$$\begin{aligned} &\Rightarrow \quad n \text{ divides } (a_1 - a_2) + (b_1 - b_2) \\ &\Rightarrow \quad n \text{ divides } (a_1 + b_1) - (a_2 + b_2) \\ &\Rightarrow \quad a_1 + b_1 \equiv a_2 + b_2 \pmod{n}. \end{aligned}$$

The second is the same, except for suitable replacement of + signs by − signs.

For the third, the expression we want n to divide, namely, $a_1b_1 - a_2b_2$, must be written in terms of the expressions we know are divisible by n, namely $a_1 - a_2$ and $b_1 - b_2$. Some experimentation gives

$$a_1b_1 - a_2b_2 = a_1(b_1 - b_2) + b_2(a_1 - a_2),$$

so a proof of the third congruence is:

$$\begin{aligned} &a_1 \equiv b_1 \pmod{n} \text{ and } a_2 \equiv b_2 \pmod{n} \\ \Rightarrow \quad &n \text{ divides } a_1 - a_2 \text{ and } n \text{ divides } b_1 - b_2 \\ \Rightarrow \quad &n \text{ divides } a_1(b_1 - b_2) + b_2(a_1 - a_2) \\ \Rightarrow \quad &n \text{ divides } a_1b_1 - a_2b_2 \\ \Rightarrow \quad &a_1b_1 \equiv a_2b_2 \pmod{n}. \end{aligned}$$

□

The arithmetic equivalence of congruent numbers means that some common manipulations with equations are also valid for congruences.

1. We can add them: if $a_1 \equiv a_2 \pmod{n}$ and $b_1 \equiv b_2 \pmod{n}$ then $a_1 + b_1 \equiv a_2 + b_2 \pmod{n}$.
2. We can subtract them: if $a_1 \equiv a_2 \pmod{n}$ and $b_1 \equiv b_2 \pmod{n}$ then $a_1 - b_1 \equiv a_2 - b_2 \pmod{n}$.
3. We can multiply them: if $a_1 \equiv a_2 \pmod{n}$ and $b_1 \equiv b_2 \pmod{n}$ then $a_1b_1 \equiv a_2b_2 \pmod{n}$.

Exercises

An old rule of arithmetic, called *casting out nines*, has a nice explanation in terms of arithmetic mod 9. The rule says that a number is divisible by 9 if the sum of its (base 10) digits is divisible by 9. For example, 774 is divisible by 9 because $7 + 7 + 4 = 18$ is divisible by 9.

6.2.1. AM radio frequencies in Melbourne are 621, 693, 774, 855, 927, 1116, 1224, 1278, 1377, 1422, 1503, 1593 (kHz). What do you notice about these numbers?

Casting out nines is easily understood when one recalls how base 10 numerals are built from their digits and powers of 10.

6.2.2. A base 10 numeral $a_k a_{k-1} \ldots a_1 a_0$ stands for $a_k 10^k + a_{k-1} 10^{k-1} + \cdots + a_1 10 + a_0$. Explain.

6.2.3. Notice that $10 \equiv 1 \pmod 9$, and hence

$$10 \times 10 \equiv 1 \times 1 \pmod 9, \quad \ldots, \quad 10^k \equiv 1^k \pmod 9.$$

Deduce from Exercise 6.2.2 that

$$a_k a_{k-1} \ldots a_1 a_0 \equiv a_k + a_{k-1} + \cdots + a_1 + a_0 \pmod 9.$$

Thus $a_k a_{k-1} \ldots a_1 a_0$ is divisible by 9 if and only if $a_k + a_{k-1} + \cdots + a_1 + a_0$ is; in fact they both have the same remainder on division by 9.

6.2.4. Notice also that $10 \equiv 1 \pmod 3$, and hence show similarly that

$$a_k a_{k-1} \ldots a_1 a_0 \equiv a_k + a_{k-1} + \cdots + a_1 + a_0 \pmod 3.$$

This gives a test for divisibility by 3 by "casting out threes." The next simplest is a test for divisibility by 11. Here again one needs to sum the digits, but now *taken with alternate + and − signs*. For example, 11 divides 16577, because 11 divides $1 - 6 + 5 - 7 + 7 = 0$.

6.2.5. Use the fact that $10 \equiv -1 \pmod{11}$ to show that

$$a_k a_{k-1} \ldots a_1 a_0 \equiv (-1)^k a_k + (-1)^{k-1} a_{k-1} + \cdots - a_1 + a_0 \pmod{11}.$$

Of course it is easy to find powers of 10 when 10 is congruent to 1 or -1. But in fact powers of any number are quite easy to evaluate, modulo any n. We need work only with remainders, and therefore we only need to multiply numbers $< n$. Also, large powers can be reached quickly by squaring wherever possible.

For example, to find large powers of 3, mod 19, we evaluate $3^2, 3^4, 3^8, 3^{16}, \ldots$ in succession by repeatedly finding the remainder on division by 19 and squaring it. The first step that involves a genuine remainder is where $3^4 = 81 \equiv 5 \pmod{19}$, and therefore $3^8 \equiv 5^2 \equiv 25 \equiv 6 \pmod{19}$. When enough powers $3^2, 3^4, 3^8, 3^{16}, \ldots$ have been found, other powers can be found as products of them. For example, $3^{100} = 3^{64} 3^{32} 3^4$ because $100 = 64 + 32 + 4$.

6.2.6. Find 3^{16}, 3^{32}, and 3^{64} mod 19, and show $3^{100} \equiv 16 \pmod{19}$.

6.2.7.* The method of repeated squaring depends on the fact that every natural number is a sum of powers of 2. Explain the dependence, and explain why the fact is true.
(*Hint*: Subtract the largest power of 2 and use descent.)

6.3 The Ring $\mathbb{Z}/n\mathbb{Z}$

The numbers congruent to a given integer a modulo n form the set

$$\{a + nk : k \in \mathbb{Z}\} = \{\ldots, a - 2n, a - n, a, a + n, a + 2n, \ldots\}.$$

We call this set the *congruence class* of a, mod n, and denote it by $a + n\mathbb{Z}$ for short. In particular, the set of all multiples of n,

$$n\mathbb{Z} = \{nk : k \in \mathbb{Z}\} = \{\ldots, -2n, -n, 0, n, 2n, \ldots\},$$

is the congruence class of 0. There are n different congruence classes, one for each remainder on division by n.

For example, there are two congruence classes mod 2:

$$2\mathbb{Z} = \{2k : k \in \mathbb{Z}\} = \{\text{even integers}\}$$

and

$$1 + 2\mathbb{Z} = \{1 + 2k : k \in \mathbb{Z}\} = \{\text{odd integers}\}.$$

We can now give a precise meaning to equations such as "odd + even = odd," "odd − even = odd" and "odd × even = even" by defining the sum, difference, and product of congruence classes.

For any modulus, the definitions say that

$$\begin{aligned}(\text{class of } a) + (\text{class of } b) &= \text{ class of } (a + b)\\ (\text{class of } a) - (\text{class of } b) &= \text{ class of } (a - b)\\ (\text{class of } a) \times (\text{class of } b) &= \text{ class of } ab\end{aligned}$$

or, in the notation we have just introduced,

$$\begin{aligned}(a + n\mathbb{Z}) + (b + n\mathbb{Z}) &= (a + b) + n\mathbb{Z}\\ (a + n\mathbb{Z}) - (b + n\mathbb{Z}) &= (a - b) + n\mathbb{Z}\\ (a + n\mathbb{Z}) \times (b + n\mathbb{Z}) &= ab + n\mathbb{Z}.\end{aligned}$$

The first of these should be compared with the very similar definition of the sum of angles in Section 5.2. There, and here, the class of a plus the class of b is the class of $a + b$.

To be careful, we should check that the class of $a+b$ is *well defined*, that it does not depend on the numbers we choose to represent the class of a and the class of b. Suppose on one occasion we take a_1 from the first class and b_1 from the second and we form the class of $a_1 + b_1$. If, on another occasion we take a_2 from the first class and b_2 from the second and we form the class of $a_2 + b_2$, is this the same as the class of $a_1 + b_1$? Yes! In fact, this is precisely the first property of arithmetic equivalence, proved in the previous section: if $a_1 \equiv a_2 \pmod{n}$ and $b_1 \equiv b_2 \pmod{n}$, then $a_1 + b_1 \equiv a_2 + b_2 \pmod{n}$.

Similarly, the difference and product of congruence classes are well defined by the second and third properties of arithmetic equivalence.

We use the symbols $+$, $-$, and $\times$ (or juxtaposition) for sum, difference, and product of congruence classes because they have the same properties as ordinary $+$, $-$, and $\times$. In fact, all the ring properties of $+$, $-$, and $\times$ on $\mathbb{Z}$ (Section 1.4) are "inherited" by the operations on congruence classes.

Here is how the $+$ on congruence classes inherits commutativity from ordinary $+$ on $\mathbb{Z}$:

$$\begin{aligned}(a + n\mathbb{Z}) + (b + n\mathbb{Z}) &= (a + b) + n\mathbb{Z} \\ &\quad \text{by definition of } + \text{ for congruence classes} \\ &= (b + a) + n\mathbb{Z} \\ &\quad \text{by commutativity of } + \text{ for } \mathbb{Z} \\ &= (b + n\mathbb{Z}) + (a + n\mathbb{Z}) \\ &\quad \text{by definition of } + \text{ for congruence classes.}\end{aligned}$$

It is equally easy to check that all the ring properties of $\mathbb{Z}$ are inherited by congruence classes; hence *the set of congruence classes* mod n *is a ring*. We denote this ring by $\mathbb{Z}/n\mathbb{Z}$. Informally speaking, $\mathbb{Z}/n\mathbb{Z}$ is what $\mathbb{Z}$ becomes when we pretend that $n = 0$.

Putting it a little more formally, $\mathbb{Z}/n\mathbb{Z}$ is what $\mathbb{Z}$ looks like when we focus on remainders mod n. Arithmetic equivalence mod n allows us to ignore all multiples of n and consistently replace each integer

by its remainder. As we saw at the beginning of this chapter, this is often the way to avoid being confused by the irrelevant vastness of $\mathbb{Z}$.

Exercises

The question whether an operation is well defined actually arises in around fifth grade, though you were probably not asked to worry about it then. The product of fractions is defined by

$$\frac{a}{b} \times \frac{c}{d} = \frac{ac}{bd},$$

but what we really want is the product of rational numbers, and a rational number is an infinite set of fractions. For example, what we call the "rational number $\frac{1}{2}$" is really the set

$$\left\{\frac{1}{2}, \frac{-1}{-2}, \frac{2}{4}, \frac{-2}{-4}, \frac{3}{6}, \frac{-3}{-6}, \dots\right\}$$

of fractions $k/2k$ for all nonzero integers k. To make sure the product of rationals is well defined, we have to check that the fractions ka/kb and lc/ld give the same product rational for any nonzero integers k and l, and of course they do.

6.3.1. Check that the sum of fractions $\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd}$ is well defined for rationals.

6.3.2. Also check that $\frac{a+c}{b+d}$ is *not* a well-defined "sum" of $\frac{a}{b}$ and $\frac{c}{d}$.

Inheritance of ring properties from $\mathbb{Z}$ implies that if an equation involving $+$, $-$, and $\times$ has a solution $x, y, \dots$ in $\mathbb{Z}$ then it has a solution in any $\mathbb{Z}/n\mathbb{Z}$. In fact, the solution in $\mathbb{Z}/n\mathbb{Z}$ comes from replacing $x, y, \dots$ by their congruence classes mod n. This sometimes enables us to prove that an equation has no solution in $\mathbb{Z}$ by showing that it has no solution in a suitably chosen $\mathbb{Z}/n\mathbb{Z}$. This is the basis of the idea, described in Section 6.1, that certain things are impossible in the integers merely because they are impossible in a certain set of remainders.

For example, if the equation $x^2 + y^2 = 4n + 3$ has a solution in $\mathbb{Z}$ then it has a solution in $\mathbb{Z}/4\mathbb{Z}$, where it becomes $x^2 + y^2 = 3$, because $4n + 3 \equiv 3 \pmod 4$. But we can see whether the equation $x^2 + y^2 = 3$ has

any solutions in $\mathbb{Z}/4\mathbb{Z}$ simply by trying the four possible values 0, 1, 2, and 3 for x and y.

6.3.3. Try this, and compare what happens with your previous solution (to Exercise 6.1.5) in terms of remainders.

There are some similar theorems about numbers of the form $x^2 + 2y^2$ and $x^2 + 3y^2$, and their possible remainders on division by 8 and 3, respectively.

6.3.4. Show that the equation $x^2 + 2y^2 = 8n + 5$ has no integer solution by considering it in $\mathbb{Z}/8\mathbb{Z}$. Discuss what happens with other numbers in place of 5.

6.3.5. Show that the equation $x^2 + 3y^2 = 3n + 2$ has no integer solution.

6.4 Inverses mod n

We have not said anything about division mod n so far, with good reason: it doesn't always work. In particular, if

$$ab \equiv ac \pmod{n} \quad \text{and} \quad a \not\equiv 0 \pmod{n}$$

it is not necessarily true that $b \equiv c \pmod{n}$. An example is

$$2 \times 1 \equiv 2 \times 3 \equiv 2 \pmod{4} \quad \text{but} \quad 1 \not\equiv 3 \pmod{4}.$$

The reason that division by 2 does not work, mod 4, is that 2 does not have an *inverse* mod 4. There is no number m such that $2m \equiv 1 \pmod{4}$, as can be seen by trying $m = 1, 2, 3$. The numbers 1 and 3 do have inverses mod 4. In fact each is its own inverse: $1 \times 1 \equiv 1 \pmod{4}$ and $3 \times 3 = 9 \equiv 1 \pmod{4}$. This means that division by 1 and 3 are valid mod 4. For example, from

$$3 \times b \equiv 3 \times c \pmod{4}$$

we can conclude that

$$3 \times 3 \times b \equiv 3 \times 3 \times c \pmod{4},$$

multiplying both sides by 3. This is the same as

$$b \equiv c \pmod 4$$

because $3 \times 3 \equiv 1 \pmod 4$.

In general, division by a in mod n arithmetic is possible precisely when a has an *inverse* mod n, a number m such that $am \equiv 1 \pmod n$. It is therefore a question of knowing which numbers have inverses mod n, and this question has a very neat answer.

Criterion for inverses mod n. *The number a has an inverse mod n if and only if* $\gcd(a, n) = 1$.

Proof To show this proof concisely, we use the symbol $\Leftrightarrow$ for "if and only if."

$$\begin{aligned} a \text{ has an inverse mod } n &\Leftrightarrow am \equiv 1 \pmod n \text{ for some integer } m \\ &\Leftrightarrow n \text{ divides } am - 1 \\ &\Leftrightarrow am + nl = 1 \text{ for some integers } l \text{ and } m \\ &\Leftrightarrow \gcd(a, n) = 1. \end{aligned}$$

The last $\Leftrightarrow$ follows from the results about the gcd in Section 1.5.

$$am + nl = 1 \Rightarrow \gcd(a, n) = 1$$

because any divisor of a and n divides the left hand side, and hence divides 1;

$$\gcd(a, n) = 1 \Rightarrow am + nl = 1 \text{ for some integers } l \text{ and } m$$

because $\gcd(a, n) = am + nl$ for integers l and m. □

This criterion gives a more general explanation why 2 has no inverse in arithmetic mod 4. It cannot have an inverse because $\gcd(2,4) = 2$. We also see that nonzero numbers without inverses will occur for any nonprime modulus n.

But if we have a *prime* modulus p the condition $a \not\equiv 0 \pmod p$ means $\gcd(a, p) = 1$, because the only numbers having a larger common divisor with p are the multiples of p, that is, the numbers $\equiv 0 \pmod p$. Translating this into the language of congruence classes, as in the previous section, we find: *in the ring $\mathbb{Z}/p\mathbb{Z}$, for prime p, every nonzero element has an inverse.* Now recall from Section 1.4 that a ring in which every nonzero element has an inverse is a field, and we can conclude that *$\mathbb{Z}/p\mathbb{Z}$ is a field.*

This means that some familiar arguments about numbers can also be applied to $\mathbb{Z}/p\mathbb{Z}$. For example, in ordinary algebra we have a theorem that at most n different numbers can satisfy a polynomial equation of degree n (Exercise 5.2.3). Applying the same argument to $\mathbb{Z}/p\mathbb{Z}$ gives the following.

Lagrange's polynomial theorem. *If $P(x)$ is a polynomial of degree n, then the congruence $P(x) \equiv 0 \pmod{p}$ has at most n solutions mod p.*

Proof Suppose $P(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x_1 + a_0$. Because

$$x^k - a^k = (x - a)(x^{k-1} + ax^{k-2} + \cdots + a^{k-2}x + a^{k-1})$$

(as can be checked by multiplying out the right-hand side), it follows that $(x - a)$ is a factor of

$$P(x) - P(a) = a_n(x^n - a^n) + a_{n-1}(x^{n-1} - a^{n-1}) + \cdots + a_1(x - a).$$

Thus

$P(x)-P(a) = (x-a)Q(x)$ for some polynomial $Q(x)$ of degree $n - 1$.

Then if $P(a) \equiv 0 \pmod{p}$ we have

$$P(x) \equiv (x - a)Q(x) \pmod{p}.$$

If also $P(b) \equiv 0 \pmod{p}$ for some $b \not\equiv a \pmod{p}$, we have

$$0 \equiv P(b) \equiv (b - a)Q(b) \pmod{p}.$$

Multiplying both sides of this by the inverse of $b - a \bmod p$ gives

$$Q(b) \equiv 0 \pmod{p},$$

which similarly implies

$$Q(x) \equiv (x - b)R(x) \pmod{p} \quad \text{for some polynomial } R(x) \text{ of degree } n - 2.$$

Thus the degree of the quotient polynomial falls by 1 for each distinct solution of $P(x) \equiv 0 \pmod{p}$, and hence the latter congruence has at most n different solutions $\pmod{p}$. □

Exercises

The proof of the criterion for inverses shows them connected to the gcd via the fact that $\gcd(a, n) = ma + ln$ for integers l and m. We know from Section 1.5 that an efficient way to find such integers is by the Euclidean algorithm. Thus the Euclidean algorithm is also ideal for finding inverses mod n. To find an inverse m of a, mod n, find $\gcd(a, n)$ as an explicit linear combination of a and n, and the answer $ma + nl$ contains the desired inverse m.

6.4.1. Find an inverse of 13, mod 31, by this method.

The efficiency of the Euclidean algorithm makes it feasible to find inverses for very large numbers and moduli, say, with hundreds of digits. A more difficult computational problem is finding *how many* numbers have inverses, for a given modulus n. By the criterion for inverses, the problem is to find how many of the numbers $1, 2, 3, \ldots, n-1$ have gcd 1 with n. The number of them is denoted by $\varphi(n)$, and φ is called the *Euler phi function*. The problem of computing $\varphi(n)$ is at least as hard as recognizing whether n is prime, for the following simple reason.

6.4.2. Show that $\varphi(n) = n - 1$ if and only if n is prime.

If n is known to be prime or a prime power, then $\varphi(n)$ is easier to compute.

6.4.3. If p is prime, show that the number of multiples of p among $1, 2, 3, \ldots, p^k - 1$ is $p^{k-1} - 1$, and deduce that $\varphi(p^k) = p^{k-1}(p-1)$.

6.4.4. Check this formula by finding $\varphi(27)$ from first principles.

Finally, if the full prime factorization of n is known, $\varphi(n)$ can be computed using the fact that $\varphi(rs) = \varphi(r)\varphi(s)$ when $\gcd(r, s) = 1$. This fact can be proved by elementary methods, but it falls more naturally out of the Chinese remainder theorem, which will be discussed in Section 6.6.

So far we have been saying *an* inverse, but in $\mathbb{Z}/n\mathbb{Z}$ we can say *the* inverse, because a number with an inverse has only one congruence class of them.

6.4.5. If $am_1 \equiv 1 \pmod{n}$ and $am_2 \equiv 1 \pmod{n}$ show that $m_1 \equiv m_2 \pmod{n}$.

Another important theorem about polynomials mod p is the "mod p binomial theorem" $(1 + x)^p \equiv 1 + x^p \pmod{p}$. First recall the ordinary binomial theorem.