This is a set of notes for the number theory unit of Math 55, which are mostly taken from Niven's *Introduction to the Theory of Numbers.* They will be continuously updated throughout the number theory unit (so there will be quite a lot of loose ends until the unit is finished and the notes are finalized, and the statements in the notes will not be in the same order that I discuss them in class).

Please send any questions/comments/corrections to hhao@berkeley.edu.

1 Divisibility

The fundamental object of study of number theory is the integers $\mathbf{Z}$. Therefore we introduce some basic concepts about the integers that allow us to find deeper relationships among its elements.

Definition 1.1. Let a and b be integers. If there is an integer x such that $b = ax$, then b is *divisible* by a, or a divides b. We write $a|b$. If a does not divide b, then we write $a \nmid b$. In the case that a, b are positive integers and $0 < a < b$, then we say that a is a *proper divisor* of b.

Example 1.1. Every integer a divides 0. Conversely, 0 does not divide any integer besides 0.

Theorem 1.1. (1) If $a|b$, then $a|bc$ for any integer c.

(2) If $a|b$ and $b|c$, then $a|c$.

(3) If a divides each b_i for a finite set of integers $b_1, \ldots, b_n$, then a divides $\sum_{i=1}^{n} b_i x_i$ for any integers x_i.

(4) If $a|b$ and $b|a$, then $a = \pm b$.

(5) If $a|b$ and a, b are positive integers, then $a \leq b$.

(6) If $m \neq 0$, then $a|b$ if and only if $ma|mb$.

Proof. We prove (4), and leave the rest as exercises. If $a|b$ and $b|a$, then $a = 0$ if and only if $b = 0$, in which case the proposition is true. So suppose a and b are nonzero, and there exist integers x, y such that $ax = b$ and $by = a$. Then $axy = a$, and since $a \neq 0$, we have $1 = xy$. Then x and y are either both positive or both negative. Assuming that they are both positive, then because 1 is the smallest positive integer (an axiom of $\mathbf{N}$), we have $x \geq 1, y \geq 1$, so $xy \geq 1$ with equality if and only if $x = 1$ and $y = 1$. Therefore $a = b$. In the other case when x and y are both negative, the same argument shows $-x = -y = 1$, so $a = -b$. □

The next statement is the *division algorithm*, which proves that division in $\mathbf{Z}$ is possible and uniquely defined, up to some conditions.

Theorem 1.2 (Division algorithm)**.** Let a and b be integers, where $a > 0$. Then there exist unique integers q (the quotient) and r (the remainder) such that $b = aq + r$ and $0 \leq r < a$. If $a \nmid b$, then $r \neq 0$.

Remark 1.1. The class textbook defines b **div** a to be the quotient q, and b **mod** a to be the remainder r. We will *not* use this notation in this class.

Proof. Consider the set S of all *nonnegative* integers of the form $b - ka$ for some $k \in \mathbf{Z}$. This set is nonempty (why?), so we invoke the *well-ordering property* of the nonnegative integers $\mathbf{N} \cup \{0\}$, and pick the smallest integer r from S. Then if $r = b - aq$, we must have $0 \leq r < a$, as otherwise $r - a = b - a(q+1)$ is also in S but strictly less than r, contradicting how we chose r. Therefore the pair (q, r) satisfy the conditions of the theorem.

We now prove uniqueness of q and r. Suppose there is another pair (q', r') satisfying the same conditions. We claim that $r = r'$. If not, then WLOG $r < r'$, so $0 < r' - r < a$, and then $r' - r = (b - aq') - (b - aq) = a(q - q')$. So by definition, $a | (r' - r)$, a contradiction to item (5) of Theorem 1.1. Therefore $r' = r$, and from this it follows that $q = q'$.

The last statement of the theorem is left as an exercise. □

This theorem probably expresses a fact you already knew: you can always divide an integer by a positive integer a, and obtain a positive remainder strictly smaller than a. But it is useful to record this result rigorously, because we will need it very soon.

Example 1.2. The quotient and remainder then -13 is divided by 4 are $q = -4$ and $r = 3$.

Definition 1.2. We say a is a *common divisor* of integers b and c if $a|b$ and $a|c$. If at least one of b and c is not zero (say, $b \neq 0$), then b only has finitely many divisors, so *a fortiori* there are only finitely many common divisors of b and c. We say the greatest integer among the set of common divisors of b and c is the *greatest common divisor (gcd)* of b and c, denoted $\gcd(b, c)$ or (b, c).

We can similarly extend this notion to any finite set $b_1, \ldots, b_n$ of integers, not all 0. We write their (mutual) gcd as $(b_1, b_2, \ldots, b_n)$. Also, note that by definition, the gcd of any two integers is positive.

Example 1.3. We have $\gcd(6, 9) = 3$, $\gcd(-4, -8) = 4$, and $\gcd(6, 10, 15) = 1$. Also, $\gcd(a, 0) = a$ for any positive integer a.

The following result relates the gcd, which is a notion stemming from divisibility, to *linear expressions* with integer coefficients.

Theorem 1.3 (Bezout's Lemma)**.** Let $g = (b, c)$. Then there exist integers x_0, y_0 such that $g = bx_0 + cy_0$.

Proof. Consider the set $S = \{bx + cy : x, y \in \mathbf{Z}\}$ of all possible integer linear combinations of x and y. This set includes some positive integers, so choose x_0, y_0 such that $bx_0 + cy_0 = l$ is the smallest positive integer in S. We claim $l|b$. If not, then there exist unique integers q and r such that $b = lq + r$ and $0 < r < l$ by Theorem 1.2. Then

$$r = b - lq = b - q(bx_0 + cy_0) = b(1 - qx_0) + c(-qy_0) \in S$$

contradicting r being the smallest positive integer in S. Therefore we must have $l|b$, and similar $l|c$. Now, since g is the *greatest* common divisor of b and c, we may write $b = gB$, $c = gC$, and so $l = bx_0 + cy_0 = g(Bx_0 + Cy_0)$. Therefore $g|l$, but g and l are positive integers, so $g \leq l$. By maximality of g, we must have $g = l$. □

Notice that the proof of this theorem gives an alternate description of $\gcd(b, c)$: it is the least positive integer than can be written as an integral linear combination of b and c. In particular:

Corollary 1.1. The gcd of b, c is 1 if and only if there are integers x, y such that $bx + cy = 1$.

Proof. The "only if" statement is immediate. For the "if" statement, if there are integers such that $bx + cy = 1$, then the least positive integer that can be written as an integral linear combination of b and c must be 1 (as any positive integer is at least 1), so $(b, c) = 1$. □

This case is so special that it deserves a name:

Definition 1.3. We say that two integers a, b are *relatively prime*, or *coprime*, if $(a, b) = 1$. More generally, a list $b_1, \ldots, b_n$ of integers is *coprime* if $(b_1, \ldots, b_n) = 1$. We say such a list is *pairwise coprime* if $(b_i, b_j) = 1$ for all $i \neq j$.

As we will see later on, the intuition for coprime integers is that divisibility conditions with respect to those integers "behave independently," precisely because those integers share no nontrivial factors in common.

Example 1.4. The integers 6 and 35 are coprime. The set of integers $\{6, 10, 15\}$ is coprime, but not pairwise coprime, as $(6, 10) = 2$, $(6, 15) = 3$, $(10, 15) = 5$. In particular, a set of integers being pairwise coprime implies they are coprime, but not the other way around.

Corollary 1.2. The gcd g of integers b, c is the unique positive common divisor of b that is divisible by every common divisor.

Proof. If d is any common divisor of b and c, then by (3) of Theorem 1.1 and Theorem 1.3, we conclude that $d|g$. To show uniqueness of g, if another positive common divisor g' satisfies this property, then we have $g|g'$ and $g'|g$, so $g = g'$ as both are positive (see (4) of Theorem 1.1). □

Theorem 1.3 can be generalized to the gcd of an arbitrarily large list $b_1, \dots, b_n$ of positive integers, but this is not necessary for any part of our discussion, so we omit this. Instead, we prove a bunch of properties about the gcd.

Proposition 1.1. For any $m \in \mathbf{N}$, $(ma, mb) = m(a, b)$.

Proof. We know that (ma, mb) is the least positive value of $max + mby$ as x, y range over all integers, and this least value is precisely $m \cdot \min\{ax + by : ax + by > 0, x, y \in \mathbf{Z}\}$, which is equal to $m \cdot (a, b)$. □

Proposition 1.2. If $d|a$ and $d|b$ with d a positive integer, then $(a/d, b/d) = (a, b)/d$. If $(a, b) = g$, then $(a/g, b/g) = 1$.

Proof. The first assertion follows from Proposition 1.1 when we replace m, a, and b by d, a/d, and b/d. The second assertion follows from the first: it is the special case $d = (a, b)$. □

Proposition 1.3. If $(a, m) = (b, m) = 1$, then $(ab, m) = 1$.

Proof. By Theorem 1.3, there are integers x, y, x', y' such that $ax+my = 1$ and $bx'+my' = 1$. Then $axbx' = (1 - my)(1 - my')$, and the right-hand side is of the form $1 - my''$ for some integer y''. Therefore $abxx' + my'' = 1$, so $(ab, m) = 1$ by Corollary 1.1. □

Proposition 1.4 (Euclid's Lemma)**.** If $c|ab$ and $(b, c) = 1$, then $c|a$.

Proof. By Proposition 1.1, we have $(ab, ac) = a(b, c) = a$. Since $c|ab$ by assumption, and we certainly have $c|ac$, we conclude that $c|a$, since the gcd a of ab and ac is divisible by all other common divisors by Corollary 1.2. □

We now come to the problem of actually *computing* the gcd. Note that we do not have the tool of unique factorization yet, so that cannot be used, even though we are able to prove the relevant theorem now. Moreover, the actual computation of the factorization of an integer is difficult (this is the assumption on which almost all of modern cryptography rests!), and so any method using prime factorization to calculate gcd is not feasible. Moreover, it is not possible to directly use the result from the proof of Theorem 1.3, that $\gcd(b, c)$ is the smallest positive integer that is an integer linear combination of b and c: this would require some way of picking the smallest element out of an infinite set (that is most likely not in increasing order). Fortunately, the next two results allow us to use the division algorithm to quickly compute the gcd (which is actually what computers use, especially for small inputs!).

Theorem 1.4. For any integer x, $(a, b) = (b, a) = (a, -b) = (a, b + ax)$.

Proof. The first two equalities are clear.

Write $d = (a, b)$ and $g = (a, b + ax)$. By Theorem 1.3, find integers x_0, y_0 such that $d = ax_0 + by_0$. Then we have

$$d = a(x_0 - xy_0) + (b + ax)y_0,$$

and since $g = (a, b + ax)$ divides both a and $b + ax$, we also have $g|d$. Conversely, since $d|a$ and $d|b$, we have $d|(b+ax)$, and from Corollary 1.2, we conclude that $d|g$ since g is divisible by every common divisor of a and $b + ax$. Hence $d = \pm g$, and since both are positive by definition of gcd, we must have $d = g$. □

Before we give the formal *Euclidean algorithm*, let's first do a numerical example. Cpnsider $b = 963, c = 657$. We can divide b by c as normal: we get a quotient $q = 1$ and remainder $r = 306$. Then $(b, c) = (b - cq, c)$ by Theorem 1.4, so

$$(963, 657) = (963 - 1 \cdot 657, 657) = (306, 657) = (657, 306).$$

So we have reduced the calculation to a pair of *smaller* integers: we replaced (b, c) with (c, r), where $c \leq b$ and $r < c$. This is a great achievement, because upon repeating the same procedure, we can reduce the calculation to an even smaller pair of integers, and so on!

So upon dividing 657 by 306, we get $q = 2$ and $r = 45$, so that

$$(657, 306) = (657 - 2 \cdot 306, 306) = (306, 45).$$

Continuing in this fashion, we end up getting

$$(963, 657) = (657, 306) = (306, 45) = (45, 36) = (36, 9) = (9, 0) = 9.$$

This process can be reversed in order to write 9 as an integer linear combination of 963 and 657. Using the quotients and remainders given by the division algorithm, we have

$$\begin{aligned}
306 &= 963 - 657, \\
45 &= 657 - 2 \cdot 306 = 657 - 2 \cdot (963 - 657) \\
&= 3 \cdot 657 - 2 \cdot 963, \\
36 &= 306 - 6 \cdot 45 = (963 - 657) - 6 \cdot (3 \cdot 657 - 2 \cdot 963) \\
&= 13 \cdot 963 - 19 \cdot 657, \\
9 &= 45 - 36 = (3 \cdot 657 - 2 \cdot 963) - (13 \cdot 963 - 19 \cdot 657) \\
&= 22 \cdot 657 - 15 \cdot 963.
\end{aligned}$$

Let us generalize this procedure to arbitrary integers b, c. If $c = 0$, then we know that $(b, 0) = |b|$. Also, because $(b, c) = (b, -c)$, we may assume that $c > 0$.

Theorem 1.5 (Euclidean algorithm)**.** Given integers b, c with c positive, we may repeat the division algorithm, Theorem 1.2, to obtain a finite series of equations

$$\begin{aligned}
b &= cq_1 + r_1, \quad 0 < r_1 < c,\\
c &= r_1 q_2 + r_2, \quad 0 < r_2 < r_1,\\
r_1 &= r_2 q_3 + r_3, \quad 0 < r_3 < r_2,\\
&\ldots\\
r_{j-2} &= r_{j-1} q_j + r_j, \quad 0 < r_j < r_{j-1},\\
r_{j-1} &= r_j q_{j+1}.
\end{aligned}$$

Then $(b, c) = r_j$, where r_j is the last nonzero remainder in the division process. We may find $x, y \in \mathbf{Z}$ such that $(b, c) = bx + cy$ by writing each r_i as a linear combination of b and c, and "back-substituting."

Proof. Note that at each step, the kth remainder r_k is a positive number *strictly less* than r_{k-1} (set r_0 to be the given integer $c > 0$ for convenience), and there are only finitely many nonnegative integers less than c. So at some step l, the lth remainder r_l must be 0 (we cannot have an infinite chain of decreasing remainders $r_0 > r_1 > r_2 > \ldots > 0$), in which case the algorithm terminates immediately.

To show that r_j, the last nonzero remainder in the division process, is equal to (b, c), we compute

$$(b, c) = (b - cq_1, c) = (c, r_1) = (c - r_1 q_2, r_1) = (r_1, r_2) = (r_2, r_1 - r_2 q_3)$$

$$= (r_2, r_3) = \ldots = (r_{j-1}, r_j) = (r_j, 0) = r_j.$$

To be rigorous, the ... requires some sort of inductive argument, but the idea is clear enough.

For the last part, we note that r_1 is an integer linear combination of b and c, by construction. Then r_2, being an integer linear combination of c and r_1, is an integer linear combination of b and c via substitution for r_1. In general, r_i is an integer linear combination of r_{i-1} and r_{i-2}, so if the latter two are integer linear combinations of b and c, then so is r_i. So via this recursive/inductive back-substitution procedure, $r_j = (b, c)$ can be written as an integer linear combination of b and c, derived from the series of equations obtained from the division algorithm. □

Exercise 1.1. Use the Euclidean algorithm to find $(42823, 6409)$, and write this gcd g as an integer linear combination of 42823 and 6409 (you should get $g = 17$ and $17 = -22 \cdot 42823 + 147 \cdot 6409$).

2 Modular Arithmetic

The notion of divisibility allows us to introduce the various useful notion of *modular arithmetic*. As a motivating example, consider the following question:

Question 2.1. Do there exist positive integers x, y, z such that $3x+5\cdot 6^y+1 = z(z+1)(z+2)$?

After a moment's reflection, we see that the answer is *no*, because the left-hand side is not divisible by 3 (why?), while the right-hand side is divisible by 3 (because one of z, $z+1$, or $z+2$ is always divisible by 3, for any integer z). This is an example of the power of modular arithmetic: we focus on a single *divisibility condition* with respect to 3, stripping away extraneous details from the equation.

Definition 2.1. If a and b are integers and m is a positive integer, we say that *a is equivalent to b modulo (mod) m*, written $a \equiv b \bmod m$, if $m|(a-b)$. We sometimes also say that b is a *residue* of a mod m (this will be made more clear below). In other words, $a \equiv b \bmod m$ if and only if there is an integer k such that $a = b + mk$. If $m \nmid (a-b)$, then we write $a \not\equiv b \bmod m$.

Some basic facts:

Theorem 2.1. (1) The following are equivalent: $a \equiv b \bmod m$, $b \equiv a \bmod m$, $a - b \equiv 0 \bmod m$, a and b leave the same remainder upon division by m.

(2) If $a \equiv b \bmod m$ and $b \equiv c \bmod m$, then $a \equiv c \bmod m$.

(3) If $a \equiv b \bmod m$ and $c \equiv d \bmod m$, then $a + b \equiv c + d \bmod m$.

(4) If $a \equiv b \bmod m$ and $c \equiv d \bmod m$, then $ab \equiv cd \bmod m$.

(5) If $a \equiv b \bmod m$ and $d|m$ with $d > 0$, then $a \equiv b \bmod d$.

(6) If $a \equiv b \bmod m$, then $ac \equiv bc \bmod mc$ for any positive integer c.

Proof. We will prove (4), and leave the rest as exercises. By assumption, there are integers k, l such that $a = b+mk$, $c = d+ml$. Then $ac = (b+mk)(d+ml) = bd+m(kd+bl+kml)$, so $ac \equiv bd \bmod m$. □

Example 2.1. If $m = 5$, then $7 \equiv 2 \bmod 5$ and $11 \equiv 1 \bmod 5$, so $18 \equiv 3 \bmod 5$ and $77 \equiv 2 \bmod 5$.

Theorem 2.2. Let $f(x) = c_n x^n + c_{n-1}x^{n-1} + \ldots + c_0$ be a polynomial with integral coefficients c_i. If $a \equiv b \bmod m$, then $f(a) \equiv f(b) \bmod m$.

Proof. By item (4) of Theorem 2.1, we have $a^2 \equiv b^2 \bmod m$, $a^3 \equiv b^3 \bmod m$, and in general $a^k \equiv b^k \bmod m$ for all $k \geq 0$. Then $c_k a^k \equiv c_k b^k \bmod m$, and hence

$$c_n a^n + c_{n-1} a^{n-1} + \ldots + c_0 \equiv c_n b^n + c_{n-1} b^{n-1} + \ldots + c_0 \bmod m,$$

by various applications of Theorem 2.1. □

A slight generalization of the above argument is used to derive the following:

Theorem 2.3. Assume the same hypotheses as in Theorem 2.2, but now consider another polynomial $g(x) = d_n x^n + d_{n-1} x^{n-1} + \ldots + d_0$ with integer coefficients, such that $c_i \equiv d_i \bmod m$ for all $0 \leq i \leq n$. Show that $f(a) \equiv g(b) \bmod m$.

Example 2.2. We claim that if $a = 6^{10001}$, then $f(a)$ for $f(x) = 8x^2 - 5x + 15$ is divisible by 7. Indeed, notice that $6^{10001} \equiv (-1)^{10001} = -1 \bmod 7$, and if $b = -1$ and $g(x) = x^2 + 2x + 1 = (x+1)^2$, then a, b, $f(x)$, and $g(x)$ satisfy the hypotheses of Theorem 2.3. Hence $f(a) \equiv g(b) \bmod 7$, and $g(b) = 0$.

Example 2.3. We show that 41 divides $2^{20} - 1$. Since $2^5 = 32 \equiv -9 \bmod 41$, we have $2^{20} = (2^5)^4 \equiv (-9)^4 \bmod 41$. Since $(-9)^4 = 81 \cdot 81$ and $81 = 2 \cdot 42 - 1$, we conclude that $2^{20} \equiv 81 \cdot 81 \equiv (-1) \cdot (-1) = 1 \bmod 41$. Therefore $2^{20} - 1 \equiv 1 - 1 = 0 \bmod 41$.

You will see more examples of this type of problem on your homework.

We make the following observation:

Proposition 2.1. Let m be a positive integer. Then every integer is congruent mod m to exactly one of the integers $0, 1, \ldots, m-1$. Moreover, no two of these integers are congruent mod m.

This explains the use of the word "residue" from Definition 2.1: if a is an integer, and $k \in \{0, 1 \ldots, m-1\}$ is the unique element such that $a \equiv k \bmod m$, then k is the *residue* (remainder) upon dividing a by m.

This construction is so important that it deserves a name.

Definition 2.2. A set $\{x_1, \ldots, x_r\}$ is called a *complete residue system mod m* if, for every integer y, there is exactly one x_j such that $y \equiv x_j \bmod m$. The elements of a complete residue system mod m are called *residues*, or *residue classes*, mod m.

So, the set $\{0, 1 \ldots, m-1\}$ is a complete residue system. We call this the *canonical residue system mod m* (this is my own terminology and is not standard).

Example 2.4. As another example, the set $\{1 \ldots, m-1, m\}$ is a complete residue system.

Exercise 2.1. Prove that every complete residue system mod m consists of exactly m elements. [Hint: use the canonical residue system.]

We will now use the canonical residue system mod m, along with items (3) and (4) of Theorem 2.1, to construct a "system" in which we can do many the basic arithmetic operations. In later algebra courses, you will learn that this "system" is really an example of an *(abelian) group* and a *commutative ring.*

Definition 2.3. We define $\mathbf{Z}/m\mathbf{Z}$, as a *set*, to be the canonical residue system mod m. We define *addition mod* m in $\mathbf{Z}/m\mathbf{Z}$ as follows: given elements $a, b \in \mathbf{Z}/m\mathbf{Z}$, we define $a +_m b$ to be the *unique* element c in $\mathbf{Z}/m\mathbf{Z}$ such that $a + b \equiv c \bmod m$ (notice that the addition on the left-hand side is addition of *integers*). We define *multiplication mod* m, $a \cdot_m b$, similarly.

Remark 2.1. The notations $+_m$ and $\cdot_m$ are nonstandard, but are useful in the beginning. Later on, I may drop the subscripts if the modulus m is clear.

Example 2.5. In $\mathbf{Z}/1\mathbf{Z}$, the only element is 0.

Example 2.6. Consider $m = 6$. Then in $\mathbf{Z}/6\mathbf{Z}$, $3 +_6 4 = 1 \bmod 6$ (we write "mod 6" to distinguish our operation from the usual addition of integers) and $3 \cdot_6 5 = 3 \bmod 6$.

Let's see how $\mathbf{Z}/m\mathbf{Z}$ is similar to and is different from the integers $\mathbf{Z}$. More or less by construction, addition and multiplication in $\mathbf{Z}/m\mathbf{Z}$ are associative and commutative, and multiplication distributes over addition (although these do require short proofs). There is an *additive identity* in $\mathbf{Z}/m\mathbf{Z}$: it is 0, because for any $a \in \mathbf{Z}/m\mathbf{Z}$, $0 + a = a$ in the integers, and a is an element of the canonical residue system by definition, so $0 +_m a = a$. Similarly, there is an *multiplicative identity* in $\mathbf{Z}/m\mathbf{Z}$: it is 1. We also have additive inverses (and therefore subtraction) in $\mathbf{Z}/m\mathbf{Z}$, where the inverse $-a$ of $a \in \mathbf{Z}/m\mathbf{Z}$ is defined as the unique element in the canonical residue system congruent mod m to the integer $0 - a$ (e.g. $-4 = 2$ in $\mathbf{Z}/6\mathbf{Z}$).

On the other hand, $\mathbf{Z}/m\mathbf{Z}$ has a few key differences. First, we might not have the "zero product property." In $\mathbf{Z}$, we know that if $ab = 0$, then either $a = 0$ or $b = 0$. But this might not be true in certain $\mathbf{Z}/m\mathbf{Z}$: if $m = 6$, for instance, then $2 \cdot_6 3 = 0$, but neither 2 nor 3 equal 0 in $\mathbf{Z}/6\mathbf{Z}$. Similarly, we have a "cancellation property" in $\mathbf{Z}$: if $ab = ac$ with $a \neq 0$, then $b = c$. But this can fail in certain $\mathbf{Z}/m\mathbf{Z}$ (can you find such an example when $m = 6$?).

Somewhat conversely, we know that not all nonzero integers have multiplicative inverses: in fact, only 1 and -1 do. On the other hand, there are certain m for which every nonzero element in $\mathbf{Z}/m\mathbf{Z}$ has a multiplicative inverse. For instance, in $\mathbf{Z}/7\mathbf{Z}$, we have $1 \cdot_7 1 = 1$, $2 \cdot_4 = 1$, $3 \cdot_7 5 = 1$, and $6 \cdot_7 6 = 1$, so every nonzero element in $\mathbf{Z}/7\mathbf{Z}$ has a multiplicative inverse![1] The great theorem is this:

Theorem 2.4. For $m \geq 2$ ($m = 1$ is excluded for technical reasons), the following are equivalent:

[1] A set in which we have addition and multiplication satisfying the usual properties, and in which every nonzero element has a multiplicative inverse, is called a *field.*

(1) The “system of modular arithmetic” $\mathbf{Z}/m\mathbf{Z}$ has the “zero product property.”

(2) The “system of modular arithmetic” $\mathbf{Z}/m\mathbf{Z}$ has the “field property”: every nonzero element has a multiplicative inverse.

(3) m is a prime number (to be defined later).

Remark 2.2. Notice that we have to take care when discussing modular arithmetic, because although we would like to treat, say, “2 mod 3” as if it were a *number*, we cannot quite do so yet. For instance, it makes no sense to say “2 mod 3 is even”, because $5 \equiv 2 \bmod 3$, and 5 is not even. For a similar reason, it does not make sense to say “3 mod 5 times 4 mod 5 equals 12” (why?). In reality, “2 mod 3” is really a *set*: it is the set $\{\dots, -4, -1, 2, 5, \dots\}$ containing all integers that leave a remainder of 2 upon division by 3. In an abstract algebra class, you will learn about, *quotient groups* (of which $\mathbf{Z}/m\mathbf{Z}$ is an example) and *cosets*, which allow you to treat a set of elements as a *single element* in its own right.

3 Prime Numbers

In this section we introduce the fundamental building blocks of the integers: the prime numbers. They will be the integers that are “indivisible” into others.

Definition 3.1. A positive integer $p > 1$ is a *prime number* if there is no divisor d of p satisfying $1 < d < p$. In other words, the only positive divisors of p are 1 and p itself. If an integer $a > 1$ is not prime, then it is *composite*.

Note that we deliberately exclude 1 from being a prime number. This is for many very good reasons, and we will touch on some of them later.

Theorem 3.1 (Fundamental Theorem of Arithmetic: Existence)**.** Every positive integer $n > 1$ can be expressed as a product of primes, perhaps trivially.

Proof. If n is a prime, then it is already expressed as a “trivial product” with only one factor. If not, then n must be able to be factored into a product $n_1 n_2$, where $1 < n_1, n_2 < n$. Repeating this process with n_1, n_2 (really an inductive argument), we can continue the factorization: if n_1 is prime, we stop, and if n_1 is not prime, then we write $n_1 = n_3 n_4$ with $1 < n_3, n_4 < n_1$. This process must stop eventually because the factors get strictly smaller at every step, yet are integers strictly greater than 1. Therefore we may write n as a product of primes $n = p_1^{a_1} p_2^{a_2} \cdot \ldots \cdot p_r^{a_r}$, where the p_i are distinct primes and the a_i are positive integers. □

Notice that this does not prove that the factorization is *unique.* Indeed, this is not immediate and must be proved. As a cautionary example of what could happen, in 1847, the French mathematician Lamé announced a proof of Fermat's Last Theorem, perhaps the most important open problem of number theory from the 17th to the 20th century. In the proof, Lamé assumed without proof a unique factorization property in a different "system of arithmetic" that he had constructed, but the German mathematician Kummer showed this to be false. This eventually led to the creation of "ideal numbers" (the precursor to the modern abstract algebraic notion of "ideal") by Kummer and Dedekind, which allows one to "save unique factorization" in a way.

For hand-calculation purposes, the following result is useful:

Exercise 3.1. If n is composite, then n has a prime divisor less than or equal to $\sqrt{n}$.

Example 3.1. To verify that 101 is prime, we only need to show that it is not divisible by 2, 3, 5, or 7, because those are the only primes less than or equal to $\sqrt{101} < 11$.

To prove uniqueness of prime factorization in $\mathbf{Z}$, the following result (really a corollary of an earlier result) is useful.

Proposition 3.1. *If $p|ab$, then either $p|a$ or $p|b$. In general, if $p|a_1a_2 \cdot \ldots \cdot a_n$, then p divides at least one of the a_i.*

Proof. If $p \nmid a$, then $(p,a) = 1$ since the *only* positive divisors of p are 1 and p. So by Proposition 1.4, $p|b$. The second statement follows by induction: if p divides $a_1a_2 \cdot \ldots \cdot a_n = a_1(a_2 \cdot \ldots \cdot a_n)$, then it either divides a_1 or $a_2 \cdot \ldots \cdot a_n$. If it divides a_1, then we are done, and if it divides $a_2 \cdot \ldots \cdot a_n$, then it divides either a_2 or $a_3 \cdot \ldots \cdot a_n$, and so on. □

Theorem 3.2 (Fundamental Theorem of Arithmetic: Uniqueness)**.** *Every positive integer $n > 1$ can be expressed as a product of primes, which is unique up to reordering the prime factors of n.*

Proof. Suppose n has two prime factorizations

$$n = p_1p_2 \cdot \ldots \cdot p_r = q_1q_2 \cdot \ldots \cdot q_s,$$

where without loss of generality, $r \leq s$. Then $p_1|q_1q_2 \ldots q_s$, so p_1 divides one of the q_i by Proposition 3.1. Renaming the q_i if necessary, suppose that $q_i = q_1$. Then $p_1|q_1$, but q_1 is prime, so the only positive factors it could have are 1 and itself. Since $p_1 > 1$ by definition of prime, we conclude that $p_1 = q_1$, so we can cancel them and obtain the equality

$$p_2 \cdot \ldots \cdot p_r = q_2 \cdot \ldots \cdot q_s.$$

Continuing in this fashion with $p_2, p_3, \ldots$, and using the assumption that $r \leq s$, we arrive at

$$1 = q_{r+1} \cdot \ldots \cdot q_s,$$

where the product on the right-hand side is empty (i.e. equals 1) if $r = s$. This must be the case, since the only positive integer dividing 1 is 1 itself, by (5) of Theorem 1.1. It follows that, in the original equality

$$p_1 p_2 \cdot \ldots \cdot p_r = q_1 q_2 \cdot \ldots \cdot q_r,$$

we have $p_i = q_i$ for all i after possibly having reordered the q_i's. □

Corollary 3.1. Every integer $n \neq 0, 1, -1$ can be expressed as a product of primes up to sign, and this product is unique up to reordering the prime factors of n.

Proof. If $n < -1$, then apply Theorem 3.2 to $-n > 1$, which is uniquely a product of primes. In particular, the (positive) prime factors of n multiply to $-n$. □

As an application of the (existence part of the) fundamental theorem of arithmetic:

Theorem 3.3 (Euclid)**.** There are infinitely many prime numbers.

Proof. If not, then suppose there are only finitely many: $p_1, \ldots, p_r$ (there is at least 1 prime: $p_1 = 2$). Consider $n = 1 + p_1 p_2 \cdot \ldots \cdot p_r$. Then none of the p_i divide n, contradicting any $n > 1$ having a prime factor. □

Be aware of the common misconception that this proof implies that the product of the first r primes, plus 1, is prime. This is not true: $1 + 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 = 59 \cdot 509$.

Let's see how we may exploit the *uniqueness* aspect of the fundamental theorem of arithmetic. By the existence aspect, every integer $a \geq 1$ may be written in the shorthand form

$$a = \prod_p p^{\alpha(p)},$$

where the product ranges over all primes p (infinitely many!), each $\alpha(p)$ is a nonnegative integer, and $\alpha(p) = 0$ for all large enough primes p. Note that if $a = 1$, then $\alpha(p) = 0$ for all p.

Example 3.2. We have $140 = \prod_p p^{\alpha(p)}$, where $\alpha(2) = 2$, $\alpha(3) = 0$, $\alpha(5) = 1$, $\alpha(7) = 1$, and $\alpha(p) = 0$ for all $p \geq 11$.

Now, if $a = \prod_p p^{\alpha(p)}$, $b = \prod_p p^{\beta(p)}$, and $c = \prod_p p^{\gamma(p)}$, and we moreover have $ab = c$, then the *uniqueness* of the factorization implies that $\alpha(p) + \beta(p) = \gamma(p)$ for every prime p. In particular, if $a|c$, then $\alpha(p) \leq \gamma(p)$ for all p. Conversely, if $\alpha(p) \leq \gamma(p)$ for all p, then you can show that $a|c$. Therefore divisibility relations may be expressed in terms of the exponents appearing in the unique prime factorizations, and in fact we have:

Proposition 3.2. If $a = \prod_p p^{\alpha(p)}$ and $b = \prod_p p^{\beta(p)}$, then $(a, b) = \prod_p p^{\min(\alpha(p), \beta(p))}$. In partricular, $(a, b) = 1$ if and only if a and b have no common prime factor p.

The proof is left as exercise.

We could also have defined:

Definition 3.2. Given nonzero integers a and b, the *least common multiple* of a and b, denoted $\operatorname{lcm}(a, b)$, is the smallest positive integer in the set $S = \{s > 0 : a|s, b|s\}$. Note that S is nonempty because it contains $|ab|$.

Then given positive integers $a = \prod_p p^{\alpha(p)}$ and $b = \prod_p p^{\beta(p)}$, we conclude that

$$\operatorname{lcm}(a, b) = \prod_p p^{\max(\alpha(p),\beta(p))},$$

and moreover

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

3.1 Distribution of Primes in the Integers

This informal section gives some discussion about the distribution of primes in the integers, since some of the material appears in our (Rosen's) textbook.

The distribution of primes is more or less the central question in the field of analytic number theory. In some sense, there are simultaneously "a lot of primes" and "not a lot of primes." For instance:

Exercise 3.2. Fix a positive integer k. Then there exist k consecutive positive integers, none of which are prime. [Hint: think about what numbers divide $(k + 1)!$.]

On the other hand, one of Euler's most famous results was that there are "enough" primes for the sum of the reciprocals of primes to diverge. More precisely, it can be proved that

Theorem 3.4. For every real number $y \geq 2$,

$$\sum_{p \leq y} \frac{1}{p} > \log\log y - 1.$$

In particular, letting $y \to \infty$ shows that

$$\sum_{p \text{ a prime}} \frac{1}{p} = \infty.$$

Remark 3.1. Note that there do exist infinite sets of positive integers, where the sum of the reciprocals of all elements in that set converges. For example, $S = \{2^0, 2^1, 2^2, \ldots\}$.

Remark 3.2. Theorem 3.4 gives another, *non-circular*, proof of the infinitude of primes.

Two much deeper results regarding the distribution of primes are as follows. First, one might be interested in primes that are contained in certain sequences, such as the sequence of positive integers congruent to $a \bmod m$ for some positive integers a, m. Assuming that $(a, m) = 1$ (this is certainly necessary—do you see why?), then the following theorem holds:

Theorem 3.5 (Dirichlet's Theorem on Arithmetic Progressions)**.** Given positive integers (a, m) such that $(a, m) = 1$, there are infinitely many prime numbers congruent to $a \bmod m$.

For certain values of a and m, such as $a = 3$ and $m = 4$, the corersponding special case of this theorem can be proved using a method very similar to that of Euclid's Theorem 3.3. The case when $a = 1$ is slightly harder, but can be proved along the same lines using special polynomials called *cyclotomic polynomials*. But the general proof requires some complex-analytic machinery, which is well beyond the scope of this course.

Even more difficult is the celebrated *prime number theorem*:

Theorem 3.6 (Prime Number Theorem)**.** For a real number $x > 1$, let $\pi(x)$ be the number of primes p less than or equal to x. Then

$$\lim_{x\to\infty} \frac{\pi(x)}{x/\log(x)} = 1.$$

That is, the number of primes less than or equal to a real number x "grows asymptotically" according to the function $x/\log(x)$.

The proof requires many subtle complex-analytic estimates. On the other hand, here is some numerical data (taken from our textbook) that showcases the plausibility of the theorem.

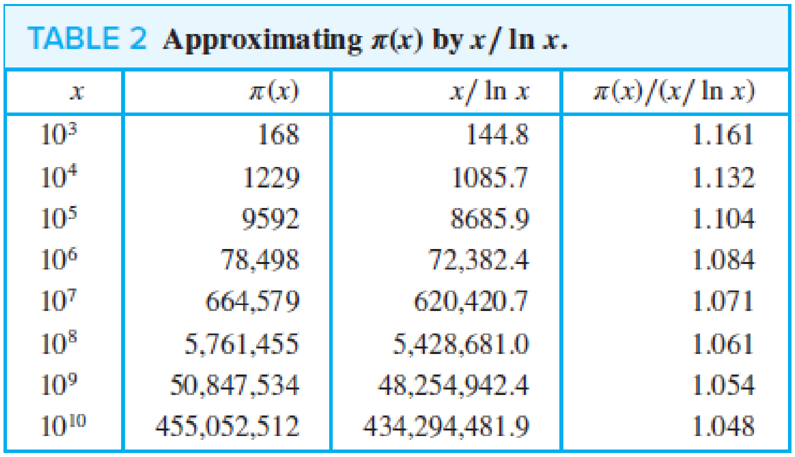

TABLE 2 **Approximating $\pi(x)$ by $x/\ln x$.**

x	$\pi(x)$	$x/\ln x$	$\pi(x)/(x/\ln x)$
10^3	168	144.8	1.161
10^4	1229	1085.7	1.132
10^5	9592	8685.9	1.104
10^6	78,498	72,382.4	1.084
10^7	664,579	620,420.7	1.071
10^8	5,761,455	5,428,681.0	1.061
10^9	50,847,534	48,254,942.4	1.054
10^{10}	455,052,512	434,294,481.9	1.048

Of course, there is a lot we don't know about the distribution of primes. Two of the most famous open problems in this area are the *twin prime conjecture* and the *Riemann hypothesis* (for which a solution comes with a 1 million dollar prize), and I'm happy to explain more about these if asked to.

4 The structure of $\mathbf{Z}/m\mathbf{Z}$

4.1 The multiplicative structure

Let $m \geq 2$ be an integer; we will keep this running assumption for this entire section. Recall that we defined $\mathbf{Z}/m\mathbf{Z}$ as a set to be the canonical residue system mod m. Therefore if a is any integer, then there is a unique element $r \in \mathbf{Z}/m\mathbf{Z}$ such that $a \equiv r \bmod m$; r is just the remainder of a upon division by m. Let's call r the *reduction of a mod m* (we will omit the "mod m" if m is clear from context). Since $a = qm + r$ for some integer q, Theorem 1.4 tells us that $(a, m) = (a - qm, m) = (r, m)$. Therefore gcd properties of a with respect to m, are the same as gcd properties of its reduction with respect to m (e.g. statements such as "if a is an integer coprime to m, then the reduction of a is also coprime to m").

This fact allows us to translate many statements regarding modular arithmetic mod m between the integers $\mathbf{Z}$ and $\mathbf{Z}/m\mathbf{Z}$. Each proposition that follows has a version pertaining to all integers, and a version pertaining to elements of $\mathbf{Z}/m\mathbf{Z}$, and they express the exact same mathematical content. At the start, we will try to state the proposition in both settings, but later on as this gets tedious, the reader should be able to supply the various "translations" themselves.

Recall that we have defined operations $+_m$ and $\times_m$ on $\mathbf{Z}/m\mathbf{Z}$, which are just "modulo m versions" of the usual operations on integers (later on, we will drop the subscript m when the context that we are working in $\mathbf{Z}/m\mathbf{Z}$ is clear). These operations satisfy many of the same properties as our regular $+$ and $\times$ operations on the integers. We will now discuss a property that almost all of the integers certainly do not have: the existence of multiplicative inverses. This turns out to be the key to unlocking the properties of $\mathbf{Z}/m\mathbf{Z}$.

Proposition 4.1. Let a be an integer coprime to m. Then there is a unique multiplicative inverse of a mod m: i.e. there is an integer b such that $ab \equiv 1 \bmod m$, and any other integer b' also satisfying $ab' \equiv 1 \bmod m$ must be congruent to b mod m. Translation in $\mathbf{Z}/m\mathbf{Z}$: if $r \in \mathbf{Z}/m\mathbf{Z}$ is coprime to m, then there is a unique $s \in \mathbf{Z}/m\mathbf{Z}$ such that $r \times_m s = 1$.

Proof. By Theorem 1.3, there exist integers x, y such that $ax + my = 1$. Then $ax \equiv 1 \bmod m$, so x is a multiplicative inverse of a mod m. To show uniqueness, suppose an integer x' is another multiplicative inverse of a mod m, so $ax' \equiv 1 \bmod m$. Then $x'ax = (x'a)x \equiv x \bmod m$, but also $x'ax = x'(ax) \equiv x' \bmod m$, so that $x \equiv x' \bmod m$.

Translation in $\mathbf{Z}/m\mathbf{Z}$: let x be an integer such that $rx \equiv 1 \bmod m$. Then if s is the reduction of x mod m, then $r \times_m s = 1$, because $rs \equiv rx \equiv 1 \bmod m$. s is unique, because if s' also has $r \times_m s' = 1$, then $rs' \equiv 1 \bmod m$, so by what we proved above, $x \equiv s' \bmod m$. Because s is the unique integer in $[0, m-1]$ that is congruent to $x \bmod m$, we have $s = s'$. □

Notice how the proof of the "translation" didn't really involve any new ideas; it was just a reinterpretation of what we knew to be true in the integer setting to the "language" of $\mathbf{Z}/m\mathbf{Z}$. On the other hand, the translation in $\mathbf{Z}/m\mathbf{Z}$ allowed us a slightly cleaner version of our statement: instead of "unique integer up to congruence mod p", we can simply say "unique element in $\mathbf{Z}/m\mathbf{Z}$." This uniqueness allows us to make the following definition:

Definition 4.1. Suppose $r \in \mathbf{Z}/m\mathbf{Z}$ is coprime to m. Then the unique multiplicative inverse of r is denoted r^{-1}, and we call r an *invertible element* or a *unit* (in $\mathbf{Z}/m\mathbf{Z}$). If r is a unit, and $k < 0$ is a negative integer, then we define r^k as $(r^{-1})^{-k}$ (which makes sense as $-k$ is positive, and we already know how to take an element to a positive power).

One can check that the usual exponent properties hold, even with negative exponents. For instance, if a and b are integers (perhaps negative), and $r \in \mathbf{Z}/m\mathbf{Z}$ is invertible, then $r^{ab} = (r^a)^b = (r^b)^a$. In particular, $(r^{-1})^{-1} = r$. As another example, a product of invertible elements $r, s \in \mathbf{Z}/m\mathbf{Z}$ is invertible, since $(rs)^{-1} = r^{-1} \times_m s^{-1}$.

Corollary 4.1. If p is a prime and $p \nmid a$, then a has a (unique) multiplicative inverse mod p. Translation in $\mathbf{Z}/p\mathbf{Z}$: if r is a nonzero element in $\mathbf{Z}/p\mathbf{Z}$, then r has a unique multiplicative inverse r^{-1} in $\mathbf{Z}/p\mathbf{Z}$.

Proof. Indeed, if p is not a prime factor of a, then p and a share no prime factors in common, so $(a, p) = 1$. □

Corollary 4.2. We can now prove Theorem 2.4: The following are equivalent:

(1) The "system of modular arithmetic" $\mathbf{Z}/m\mathbf{Z}$ has the "zero product property": if $a, b \in \mathbf{Z}/m\mathbf{Z}$ satisfy $ab = 0$, then either $a = 0$ or $b = 0$.

(2) The "system of modular arithmetic" $\mathbf{Z}/m\mathbf{Z}$ has the "field property": every nonzero element has a unique multiplicative inverse.

(3) m is a prime number.

Proof. We first show that (1) implies (3). Assume (1), and suppose for contradiction that m is not prime. Then because $m \geq 2$ (running assumption), m is a product of integers kl, where $1 < k, l < m$. Then k and l are nonzero elements in $\mathbf{Z}/m\mathbf{Z}$ and $k \times_m l = 0$ in $\mathbf{Z}/m\mathbf{Z}$, contradicting the assumption (1). Therefore m must be prime.

Next, that (3) implies (2) is Corollary 4.1. Finally, to show (2) implies (1), suppose $a \times_m b = 0$ where a, b are elements in $\mathbf{Z}/m\mathbf{Z}$. If $a = 0$, then we are done. If not, then by assumption a has a multiplicative inverse a^{-1}, so $0 = a^{-1} \times_m 0 = a^{-1} \times_m a \times_m b = 1 \times_m b = b$. Therefore $b = 0$, so at least one of a or b is 0. □

Example 4.1. Using Corollary 4.2, we can make the following observation. Suppose $a, b, c \in \mathbf{Z}/m\mathbf{Z}$ are such that $ab = ac$ and $(a, m) = 1$. Then we can conclude that $b = c$ by multiplying by a^{-1} on both sides of the preceding equality. Note that this could be false if we don't assume that a is invertible in $\mathbf{Z}/m\mathbf{Z}$: for instance, $3 \times_6 1 = 3 \times_6 3$, and 3 is not invertible in $\mathbf{Z}/6\mathbf{Z}$ as $(3, 6) = 3$.

Example 4.2. Suppose p is a prime number, and $x \in \mathbf{Z}/p\mathbf{Z}$ satisfies $x^2 = 1$. Then $(x + 1)(x - 1) = 0$, so $x + 1 = 0$ or $x - 1 = 0$. In other words, $x = 1$ or $x = p - 1$. Translation into $\mathbf{Z}$: if a is an integer such that $a^2 \equiv 1 \bmod p$, then either $a \equiv 1 \bmod p$ or $a \equiv -1 \bmod p$.

Similarly, if $x \in \mathbf{Z}/p\mathbf{Z}$ satisfies $x^n = 0$ for some positive integer n, then $x = 0$. Indeed, $x^n = x(x^{n-1})$, so either $x = 0$ or $x^{n-1} = 0$. If the former is true, then we are done; if the latter is true, then either x or x^{n-2} is 0, etc. Translation into $\mathbf{Z}$: if a is an integer such that $p|a^n$, then $p|a$ (compare this to Proposition 3.1).

Example 4.3. As another application, suppose we have a congruence $ax \equiv b \bmod m$, and we would like to find all possible integers x that make this congruence true. Suppose for simplicity that $(a, m) = 1$ (the case when $(a, m) \neq 1$ is more annoying and is better dealt with using gcd conditions or the Chinese remainder theorem, the latter of which we will discuss later). Then if c is an integer that is an inverse to a mod m (so c is unique up to congruence mod m), we have $ax \equiv b \bmod m$ if and only if $x \equiv bc \bmod m$, which solves the congruence. For instance, we would like to solve

$$3x \equiv 4 \bmod 16.$$

We have $(3, 16) = 1$, and indeed $3 \cdot (-5) + 16 \cdot 1 = 1$. Therefore -5 is an inverse to 3 mod 16, so the solutions to the above congruence are any integer x congruent to $(-5) \cdot 4 = -20$ mod 16. Or, if one prefers to take the canonical residue class mod 16 instead, then because $-20 \equiv 12 \bmod 16$ and $12 \in [0, 15]$, we can say that the solutions are exactly of the form 12 mod 16.

Let's now specialize to the case when m is a prime p, so every nonzero element in $\mathbf{Z}/p\mathbf{Z}$ has a multiplicative inverse. Consider some nonzero $x \in \mathbf{Z}/p\mathbf{Z}$, and consider the elements $1 = x^0, x, x^2, x^3, \ldots, x^{p-1}$. There are p of these elements, and none of them are 0 as discussed in Example 4.2, so they must be elements in $\{1, 2, \ldots, p - 1\}$. Therefore two of them must be the same by cardinality considerations, so suppose $x^k = x^l$ with $0 \leq k < l \leq p - 1$. Then multiplying by $x^{-k} = (x^k)^{-1}$ on both sides, we get $1 = x^{l-k}$ with $1 \leq l - k \leq p - 1$. The conclusion is that:

Theorem 4.1. For every nonzero $x \in \mathbf{Z}/p\mathbf{Z}$, there is a positive integer $1 \leq d \leq p - 1$ such that $x^d = 1$.

Of course, the exponent d in question may depend on x. Therefore it is natural to ask if there is a *common* exponent d, $1 \leq d \leq p - 1$, such that $x^d = 1$ for *every* nonzero $x \in \mathbf{Z}/p\mathbf{Z}$. The answer is yes:

Theorem 4.2 (Fermat's little theorem)**.** Let p be a prime and $x \in \mathbf{Z}/p\mathbf{Z}$ be nonzero. Then $x^{p-1} = 1$. Translation to $\mathbf{Z}$: if $a \in \mathbf{Z}$ is coprime to p, then $a^{p-1} \equiv 1 \bmod p$.

Proof. There are many ways to prove this theorem. The most basic proof comes from a "necklace-counting argument," while the most abstract proof is to deduce it as a corollary of *Lagrange's theorem* from group theory. However, all of these proofs really share the same core idea. So we will provide a proof that mimics the most abstract proof, but stays within the realm of modular arithmetic in $\mathbf{Z}/p\mathbf{Z}$.

Consider the list $x, 2x, \ldots, (p-1)x$ of elements of $\mathbf{Z}/p\mathbf{Z}$, where the multiplication is really $\times_p$. We do not yet call this a *set*, because we have not yet proven that there are no repeated elements in this list. First, because $x \neq 0$, none of the elements we wrote down are 0, by the zero-product property of $\mathbf{Z}/p\mathbf{Z}$ (Theorem 4.2). Next, suppose $ax = bx$ for some $a, b \in \mathbf{Z}/p\mathbf{Z}$. Then because $x \neq 0$, it is invertible, so multiplying both sides by x^{-1} produces the equality $a = b$. In other words, if $a, b \in \mathbf{Z}/p\mathbf{Z}$ are distinct, then $ax \neq bx$. Therefore the elements $x, 2x, \ldots, (p-1)x$ are all distinct, because the "coefficients" appearing in front of x, i.e. $1, 2, \ldots, p-1$, are distinct elements in $\mathbf{Z}/p\mathbf{Z}$ (formally, what we've done is construct a map from $[p-1]$ to $\{x, 2x, \ldots, (p-1)x\}$, and shown that it is injective).

Therefore $\{x, 2x, \ldots, (p-1)x\}$ is a subset of $p-1$ *distinct* elements of $\mathbf{Z}/p\mathbf{Z}$, none of which are 0. Hence it is a subset of $\mathbf{Z}/p\mathbf{Z} - \{0\} = \{1, 2, \ldots, p-1\}$. But this subset only contains $p-1$ elements, so we conclude that $\{x, 2x, \ldots, (p-1)x\}$ (which has $p-1$ elements) and $\{1, 2, \ldots, p-1\}$ (which also has $p-1$ elements) are the same! In other words, the elements $x, 2x, \ldots, (p-1)x$ of $\mathbf{Z}/p\mathbf{Z}$ are simply a rearrangement of the elements $1, 2, \ldots, (p-1)$.

Let's now multiply all the numbers in both sets, with multiplication done in $\mathbf{Z}/p\mathbf{Z}$. Since they are the exact same list of numbers (with no repetitions in either list!),[2] and the order in which we multiply doesn't affect the final result, we have

$$1 \cdot 2 \cdot \ldots \cdot (p-1) = x \cdot 2x \cdot \ldots \cdot (p-1)x = x^{p-1} \cdot (1 \cdot 2 \cdot \ldots \cdot (p-1)).$$

Write t for $1 \cdot 2 \cdot \ldots \cdot (p-1)$. Because none of the factors in t are 0, by the zero-product property, t is nonzero in $\mathbf{Z}/p\mathbf{Z}$. Therefore t is invertible, and we have $t = x^{p-1}t \Rightarrow 1 = x^{p-1}$. □

Corollary 4.3. If $x \in \mathbf{Z}/p\mathbf{Z}$, then $x^p = x$.

Proof. If x is nonzero, this follows from the equality $x^{p-1} = 1$. If x is 0, then $x^p = 0 = x$. □

Corollary 4.4. If $x \in \mathbf{Z}/p\mathbf{Z}$ is nonzero, then $x^{p-2} = x^{-1}$.

[2] We are very careful to emphasize this point about no repetitions, because things can go wrong if we're sloppy about it. For instance, $\{1, 2, 2\}$ and $\{1, 2\}$ are the same as *sets* of integers, because sets disregard repetition, but if we multiply the integers in the first list, we get 4, while if we multiply the numbers in the second list, we get 2. Of course, the issue was that when we listed out the elements of our sets, we did not insist that those lists did not contain repetitions.

Remark 4.1. The *converse* of this theorem is not true: if $n \geq 2$ is an integer, and $a^{n-1} \equiv 1 \bmod n$ for all integers a coprime to n, it does *not* imply that n is prime. Such "pseudoprimes" are called *Carmichael numbers*, and the smallest one is $561 = 3 \cdot 11 \cdot 17$. After we discuss the Chinese remainder theorem, we will see that this property is due to the coincidence that $560 = 561 - 1$ is divisible by each of $2 = 3 - 1$, $10 = 11 - 1$, and $16 = 17 - 1$.

Here are some applications of Fermat's little theorem:

Example 4.4. We will find the unique integer $0 \leq a \leq 10$ such that $3^{333} \equiv a \bmod 11$. By Fermat's little theorem, we know that $3^{10} \equiv 1 \bmod 11$, so $3^{330} = (3^{10})^{33} \equiv 1^{33} = 1 \bmod 11$. Therefore $3^{333} = 3^{330} \cdot 3^3 \equiv 3^3 \equiv 5 \bmod 11$.

Example 4.5. We claim that $n^5/5 + n^3/3 + 7n/15$ is an integer, whenever n is an integer. A clever rearrangement of the expression gives

$$\frac{n^5}{5} + \frac{n^3}{3} + \frac{7n}{15} = \frac{n^5 - n}{5} + \frac{n^3 - n}{3} + n.$$

By Corollary 4.3, $n^5 - n$ is 0 mod 5 for any integer n, and likewise $n^3 - n$ is 0 mod 3 for any integer n. Therefore the right-hand side of the above equality is an integer for any integer n.

Notice that during the course of the proof of Theorem 4.2, we needed the fact that $(p-1)!$ is an integer coprime to p, but it was not necessary to actually find the reduction of $(p-1)!$ mod p. But this is not too hard:

Theorem 4.3 (Wilson)**.** If p is a prime number, then $(p-1)! \equiv -1 \bmod p$. If $m > 1$ is a composite number, then $(m-1)! \not\equiv 1 \bmod m$.

Therefore this theorem gives a (very, very, very slow) test to determine whether a given positive integer is prime.

Proof. The result for prime p can be immediately verified when $p = 2$ or $p = 3$. So now assume $p \geq 5$, and $p - 2 \geq 3$. We first claim that there is no integer x in $\{2, 3, \dots, p-2\}$ such that x is congruent to $x^{-1} \bmod p$. Indeed, such an x would satisfy $x^2 \equiv 1 \bmod p$, and we saw in Example 4.2 that any such integer is congruent to 1 or $p - 1 \bmod p$, and none of the integers in $\{2, 3, \dots, p-2\}$ are congruent to 1 or $p - 1 \bmod p$. Therefore we may pair up the integers in $\{2, 3, \dots, p-2\}$ into $(p-3)/2$ different pairs (a, a'), where $a < a'$ and $aa' \equiv 1 \bmod p$ (so we consider (a, a') and (a', a) to be the same pair). This is possible because none of those integers are their own inverses mod p, and none of them have inverse 1 or $p - 1$ (as those are their own inverses). Therefore the product

$$2 \cdot 3 \cdot \ldots \cdot (p-2)$$

can be rearranged into a product as follows: for each of the aforementioned $(p-3)/2$ pairs (a, a'), we multiply a' immediately after a in the product, so that a and a' will "cancel mod p." Somewhat more formally,

$$2 \cdot 3 \cdot \ldots \cdot (p-2) = \prod_{(a,a'):2 \leq a < a' \leq p-2, aa' \equiv 1 \bmod p} aa' \equiv \prod_{(a,a'):2 \leq a < a' \leq p-2, aa' \equiv 1 \bmod p} 1 = 1 \bmod p.$$

Therefore

$$(p-1)! = 1 \cdot (2 \cdot 3 \cdot \ldots \cdot (p-2)) \cdot (p-1) \equiv 1 \cdot 1 \cdot (p-1) \equiv -1 \bmod p.$$

We now deal with the case when m is composite. By definition, m is divisible by some integer a such that $1 < a < m$. Then the factorial $(m-1)!$ will contain a factor of a, so $a|(m-1)!$. Therefore $a > 1$ is a common divisor of $(m-1)!$ and m, and so $(m-1)!$ cannot be congruent to -1 mod m, because that would imply that $\gcd((m-1)!, m) = 1$. □

Exercise 4.1. As an exercise, you can show that if m is composite, then $(m-1)! \equiv 0 \bmod m$, unless $m = 4 = 2^2$ (in which case $3! = 6 \equiv 2 \bmod 4$). [Hint: consider the prime factorization of m, and show that unless $m = 2^2$, then there are two distinct integers a, b in $\{1, 2, \ldots, m-1\}$ whose product is divisible by m.]

As an example, let's see the method of the proof of Wilson's theorem in action when $p = 7$. The integers $\{2, 3, 4, 5\}$ can be paired up according to multiplicative inverses: the pairs are $(2, 4)$ and $(3, 5)$ as $2 \cdot 4 = 3 \cdot 5 \equiv 1 \bmod 7$. Then

$$6! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 = 1 \cdot (2 \cdot 4) \cdot (3 \cdot 5) \cdot 6 \equiv 1 \cdot 1 \cdot 1 \cdot (-1) \equiv -1 \bmod 7.$$

Remark 4.2. On the homework, you will encounter the *Euler totient function* φ, where $\varphi(n)$ for a positive integer n is the number of positive integers $1 \leq m \leq n$ such that $(m, n) = 1$. Then the method we used to prove Theorem 4.2 can be used to prove the following generalization of Fermat's little theorem, called *Euler's theorem*:

Theorem 4.4 (Euler)**.** Let $n > 1$ be an integer, and a an integer such that $(a, n) = 1$. Then $a^{\varphi(n)} \equiv 1 \bmod n$.

Indeed, if n is prime, then $\varphi(n) = n - 1$ (why?), which recovers Fermat's little theorem.

The final theorem regarding the internal (multiplicative) structure of $\mathbf{Z}/m\mathbf{Z}$ is motivated by analogy with the additive structure of the integers. Observe that the integer 1 is a "generator" of the integers, in the following sense: every integer k is equal to 1 added to itself k times, where if $k < 0$, we interpret this as adding the *additive inverse* -1 to itself k times. We can notice something similar occurring in the invertible elements of $\mathbf{Z}/m\mathbf{Z}$ when m is prime (so the invertible elements are the nonzero ones). For example:

- In $\mathbf{Z}/3\mathbf{Z}$, every nonzero element is a power of 2, since $2^1 = 2$ and $2^2 = 1$.
- In $\mathbf{Z}/5\mathbf{Z}$, every nonzero element is a power of 2, since $2^1 = 2$, $2^2 = 4$, $2^3 = 3$, and $2^4 = 1$.
- In $\mathbf{Z}/7\mathbf{Z}$, every nonzero element is a power of 3, since $3^1 = 3, 3^2 = 2, 3^3 = 6, 3^4 = 4, 3^5 = 5, 3^6 = 1$.

However, if m is not prime, this can fail. For instance, the invertible elements of $\mathbf{Z}/8\mathbf{Z}$ are $\{1, 3, 5, 7\}$. But all of them square to 1, so there is no invertible element $a \in \mathbf{Z}/8\mathbf{Z}$ such that every invertible element of $\mathbf{Z}/8\mathbf{Z}$ is a power of a.

We give such a "generating element" a name:

Definition 4.2. Suppose a is an invertible element of $\mathbf{Z}/m\mathbf{Z}$ such that if b is some invertible element $\mathbf{Z}/m\mathbf{Z}$, then $b = a^k$ for some positive integer k. Then we call a a *primitive root* (mod m).

Here is the key fact:

Theorem 4.5. If p is a prime, then $\mathbf{Z}/p\mathbf{Z}$ has a primitive root. In other words, there is a single nonzero element in $\mathbf{Z}/p\mathbf{Z}$ that generates all other nonzero elements through its powers.

We will not prove this theorem, since our current setup for $\mathbf{Z}/p\mathbf{Z}$ is slightly clunky, and the proof is stated much more easily in the language of abstract algebra.[3] However, on the HW, you will look at some consequences of Theorem 4.5, and you will also see some ideas used in its proof.

4.2 Chinese remainder theorem

In Example 4.3, we found that we were able to solve a single congruence $ax \equiv b \bmod m$, as long as certain conditions are met (i.e. a is coprime to m). We now look at the problem of satisfying *multiple* congruences simultaneously. Let's restrict ourselves to the simplest possible case: that of a system of congruences

$$x \equiv a_1 \bmod m_1, \quad x \equiv a_2 \bmod m_2, \quad \ldots, \quad x \equiv a_n \bmod m_n \tag{1}$$

for integers $m_1, \ldots, m_n > 1$. We would like to find all possible x that satisfy these congruences simultaneously. As an initial observation, notice that if x satisfies (1), and if m is the product of all the m_i, then $x + km$ also satisfies (1) for any integer k. Therefore one solution x gives rise to an infinite set of solutions: all integers equivalent to x mod m.

[3] For instance, in an abstract algebra class, we would be able to discuss all of these ideas in the general context of *cyclic groups*.

Certainly not all possible congruence systems have solutions. For instance, the system

$$x \equiv 1 \bmod 2, \quad x \equiv 2 \bmod 4$$

has no solutions, because the first congruence implies x would have to be odd, while the second implies x would have to be even: an obvious contradiction. For the same reason, the system

$$x \equiv 2 \bmod 6, \quad x \equiv 3 \bmod 9$$

cannot be solved (think about divisibility conditions at 3).

If we look at more examples of this form, we see that a problem might arise when our m_i have common prime factors. In that case, the conditions $x \equiv a_i \bmod m_i$ might "clash", as we saw in the example with $x \equiv 1 \bmod 2, x \equiv 2 \bmod 4$: 2 is a common divisor of 2 and 4, and indeed there was a contradiction arising from divisibility conditions at 2. So suppose we stipulate that the m_i are pairwise coprime, and perhaps we expect that the system (1) has some solution x, no matter what the a_i's are. For instance, the system

$$x \equiv 1 \bmod 5, \quad x \equiv 2 \bmod 6$$

has a solution $x = 26$, and hence any integer equivalent to 26 mod 30 is also a solution.

Our guess turns out to be true: it is the content of the Chinese remainder theorem.

Theorem 4.6 (Chinese remainder theorem)**.** Let $m_1, \ldots, m_n$ be *pairwise coprime* integers greater than 1, and let m be their product. Let $a_1, \ldots, a_n$ be arbitrary integers. Then the system of congruences

$$x \equiv a_1 \bmod m_1, \quad x \equiv a_2 \bmod m_2, \quad \ldots, \quad x_n \equiv a_n \bmod m_n$$

has a unique solution mod m. In other words, there is a solution x in $[0, m-1]$, and any other solution x' is congruent to x mod m.

Proof. The idea is as follows: we want to build x as an integral linear combination

$$x = a_1 b_1 + \ldots + a_n b_n, \tag{2}$$

where for a given $1 \leq i \leq n$, $b_i \equiv 1 \bmod m_i$ and $b_i \equiv 0 \bmod m_j$ for all $j \neq i$ (so far we haven't shown that b_i satisfying these properties exist: this is just wishful thinking as of now). If you've taken linear algebra, this idea may make a bit more sense: the b_i act as "standard basis vectors for modular arithmetic" in a sense (and this idea is not far off: if we had the language of groups, we would see that this description is exactly what we are doing!). Then given this x, for any $1 \leq i \leq n$, we have

$$x \equiv a_i \cdot 1 = a_i \bmod m_i,$$

since all the other b_j are 0 mod m_i, so those terms $a_j b_j$ vanish mod m.

It remains to produce these b_i. Recall that m is the product of the m_i. Let t_i be the quotient m_i/m, so t_i is the product of all the m_j except m_i. Since the m_i are *pairwise* coprime, Proposition 1.3 shows that $(m_i, t_i) = 1$ for all i. Therefore t_i is invertible mod m_i, so find y_i such that $y_i t_i \equiv 1 \bmod m_i$.

We claim that if we set $b_i = y_i t_i$ for each i, then we have the desired properties of the b_i. Since $m_j | t_i$ for all $j \neq i$ by construction, we have $b_i = y_i t_i \equiv 0 \bmod m_j$ for all $j \neq i$. The remaining property, $b_i \equiv 1 \bmod m_i$, is true by definition of t_i! So we have constructed the desired b_i, and x as in (2) solves the congruences simultaneously.

It remains to show uniqueness of x mod m. Suppose x' is another solution to the same system of congruences, so $x - x' \equiv 0 \bmod m_i$ for each i. In other words, m_i divides $x - x'$ for each i, and we aim to show that m divides $x - x'$. Since m is the product of the *pairwise coprime* m_i, we are done once we prove the following Lemma 4.1. □

Lemma 4.1. Let c, d be relatively prime integers. If k is an integer with $c|k$ and $d|k$, then $cd|k$. By induction, if $c_1, \ldots, c_n$ are pairwise relatively prime integers and k is divisible by each c_i, then k is divisible by their product.

Proof. Let d' be such that $d'd = k$. Then as $c|k$, we have $cd|kd$ and thus $cd|d'd^2$. Hence $c|d'd^2$. But c and d are coprime, so c and d^2 are as well (Proposition 1.3), so by Euclid's Lemma (Proposition 1.4), $c|d'$. Hence $cd|d'd \Rightarrow cd|k$. □

Let's end with some numerical applications of the Chinese remainder theorem.

Example 4.6. Suppose we want to solve the system

$$x \equiv 4 \bmod 6, x \equiv 11 \bmod 35.$$

The integers 6 and 35 are coprime, and we have $6 \cdot 6 + 35 \cdot (-1) = 1$. Via this equation, we see that 6 is an inverse of 6 mod 35, and -1 is an inverse of 35 mod 6. So the proof of Theorem 4.6 shows that

$$x = 4(-1 \cdot 35) + 11 \cdot (6 \cdot 6) = 256$$

solves the given congruences. Moreover, by the *uniqueness* statement of the Chinese remainder theorem, the solutions to the system are precisely the integers equivalent to $256 \equiv 46 \bmod 210$, where $210 = 6 \cdot 35$.

Example 4.7. Here is the Chinese mathematician Sunzi's original problem, which lends Theorem 4.6 its name: there are certain things whose number is unknown. If we count them by threes, we have two left over; by fives, we have three left over; and by sevens, two are left over. How many things are there?

The question is equivalent to solving the system of congruences

$$x \equiv 2 \bmod 3, \quad x \equiv 3 \bmod 5, \quad x \equiv 2 \bmod 7.$$

With notation as in the proof of Theorem 4.6, set $m_1 = 3$, $m_2 = 5$, and $m_3 = 7$, so $m = 105$. Then $t_1 = m_2 m_3 = 35$, $t_2 = m_1 m_3 = 21$, and $t_3 = m_1 m_2 = 15$. A straightforward computation shows that $y_1 = 2$, $y_2 = 1$, and $y_3 = 1$ are inverses of t_1, t_2, t_3 mod m_1, m_2, m_3 respectively. Therefore

$$x = 2t_1y_1 + 3t_2y_2 + 2t_3y_3 = 2 \cdot 2 \cdot 35 + 3 \cdot 1 \cdot 21 + 2 \cdot 1 \cdot 15 = 233$$

solves the system of congruences. Since $233 \equiv 23 \bmod 105$, and *every* solution to the system is congruent to 233 mod 105, we conclude that 23 is the smallest positive integer solution to the question.

Here is one more example, combining the Chinese remainder theorem with Fermat's little theorem:

Example 4.8. In Remark 4.1, we noted that if $n \geq 2$ is an integer (not necessarily prime), and $a^{n-1} \equiv 1 \bmod n$ for all integers a coprime to n, then this does *not* imply that n is prime (i.e. any possible formulation of a converse to Fermat's little theorem is false). Composite integers n satisfying this annoying property are called *Carmichael numbers*, and we claimed that 561 is such a Carmichael number. We now have the tools to demonstrate this.

Note that $561 = 3 \cdot 11 \cdot 17$. Suppose $(a, 561) = 1$, so in particular a is not equivalent to 0 mod 3, 11, or 17. So by Fermat's little theorem, we have $a^2 \equiv 1 \bmod 3$, $a^{10} \equiv 1 \bmod 11$, and $a^{16} \equiv 1 \bmod 17$. Since 560 is divisible by each of 2, 10, and 16, we conclude that $a^{560} \equiv 1 \bmod m$ for each of $m = 3, 11, 17$. But 1 is another integer satisfying the *same* system of congruences! So by the *uniqueness* statement of the Chinese remainder theorem, we conclude that a^{560} is congruent to 1 mod 561, as $3 \cdot 11 \cdot 17 = 561$.

The same tools—Fermat's little theorem and the Chinese remainder theorem—can be used to prove *Korselt's criterion*: a composite integer $n > 2$ is a Carmichael number if and only if the following two conditions hold:

(i) n is squarefree (i.e. n is not divisible by any perfect square other than 1).

(ii) For every prime p dividing n, we also have $(p-1)|(n-1)$.

In fact, our example with $n = 561$ actually supplies most of the ideas needed in this proof, although some slight generalizations of Fermat's little theorem are needed.

We have only touched the surface of the Chinese remainder theorem (there are many generalizations of this fact, including *geometric* interpretations!). Indeed, we have only scratched the surface of what number theory has to offer, and I'm always happy to discuss the mathematics that can pop out of even the most innocuous number-theoretic questions.