6.4.6. Show by induction on n that

$$(1+x)^n = 1 + \binom{n}{1}x + \cdots + \binom{n}{n-1}x^{n-1} + x^n,$$

where $\binom{n}{j} = \frac{n(n-1)\cdots(n-j+1)}{j(j-1)\cdots 2\cdot 1}$ is the number of ways of choosing j things from n things.

6.4.7. Show that p divides $\binom{n}{j}$ for $1 \le j \le n-1$, and hence conclude that $(1+x)^p \equiv 1 + x^p \pmod{p}$.

6.5 The Theorems of Fermat and Wilson

Suppose a is any positive integer and we form the sequence of its powers: $a, a^2, a^3, \ldots$. If we reduce these powers to their values mod p, then some value must eventually repeat, because there are only p different values available. Trials with actual values of a and p suggest that the sequence of powers $a^m \bmod p$ is actually periodic, and that it always includes the number 1.

For example, the sequence of powers of 2, mod 5, is

$$2, 4, 3, 1, 2, 4, 3, 1, 2, 4, 3, 1, \ldots,$$

which strongly suggests that the sequence has period 2, 4, 3, 1. Indeed it must, because the first 1 shows that $2^4 \equiv 1 \pmod 5$, in which case $2^5 \equiv 2^1$, $2^6 \equiv 2^2$, $2^7 \equiv 2^3$, and so on, (mod 5).

It is clear from this example that the behavior of powers in arithmetic mod p depends on whether there is a power congruent to 1. Fermat's little theorem tells us that such a power always occurs. It is called Fermat's "little" theorem to distinguish it from the much more difficult "Fermat's last theorem." However, it deserves a place of its own, for both its elegance and historical importance.

Fermat's little theorem. *If p is prime and $\gcd(a,p) = 1$, then $a^{p-1} \equiv 1 \pmod p$.*

Proof The condition $\gcd(a,p) = 1$ says that a has an inverse mod p, by the criterion for inverses. The numbers $1, 2, \ldots, p-1$ also have inverses mod p, because p is prime.

Now consider the remainders of $a, 2a, \ldots, (p-1)a$ on division by p:

$$a \bmod p, \quad 2a \bmod p, \quad \ldots, \quad (p-1)a \bmod p.$$

These remainders are the numbers $1, 2, \ldots, p-1$ again (in a different order), because they are nonzero and unequal (mod p): $ja \equiv ka \pmod p$ implies $j \equiv k \pmod p$, multiplying both sides by the inverse of a. It follows that

$$a \times 2a \times \cdots \times (p-1)a \equiv 1 \times 2 \times \cdots \times (p-1) \pmod p,$$

that is,

$$a^{p-1} \times 1 \times 2 \times \cdots \times (p-1) \equiv 1 \times 2 \times \cdots \times (p-1) \pmod p,$$

and therefore

$$a^{p-1} \equiv 1 \pmod p,$$

multiplying both sides by the inverses of $1, 2, \ldots, p-1$. □

Fermat's little theorem is proved by equating two different expressions for $1 \times 2 \times 3 \times \cdots \times (p-1)$. The actual value of this product, mod p, can be found by pairing factors with their inverses. The result is known as Wilson's theorem, and the following proof was given by Gauss (1801).

Wilson's theorem. *If p is prime, then*

$$1 \times 2 \times 3 \times \cdots \times (p-1) \equiv -1 \pmod p.$$

Proof Before pairing factors with their inverses, we have to weed out the factors that are inverse to themselves. One such factor is obviously 1, and another is -1 (which is $p-1$, mod p). To see that these are the only self-inverse factors, mod p, we note that self-inverse numbers x satisfy the quadratic equation mod p:

$$x^2 \equiv 1 \pmod p.$$

By Lagrange's polynomial theorem (Section 6.4) this equation has at most two solutions; hence $x \equiv 1 \pmod p$ and $x \equiv -1 \pmod p$ are the only ones.

Thus the factors of $1 \times 2 \times 3 \times \cdots \times (p-1)$ include exactly two that are self-inverse, 1 and $p-1$. Canceling the remaining inverse

pairs leaves

$$1 \times 2 \times 3 \times \cdots \times (p-1) \equiv 1 \times (p-1) \equiv -1 \pmod{p},$$

as required. □

This theorem has a striking and unexpected corollary.

Wilson's primality criterion. *A natural number n is prime if and only if*

$$1 \times 2 \times 3 \times \cdots \times (n-1) \equiv -1 \pmod{n}.$$

Proof As we have just seen, if n is prime then $1 \times 2 \times 3 \times \cdots \times (n-1) \equiv -1 \pmod{n}$.

Conversely, if n is *not* prime then the numbers $2, 3, \ldots, n-1$ include a divisor d of n, and they also include n/d. But then $1 \times 2 \times 3 \times \cdots \times (n-1)$ is a multiple of n and hence

$$1 \times 2 \times 3 \times \cdots \times (n-1) \equiv 0 \not\equiv -1 \pmod{n}.$$

□

It is extremely surprising to find such a simply stated criterion for n to be prime, but unfortunately the criterion seems to have no practical value. When n is large enough for its primality to be worth asking about, it is also large enough to make $1 \times 2 \times 3 \times \cdots \times (n-1)$ impossible to compute.

Exercises

Fermat discovered his little theorem in around 1640. As mentioned in Section 1.6, he was looking for a way to find factors of numbers of the form $2^p - 1$. His theorem can detect such factors with surprising ease, if they exist.

6.5.1. Suppose a prime $q > p$ divides $2^p - 1$, so $2^p \equiv 1 \pmod{q}$. Show the following, in turn, using Fermat's little theorem for the third step:

- If $2^a \equiv 1 \pmod{q}$ and $2^b \equiv 1 \pmod{q}$ with $a > b$ then $2^{a-b} \equiv 1 \pmod{q}$.
- If $2^a \equiv 1 \pmod{q}$ and $2^b \equiv 1 \pmod{q}$ then $2^{\gcd(a,b)} \equiv 1 \pmod{q}$.

- If $2^p \equiv 1 \pmod{q}$ for a prime p, then p divides $q - 1$.
- q is of the form $kp + 1$ for some integer k.

The first number on Fermat's hit list was $2^{37} - 1$. According to Exercise 6.5.1, any prime divisor > 37 must be one of $37 + 1$, $2 \times 37 + 1$, $3 \times 37 + 1$, ... The first prime in this sequence is $6 \times 31 + 1 = 223$ and ... bingo!

6.5.2. Check that 223 divides $2^{37} - 1$.

If the prime q divides $2^n - 1$ and n is *not* prime, then n does not necessarily divide $q - 1$.

6.5.3. Find an n such that 31 divides $2^n - 1$ but n does not divide 30.

However, if m is the *least* positive exponent such that the prime q divides $2^m - 1$ it is true that m divides $q - 1$.

6.5.4. If m is the least positive exponent such that $2^m \equiv 1 \pmod{q}$, show that m divides any positive n such that $2^n \equiv 1 \pmod{q}$ (in particular, m divides the exponent $q - 1$ given by Fermat's little theorem).

This fact greatly shortens the search for divisors of the Fermat number $2^{2^5} + 1 = 2^{32} + 1$. Any prime divisor q of $2^{32} + 1$ also divides $(2^{32} + 1)(2^{32} - 1) = 2^{64} - 1$, and 64 is the least m such that q divides $2^m - 1$, for the following reason.

6.5.5. Show that the least positive m such that $2^m \equiv 1 \pmod{q}$ is a divisor of 64. Conclude, using the fact that $2^{32} \equiv -1 \pmod{q}$, that $m = 64$.

6.5.6. Deduce from Exercise 6.5.5 that any prime divisor of $2^{2^5} + 1$ is of the form $64k + 1$.

If Fermat had followed his own train of thought this far he would not have made the mistake of thinking that all the numbers $2^{2^n} + 1$ are prime. In fact, this is precisely how Euler discovered the divisor 641 of $2^{2^5} + 1$.

Wilson's theorem was first published without proof, in a book by Edward Waring in 1770. The first proof was given by Lagrange in 1771, and he also used it to find the primes p for which -1 is a square, mod p.

6.5.7. If $p = 2$, then -1 is certainly a square mod p. Why?

6.5.8. If $p = 4n + 3$, use congruences mod 4 to show that -1 is not a square mod p.

The most challenging case is where $p = 4n+1$, in which case Wilson's theorem is helpful, combined with the fact that $1 \times 2 \times \cdots \times (p-1) = 1 \times 2 \times \cdots \times 4n$.

6.5.9. Show that $1 \times 2 \times \cdots \times 4n \equiv (1 \times 2 \times \cdots \times 2n)^2 \pmod{4n+1}$, and hence conclude from Wilson's theorem that $-1 \equiv (1 \times 2 \times \cdots \times 2n)^2 \pmod{p}$ when $p = 4n+1$ is prime.

6.6 The Chinese Remainder Theorem

The behavior of numbers mod n is quite complicated when n is not prime. As we have seen, there are nonzero numbers without inverses, and finding all the numbers with inverses is tied up with the hard problem of computing $\varphi(n)$. Some relief from this situation is obtained by "factorizing" the ring $\mathbb{Z}/n\mathbb{Z}$ into smaller and simpler rings. The germ of this idea was discovered by Chinese mathematicians around 300 A.D., and various generalizations of it are now called the *Chinese remainder theorem*.

The theorem grows out of the discovery that a number can be known modulo lm if it is known modulo l and m. For example, the number 25 is completely determined, mod 77, by the two remainders $25 \bmod 7 = 4$ and $25 \bmod 11 = 3$. The reason no other number < 77 gives these remainders is that *all* 77 *pairs of remainders occur*, so there is exactly one pair for each of the numbers $0, 1, 2, 3, \ldots, 76$. This is proved by an algorithm that actually obtains the natural number x with a given pair of remainders. To do this, the Chinese used what they called the *method of finding 1*.

The method (in its basic form) assumes we have a relatively prime pair l and m, so that $\gcd(l, m) = 1$, and the Euclidean algorithm can be used to express 1 as a linear combination of l and m.

Example. To obtain x with $x \bmod 7 = 4$ and $x \bmod 11 = 2$.

- First express $1 = \gcd(11, 7)$ as a linear combination of 11 and 7, say,

$$1 = 2 \times 11 - 3 \times 7.$$

- Then express 4 and 2 as multiples of this combination:

$$4 = 8 \times 11 - 12 \times 7 \quad \text{and} \quad 2 = 4 \times 11 - 6 \times 7.$$

- This gives 4 and 2 as remainders on division by 7 and 11, respectively,

$$4 = 8 \times 11 \bmod 7 \quad \text{and} \quad 2 = -6 \times 7 \bmod 11.$$

- And the *sum* of these multiples of 11 and 7 has the same remainders,

$$4 = (8 \times 11 - 6 \times 7) \bmod 7 \quad \text{and} \quad 2 = (8 \times 11 - 6 \times 7) \bmod 11.$$

- Hence the solution is $x = (8 \times 11 - 6 \times 7) = 88 - 42 = 46$.

The traditional Chinese remainder theorem is about *determining* a number by a pair (or triple, quadruple, etc.) of smaller numbers. However, we can also add and multiply numbers by adding and multiplying the corresponding pairs. We want the sum of the pairs for x_1 and x_2 to be the pair for $x_1 + x_2$, so the rule for adding pairs is

$$(x_1 \bmod l, x_1 \bmod m) + (x_2 \bmod l, x_2 \bmod m) = (x_1 + x_2 \bmod l, x_1 + x_2 \bmod m),$$

and similarly the rule for multiplying pairs is

$$(x_1 \bmod l, x_1 \bmod m)(x_2 \bmod l, x_2 \bmod m) = (x_1x_2 \bmod l, x_1x_2 \bmod m).$$

The fully fledged Chinese remainder theorem includes this arithmetic of pairs by describing the ring $\mathbb{Z}/lm\mathbb{Z}$ as a "product" of the rings $\mathbb{Z}/l\mathbb{Z}$ and $\mathbb{Z}/m\mathbb{Z}$. It is called the *direct product* $\mathbb{Z}/l\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$ and its members are pairs of congruence classes $(x + l\mathbb{Z}, x + m\mathbb{Z})$, added and multiplied according to the rules

$$(x_1 + l\mathbb{Z}, x_1 + m\mathbb{Z}) + (x_2 + l\mathbb{Z}, x_2 + m\mathbb{Z}) = (x_1 + x_2 + l\mathbb{Z}, x_1 + x_2 + m\mathbb{Z}),$$
$$(x_1 + l\mathbb{Z}, x_1 + m\mathbb{Z})(x_2 + l\mathbb{Z}, x_2 + m\mathbb{Z}) = (x_1x_2 + l\mathbb{Z}, x_1x_2 + m\mathbb{Z}).$$

These rules are just a translation, into the language of congruence classes, of the rules just stated for pairs of remainders mod l and mod m.

$\mathbb{Z}/l\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$ is not strictly identical with $\mathbb{Z}/lm\mathbb{Z}$, because its members are pairs rather than single congruence classes, but it *be-*

haves the same in a sense that will be explained in the proof of the theorem. We say that $\mathbb{Z}/l\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$ is *isomorphic* to $\mathbb{Z}/lm\mathbb{Z}$ (from the Greek for "same form") and write

$$\mathbb{Z}/lm\mathbb{Z} \cong \mathbb{Z}/l\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}.$$

In practice, there is no harm in saying that the ring $\mathbb{Z}/lm\mathbb{Z}$ is the direct product $\mathbb{Z}/l\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$.

Chinese remainder theorem. *If* $\gcd(l, m) = 1$ *then*

$$\mathbb{Z}/lm\mathbb{Z} \cong \mathbb{Z}/l\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}.$$

Proof We begin by letting the congruence class $x + lm\mathbb{Z}$ correspond to the pair $(x+l\mathbb{Z}, x+m\mathbb{Z})$. Because there are lm classes $x+lm\mathbb{Z}$, also l classes $a+l\mathbb{Z}$ and m classes $b+m\mathbb{Z}$, the latter form lm pairs. Thus the correspondence will be one-to-one between $\mathbb{Z}/lm\mathbb{Z}$ and $\mathbb{Z}/l\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$ provided each pair $(a + l\mathbb{Z}, b + m\mathbb{Z})$ is $(x + l\mathbb{Z}, x + m\mathbb{Z})$ for some x.

It suffices to find integers u and v with $um \bmod l = a$ and $vl \bmod m = b$, because $x = um + vl$ will then give $x \bmod l = a$ and $x \bmod m = b$, as required. This is where we use the fact that $\gcd(l, m) = 1$. By Section 1.5, and the "method of finding 1,"

$$\begin{aligned} \gcd(l, m) = 1 &\Rightarrow 1 = rl + sm \text{ for some integers } r \text{ and } s, \\ &\Rightarrow a = arl + asm \quad \text{and} \quad b = brl + bsm, \\ &\Rightarrow asm = a - arl \quad \text{and} \quad brl = b - bsm, \\ &\Rightarrow asm \bmod l = a, \quad \text{and} \quad brl \bmod m = b. \end{aligned}$$

Hence suitable integers are $u = as$ and $v = br$, and $x = asm + brl$.

By definition of the sum of pairs, the pair in $\mathbb{Z}/l\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$ corresponding to a sum in $\mathbb{Z}/lm\mathbb{Z}$ is the sum of the corresponding pairs. Products similarly correspond to products, so we have a one-to-one correspondence between $\mathbb{Z}/lm\mathbb{Z}$ and $\mathbb{Z}/l\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$ that preserves sums and products. This is precisely what we mean by $\mathbb{Z}/lm\mathbb{Z} \cong \mathbb{Z}/l\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$. □

Exercises

It is helpful to work out the actual pairs for a small case of $\mathbb{Z}/lm\mathbb{Z}$, for example:

6.6.1. For each x in $\mathbb{Z}/15\mathbb{Z}$, work out the corresponding pair (x mod 3, x mod 5) in $\mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z}$. What do you notice about the pairs for invertible x?

The isomorphism between $\mathbb{Z}/l\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$ and $\mathbb{Z}/lm\mathbb{Z}$ looks like a formality once the one-to-one correspondence has been discovered, but the structure it gives to the set of pairs $(a+l\mathbb{Z}, b+m\mathbb{Z})$ is surprisingly helpful. For example, because classes $x+lm\mathbb{Z}$ behave the same as the corresponding pairs $(a+l\mathbb{Z}, b+m\mathbb{Z})$, it follows in particular that classes with inverses correspond to pairs with inverses.

6.6.2. Show that a pair $(a + l\mathbb{Z}, b + m\mathbb{Z})$ with an inverse corresponds in turn to an $a + l\mathbb{Z}$ with an inverse and a $b + m\mathbb{Z}$ with an inverse.

Now, by definition of the Euler φ function, there are $\varphi(l)$ such classes $a+l\mathbb{Z}$, and $\varphi(m)$ such classes $b+m\mathbb{Z}$. This gives the *multiplicative property* of φ.

6.6.3. Deduce that there are $\varphi(l)\varphi(m)$ invertible pairs $(a+l\mathbb{Z}, b+m\mathbb{Z})$ and hence conclude: if $\gcd(l, m) = 1$ then $\varphi(lm) = \varphi(l)\varphi(m)$.

As mentioned in the exercises to Section 6.4, this property of the Euler φ function enables us to compute $\varphi(n)$ when the prime factorization of n is known.

6.6.4. Show that if $n = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}$ is the prime factorization of n then

$$\varphi(n) = p_1^{e_1-1}(p_1 - 1)p_2^{e_2-1}(p_2 - 1)\cdots p_k^{e_k-1}(p_k - 1).$$

6.7 Squares mod p

In arithmetic mod p the analogs of linear and quadratic equations are linear and quadratic congruences, and they are solved in an analogous way—up to a point. Because the ordinary operations of arithmetic are valid mod p, we can solve the linear congruence

$$ax + b \equiv 0 \pmod{p}$$

by subtracting b from both sides, then multiplying both sides by the mod p inverse of a. If we write this inverse as $1/a$, then the solution looks the same as in ordinary algebra: $x = -b/a$. The difference, of