# 3 Modular arithmetic

For further reading on the material in this subsection, consult **Rosen 4.1, PMF 7.1-2, Stein 2.1, Shoup 2.1-2,2.5**.

Modular arithmetic is a powerful tool that lets us do arithmetic while preserving information about divisibility, and has a broad range of number theory applications. We'll be studying various aspects of it for the next few sections.

## 3.1 Congruences

**Definition 3.1.** Let $m$ be a positive integer. If $a, b$ are integers, we say that $a$ is congruent to $b$ modulo $m$, and write $a \equiv b \mod m$, if $m$ divides $a - b$.

**Proposition 3.2.** *The congruence* $\mod m$ *relation is an equivalence relation; that is:*

- *(Reflexive Property) If $a$ is an integer, then $a \equiv a \mod m$.*
- *(Symmetric Property) If $a, b$ are integers and $a \equiv b \mod m$ then $b \equiv a \mod m$.*
- *(Transitive property) If $a, b, c$ are integers, and $a \equiv b \mod m$ and $b \equiv c \mod m$, then $a \equiv c \mod m$.*

*Proof.*

- We see that $m|0 = (a - a)$ so $a \equiv a \mod m$.
- If $a \equiv b \mod m$ then $m|a - b$, which means that there is an integer $k$ such that $km = a - b$. Then $(-k)m = b - a$ so $m|b - a$ so $b \equiv a \mod m$.
- If $a \equiv b \mod m$ and $b \equiv c \mod m$ then $m|a - b$ and $m|b - c$. By the lemma on linear combinations, $m|(a - b) + (b - c) = a - c$, so $a \equiv c \mod m$.

□

You should recall from Math 210 that an equivalence relationship partitions a set into *equivalence classes*, which in this case are called *congruence classes* $\mod m$. Two integers are in the same congruence class $\mod m$ if they are congruent to each other. For a fixed integer $m$, there are precisely $m$ congruence classes $\mod m$. For example, if $m = 2$, the two congruence classes are even integers (congruent to $0 \mod m$) and odd integers (congruent to $1 \mod m$).

We'd like to pick canonical representatives of each equivalence class. There are a few different ways to do this.

**Definition 3.3.** Let $a$ be an integer and $m$ a positive integer. By the Division Algorithm, there is a unique $r$ with $0 \leq r < m$ such that $a = mq + r$ for some integer $q$. We call this $r$ the *reduction of $a \mod m$.*

**Proposition 3.4.**
- *The reduction of $a \mod m$ is congruent to $a \mod m$.*
- *The reduction of $a \mod m$ is an element of the set $\{0, 1, \ldots, m-1\}$.*

*Proof.*
- Let $r$ be the reduction of $a \mod m$. Then $a = mq + r$ for some integer $q$, so $a - r = mq$ is divisible by $m$. Thus $a \equiv r \mod m$.
- We know that $r$ is an integer with $0 \leq r < m$ from the division algorithm.

□

**Definition 3.5.** We say a set $S$ is a *complete system of residues* $\mod m$ if any integer $a$ is congruent $\mod m$ to exactly one element of $S$.

**Corollary 3.6.** *The set $\{0, 1, \ldots, m-1\}$ is a complete system of residues* $\mod m$. *We sometimes call it the set of* least nonnegative residues $\mod m$. *I will sometimes write $\mathbb{Z}/m\mathbb{Z}$ for this set.*

**Example 3.7.** The set $\{1, 5, 9\}$ is a complete system of residues $\mod 3$.

**Example 3.8.** If $m$ is an odd positive integer, then the set

$$\left\{-\frac{m-1}{2}, -\frac{m-2}{2}, \ldots, -1, 0, 1, \ldots, \frac{m-1}{2}\right\}$$

is a complete system of residues $\mod m$. We sometimes call it the set of *absolute least residues* $\mod m$.

**Exercise 3.9.** *Let $S$ be a set of $m$ integers such that no element of $S$ is congruent to any other element of $S \mod m$. Prove that $S$ is a complete system of residues.*

Possibly the most useful fact about congruences is that they behave well with respect to basic arithmetic operations.

**Theorem 3.10.** *If $a, b, c, d, m$ are integers with $m > 0$, and $a \equiv b \mod m$ and $c \equiv d \mod m$, then*

1. $a + c \equiv b + d \mod m$
2. $a - c \equiv b - d \mod m$

*3.* $ac \equiv bd \mod m$.

*Proof.* Since $m|a-b$ and $m|c-d$, there are integers $k, \ell$ such that $km = a-b$ and $\ell m = c-d$. Then

1. $(a+c)-(b+d) = (a-b)+(c-d) = km + \ell m = (k+\ell)m$. Thus $m|(a+c)-(b+d)$ and $a+c \equiv b+d \mod m$ by definition.
2. The same proof holds here.
3. $ac - bd = ac - bc + bc - bd = c(a-b) + b(c-d) = ckm + b\ell m = m(ck + b\ell)$. Thus $m|ac-bd$ and $ac \equiv bd$ by definiton.

□

*Remark* 3.11. For those of you who have taken algebra: we can interpret these results as showing that the set of integers $\mod m$ form an abelian group under addition, and in fact form a ring under the usual addition and multiplication. Then the reduction modulo $m$ map is a group (or ring) homomorphism from the integers to the integers modulo $m$. And in this case we can interpret $\mathbb{Z}/m\mathbb{Z}$ as the quotient of the integers with respect to the kernel of this homomorphism. See PMF 8.1-2, or Stein or Shoup for more on this perspective.

*Remark* 3.12. Note that division is *not* on this list. Division in modular arithmetic is in fact somewhat subtle, in contrast to the straightforwardness of addition and multiplication.

For instance, we see that $7 \cdot 2 = 14 \equiv 8 = 4 \cdot 2 \mod 6$. But it is not true that $7 \cdot 4 \mod 6$.

Exponentiation, however, works as well as we'd like.

**Lemma 3.13.** *Let $a, b, k, m$ be integers with $k, m > 0$, and $a \equiv b \mod m$. Then $a^k = b^k \mod m$.*

*Proof.* Recall that

$$a^k - b^k = (a-b)(a^{k-1} + a^{k-2}b + \cdots + ab^{k-2} + b^{k-1}) = (a-b)\prod_{i=0}^{k-1} a^{k-1-i}b^i.$$

Thus since $m|(a-b)$, we also know that $m|a^k - b^k$ and so $a^k \equiv b^k \mod m$ by definition. □

*Remark* 3.14. The converse is not actually true, which is easy to see: e.g. $1^2 \equiv 2^2 \mod 3$.

*Remark* 3.15. When computing the reduction mod $p$ of some large exponent, it's hopelessly inefficient to do the entire exponentiation and then do a division. It's moderately more efficient to see what power you need to raise the base to to get a reduction, and then iterate: e.g. if I want to compute $2^{100} \mod 5$ I will observe that $2^4 = 16 \equiv 1 \mod 5$, and thus $2^{100} = (2^4)^{20} \equiv 1^{20} \equiv 1 \mod 5$.

Rosen covers a more efficient way still towards the end of §4.1, which turns the problem mostly into a huge bitwise XOR.

**Corollary 3.16.** *Let $n$ be an integer. Then $3|n$ if and only if 3 divides the sum of the (base ten) digits of $n$.*

*Proof.* Write $n = n_0 + n_1 \cdot 10^1 + n_2 \cdot 10^2 + \cdots + n_k \cdot 10^k$ (with $0 \leq n_i \leq 9$). We notice that $10 \equiv 1 \mod 3$ and thus for any $\ell > 0$, $10^\ell \equiv 1^\ell = 1 \mod 3$. Thus

$$n \equiv n_0 + n_1 + \cdots + n_k \mod 3$$

and the right-hand side is the sum of the decimal digits.

Then in particular $n \equiv 0 \mod 3$ if and only if the sum of the digits is congruent to 0 mod 3, as desired. □

*Remark* 3.17. Similar arguments can be made for divisibility by other integers, like 9 or 11.

## 3.2 Linear Congruences and Modular Division

For further reading on the material in this subsection, consult **Rosen 4.1-2, PMF 8.3, Stein 2.1.1, Shoup 2.3**.

Modular division is a bit trickier to understand than other modular arithmetic. Recall our earlier example:

**Example 3.18.** $8 \equiv 2 \mod 6$ but $4 \not\equiv 1 \mod 6$.

However, you might notice that $4 \equiv 1 \mod 3$; we have essentially divided the *modulus* by 2 as well as the two equivalent integers. It turns out that this is basically what's going on.

**Proposition 3.19** (Modular cancellation law)**.** *Let $a, b, c, m$ be integers with $m > 0$, and set $d = (c, m)$. Then if $ac \equiv bc \mod m$ then $a \equiv b \mod m/d$.*

*Proof.* If $ac \equiv bc \mod m$ then $m|ac - bc = c(a - b)$, so there is a $k$ with $km = c(a - b)$. Dividing through by $d$ gives $km/d = c/d(a - b)$. Thus $m/d|(c/d)(a - b)$.

But $(m/d, c/d) = 1$, so we know that $m/d|a - b$ and thus $a \equiv b \mod m/d$. □

**Corollary 3.20.** *Let $a, b, c, m$ are integers with $m > 0$ and $(c, m) = 1$. Then if $ac \equiv bc \mod m$ then $a \equiv b \mod m$.*

But often we would really like not to change the modulus. Thus we ask ourselves how division works with respect to a fixed modulus.

To answer this question, let's think about what division really means. Division is in essence undoing multiplication. So when we compute $b/a$, we are actually solving the equation $ax = b$ for $x$. Similarly, if we want to understand modular division, we should study the congruence $ax \equiv b \mod m$.

**Definition 3.21.** A congruence of the form $ax \equiv b \mod m$ , where $a, b$ are constant integers and $x$ is unknown, is a *linear congruence in one variable.*

First notice that if this congruence has any solution, it has infinitely many. Suppose $x_0$ is a solution to this equation, and $x_1$ is an integer such that $x_1 \equiv x_0 \mod m$. Then $ax_1 \equiv ax_0 \equiv b \mod m$ so $x_1$ is another solution.

Thus if $x_0$ solves a linear congruence $\mod m$, any element of its congruence class will. (This will generally be the case for solving congruences, whether linear or not). But in modular arithmetic we tend to want to treat elements of the same congruence class as being the same.

The next question we might ask is *how many* solutions a given congruence has? Non-modularly, a linear equation $ax = b$ has either zero solutions (when $b$ is not divisible by $a$), or exactly one solution (when it is). The modular situation is a bit more complicated; fortunately, we have already done some work towards solving this problem in another form.

**Lemma 3.22** (Linear Diophantine Equations)**.** *Let $a, b$ be integers with $d = (a, b)$. Then the equation $ax + by = c$ has solutions if and only if $d \mid c$.*

*Further, if $d|c$ then there are infinitely many solutions, all of which have the following form: if $x_0, y_0$ is some particular solution, the set of all solutions is given by*

$$x = x_0 + (b/d)t, \qquad y = y_0 - (a/d)t$$

*where $t$ is an integer.*

*Proof.* The first statement was show in Homework 2 problem 1.

Suppose $d|c$ and let $(x_0, y_0)$ be a solution. Then

$$\begin{aligned} a(x_0 + (b/d)t) + b(y_0 - (a/d)t) &= ax_0 + abt/d + by_0 - abd/t \\ &= ax_0 + by_0 = c. \end{aligned}$$

Thus every pair of this form is a solution, and there are infinitely many solutions.

Finally, we prove every solution is of this form. Suppose $x, y$ are integers with $ax+by = c$. Then subtracting $ax_0 + by_0 = c$ from this equation gives

$$\begin{aligned} a(x - x_0) + b(y - y_0) &= 0 \\ a(x - x_0) &= b(y - y_0) \\ (a/d)(x - x_0) &= (b/d)(y - y_0). \end{aligned}$$

Since $(a, b) = d$, we know that $(a/d, b/d) = 1$ and thus $(a/d)|(y - y_0)$. Thus there is an integer $t$ with $(a/d)t = y - y_0$, which we can rewrite $y = y_0 + (a/d)t$.

Plugging this in to our earlier equation gives

$$\begin{aligned} a(x - x_0) &= b(a/d)t \\ x - x_0 &= (b/d)t \\ x &= x_0 + (b/d)t \end{aligned}$$

as desired. □

**Proposition 3.23.** *Let $a, b, m$ be integers with $m > 0$, and set $(a, m) = d$. If $d \not| b$, then $ax \equiv b \mod m$ has no solutions. If $d|b$, then $ax \equiv b \mod m$ has exactly $d$ "distinct" or incongruent solutions modulo $m$.*

*Proof.* The key step here is to turn our linear congruence in one variable into a linear equation in two variables. We know that $ax \equiv b \mod m$ if $m|ax - b$, which is true precisely when there is some integer $y$ such that $my = ax - b$. Thus $x$ is a solution to $ax \equiv b \mod y$ if and only if there is a $y$ such that $(x, y)$ is a solution to $ax - my = b$.

We showed in homework 2 problem 1 that $ax - my = b$ has solutions if and only if $d = (a, m)|b$. This proves the first claim.

If $d|b$, then $ax - my = b$ has infinitely many solutions, all of which are given by the formulas

$$x = x_0 + (m/d)t, \qquad y = y_0 - (a/d)t$$

Thus the values $x = x_0 + (m/d)t$ are the infinitely many solutions of the linear congruence.

Finally, we wish to prove that there are $d$ incongruent solutions. Let $x_1 = x_0 + (m/d)t_1$ and $x_2 = x_0 + (m/d)t_2$ be two solutions; they are congruent modulo $m$ if and only if $(m/d)t_1 \equiv (m/d)t_2 \mod m$.

But $(m, m/d) = m/d$, so by the modular cancellation law of proposition 3.19, this holds if and only if $t_1 \equiv t_2 \mod (m/(m/d)) = d$.

Thus we have two incongruent solutions when we have solutions whose $t$ are congruent modulo $d$ but not modulo $m$. In particular, the set of solutions $\{x = x_0 + (m/d)t : 0 \leq t < d\}$ is a set of $d$ mutually incongruent solutions, □

**Corollary 3.24.** *If $a, b, m$ are integers with $m > 0$ and $(a, m) = 1$, then the linear congruence $ax \equiv b \mod m$ has a unique solution modulo $m$.*

This tells us that division is not, in general, unique modulo $m$; sometimes $b$ is not divisible by $a$, but sometimes $b/a$ gives multiple reasonable answers. But if $(a, m) = 1$, then division by $a$ modulo $m$ *always* gives one solution, thus every number is uniquely divisble by $a$.

In even more particular, if $p$ is a prime number, then every number is uniquely divisible by every non-zero number modulo $p$. We will use this fact a lot during the rest of the course.

**Example 3.25.** Let's compute $10/4 \mod 14$. I.e. let's find solutions to $4x \equiv 10 \mod 14$. We first note that $(10, 14) = 2$ and $2|10$, so there are exactly 2 incongruent solutions.

We first need to find a particular solution. In small cases like this it's easy enough to just plug numbers in; a more general approach is to use the Euclidean algorithm to write the $2 = -3 \cdot 4 + 1 \cdot 14$ as a linear combination of 4 and 14. We then use this to solve the equation $4x - 14y = 10$, giving us

$$10 = 5 \cdot 2 = 5 \cdot (-3 \cdot 4 + 1 \cdot 14) = -15 \cdot 4 + 5 \cdot 14$$

and thus one solution is given by $x_0 = -15 \equiv 13 \mod 14$ and $y_0 = 10$.

Then a complete set of incongruent solutions is given by

$$x = x_0 + (m/d)t = 13 + (14/2)t = 13 + 7t$$

and thus we have $x = x_0 = 13$ and $x = x_0 + 7 = 20 \equiv 6$.

We can somewhat simplify this process by understanding how reciprocals work.

**Definition 3.26.** Given an integer $a$ with $(a, m) = 1$, an integer solution $x$ to $ax \equiv 1 \mod m$ is called an *inverse of $a$ modulo $m$*.

*Remark* 3.27. Note that we never have a modular inverse if $(a, m) \neq 1$.

**Example 3.28.** What is a modular inverse of $9 \mod 29$?

We use the Euclidean algorithm:

$$\begin{aligned} 29 &= 3 \cdot 9 + 2 \\ 9 &= 4 \cdot 2 + 1 = 4(29 - 3 \cdot 9) + 1 \\ 1 &= 9 - (4 \cdot 29 - 12 \cdot 9) = 13 \cdot 9 - 4 \cdot 29 \end{aligned}$$

thus $(13, 4)$ is a solution to $9x - 29y = 1$ and we see that $9 \cdot 13 \equiv 1 \mod 29$. (In particular, $9 \cdot 13 = 117 = 116 + 1 = 4 \cdot 29 + 1$).

Importantly, if we have a an inverse of $a \mod m$ we can use this to solve other linear congruences of the form $ax \equiv b \mod m$. In particular, if $a^{-1}$ is an inverse of $a \mod m$, then $aa^{-1} \equiv 1 \mod m$ and thus $a(a^{-1}b) \equiv b \mod m$.

**Example 3.29.** Solve the linear congruence $9x \equiv 5 \mod 29$.

We know that 13 is an inverse for $9 \mod 29$. Thus the unique up to congruence solution for this congruence is $13 \cdot 5 = 65 \equiv 7 \mod 29$.

Notice that every (non-zero) number has a modular inverse modulo $p$ if $p$ is a prime number. We state one result for these special cases.

**Lemma 3.30.** *Let $p$ be prime. The positive integer $a$ is its own inverse modulo $p$ if and only if $a \equiv 1 \mod p$ or $a \equiv -1 \mod p$.*

*Proof.* If $a \equiv \pm 1 \mod p$ then $a^2 \equiv (\pm 1)^2 = 1 \mod p$.

Conversely, suppose $a^2 \equiv 1 \mod p$. Then $p | a^2 - 1 = (a+1)(a-1)$, and since $p$ is prime, either $p|(a+1)$ or $p|(a-1)$. In the first case, $a \equiv -1 \mod p$, and in the second case, $a \equiv 1 \mod p$. □

## 3.3 Multiple Moduli and the Chinese Remainder Theorem

For further reading on the material in this subsection, consult **Rosen 4.1,4.3, PMF 7.4, Stein 2.2, Shoup 2.4**.

All of our results so far, except for a few results on division, have kept the modulus $m$ unchanged. But it is useful sometimes to combine congruences with different moduli, and this is in fact quite possible.

**Exercise 3.31.** *Let $a, b, m, n$ be integers with $m, n > 0$ and $m|n$. If $a \equiv b \mod n$, then $a \equiv b \mod m$.*

**Proposition 3.32.** *If $a \equiv b \mod m_1, a \equiv b \mod m_2, \dots, a \equiv b \mod m_k$, then*

$$a \equiv b \mod \operatorname{lcm}(m_1, m_2, \dots, m_k).$$

*Proof.* We know that $m_1|a-b, m_2|a-b, \dots, m_k|a-b$. Thus the LCM of $m_1, \dots, m_k$ also divides $a-b$ (by e.g. Exercise 5 of HW2), and $a \equiv b \mod \operatorname{lcm}(m_1, \dots, m_k)$ by definition. □

In the past section we solved single congruences. Now we want to turn our attention to solving systems of multiple congruences. The first known discussion of this problem comes from Sunzi Suanjing (also known as Sun Tzu, but not the one who wrote *The Art of War*) in the third century:

"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? "

We can rephrase this in our language:

**Question 3.33.** Suppose we have the following system of congruences:

$$\begin{aligned} x &\equiv 2 \mod 3 \\ x &\equiv 3 \mod 5 \\ x &\equiv 2 \mod 7. \end{aligned}$$

What are the possible values of $x$?

This question was studied by many mathematicians in China, India, the Middle East, and Europe; the first known algorithm to solve the question is due to Aryabhata in the sixth century, and the first known complete solution is due to Qin Jiushao in 1247.

**Theorem 3.34** (Chinese Remainder Theorem)**.** *Let $m_1, m_2, \dots, m_r$ be pairwise prime positive integers. Then the system*

$$\begin{aligned} x &\equiv a_1 \mod m_1 \\ x &\equiv a_2 \mod m_2 \\ &\vdots \\ x &\equiv a_r \mod m_r \end{aligned}$$

*has a unique solution modulo $M = m_1 m_2 \dots m_r$. Further there is an algorithm for finding the solution.*

*Proof.* First we will present an algorithm that will always find a solution, proving existence. Then we will prove uniqueness modulo $M$.

For each $k$, set $M_k = M/m_k = m_1 m_2 \dots m_{k-1} m_{k+1} \dots m_r$. We see that $(M_k, m_k) = 1$ since $m_k$ is relatively prime to $m_j$ for each $j \neq k$. Thus by modular division we can find an inverse of $M_k$ modulo $m_k$, which we shall call $y_k$. Thus $y_k M_k \equiv 1 \mod m_k$.

Now set

$$x = a_1M_1y_1 + a_2M_2y_2 + \cdots + a_rM_ry_r.$$

We know that if $i \neq j$, then $M_j \equiv 0 \mod m_i$. Thus we see that for each $i$, we have

$$\begin{aligned} x &= a_1M_1y_1 + a_2M_2y_2 + \cdots + a_rM_ry_r \\ &\equiv 0 + 0 + \cdots + 0 + a_iM_iy_i + 0 + \cdots + 0 \mod m_i \\ &\equiv a_i(M_iy_i) \mod m_i \\ &\equiv a_i \cdot 1 \mod m_i \end{aligned}$$

and thus $x$ satisfies each congruence.

Now we prove that this solution is unique modulo $M$. Suppose $x_0$ and $x_1$ are both solutions to the system of congruences. Then for each $i$ we know that $x_0 \equiv a_i \mod m_i$ and $x_1 \equiv a_i \mod m_i$ and thus $x_0 \equiv x_1 \mod m_i$.

But then $x_0 \equiv x_1 \mod \operatorname{lcm}(m_1, m_2, \ldots, m_r)$, and since the $m_i$ are all relatively prime $\operatorname{lcm}(m_1, m_2, \ldots, m_r) = m_1m_2 \ldots m_r = M$. □

**Example 3.35.** Let's solve the system of congruences given earlier:

$$\begin{aligned} x &\equiv 2 \mod 3 \\ x &\equiv 3 \mod 5 \\ x &\equiv 2 \mod 7. \end{aligned}$$

Then we have $M = 3 \cdot 5 \cdot 7 = 105$. Then we compute

$$\begin{aligned} M_1 &= 5 \cdot 7 = 35 \equiv 2 \mod 3 & y_1 &= 2 \\ M_2 &= 3 \cdot 7 = 21 \equiv 1 \mod 5 & y_2 &= 1 \\ M_3 &= 3 \cdot 5 = 15 \equiv 1 \mod 7 & y_3 &= 1 \end{aligned}$$

and thus

$$x = 2 \cdot 35 \cdot 2 + 3 \cdot 21 \cdot 1 + 2 \cdot 15 \cdot 1 = 140 + 63 + 30 = 233.$$

We can check that 233 satisfies the three equivalences (and in particular, $3 \cdot 77 = 231, 5 \cdot 46 = 230, 7 \cdot 33 = 231$).

A possible last step is to notice that the solution is unique modulo $M = 105$. Thus the least nonnegative solution is $233 = 210 = 23$, which we can easily see satisfies the congruences.

## 3.4 Systems of linear congruences

For further reading on the material in this subsection, consult **Rosen 4.5, PMF 8.3, Stein 2.3**.

So far we have solved congruences in one variable; we can also try to solve systems of congruences in two or more variables.

In this subsection we will study systems of two linear congruences in two variables. We can generalize this theory to larger systems of linear congruences in more variables, but this requires a substantial dose of linear algebra (similar to the linear algebra involved in solving large systems of linear equations), so we shall pass over it here. If you're interested, this can make a good course paper.

**Theorem 3.36.** *Let $a, b, c, d, e, f, m$ be integers with $m > 0$. Set $\Delta = ad - bc$ (notice this is the determinant of a $2 \times 2$ matrix ). Then if $(\Delta, m) = 1$, the system of congruences*

$$\begin{aligned} ax + by &\equiv e \mod m \\ cx + dy &\equiv f \mod m \end{aligned}$$

*has a unique solution modulo $m$, given by*

$$\begin{aligned} x &\equiv \Delta^{-1}(de - bf) \mod m \\ y &\equiv \Delta^{-1}(af - ce) \mod m \end{aligned}$$

*where $\Delta^{-1}$ is a multiplicative inverse of $\Delta$ modulo $m$.*

*Proof.* We can solve these just like we normally solve linear equations. To remove $y$ we see

$$\begin{aligned} adx + bdy &\equiv de \mod m \\ bcx + bdy &\equiv bf \mod m \\ adx - bcx &\equiv de - bf \mod m \\ \Delta x &\equiv de - bf \mod m \end{aligned}$$

Then since $(\Delta, m) = 1$, we know $\Delta$ has a multiplicative inverse $\Delta^{-1}$; multiplying by this gives

$$x \equiv \Delta^{-1}(de - bf) \mod m.$$

Thus any solution must satisfy this relation, as claimed.

We can similarly eliminate $y$ from these equations via

$$\begin{aligned} acx + bcy &\equiv ce \mod m \\ acx + ady &\equiv af \mod m \\ ady - bcy &\equiv af - ce \mod m \\ \Delta y &\equiv af - ce \mod m \\ y &\equiv \Delta^{-1}(af - ce) \mod m. \end{aligned}$$

Again, this is the relation we claimed was necessary. Thus we have proven uniqueness.

Now we just need to check that any pair like this is a solution. But then

$$\begin{aligned} ax + by &\equiv a\Delta^{-1}(de - bf) + b\Delta^{-1}(af - ce) \\ &\equiv \Delta^{-1}(ade - abf + abf - bce) \\ &\equiv \Delta^{-1}(ad - bc)e \\ &\equiv \Delta^{-1}\Delta e \equiv e \mod m \end{aligned}$$

and similarly

$$\begin{aligned} cx + dy &\equiv c\Delta^{-1}(de - bf) + d\Delta^{-1}(af - ce) \\ &\equiv \Delta^{-1}(cde - bcf + adf - cde) \\ &\equiv \Delta^{-1}(ad - bc)e \\ &\equiv \Delta^{-1}\Delta e \equiv e \mod m. \end{aligned}$$

□

This analysis could be extended to solve systems with more equations and more variables, but it's mostly an exercise in linear algebra, so we won't do it here.

**Example 3.37.** Consider the system

$$\begin{aligned} 2x + 3y &\equiv 7 \mod 13 \\ 5x + 2y &\equiv 3 \mod 13 \end{aligned}$$

We have $\Delta = 2 \cdot 2 - 3 \cdot 5 = -11 \equiv 2 \mod 13$, and thus $\Delta^{-1} \equiv 7 \mod 13$. Then the system has a unique solution $\mod 13$, given by

$$\begin{aligned} x &\equiv \Delta^{-1}(de - bf) \equiv 7(2 \cdot 7 - 3 \cdot 3) \equiv 35 \equiv 9 \mod 13 \\ y &\equiv \Delta^{-1}(af - ce) \equiv 7(2 \cdot 3 - 5 \cdot 7) \equiv 7 \cdot (-29) \equiv -21 \equiv 5 \mod 13. \end{aligned}$$

We can check our work by plugging these in:

$$2 \cdot 9 + 3 \cdot 5 \equiv 18 + 15 \equiv 33 \equiv 7 \mod 13$$
$$5 \cdot 9 + 2 \cdot 5 \equiv 45 + 5 \equiv 55 \equiv 3 \mod 13.$$

## 3.5 Solving Polynomial Congruences

For further reading on the material in this subsection, consult **Rosen 4.4**.

So far we've been looking only at linear congruences. But we can also solve congruences with polynomial equations in the variable.

We can break this problem up into two pieces (as we can with all congruence problems in the future). If we have a congruence modulo $m = p_1^{n_1} \dots p_r^{n_r}$, we can use the Chinese Remainder theorem to split this up into a system of $r$ congruences, modulo each prime power

**Example 3.38.** Suppose we want to solve the congruence

$$2x^3 + 12x + 4 \equiv 0 \mod 100.$$

We see that $200 = 2^3 5^2$ so we need to solve

$$2x^3 + 12x + 4 \equiv 0 \mod 4$$
$$2x^3 + 12x + 4 \equiv 0 \mod 25.$$

The first we can solve easily enough by testing numbers; we see that it holds when $x \equiv 0$ or $x \equiv 2 \mod 4$. As we shall see below, the second equivalence holds when $x \equiv 19 \mod 25$.

Then by the Chinese Remainder theorem, the congruence $\mod 100$ holds if and only if

$$x = 0 \cdot 25 \cdot 1 + 19 \cdot 4 \cdot 19 = 1444 \equiv 44 \mod 100$$
$$x2 \cdot 25 \cdot 1 + 19 \cdot 4 \cdot 19 = 1494 \equiv 94 \mod 100$$

So we can reduce the problem of solving congruences in general to the problem of solving congruences modulo a prime power (i.e. $p^n$ for some integer $n$). Fortunately, we can approach these congruences $\mod p^n$ by an even simpler step, by simply solving them $\mod p$.

**Example 3.39.** We wish to find the solutions to

$$2x^3 + 12x + 4 \equiv 0 \mod 25.$$

First we solve

$$2x^3 + 12x + 4 \equiv 0 \mod 5$$

which we can do by the guess-and-check method:

$$\begin{aligned}
2(0)^3 + 12(0) + 4 &\equiv 4 \mod 5 \\
2(1)^3 + 12(1) + 4 &\equiv 13 \equiv 3 \mod 5 \\
2(2)^3 + 12(2) + 4 &\equiv 1 + 4 - 1 \equiv 4 \mod 5 \\
2(3)^3 + 12(3) + 4 &\equiv 4 + 6 - 1 \equiv 4 \mod 5 \\
2(4)^3 + 12(4) + 4 &\equiv -2 + 8 - 1 \equiv 0 \mod 5
\end{aligned}$$

So this has a solution if and only if $x \equiv 4 \mod 5$.

So what about $\mod 25$? Well, we know any solution must be equivalent to $4 \mod 5$, so we only need to test solutions of the form $4 + 5t$. Plugging this in gives

$$\begin{aligned}
2(4 + 5t)^3 + 12(4 + 5t) + 4 &\equiv 0 \mod 25 \\
2(64 + 240t + 300t^2 + 125t^3) + 48 + 60t + 4 &\equiv 0 \mod 25 \\
3 + 5t + 23 + 10t + 4 &\equiv 0 \mod 25 \\
15t + 5 &\equiv 0 \mod 25
\end{aligned}$$

which holds only if $t \equiv 3 \mod 5$. Thus the only solution $\mod 25$ is 19.

We call this process "lifting", where we start with a solution in some small modulus, and then lift it up to a power of that modulus. We'd like to systematize this, which will require some tools that are annoyingly familiar.

**Definition 3.40.** Let $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$. We define the *derivative* of $f(x)$ to be

$$f'(x) = na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + a_1.$$

We use the notation $f^{(k)}(x)$ to denote the result of repeating the derivative $k$ times.

*Remark* 3.41. This should be familiar from calculus, but the definition doesn't actually require calculus; this is useful, because it means we can use it even when we're not working with real numbers. We can do the same trick with functions like log and exp that can be represented by power series, but we won't worry about those here.

**Lemma 3.42.** *If $f(x), g(x)$ are polynomials and $c$ is a constant, then $(f+g)'(x) = f'(x) + g'(x)$ and $(cf)'(x) = cf'(x)$. Further, $(f+g)^{(k)}(x) = f^{(k)}(x) + g^{(k)}(x)$ and $(cf)^{(k)}(x) = cf^{(k)}(x)$.*

**Lemma 3.43.** *If $m, k$ are positive integers, and $f(x) = x^m$, then*

$$f^{(k)} = m(m-1)\dots(m-k+1)x^{m-k} = \frac{m!}{(m-k)!}x^{m-k}.$$

**Lemma 3.44** (Taylor expansions)**.** *If $f(x)$ is a polynomial of degree $n$, and $a, b \in \mathbb{R}$, then*

$$f(a+b) = \sum_{k=0}^{n} \frac{f^{(k)}(a)b^k}{k!} = f(a) + f'(a)b + \frac{f''(a)b^2}{2} + \dots + \frac{f^{(n)}(a)b^n}{n!}.$$

*and this is a polynomial in $b$ whose coefficients are polynomials in $a$ with integer coefficients.*

*Proof.* We will prove this for $f_m(x) = x^m$. This is sufficient to prove it for any polynomial $f(x)$, since any polynomial is the sum $a_0 f_0(x) + \dots + a_n f_n(x)$ of scalar multiples of $x^m$ for various $m$, and derivatives commute with addition and scalar multiplication.

By the binomial theorem,

$$(a+b)^m = \sum_{k=0}^{m} \binom{m}{k} a^{m-k}b^k.$$

But $f_m^{(k)}(a) = \frac{m!}{(m-k)!}a^{m-k} = k!\binom{m}{k}a^{m-k}$, so

$$f_m(a+b) = \sum_{k=0}^{m} \binom{m}{k} a^{m-k}b^k = \frac{f_m^{(k)}(a)b^k}{k!}.$$

We can see these coefficients must be integers because $\frac{f^{(k)}(a)}{k!} = \binom{m}{k}a^{m-k}$ and $\binom{m}{k}$ is an integer. □

Now we're ready to prove the key lemma about lifting, known as Hensel's Lemma after Kurt Hensel, who studied a field known as $p$-adic analysis. This lemma is complicated to state and annoyingly technical (that's why we say it's a lemma!), but is exceptionally useful for studying polynomial congruences.

**Theorem 3.45** (Hensel's Lemma)**.** *Suppose $f(x)$ is a polynomial with integer coefficients, $k$ is an integer with $k \geq 2$, and $p$ is a prime. Suppose $r$ is a solution to $f(x) \equiv 0 \mod p^{k-1}$ (that is, $f(r) \equiv 0 \mod p^{k-1}$). Then*

1. *If $f'(r) \not\equiv 0 \mod p$, then there is a unique integer $t$ with $0 \leq t < p$ such that $f(r + tp^{k-1}) \equiv 0 \mod p^k$. Further, $t$ is given by the formula*

$$t \equiv -(f'(r))^{-1}(f(r)/p^{k-1}) \mod p$$

*where $(f'(r))^{-1}$ is an inverse of $f'(r)$ modulo $p$.*

2. *If* $f'(r) \equiv 0 \mod p$ *and* $f(r) \equiv 0 \mod p^k$*, then* $f(r + tp^{k-1}) \equiv 0 \mod p^k$ *for any integer* $t$*.*

3. *If* $f'(r) \equiv 0 \mod p$ *and* $f(r) \not\equiv 0 \mod p^k$*, then* $f(x) \equiv 0 \mod p^k$ *has no solutions with* $x \equiv r \mod p^{k-1}$*.*

In particular, this tells us that if the derivative is 0, then a solution modulo $p^{k-1}$ lifts to a unique solution modulo $p^k$; if the derivative is not zero, it lifts either to $p$ different solutions modulo $p^k$, or to none at all.

*Proof.* Notice that every solution modulo $p^k$ is also a solution modulo $p^{k-1}$. That is, if $f(x) \equiv 0 \mod p^k$ then $f(x) \equiv 0 \mod p^{k-1}$. Therefore, if $f(x) \equiv 0 \mod p^k$, then $x = r + tp^{k-1}$ for some $r$ which satisfies $f(r) \equiv 0 \mod p^{k-1}$, and some integer $t$. We just need to determine the conditions on $t$.

So suppose $f(r + tp^{k-1}) \equiv 0 \mod p^k$. By the lemma on Taylor expansions,

$$f(r+tp^{k-1}) = \sum_{i=1}^{n} \frac{f^{(i)}(r)(tp^{k-1})^i}{i!} = f(r) + f'(r)tp^{k-1} + \frac{f''(r)}{2}(tp^{k-1})^2 + \cdots + \frac{f^{(n)}(r)}{n!}(tp^{k-1})^n,$$

with $f^{(i)}(r)/i!$ an integer for $1 \leq i \leq n$. But when we reduce this modulo $p^k$, we see that all but the first two terms will disappear, since $i(k-1) \geq k$ for $i \geq 2$, and thus $p^k | p^{i(k-1)}$. So we have

$$f(r + tp^{k-1}) \equiv f(r) + f'(r)tp^{k-1} \mod p^k.$$

But since $r + tp^{k-1}$ is a solution to $f(x) \equiv 0 \mod p^k$, we thus know that

$$f'(r)tp^{k-1} \equiv -f(r) \mod p^k.$$

Since by hypothesis $f(r) \equiv 0 \mod p^{k-1}$, we know that $p^{k-1} | f(r)$ (considered as integers). So we can cancel out the $p^{k-1}$ on both sides of the congruence by the cancellation law, and get

$$f'(r)t \equiv -f(r)/p^{k-1} \mod p. \tag{1}$$

So far we've shown that if $f(r) \equiv 0 \mod p^{k-1}$, and $r + tp^{k-1}$ is a lift of $r$ to a solution modulo $p^k$ so that $f(r + tp^{k-1}) \equiv 0 \mod p^k$, then $t$ must satisfy this (linear!) congruence in equation (1). We know proceed to analyze it in the three cases given in the lemma.

1. Suppose $f'(r) \not\equiv 0 \mod p$. Then $f'(r), p) = 1$, there is a unique $t$ modulo $p$ satisfying equation (1), given by

$$t \equiv (-f(r)/p^{k-1})(f'(r))^{-1} \mod p.$$

2. Suppose $f'(r) \equiv 0 \mod p$, so that $(f'(r), p) = p$. Suppose further that $f(r) \equiv 0 \mod p^k$, implying that $p | f(r)/p^{k-1}$. Again by our results on linear congruences, if $p | f(r)/p^{k-1}$ then the equation (1) has $p$ solutions, and thus all values of $t$ are solutions.

3. Finally, suppose $f'(r) \equiv 0 \mod p$, so that $(f'(r), p) = p$, but that $f(r) \not\equiv 0 \mod p^k$ so that $p \nmid f(r)/p^{k-1}$. Then equation (1) has no solutions.

□

**Example 3.46.** Find the solutions of $f(x) = x^3 + x^2 + 29 \equiv 0 \mod 25$.

We first find the solutions $\mod 5$. By plugging in values we see the only solution is $x \equiv 3 \mod 5$. We compute $f'(3) = 27 + 6 = 33 \equiv 3 \not\equiv 0 \mod 5$, and thus Hensel's lemma tells us there is a unique solution $\mod 25$ given by $x_2 \equiv 3 + 5t$ where

$$t \equiv -(f'(3))^{-1}(f(3)/5) \equiv -3^{-1}(65/5) \equiv -2 \cdot 13 \equiv -26 \equiv 4 \mod 5$$

and thus $x_2 \equiv 23 \mod 25$ is a unique solution $\mod 25$.

**Example 3.47.** Find the solutions of $x^2 + x + 7 \equiv 0 \mod 27$.

Let $x(x) = x^2 + x + 7$. We check for solutions modulo 3 and find the only one is when $x \equiv 1 \mod 3$. We compute $f'(1) = 3 \equiv 0 \mod 3$. Thus $f(x)$ will have either 3 or 0 solutions modulo 9; we see that $f(1) = 9 \equiv 0 \mod 9$ and thus $1 + 3t$ is a solution modulo 9 for all integers $t$; thus the solutions modulo 9 are $1, 4, 7$.

We may try to lift again to 27, but we still have $f'(1) = 3 \equiv 0 \mod 3$, so each solution either lifts to three solutions or does not lift at all. $f(1) = 9 \not\equiv 0 \mod 27$ so $1 + 9t$ is not a solution modulo 27 for any intger $t$. $f(4) = 27 \equiv 0 \mod 27$ so $4 + 9t$ is a solution modulo 27 for any integer $t$. $f(7) = 63 \not\equiv 0 \mod 27$ so $7 + 9t$ is not a solution modulo 27 for any integer $t$.

Thus the solutions modulo 27 are $x \equiv 4, 13, 22$.

**Example 3.48.** Find solutions of $x^3 + x^2 + 23 \equiv 0 \mod 125$.

We first find solutions modulo 5. We observe that the solutions are all congruent to 1 or 2 modulo 5. We calculate $f'(x) = 3x^2 + 2x$ so $f'(1) = 3 + 2 \equiv 0 \mod 5$ and $f'(2) = 12 + 4 \equiv 1 \mod 5$; we have to handle these two cases separately.

First let's try to lift 1. We see that $f'(1) \equiv 0 \mod 5$ so either $f(1) \equiv 0 \mod 25$ or 1 has no lifts to roots $\mod 25$. In fact we see that $f(1) = 1 + 1 + 23 \equiv 0 \mod 25$, so every possible lift of 1 is a root modulo 25. Thus $1, 6, 11, 16, 21$ are all solutions to $f(x) \equiv 0 \mod 25$.

Now we need to test if each of these lifts to a root modulo 125. We know the derivatives are all equivalent to 0 modulo 5, so each solution modulo 25 has either 5 lifts or zero. We compute

$$\begin{aligned} f(1) &= 25 \equiv 25 \mod 125 \\ f(6) &= 275 \equiv 25 \mod 125 \\ f(11) &= 1475 \equiv 100 \mod 125 \\ f(16) &= 4375 \equiv 0 \mod 125 \\ f(21) &= 9725 \equiv 100 \mod 125 \end{aligned}$$

Thus 1, 6, 11, and 21 all lack lifts, but 16 has a lift and thus has five lifts. So $16, 41, 66, 91, 116$ are all solutions of $f(x) \equiv 0 \mod 125$.

Now let's return to considering the root 2. We have $f'(2) = 12 + 4 = 16 \equiv 1 \mod 5$ which has 1 as an inverse modulo 5. Thus there is a unique root modulo 25, and by Hensel's lemma we have a lift $r_2 \equiv 2 + t \cdot 5 \mod 25$ with

$$t \equiv -(f'(2))^{-1} \left( \frac{f(r)}{5} \right) \equiv -1 \cdot \frac{35}{5} \equiv -2 \equiv 3 \mod 5$$

and thus our lift $r_2 \equiv 2 + 3 \cdot 5 = 17 \mod 25$ is a root of $f(x) \equiv 0 \mod 25$. Now we wish to lift this again, but our derivative modulo 5 is still 1, so by Hensel's lemma we have a unique root modulo 125 given by $r_3 \equiv 17 + 25t \mod 125$ with

$$t \equiv -(f'(17))^{-1} \frac{f(17)}{25} \equiv -1 \cdot \frac{5225}{25} \equiv -249 \equiv 1 \mod 5$$

so $r_3 \equiv 17 + 25 \equiv 42 \mod 125$.

Thus the complete set of solutions to $f(x) \equiv 0 \mod 125$ is the set $\{x : x \equiv 16, 41, 42, 66, 91, 116 \mod 125\}$.

**Corollary 3.49.** *Suppose $f(x)$ is a polynomial and $r$ is a solution to the polynomial congruence $f(x) \equiv 0 \mod p$ for a prime number $p$. If $f'(r) \not\equiv 0 \mod p$, then for each integer $k \geq 1$ there is a solution $r_k$ to the congruence $f(x) \equiv 0 \mod p^k$, such that $r_k \equiv r \mod p$. Further, this $r_k$ is unique modulo $p^k$. In particular, $r_1 = 1$ and for $k > 1$ we have*

$$r_k = r_{k-1} - f(r_{k-1})(f'(r))^{-1}.$$

*Proof.* We prove this by induction. For $k = 1$ it is given that there is a unique solution that is equivalent to $r \mod p$, and $f'(r) \not\equiv 0 \mod p$.

Suppose we have proven that this property for $n$–that is there is a $r_n$, unique modulo $p^n$, such that $r_n \equiv r \mod p$ and $f(r_n) \equiv 0 \mod p^n$, and further $f'(r_n) \equiv f'(r) \not\equiv 0 \mod p$.

Then by Hensel's lemma, since $f'(r_n) \not\equiv 0 \mod p$, there is a unique integer $t$ with $0 \leq t < p$ such that $f(r_n + tp^n) \equiv - \mod p^{n+1}$, and $t \equiv -(f'(r_n))^{-1}(f(r_n)/p^{kn}) \mod p$.

Then there is a unique solution to $f(x) \equiv 0 \mod p^{n+1}$ that is equivalent to $r_n \mod p^n$, given by

$$r_{n+1} = r_n - (f'(r_n))^{-1} f(r_n).$$

Since $r_n \equiv r \mod p$ and $r_{n+1} \equiv r_n \mod p^n$ we know that $r_{n+1} \equiv r \mod p$.

Finally, we see that

$$\begin{aligned} f'(r_{n+1}) &= f'(r_n - (f'(r_n))^{-1} f(r_n)) \equiv f'(r_n - (f'(r_n))^{-1} \cdot 0) \mod p \\ &\equiv f'(r_n) \equiv f'(r) \not\equiv 0 \mod p. \end{aligned}$$

□

**Example 3.50.** Find the solutions of

$$x^3 + x^2 + 2x + 26 \equiv 0 \mod 7^k.$$

We first solve $x^3 + x^2 + 2x + 26 \equiv 0 \mod 7$, and see that the only solution is $x \equiv 2 \mod 7$. We see that $f'(x) = 3x^2 + 2x + 2$ so $f'(2) = 12 + 4 + 2 = 18 \equiv 4 \mod 7 \not\equiv 0 \mod 7$. Thus by the corollary, we can find solutions modulo $7^k$ for $k \in \mathbb{N}$.

We compute $(f'(2))^{-1} \equiv 4^{-1} \equiv 2 \mod 7$. Thus we have

$$\begin{aligned} r_2 &\equiv 2 - f(2)(f'(2))^{-1} \equiv 2 - 42 \cdot 2 = -82 \equiv 16 \mod 49 \\ r_3 &\equiv 16 - f(16)(f'(2))^{-1} = 16 - 4410 \cdot 2 = -8804 \equiv 114 \mod 343 \\ r_4 &\equiv 114 - f(114)(f'(2))^{-1} = 114 - 1494794 \cdot 2 = -2989474 \equiv 2172 \mod 2401 \quad \vdots \end{aligned}$$