1 = 1; X_{n+2} = X_n - (D_n \texttt{div} D_{n+1})(X_{n+1}); Y_{n+2} = Y_n - (D_n \texttt{div} D_{n+1})(Y_{n+1})$.

The recurrence relations for the X and Y sequences are closely related to those for D:

$D_{n+2} = D_n \texttt{mod} D_{n+1} = D_n - (D_n \texttt{div} D_{n+1})(D_{n+1})$, which is precisely parallel in form to the definitions for X and Y.

This is actually a description of the table format I used in class, of which I'll give an example below.

We prove by strong induction on n that $D_n = aX_n + bY_n$ for every n.

The basis is direct: $D_0 = a = a1 + b0 = aX_0 + bY_0$; $D_0 = b = a0 + b1 = aX_1 + bY_1$.

Assume that $D_m = aX_m + bY_m$ for every $m \leq k$ and show that $D_{k+1} = aX_{k+1} + bY_{k+1}$:

If $k = 0$ we have already shown this, so assume $k \geq 1$.

$D_{k+1} = D_{k-1} - (D_{k-1}\texttt{div}D_k)D_k$ which by ind hyp

$= (aX_{k-1} + bY_{k-1}) - (D_{k-1}\texttt{div}D_k)(aX_k + bY_k)$ which by algebra

$= a(X_{k-1} - (D_{k-1}\texttt{div}D_k)X_k) + b(Y_{k-1} - (D_{k-1}\texttt{div}D_k)Y_k)$ which by the recurrence relation for X and Y

$= aX_{k+1} + bY_{k+1}$

so we have shown that every D_n is a linear combination of a and b, and we know that the second to last D_n is $\texttt{gcd}(a, b)$, so $\texttt{gcd}(a, b)$ is a linear combination of a and b.

[when I first put this up I had $k - 1$th terms and kth terms of the sequences reversed.]

I present an example of this calculation. The proof is not just an abstract argument: it directly describes how to do the calculation.

	x	y	q
925	1	0	
851	0	1	
74	1	-1	1
37	-11	12	11
0			

This calculation shows that $\texttt{gcd}(925, 851) = 37 = (-11) \cdot 925 + 12 \cdot 851$.

The first column is the D sequence, the second is the X sequence, the third is the Y sequence, and the fourth contains the integer quotient of the two entries in the first column just above that row, used in computation of the row the quotient appears in.

I will provide a spreadsheet to download on the class web page which carries out these calculations automatically. I would use this to **check** calculations, not do them in the first place, because you will not have the spreadsheet on the exam and you will need to use this computational procedure several times, not just do one demonstration.

We will demonstrate practical uses for this computation method throughout the later parts of this unit: we will use it constantly and you need to master it.

First, we present an abstract use for it in a proof:

Theorem (Euclid's lemma): Let p be a prime and let a, b be integers. If $p|ab$ then either $p|a$ or $p|b$.

Proof: Suppose p is a prime a, b are integers and $p|ab$.

If $p|a$ we are done.

Suppose p does not go into a. Then $\texttt{gcd}(p, a) = 1$ (because the only divisors of p are p and 1). Thus there are integers x and y such that $px + ay = 1$, by the extended Euclidean algorithm theorem.

Now $b = b1 = bpx + bay$. bpx is obviously divisible by p. bay is divisible by p because $p|ab$. The sum of two numbers divisible by p is divisible by p, so $p|b$, and we have shown that in either case we either have $p|a$ or $p|b$.

This theorem will be used to prove the uniqueness of prime factorizations.

Please note that this proof or something similar might appear as a test question. Be familiar with it.

5 Homework 9, assigned 10/8/2022

1. Compute the `gcd` of each pair of numbers and present the gcd as a linear combination of the original pair of numbers.

   You can get answers (with care) using the spreadsheet, but I strongly recommend doing these (and more examples) by hand. You will need the skill of constructing these tables and reading them correctly for later activities in this unit.

   (a) 55,34

   (b) 337,216

   (c) 12076, -8976. You need to think about how to handle the fact that b is negative. The spreadsheet does **not** work correctly with negative b. Hint: I would do a calculation with positive values then fix it.

2. Write out a proof that if $d|a$ and $d|b$, then $d|(a-b)$. This is very similar to something I did in class.

3. Suppose that each of us have a large supply of 115 pound notes and a large supply of 389 pound notes. How can I pay you one pound?

4. Suppose that in my mad scientist lab, recently devastated by a monster I unwittingly created, I have a balance, but can only find a large supply of 651 gram weights and a large supply of 133 gram weights.

   Can I verify the weight of an object that is supposed to weigh 28 grams?

   What is the smallest weight I can check with these weights (remember that I can put my known weights in either pan of the balance).

6 Modular arithmetic

Let $m > 1$ be an integer. We will describe the systems of "mod m arithmetic" (one for each value of m) which are finite (though large if m is large) mathematical systems which look a lot like the integers (and some of them even like the rational numbers) in theoretically and practically interesting ways.

The objects of mod m arithmetic are written as the remainders mod m (the numbers $0, 1, \ldots, m-1$). These can be thought of as either literally those numbers, , or we can think of an element n of the system of mod m arithmetic as the set $\{x \in \mathbb{Z} : x \texttt{mod} m = n\}$. Both viewpoints have some merit.

We define an important relation.

Definition: We define a relation on integers x, y. We say that x is congruent to y mod m, which is written either $x \equiv_m y$ or $x \equiv y \texttt{mod} m$ (both usages are fairly common and I am likely to write both without thinking about it) just in case $m|(x-y)$ or, equivalently, $x \,\texttt{mod}\, m = y \,\texttt{mod}\, m$.

Theorem (verifying something I say in the definition): For any integers x, y, $m|(x-y)$ iff $x \texttt{mod} m = y \texttt{mod} m$.

We remind you that the division algorithm tells us that for any integer z, $z = m(z \texttt{div} m) + z \texttt{mod} m$ and $0 \leq z \texttt{mod} m < m$.

Suppose $m|(x-y)$. $x - y = (m(x\texttt{div}m) + x\texttt{mod}m) - (m(y\texttt{div}m) + y\texttt{mod}m) = m(x\texttt{div}m - y\texttt{div}m) + (x\texttt{mod}m - y\texttt{mod}m)$. From this we can see $x\texttt{mod}m - y\texttt{mod}m = (x-y) - m(x\texttt{div}m - y\texttt{div}m)$ is divisible by m. But $-m < x\texttt{mod}m - y\texttt{mod}m < m$, so $\texttt{x}\texttt{mod}m - y\texttt{mod}m$ can only be divisible by m if it is 0, establishing $x\texttt{mod}m = y\texttt{mod}\texttt{m}$.

Suppose that $x\texttt{mod}m = y\texttt{mod}m$. Then $x - y = (m(x\texttt{div}m) + x\texttt{mod}m) - (m(y\texttt{div}m) + y\texttt{mod}m) = m(x\texttt{div}m - y\texttt{div}m) + (x\texttt{mod}m - y\texttt{mod}m) = m(x\texttt{div}m - y\texttt{div}m)$, which is divisible by m.

Definition: A relation R on a set A is an equivalence relation if and only if it is reflexive (for all $x \in A$, xRx), symmetric (for all x, y, if xRy then yRx) and transitive (for all x, y, z, if xRy and yRz, then xRz).

Notice that equality is an equivalence relation, on any set.

Theorem: $\equiv_m$ is an equivalence relation.

Quick proof: I gave a longer one in class, which I will eventually put here. If we use the formulation of $x \equiv_m y$ as meaning $x\texttt{mod}m = y\texttt{mod}m$, this is obvious, basically because equality is an equivalence relation.

reflexive: $x \equiv_m x$ if x is an integer, because $x\texttt{mod}m = x\texttt{mod}m$.

symmetric: if $x \equiv_m y$, then $x\texttt{mod}m = y\texttt{mod}m$, so $y\texttt{mod}m = x\texttt{mod}m$, so $y \equiv_m x$.

transitive: if $x \equiv_m y$ and $y \equiv_m z$, then $x\texttt{mod}m = y\texttt{mod}m = z\texttt{mod}m$, so $x\texttt{mod}m = z\texttt{mod}m$, so $x \equiv_m z$.

The proof I gave in class uses the other equivalent form of the definition ($m|(x-y)$). It is not too much harder, and is a nice example of basic proofs about divisibility, so it will eventually be here.

As I observed above the objects of mod m arithmetic are represented by the numerals $0, \ldots m-1$ which are remainders mod m. We can think of these as literally those integers, or we can think of the object represented by n ($0 \leq n < m$) as the set $\{x \in \mathbb{Z} : x \equiv_m n\}$. These classes are called the equivalence classes under the equivalence relation $\equiv_m$.

I will generally take the view that the objects of modular arithmetic are literally the remainders, but if so I am also taking the view informally that all the integers congruent mod m to a given remainder are in some way being identified with it in our modular arithmetic calculations.

We present the addition and multiplication tables for mod 5 arithmetic.

$+$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

$\cdot$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Each of these tables is constructed by computing the usual operation on integers, then taking the remainder mod 5. Temporarily using $\oplus$ for modular addition and $\odot$ for modular multiplication (we will usually use the standard notation) we define $x \oplus y = (x + y)\texttt{mod}m$ and $x \odot y = (x \cdot y)\texttt{mod}m$.

You should be able to construct these tables for other small values of m.

The addition tables are very stereotyped. The multiplication tables are more interesting.

We state a theorem which justifies this procedure (and indicates why we might identify remainders in mod m arithmetic with whole equivalence classes).

Theorem: Let $x \equiv_m x'$ and $y \equiv_m y'$. It follows that $x + y \equiv_m x' + y'$ and $x \cdot y \equiv x' \cdot y'$.

Proof: Since $x \equiv_m x'$ and $y \equiv_m y'$, we have integers k, l such that $x + km = x'$, $y + lm = y'$.

Then $x' + y' = (x + km) + (y + lm) = x + y + (k + 1)m \equiv_m x + y$

and

$x' \cdot y' = (x + km)(y + lm) = xy + kmy + xlm + klm^2 = xy + (ky + lm + klm)m \equiv_m xy$

This means that as long as we are only interested in remainders mod m of outputs of the addition and multiplication operators we can in effect identify

numbers with the same remainder mod m where they appear as inputs (we can collapse input numbers to their smaller remainders before carrying out addition or multiplication).

The commutative, associative, and distributive laws of algebra hold in these systems because they hold in the integers.

Order properties do not hold (you can think about how they fail).

The additive inverse property holds: the additive inverse of n in mod m arithmetic will be $m - n$ (so for example the additive inverse of 2 in mod 5 arithmetic is 5-2 = 3. $2 + 3\texttt{mod}5 = 0 \ldots$.

The multiplicative cancellation property does not hold in most systems of modular arithmetic. For example, $2 \cdot 3 \equiv_6 0 \cdot 3 \equiv_6 0$: in mod 6 arithmetic, $2 \cdot 3 = 0 \cdot 3$, $3 \neq 0$, but $2 \neq 0$ (we cannot "divide both sides by 3" to see that $2 = 0$ as we could in ordinary algebra).

But the multiplicative cancellation property *does* hold in mod 5, and in fact mod 5 arithmetic has a stronger property characteristic of the rational numbers, not the integers: each nonzero remainder in mod 5 arithmetic has a multiplicative inverse, so we really can define division in this system (which is really unexpected).

There is a general theorem about when this is true.

Theorem: Mod m arithmetic satisfies the multiplicative cancellation property and in fact the existence of reciprocals of each nonzero remainder if and only if m is prime.

Proof: If m is not prime then there are a and b such that $0 < a \leq b < m$ and $ab = m$.

We then have in mod m arithmetic that $a0 = ab = 0$ (because $a0 \equiv_m 0 \equiv_m m = ab$) but $0 \neq b$, so multiplicative cancellation fails (and so a does not have a reciprocal: if a^{-1} such that $a^{-1}a = 1$ in mod m arithmetic existed we would have $0 = 10 = (a^{-1}a)0 = a^{-1}(a0) = a^{-1}0 = a^{-1}(ab) = (a^{-1}a)b = 1b = b$, which is absurd.

Now suppose that m is prime. It follows that for any nonzero remainder a, we have $\texttt{gcd}(a, m) = 1$, so there are integers x, y such that $ax+my = 1$ (in the ordinary arithmetic of the integers) so $ax \equiv_m 1$, so $x]ttmodm$ is a reciprocal of a.

Further, there is only one reciprocal of a in mod m arithmetic. Suppose $ax \equiv_m ay = 1$. It follows that $m|(ax-ay)$ so $m|a(x-y)$. Now, because

m is prime and m does not go into a, we must have $m|(x-y)$ by Euclid's Lemma, so $x \equiv_m y$, and there is only one inverse of a in mod m arithmetic, which we will write a^{-1}, or $a^{-1}\texttt{mod}m$. This is not to be confused with $\frac{1}{a}$, the usual reciprocal of a.

It then follows that if $ac = bc$ and $c \neq 0$, we have $(ac)c^{-1} = (bc)c^{-1}$ and so $a = b$, ini mod m arithmetic as in the arithmetic of the rationals or reals.

This shows us not only that there are reciprocals of nonzero remainders in mod m arithmetic, but also how to compute them using the extended Euclidean algorithm. Moreover, we can solve equations of the form $ax \equiv_m b$ in mod m arithmetic, if m is prime and a is not 0, by multiplying both sides by $a^{-1}\texttt{mod}m$.

Example: Find $23^{-1}\texttt{mod}137$

	x	y	q
137	1	0	
23	0	1	
22	1	-5	5
1	-1	6	1

We see that $(-1)137+(6)(23) = 1$ so $(23)(6) \equiv_{137} 1$ so $23^{-1}\texttt{mod}137 = 6$.

Example: Solve $23x \equiv_{137} 31$

Multiply both sides by the reciprocal we just found.

$x = 1x \equiv_{137} (6)(23)x \equiv_{137} (6)(31)$ for any x for which the original equation holds, so we must have $x \equiv_{137} 186 \equiv_{137} 49$ and indeed you can check that 49 is a solution.

7 Lecture of 10/12: Chinese Remainder Theorem, introduction to exponentiation

7.1 a remark about computing modular reciprocals

I started the lecture with a reminder about computation of modular reciprocals. To find $a^{-1}\texttt{mod}m$ when m is prime (or more generally when a is relatively prime to m: some numbers have reciprocals in mod m arithmetic

when m is composite) we use the extended Euclidean algorithm to find x and y such that $mx + ay = 1$, so $ay \equiv_m 1$, so y is the desired reciprocal of a: notice that y might turn out to be negative: in that case we add m to it to get an actual remainder (if $y = -3$, the reciprocal we report is $m - 3$).

I did an example of this in class, and Ill point this out when it happens in an example in this text.

7.2 The Chinese Remainder Theorem

Next I lectured the Chinese Remainder Theorem.

Theorem: If $m, n > 1$ and $\texttt{gcd}(m, n) = 1$, each system of equations of the form

$x \equiv_m a$

$x \equiv_n b$

has a unique solution mod mn.

Proof (and method of computation!): Because $x \equiv_m a$, there is an integer k such that $x = a + mk$.

So to find x solve for k the equation

$a + mk \equiv_n b$ [it is convenient here to let m be the larger of the two moduli, but the procedure works either way]

which is equivalent to

$mk \equiv_n (b - a)$ [notice that m and $b - a$ can be replaced with their remainders mod n since this is a calculation in mod n arithmetic]

Because $\texttt{gcd}(m, n) = 1$, there is a reciprocal $m^{-1}\texttt{mod}n$ (which we know how to compute using the extended Euclidean algorithm)

Thus we get

$k \equiv_n (m^{-1}\texttt{mod}n)(b - a)$

so for some integer l we have $k = (m^{-1}\texttt{mod}n)(b - a) + ln$

and we have $x = a + m(m^{-1}\texttt{mod}n)(b - a) + lmn$, that is

$x \equiv_{mn} a + m(m^{-1}\texttt{mod}n)(b - a)$.

Strictly speaking our reasoning shows that a solution x must be of this form. You can check directly that

$a + m(m^{-1} \mathtt{mod} n)(b-a) + lmn \equiv_m a$ [this is obvious, everything but the a is a multiple of m]

and

$a + m(m^{-1} \mathtt{mod} n)(b-a) + lmn \equiv_n a + 1(b-a) + 0 = b$,

so in fact every number of this form is a solution and we have an exact description of the solutions.

I do observe that the formula here does provide a compact way to compute solutions to these problems, which I have seen students use successfully. I will carry out the described computation step by step in examples, that being my particular practice.

I reproduce the example I did in class, which also illustrates that more than two equations of the form $x \equiv_m a$ for different moduli m can be solved at the same time, as long as the moduli are pairwise relatively prime.

Example: 23 pirates are dividing a stack of pieces of eight.

The pieces of eight are divided evenly under jealous eyes into 23 stacks...but there are four too few to make the piles even.

A scuffle ensues, cutlasses are drawn, and four pirates are interred in unmarked graves.

But now the pile of pieces of eight has to be divided again among the 19 remaining pirates.

There are two too few for an even division, and in the ensuing fracas two more pirates are despatched.

Now the 17 remaining pirates divide the loot and it comes out even! They sail off in various directions in search of more mischief...

How many pieces of eight are there?

[Disclaimer: no pirates were actually harmed in the creation of this word problem]

Solution: We have the following equations (x being the unknown number of pieces of eight):

$x \equiv_{23} -4 \equiv_{23} 19$

$x \equiv_{19} -2 \equiv_{19} 17$

$x \equiv_{17} 0$

We solve the first pair of equations.

$x \equiv_{23} \equiv_{23} 19$

$x \equiv_{19} \equiv_{19} 17$

We begin by setting $x = 19 + 23k$ using the first equation.

We then want to solve for k in

$19 + 23k \equiv_{19} 17$

which is equivalent to

$23k \equiv_{19} -2$

which is equivalent to

$4k \equiv_{19} 17$ [replace 23 and -2 with their remainders mod 19]

so we need to compute $4^{-1}\texttt{mod}19$ (which is not hard to guess but we will compute it)

$$\begin{array}{cccc} 19 & 1 & 0 & \\ 4 & 0 & 1 & \\ 3 & 1 & -4 & 4 \\ 1 & -1 & 5 & 1 \end{array}$$

so $4^{-1}\texttt{mod}19 = 5$
so $k \equiv_{19} (17)(5) \equiv_{19} 9$
and $x \equiv_{437} 19 + (23)(9) = 226$ [$437 = mn = (23)(19)$]
and now we have two equations
$x \equiv_{437} 19 + (23)(9) = 226$
$x \equiv_{17} = 0$
We set $x = 226 + 437k$
to get a new equation
$226 + 437k \equiv_{17} 0$
The remainder of 226 on division by 17 is 5
The remainder of 437 on division by 17 is 12
so we have
$5 + 12k \equiv_{17} 0$
$12k \equiv_{17} -5 \equiv_{17} 17 - 5 = 12$
so we don't have to do anything elaborate to find $k = 1$.
Thus

$x \equiv_{(437)(17)} 226 + 437 = 663$
437 times 17 is 7429
so our solution is
$x \equiv_{7429} 663$.

If I want to force an exact answer, I could ask what the smallest possible number of coins was. Or I could give extra information such as...we know they didnt have more than eight thousand coins.

7.3 Modular exponentation

The topic introduced at the end of the lecture was exponentiation in modular arithmetic.

If a and n are integers and $a \equiv_m A$, we will have $a^n \equiv_m A^n$. We can simplify the base in a modular exponentiation calcuation. The reason for this is just that exponentiation is repeated multiplication.

It is NOT TRUE that $n \equiv_m N$ implies $a^n \equiv_m a^N$. I gave a counterexample in lecture, and I invite you to construct one in a homework problem.

Nonetheless, we can compute quantities like $a^n \texttt{mod} m$ very efficiently, with far fewer than n multiplications, and without computing any numbers larger than about m^3.

We describe the technique by narrating the example we gave in class: compute $27^{100}\texttt{mod}211$.

We begin by making a table of exponents 100,50,25,12,6,3,1, where each exponent is half the previous one, possibly rounding down to get an integer. These are the only powers of 27 mod 211 that we need to compute.

I usually build this table upward, but writing it out in a LaTeX document rather suggests working from smaller exponents to larger, so Ill turn the table over :-)

27^1	27
27^3	$19683 - (93)(211) = 60$
27^6	$(27^3)^2 = 60^2 = 3600 - (17)(211) = 13$
27^{12}	$(27^6)^2 = 13^2 = 169(< 211!)$
27^{25}	$27^{24}(27) = (27^{12})^2(27) = 169^2(27) = 771147 - (3654)(211) = 153$
27^{50}	$(27^{25})^2 = 153^2 = 23409 - (110)(211) = 199$
27^{100}	$(27^{50})^2 = 199^2 = 39601 - (187)(211) = 144$

So the result is 144.

Note that we absolutely did not have to carry out 100 multiplications by 27.

This is called the method of repeated squaring. Even powers are computed simply by squaring (and reduction to remainders); for odd powers you have to multiply in an extra copy of the base of your exponential calculation (27 in this case: see what I did with the 25th power).

The number of powers of the base you need to compute is roughly the logarithm to the base two of the exponent. You never need to compute a number larger than m^3 (where m is the modulus): the worst case is where you have to square a value and multiply in an extra copy of the base, for an odd power.

This computation is manageable in size for a computer even if base, exponent, and modulus have hundreds of digits. This is noteworthy because computation of an exponentiation with an exponent with hundreds of digits is clearly simply impossible in a direct way.

We will see that the ability to compute exponentials in modular arithmetic in this way has practical uses.

8 Homework 10 assigned 10/13/2022 due Mon 10/17/2022

1. Construct the addition and multiplication tables for mod 7 arithmetic.

2. Compute the reciprocal of 21 in mod 137 arithmetic. Show all work.

3. Solve the system of equations

   $x \equiv_{111} 85$

   $x \equiv_{137} 60$

   State the general solution and the smallest positive solution.

   Show all work.

4. Pirates, again.

   137 pirates have a large stack of doubloons to dole out.

   Captain Hook doles them out evenly only to find that when he had made 137 even piles as large as possible there were 6 doubloons left over. So he casually sent two of his colleagues to walk the plank and tried again. When he made 135 equal piles as large as possible, there were 57 doubloons left over. He called for a round of rum and poisoned

two glasses, and tried again. When he made 133 equal piles as large as possible, there were 62 left over. At this point, the pirates mutinied and divided the loot with less mathematical precision (we won't discuss what happened to Captain Hook: it is hard for pirate captains to get life insurance). I can tell you that it was a scant year for pirates and they did not have millions of doubloons. How many doubloons did they have?

Show all work.

5. Find an example illustrating that it does not follow from $b \equiv_m B$ that $a^b \equiv_m a^B$ (examples are very easy to find if you understand what is being asked. Choose almost any smallish m and try some numbers).

6. Compute the last three digits of 3^{75}, using the method of repeated squaring and showing all work (hint: this is asking you to compute $3^{75} \texttt{mod} 1000$: an advantage here is that the remainder mod 1000 of a number is visually obvious).

   An advantage here for checking is that you actually can compute 3^{75} on a computer, and maybe on your calculator if it is very nice. But your calculation should show work with no numbers nearly as large as 3^{75}, and nothing like 75 multiplications.

7. Compute the last three digits of 27^{1024} using repeated squaring (why is this really enormous exponent actually rather easy to work with?) Show all work.

9 Notes from Friday 10/14/2022

In my first version, I am going to state theorems proved, and fill in the details of the description of the RSA algorithm to support the homework. I'll come back and fill in all the proofs today (Tuesday) but I want to release the homework as soon as I can.

9.1 The well-ordering principle

Definition: Let A be a set of integers. We say that $b \in \mathbb{Z}$ is a lower bound for A if and only if for every $a \in A$ we have $b \leq A$. Notice that a

smallest element of a set is a lower bound of it, and so is every integer smaller than the smallest element of the set.

Theorem (well-ordering principle): Any nonempty set of integers which is bounded below has a smallest element.

Suppose that S is a nonempty set of integers and b is a lower bound for S.

Suppose further that S has no smallest element: we then prove by strong induction that S is empty; this demonstrates that if S is nonempty and bounded below, it must be nonempty.

The induction proof: we prove by strong induction that if S is bounded below by b and has no smallest element, then for every $n \geq b$, $n \notin S$ (which shows that S is empty, because certainly nothing less than b is in S).

Basis: If $b \in S$, then b would be the smallest element of S, so $b \notin S$.

Induction: Let k be chosen arbitrarily. Suppose that for all m with $b \leq m \leq k$, we have $b \notin S$. It then follows that if $k+1 \in S$, $k+1$ is the smallest element of S, and there isn't supposed to be one. So under the assumptions, $k+1 \notin S$, completing the proof.

Alternative argument: I've never seen it done this way, but I thought of another way to prove this which doesn't involve counterfactuals. Suppose S is bounded below by b. Show by strong induction that every element of S is either a smallest element of S or larger than a smallest element of S.

Basis: If b is not in S there is nothing to check. If $b \in S$, it is the smallest element of S, and we have this case.

Induction: Let k be chosen arbitrarily, and suppose that each m with $b \leq m \leq k$, if it is an element of S, is either a smallest element of S or larger than a smallest element of S. If any such $m \in S$, we have by inductive hypothesis a smallest element of S which is $\leq m$. If no m with $b \leq m \leq k$ is in S and $k+1 \notin S$, then there is nothing to check. f no m with $b \leq m \leq k$ is in S and $k+1 \in S$, then $k+1$ is the smallest element of S.

Now use this to prove the well-ordering principle: if S is nonempty and bounded below by b, then there is $a \in S$, and by the proof just given,

a is either the smallest element of S or there is a smallest element of S less than a, and in either case S has a smallest element, which is what was to be shown.

9.2 The fundamental theorem of arithmetic (unique prime factorization)

Theorem: Each integer greater than or equal to two can be factored into primes (we already proved this) but in addition can be factored into primes in only one way (to put this exactly, each $n \geq 2$ can be expressed as the product of a sequence of primes $\prod_{i=0}^{n} p_i$ where the sequence p_i of primes is nondecreasing ($p_i > p_j \rightarrow i > j$) in exactly one way.

9.3 Fermat's little theorem; a theorem for RSA

Fermat's little theorem: If p is prime and p does not go into a, then $a^{p-1} \equiv_p 1$.

Proof: Suppose p is prime and p does not go into a. $(p-1)!$ is not divisible by p, so $(p-1)! \not\equiv_p 0$

$(p-1)!$ is the product of one copy of each b with $1 \leq b \leq p-1$.

Now consider the product of one copy of ab with $1 \leq b \leq p-1$. Clearly this is $a^{p-1}(p-1)!$ because we are just including $p-1$ additional factors of a in the previous product.

The product of all the ab's with $1 \leq b \leq p-1$ is congruent mod p to the product of all $ab \texttt{mod} p$ with $1 \leq b \leq p-1$.

Now each c with $1 \leq c \leq p-1$ is equal to $a(a^{-1} \texttt{mod} p) \texttt{mod} p$, and so appears in the previous product, and so in fact the product of all the ab's with $1 \leq b \leq p-1$ is $(p-1)!$ written in a different order.

So $a^{p-1}(p-1)! \equiv_p (p-1)!$, and we can multiply both sides by the mod p reciprocal of $(p-1)!$ to get $a^{p-1} \equiv_p 1$.

Theorem (a special case of Euler's theorem, useful for RSA): If $p \neq q$ are distinct primes, and $\texttt{gcd}(a, pq) = 1$, then $a^{(p-1)(q-1)} \equiv_{pq} 1$.

Suppose that p and q are distinct primes and $\texttt{gcd}(a, pq) = 1$.

This means that neither p nor q goes into a so $a^{p-1} \equiv_p 1$ and $a^{q-1} \equiv_q 1$.

Since any power of 1 is 1, this means that $a^{(p-1)(q-1)} \equiv_p 1$ and $a^{(p-1)(q-1)} \equiv_q 1$.

This means that $a^{(p-1)(q-1)}$ is a solution of $x \equiv_p 1; x \equiv_q 1$. By the Chinese Remainder Theorem, any two solutions of this system of equations are congruent mod pq. But $x = 1$ is also a solution so $a^{(p-1)(q-1)} \equiv_{pq} 1$.

A bonus theorem, introduced in discussion of the Rabin Miller test for primes: If $a^2 \equiv_m 1$ and m is prime, then $a \equiv_m 1$ or $a \equiv_m -1$.

This is totally unsurprising, it looks like a familiar fact of algebra.

Suppose $a^2 \equiv_m 1$. Then $a^2 - 1 \equiv_m 0$, so $m|a^2 - 1 = (a-1)(a+1)$. Then if m is prime it follows that either $m|a-1$ (so $a \equiv_m 1$) or $m|a+1$ (so $a \equiv_m -1$)

The negative form of this is what we use in the Rabin Miller test. When we are doing repeated squaring to compute a^{m-1}, we do a check whenever we compute a $b^2 \equiv_m 1$ at any stage of the repeated squaring: if b is not either 1 or $m-1$, we can stop the computation and report that m is composite. At the end, if we haven't encountered bad square roots of 1, if $a^{m-1} \not\equiv_m 1$, we find that m is composite by Fermat's little theorem (turned on its head). If neither of these things happen, m might be prime.

9.4 The RSA Algorithm

The RSA algorithm is a public key cryptosystem: this means that the method of encrypting a message to send to a user is public. For this to work, it has to be a hard mathematical problem to deduce the method of decryption from the method of encryption. In the case of RSA, the belief that this problem is hard hinges on the fact that it is apparently very hard to factor products of two large primes (large being about 350 digits currently, so 700 digit RSA keys).

What is supplied as my public key is two integers, N and r.

You send a message M (a remainder mod N) to be by computing $M' = M^r \texttt{mod} N$ and sending M' to me on open channels. Presumably only I can read it.

In the secrecy of my lab, I know that $N = pq$, where p and q are distinct primes, and further I know that $\texttt{gcd}(r, (p-1)(q-1)) = 1$.

I compute $s = r^{-1} \mathtt{mod}(p-1)(q-1)$ [it is very important to use the modulus $(p-1)(q-1)$ here, and this is the only stage in the process where it is used]. This can be done (in spite of the fact that $(p-1)(q-1)$ is composite) because r is relatively prime to $(p-1)(q-1)$, so there are integers s and t such that $rs + t(p-1)(q-1) = 1$, which we can find using the extended Euclidean algorithm, and this s is the one we are looking for [with the usual remark that we need to adjust the output of the EEA if it is negative].

Now I can decrypt $M' = M^r \mathtt{mod} N$ because $(M')^s \mathtt{mod} N = M$: we prove this.

$(M')^s \mathtt{mod} N =$ by def of M'

$(M^r)^s \mathtt{mod} N =$ by rules of exponents and reduction of bases of an exponentiation by a modulus

$M^{rs} \mathtt{mod} N =$ by the fact that $s = r^{-1} \mathtt{mod}(p-1)(q-1)$

$M^{1+k(p-1)(q-1)} \mathtt{mod} N =$ by rules of exponents

$M \cdot (M^{(p-1)(q-1)})^k =$ by the special case of Euler's theorem

$M \cdot 1^k \mathtt{mod} N = M$ (last step using the fact that M is a remainder mod N).

An Example.

Choose $p = 19, q = 23$.

we then have $N = (19)(23) = 437$.

We want to choose r relatively prime to $(18)(22) = 396$. It is conventional but certainly not mandatory to choose the smallest prime which does not go into $(p-1)(q-1)$: this convention gives $r = 5$.

I publish my key, $N = 437; r = 5$.

I privately compute $5^{-1} \mathtt{mod} 396 = 317$.

You send me the message $M = 42$.

$M' = 42^5 \mathtt{mod} 437$.

5	264	$256 = 16^2 \mathtt{mod} 437$	$264 = (256)(42) \mathtt{mod} 437$
2	16	$16 = 42^2 \mathtt{mod} 437$	16
1	42		42

The computation of M' is exhibited as the spreadsheet displays it, with comments indicating how the numbers are computed. The first column is the exponent and the second is the corresponding power of 42 mod 437. Notice that the fourth column repeats the third unless the exponent in that row is odd. In the third column what is computed is just the square of the previous entry. The idea of this format is that the bound on the numbers

with compute with is the square of the modulus instead of the cube of the modulus.

I receive the message and decrypt it by computing $264^{317} \mathtt{mod} 437$.

317	42	$169 = 82^2 \mathtt{mod} 437$	$42 = (169)(264) \mathtt{mod} 437$
158	82	$82 = 385^2 \mathtt{mod} 437$	82
79	385	$311 = 106^2 \mathtt{mod} 437$	$358 = (311)(264) \mathtt{mod} 437$
39	106	$156 = 245^2 \mathtt{mod} 437$	$106 = (156)(264) \mathtt{mod} 437$
19	245	$39 = 134^2 \mathtt{mod} 437$	$245 = (39)(264) \mathtt{mod} 437$
9	134	$123 = 358^2 \mathtt{mod} 437$	$134 = (123)(264) \mathtt{mod} 437$
4	358	$358 = 213^2 \mathtt{mod} 437$	358
2	213	$213 = 264^2 \mathtt{mod} 437$	213
1	264		264

You might want to go through this table and check for typos.

10 Homework 11, assigned Tuesday 10/18/2022 and due Monday 10/24/2022

1. In this problem you should do all calculations by hand and show all details, only using the spreadsheets to check.

   Work through the entire process of setting up your RSA key using $p = 11, q = 17, r = 7$.

   You need to state what N is, verify that r has the correct property.

   If I send you the message $M = 42$, what is the M' you will receive?

   If you receive from me the message $M' = 76$, what was the message I sent?

2. Longer messages

   Messages expressed in the usual alphabet can be coded as numbers in various ways. One method is to replace each letter by two digits, A = 01, B = 02 up to Z = 26 and then space = 27.

   Then the numbers may be larger than M. So break the numbers into chunks, all of the same length and all less than M, and encrypt them one by one.

The key $N = 28907 = (137)(211), r = 11$ which I used in class could be used to encrypt 4 digit (2 letter) blocks. Encrypt the message "RUN FOR YOUR LIFE" as a series of numbers. Show all work, but you may copy from the spreadsheet (or include snapshots of your spreadsheet calculations).

Cryptographically of course this is very weak: a cipher encoding two digit blocks is easy to solve using statistical analysis. But a larger N will allow encryption of much longer blocks of text which will not be likely to be repeated.

3. Message exchange

   I supply you with the key $N = 10403; r = 7$. Send me a coded message, a brief text sentence as in the previous problem, coding four digits at a time.

   You prepare an RSA key of your own with N at least 10000. Make sure that it is not too large to work with my spreadsheet (do some practice encryption and decryption). Send me your RSA key at the same time you send me your encrypted message for the first par and I will send you a reply to your message.

   In both parts of this problem, it makes sense to retain your work (which may use the spreadsheet of course) but you do not need to send it to me: just send me the message and your public key. I talked about this in class. You aren't going to lose points if you send work to me, but cryptographically that is your private stuff :-)

   Be aware that a point on this problem depends on replying when I send you a message using the key I send you, with the decryption of the message.

4. Verify that 49 is not a prime by finding a such that $a^{48} \texttt{mod} 49 \neq 1$.

   Find an $a < 49$ which lies and claims in effect that 49 is a prime. This indicates why you have to repeat the test with many random values of a to be certain (in practical terms) that you have a prime. But it doesn't happen often.

   Verify that 3293 is not a prime in the same way.

5. Compute $3^{1000} \texttt{mod} 23$ by repeated squaring.

You can use the spreadsheet to check, but this might be a good drill for reliably doing these calculations on an exam.

Then do this calculation using Fermat's little theorem. Depending on how good your calculator is, you might have to do a wee amount of repeated squaring, but not very much.

Try determining by experiment a similarly efficient method of computing powers mod 49. 49 is of course not a prime, so the exact approach of Fermat's theorem doesn't work. But there is a modulus you can apply to reduce the exponent in any calculation of an $a^n \texttt{mod} 49$. Find out what it is. You can use the spreadsheet to quickly compute powers mod 49 for your experimental investigation.