2.3 The RSA Cryptosystem

Exponentiation mod n

In the previous sections, we have considered encryption using modular addition and multiplication, and have seen the shortcomings of both. In this section, we will consider using exponentiation for encryption, and will show that it provides a much greater level of security.

The idea behind RSA encryption is exponentiation in Z_n . By Lemma 2.3, if $a \in Z_n$,

$$a^{j} \bmod n = \underbrace{a \cdot_{n} a \cdot_{n} \cdots \cdot_{n} a}_{j \text{ factors}}.$$
(2.10)

In other words $a^j \mod n$ is the product in Z_n of j factors, each equal to a.

The Rules of Exponents

Lemma 2.3 and the rules of exponents for the integers tell us that

Lemma 2.19 For any $a \in Z_n$, and any nonnegative integers i and j,

$$(a^{i} \bmod n) \cdot_{n} (a^{j} \bmod n) = a^{i+j} \bmod n$$

$$(2.11)$$

and

$$(a^i \mod n)^j \mod n = a^{ij} \mod n. \tag{2.12}$$

- **Exercise 2.3-1** Compute the powers of 2 mod 7. What do you observe? Now compute the powers of 3 mod 7. What do you observe?
- **Exercise 2.3-2** Compute the sixth powers of the nonzero elements of Z_7 . What do you observe?
- **Exercise 2.3-3** Compute the numbers $1 \cdot_7 2$, $2 \cdot_7 2$, $3 \cdot_7 2$, $4 \cdot_7 2$, $5 \cdot_7 2$, and $6 \cdot_7 2$. What do you observe? Now compute the numbers $1 \cdot_7 3$, $2 \cdot_7 3$, $3 \cdot_7 3$, $4 \cdot_7 3$, $5 \cdot_7 3$, and $6 \cdot_7 3$. What do you observe?
- **Exercise 2.3-4** Suppose we choose an arbitrary nonzero number *a* between 1 and 6. Are the numbers $1 \cdot_7 a$, $2 \cdot_7 a$, $3 \cdot_7 a$, $4 \cdot_7 a$, $5 \cdot_7 a$, and $6 \cdot_7 a$ all different? Why or why not?

In Exercise 2.3-1, we have that

$$2^{0} \mod 7 = 1$$

$$2^{1} \mod 7 = 2$$

$$2^{2} \mod 7 = 4$$

$$2^{3} \mod 7 = 1$$

$$2^{4} \mod 7 = 2$$

$$2^{5} \mod 7 = 4$$

$$2^{6} \mod 7 = 1$$

$$2^{7} \mod 7 = 2$$

$$2^{8} \mod 7 = 4.$$

Continuing, we see that the powers of 2 will cycle through the list of three values 1, 2, 4 again and again. Performing the same computation for 3, we have

 $3^{0} \mod 7 = 1$ $3^{1} \mod 7 = 3$ $3^{2} \mod 7 = 2$ $3^{3} \mod 7 = 6$ $3^{4} \mod 7 = 4$ $3^{5} \mod 7 = 5$ $3^{6} \mod 7 = 1$ $3^{7} \mod 7 = 3$ $3^{8} \mod 7 = 2$.

In this case, we will cycle through the list of six values 1, 3, 2, 6, 4, 5 again and again.

Now observe that in Z_7 , $2^6 = 1$ and $3^6 = 1$. This suggests an answer to Exercise 2.3-2. Is it the case that $a^6 \mod 7 = 1$ for all $a \in Z_7$? We can compute that $1^6 \mod 7 = 1$, and

$$4^{6} \mod 7 = (2 \cdot_{7} 2)^{6} \mod 7$$

= $(2^{6} \cdot_{7} 2^{6}) \mod 7$
= $(1 \cdot_{7} 1) \mod 7$
= 1.

What about 5⁶? Notice that $3^5 = 5$ in Z_7 by the computations we made above. Using Equation 2.12 twice, this gives us

$$5^{6} \mod 7 = (3^{5})^{6} \mod 7$$

= $3^{5 \cdot 6} \mod 7$
= $3^{6 \cdot 5} \mod 7$
= $(3^{6})^{5} = 1^{5} = 1$

in Z_7 . Finally, since $-1 \mod 7 = 6$, Lemma 2.3 tells us that $6^6 \mod 7 = (-1)^6 \mod 7 = 1$. Thus the sixth power of each element of Z_7 is 1.

In Exercise 2.3-3 we see that

$$1 \cdot_{7} 2 = 1 \cdot 2 \mod 7 = 2$$

$$2 \cdot_{7} 2 = 2 \cdot 2 \mod 7 = 4$$

$$3 \cdot_{7} 2 = 3 \cdot 2 \mod 7 = 6$$

$$4 \cdot_{7} 2 = 4 \cdot 2 \mod 7 = 1$$

$$5 \cdot_{7} 2 = 5 \cdot 2 \mod 7 = 3$$

$$6 \cdot_{7} 2 = 6 \cdot 2 \mod 7 = 5.$$

Thus these numbers are a permutation of the set $\{1, 2, 3, 4, 5, 6\}$. Similarly,

$$1 \cdot_7 3 = 1 \cdot 3 \mod 7 = 3$$

 $2 \cdot_7 3 = 2 \cdot 3 \mod 7 = 6$

Again we get a permutation of $\{1, 2, 3, 4, 5, 6\}$.

In Exercise 2.3-4 we are asked whether this is always the case. Notice that since 7 is a prime, by Corollary 2.17, each nonzero number between 1 and 6 has a mod 7 multiplicative inverse a^{-1} . Thus if i and j were integers in Z_7 with $i \cdot_7 a = j \cdot_7 a$, we multiply mod 7 on the right by a^{-1} to get

$$(i \cdot_7 a) \cdot_7 a^{-1} = (j \cdot_7 a) \cdot_7 a^{-1}.$$

After using the associative law we get

$$i_{7}(a_{7}a^{-1}) = j_{7}(a_{7}a^{-1}).$$
(2.13)

Since $a \cdot_7 a^{-1} = 1$, Equation 2.13 simply becomes i = j. Thus, we have shown that the only way for $i \cdot_7 a$ to equal $j \cdot_7 a$ is for i to equal j. Therefore, all the values $i \cdot_7 a$ for i = 1, 2, 3, 4, 5, 6 must be different. Since we have six different values, all between 1 and 6, we have that the values ia for i = 1, 2, 3, 4, 5, 6 are a permutation of $\{1, 2, 3, 4, 5, 6\}$.

As you can see, the only fact we used in our analysis of Exercise 2.3-4 is that if p is a prime, then any number between 1 and p-1 has a multiplicative inverse in Z_p . In other words, we have really proved the following lemma.

Lemma 2.20 Let p be a prime number. For any fixed nonzero number a in Z_p , the numbers $(1 \cdot a) \mod p$, $(2 \cdot a) \mod p$, \ldots , $((p-1) \cdot a) \mod p$, are a permutation of the set $\{1, 2, \cdots, p-1\}$.

With this lemma in hand, we can prove a famous theorem that explains the phenomenon we saw in Exercise 2.3-2.

Fermat's Little Theorem

Theorem 2.21 (Fermat's Little Theorem). Let p be a prime number. Then $a^{p-1} \mod p = 1$ in Z_p for each nonzero a in Z_p .

Proof: Since p is a prime, Lemma 2.20 tells us that the numbers $1 \cdot_p a, 2 \cdot_p a, \ldots, (p-1) \cdot_p a$, are a permutation of the set $\{1, 2, \cdots, p-1\}$. But then

$$1 \cdot_p 2 \cdot_p \cdots \cdot_p (p-1) = (1 \cdot_p a) \cdot_p (2 \cdot_p a) \cdot_p \cdots \cdot_p ((p-1) \cdot_p a).$$

Using the commutative and associative laws for multiplication in Z_p and Equation 2.10, we get

$$1 \cdot_p 2 \cdot_p \cdots \cdot_p (p-1) = 1 \cdot_p 2 \cdot_p \cdots \cdot_p (p-1) \cdot_p (a^{p-1} \mod p)$$

Now we multiply both sides of the equation by the multiplicative inverses in Z_p of 2, 3, ..., p-1 and the left hand side of our equation becomes 1 and the right hand side becomes $a^{p-1} \mod p$. But this is exactly the conclusion of our theorem. **Corollary 2.22** (Fermat's Little Theorem, version 2) For every positive integer a, and prime p, if a is not a multiple of p,

$$a^{p-1} \mod p = 1.$$

Proof: This is a direct application of Lemma 2.3, because if we replace a by $a \mod p$, then Theorem 2.21 applies.

The RSA Cryptosystem

Fermat's Little Theorem is at the heart of the RSA cryptosystem, a system that allows Bob to tell the world a way that they can encode a message to send to him so that he and only he can read it. In other words, even though he tells everyone how to encode the message, nobody except Bob has a significant chance of figuring out what the message is from looking at the encoded message. What Bob is giving out is called a "one-way function." This is a function f that has an inverse f^{-1} , but even though y = f(x) is reasonably easy to compute, nobody but Bob (who has some extra information that he keeps secret) can compute $f^{-1}(y)$. Thus when Alice wants to send a message x to Bob, she computes f(x) and sends it to Bob, who uses his secret information to compute $f^{-1}(f(x)) = x$.

In the RSA cryptosystem Bob chooses two prime numbers p and q (which in practice each have at least a hundred digits) and computes the number n = pq. He also chooses a number $e \neq 1$ which need not have a large number of digits but is relatively prime to (p-1)(q-1), so that it has an inverse d in $Z_{(p-1)(q-1)}$, and he computes $d = e^{-1} \mod (p-1)(q-1)$. Bob publishes eand n. The number e is called his *public key*. The number d is called his *private key*.

To summarize what we just said, here is a pseudocode outline of what Bob does:

```
Bob's RSA key choice algorithm
```

```
(1) Choose 2 large prime numbers p and q

(2) n = pq

(3) Choose e \neq 1 so that e is relatively prime to (p-1)(q-1)

(4) Compute d = e^{-1} \mod (p-1)(q-1).

(5) Publish e and n.

(6) Keep d secret.
```

People who want to send a message x to Bob compute $y = x^e \mod n$ and send that to him instead. (We assume x has fewer digits than n so that it is in Z_n . If not, the sender has to break the message into blocks of size less than the number of digits of n and send each block individually.)

To decode the message, Bob will compute $z = y^d \mod n$.

We summarize this process in pseudocode below:

Alice-send-message-to-Bob(x)			
Alice does	3:		
(1)	Read the public directory for Bob's keys e and n .		
(2)	Compute $y = x^e \mod n$		

```
(3) Send y to Bob
Bob does:
(4) Receive y from Alice
(5) Compute z = y^d \mod n, using secret key d
(6) Read z
```

Each step in these algorithms can be computed using methods from this Chapter. In Section 2.4, we will deal with computational issues in more detail.

In order to show that the RSA cryptosystem works, that is, that it allows us to encode and then correctly decode messages, we must show that z = x. In other words, we must show that, when Bob decodes, he gets back the original message. In order to show that the RSA cryptosystem is secure, we must argue that an eavesdropper, who knows n, e, and y, but does not know p, q or d, can not easily compute x.

Exercise 2.3-5 To show that the RSA cryptosystem works, we will first show a simpler fact. Why is

$$y^d \mod p = x \mod p?$$

Does this tell us what x is?

Plugging in the value of y, we have

$$y^d \bmod p = x^{ed} \bmod p. \tag{2.14}$$

But, in Line 4 we chose e and d so that $e \cdot_m d = 1$, where m = (p-1)(q-1). In other words,

$$ed \mod (p-1)(q-1) = 1.$$

Therefore, for some integer k,

$$ed = k(p-1)(q-1) + 1$$
.

Plugging this into Equation (2.14), we obtain

$$x^{ed} \mod p = x^{k(p-1)(q-1)+1} \mod p$$

= $x^{(k(q-1))(p-1)} x \mod p.$ (2.15)

But for any number a which is not a multiple of p, $a^{p-1} \mod p = 1$ by Fermat's Little Theorem (Theorem 2.22). We could simplify equation 2.15 by applying Fermat's Little Theorem to $x^{k(q-1)}$, as you will see below. However we can only do this when $x^{k(q-1)}$ is not a multiple of p. This gives us two cases, the case in which $x^{k(q-1)}$ is not a multiple of p (we'll call this case 1) and the case in which $x^{k(q-1)}$ is a multiple of p (we'll call this case 2). In case 1, we apply Equation 2.12 and Fermat's Little Theorem with a equal to $x^{k(q-1)}$, and we have that

$$x^{(k(q-1))(p-1)} \mod p = \left(x^{k(q-1)}\right)^{(p-1)} \mod p$$

$$= 1.$$
(2.16)

Combining equations 2.14, 2.15 and 2.17, we have that

$$y^d \mod p = x^{k(q-1)(p-1)}x \mod p = 1 \cdot x \mod p = x \mod p,$$

and hence $y^d \mod p = x \mod p$.

We still have to deal with case 2, the case in which $x^{k(q-1)}$ is a multiple of p. In this case x is a multiple of p as well since x is an integer and p is prime. Thus $x \mod p = 0$. Combining Equations 2.14 and 2.15 with Lemma 2.3, we get

$$y^{d} \mod p = (x^{k(q-1)(p-1)} \mod p)(x \mod p) = 0 = x \mod p.$$

Hence in this case as well, we have $y^d \mod p = x \mod p$.

While this will turn out to be useful information, it does not tell us what x is, however, because x may or may not equal $x \mod p$.

The same reasoning shows us that $y^d \mod q = x \mod q$. What remains is to show what these two facts tell us about $y^d \mod pq = y \mod n$, which is what Bob computes.

Notice that by Lemma 2.3 we have proved that

$$(y^d - x) \bmod p = 0 \tag{2.17}$$

and

$$(y^d - x) \bmod q = 0. (2.18)$$

- **Exercise 2.3-6** Write down an equation using only integers and addition, subtraction and multiplication in the integers, but perhaps more letters, that is equivalent to Equation 2.17, which says that $(y^d x) \mod p = 0$. (Do not use mods.)
- **Exercise 2.3-7** Write down an equation using only integers and addition, subtraction and multiplication in the integers, but perhaps more letters, that is equivalent to Equation 2.18, which says that $(y^d x) \mod q = 0$. (Do not use mods.)
- **Exercise 2.3-8** If a number is a multiple of a prime p and a different prime q, then what else is it a multiple of? What does this tell us about y^d and x?

The statement that $y^d - x \mod p = 0$ is equivalent to saying that $y^d - x = ip$ for some integer *i*. The statement that $y^d - x \mod q = 0$ is equivalent to saying $y^d - x = jq$ for some integer *j*. If something is a multiple of the prime *p* and the prime *q*, then it is a multiple of *pq*. Thus $(y^d - x) \mod pq = 0$. Lemma 2.3 tells us that $(y^d - x) \mod pq = (y^d \mod pq - x) \mod pq = 0$. But *x* and $y^d \mod pq$ are both integers between 0 and pq - 1, so their difference is between -(pq - 1) and pq - 1. The only integer between these two values that is 0 mod pq is zero itself. Thus $(y^d \mod pq) - x = 0$. In other words,

$$\begin{aligned} x &= y^d \bmod pq \\ &= y^d \bmod n \end{aligned}$$

which means that Bob will in fact get the correct answer.

Theorem 2.23 (Rivest, Shamir, and Adleman) The RSA procedure for encoding and decoding messages works correctly.

Proof: Proved above.

One might ask, given that Bob published e and n, and messages are encrypted by computing $x^e \mod n$, why can't any adversary who learns $x^e \mod n$ just compute eth roots mod n and break the code? At present, nobody knows a quick scheme for computing eth roots mod n, for an arbitrary n. Someone who does not know p and q cannot duplicate Bob's work and discover x. Thus, as far as we know, modular exponentiation is an example of a one-way function.

The Chinese Remainder Theorem

The method we used to do the last step of the proof of Theorem 2.23 also proves a theorem known as the "Chinese Remainder Theorem."

Exercise 2.3-9 For each number in $x \in Z_{15}$, write down $x \mod 3$ and $x \mod 5$. Is x uniquely determined by these values? Can you explain why?

x	$x \mod 3$	$x \bmod 5$
0	0	0
1	1	1
2	2	2
3	0	3
4	1	4
5	2	0
6	0	1
7	1	2
8	2	3
9	0	4
10	1	0
11	2	1
12	0	2
13	1	3
14	2	4

Table 2.2: The values of $x \mod 3$ and $x \mod 5$ for each x between zero and 14.

As we see from Table 2.2, each of the $3 \cdot 5 = 15$ pairs (i, j) of integers i, j with $0 \le i \le 2$ and $0 \le j \le 4$ occurs exactly once as x ranges through the fifteen integers from 0 to 14. Thus the function f given by $f(x) = (x \mod 3, x \mod 5)$ is a one-to-one function from a fifteen element set to a fifteen element set, so each x is uniquely determined by its pair of remainders.

The Chinese Remainder Theorem tells us that this observation always holds.

Theorem 2.24 (Chinese Remainder Theorem) If m and n are relatively prime integers and $a \in Z_m$ and $b \in Z_n$, then the equations

$$x \bmod m = a \tag{2.19}$$

$$x \bmod n = b \tag{2.20}$$

have one and only one solution for an integer x between 0 and mn - 1.

Proof: If we show that as x ranges over the integers from 0 to mn - 1, then the ordered pairs $(x \mod m, x \mod n)$ are all different, then we will have shown that the function given by $f(x) = (x \mod m, x \mod n)$ is a one to one function from an mn element set to an mn element

set, so it is onto as well.⁶ In other words, we will have shown that each pair of equations 2.19 and 2.20 has one and only one solution.

In order to show that f is one-to-one, we must show that if x and y are different numbers between 0 and mn - 1, then f(x) and f(y) are different. To do so, assume instead that we have an x and y with f(x) = f(y). Then $x \mod m = y \mod m$ and $x \mod n = y \mod n$, so that $(x-y) \mod m = 0$ and $(x-y) \mod n = 0$. That is, x-y is a multiple of both m and n. Then as we show in Problem 11 in the problems at the end of this section, x-y is a multiple of mn; that is, x-y = dmn for some integer d. Since we assumed x and y were different, this means x and ycannot both be between 0 and mn - 1 because their difference is mn or more. This contradicts our hypothesis that x and y were different numbers between 0 and mn - 1, so our assumption must be incorrect; that is f must be one-to-one. This completes the proof of the theorem.

Important Concepts, Formulas, and Theorems

1. Exponentiation in Z_n . For $a \in Z_n$, and a positive integer j:

$$a^j \mod n = \underbrace{a \cdot_n a \cdot_n \cdots \cdot_n a}_{j \text{ factors}}$$
.

2. Rules of exponents. For any $a \in Z_n$, and any nonnegative integers i and j,

$$(a^i \mod n) \cdot_n (a^j \mod n) = a^{i+j} \mod n$$

and

$$(a^i \mod n)^j \mod n = a^{ij} \mod n.$$

- 3. Multiplication by a fixed nonzero a in Z_p is a permutation. Let p be a prime number. For any fixed nonzero number a in Z_p , the numbers $(1 \cdot a) \mod p$, $(2 \cdot a) \mod p$, \ldots , $((p-1) \cdot a) \mod p$, are a permutation of the set $\{1, 2, \dots, p-1\}$.
- 4. Fermat's Little Theorem. Let p be a prime number. Then $a^{p-1} \mod p = 1$ for each nonzero a in Z_p .
- 5. Fermat's Little Theorem, version 2. For every positive integer a and prime p, if a is not a multiple of p, then

$$a^{p-1} \bmod p = 1.$$

6. RSA cryptosystem. (The first implementation of a public-key cryptosystem) In the RSA cryptosystem Bob chooses two prime numbers p and q (which in practice each have at least a hundred digits) and computes the number n = pq. He also chooses a number $e \neq 1$ which need not have a large number of digits but is relatively prime to (p-1)(q-1), so that it has an inverse d, and he computes $d = e^{-1} \mod (p-1)(q-1)$. Bob publishes e and n. To send a message x to Bob, Alice sends $y = x^e \mod n$. Bob decodes by computing $y^d \mod n$.

⁶If the function weren't onto, then because the number of pairs is the same as the number of possible x-values, two x values would have to map to the same pair, so the function wouldn't be one-to-one after all.

7. Chinese Remainder Theorem. If m and n are relatively prime integers and $a \in Z_m$ and $b \in Z_n$, then the equations

$$x \mod m = a$$
$$x \mod n = b$$

have one and only one solution for an integer x between 0 and mn - 1.

Problems

- 1. Compute the powers of 4 in Z_7 . Compute the powers of 4 in Z_{10} . What is the most striking similarity? What is the most striking difference?
- 2. Compute the numbers $1_{11} 5, 2_{11} 5, 3_{11} 5, \ldots, 10_{11} 5$. Do you get a permutation of the set $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$? Would you get a permutation of the set $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$? if you used another nonzero member of Z_{11} in place of 5?
- 3. Compute the fourth power mod 5 of each element of Z_5 . What do you observe? What general principle explains this observation?
- 4. The numbers 29 and 43 are primes. What is (29-1)(43-1)? What is $199 \cdot 1111$ in Z_{1176} ? What is $(23^{1111})^{199}$ in Z_{29} ? In Z_{43} ? In Z_{1247} ?
- 5. The numbers 29 and 43 are primes. What is (29-1)(43-1)? What is $199 \cdot 1111$ in Z_{1176} ? What is $(105^{1111})^{199}$ in Z_{29} ? In Z_{43} ? In Z_{1247} ? How does this answer the second question in Exercise 2.3-5?
- 6. How many solutions with x between 0 and 34 are there to the system of equations

$$x \mod 5 = 4$$
$$x \mod 7 = 5?$$

What are these solutions?

- 7. Compute each of the following. Show or explain your work, and do not use a calculator or computer.
 - (a) 15^{96} in Z_{97}
 - (b) 67^{72} in Z_{73}
 - (c) 67^{73} in Z_{73}
- 8. Show that in Z_p , with p prime, if $a^i \mod p = 1$, then $a^n \mod p = a^{n \mod i} \mod p$.
- 9. Show that there are $p^2 p$ elements with multiplicative inverses in Z_{p^2} when p is prime. If x has a multiplicative inverse in Z_p^2 , what is $x^{p^2-p} \mod p^2$? Is the same statement true for an element without an inverse? (Working out an example might help here.) Can you find something (interesting) that is true about x^{p^2-p} when x does not have an inverse?
- 10. How many elements have multiplicative inverses in Z_{pq} when p and q are primes?

- 11. In the paragraph preceding the proof of Theorem 2.23 we said that if a number is a multiple of the prime p and the prime q, then it is a multiple of pq. We will see how that is proved here.
 - (a) What equation in the integers does Euclid's extended GCD algorithm solve for us when m and n are relatively prime?
 - (b) Suppose that m and n are relatively prime and that k is a multiple of each one of them; that is, k = bm and k = cn for integers b and c. If you multiply both sides of the equation in part (a) by k, you get an equation expressing k as a sum of two products. By making appropriate substitutions in these terms, you can show that k is a multiple of mn. Do so. Does this justify the assertion we made in the paragraph preceding the proof of Theorem 2.23?
- 12. The relation of "congruence modulo n" is the relation \equiv defined by $x \equiv y \mod n$ if and only if $x \mod n = y \mod n$.
 - (a) Show that congruence modulo n is an equivalence relation by showing that it defines a partition of the integers into equivalence classes.
 - (b) Show that congruence modulo n is an equivalence relation by showing that it is reflexive, symmetric, and transitive.
 - (c) Express the Chinese Remainder theorem in the notation of congruence modulo n.
- 13. Write and implement code to do RSA encryption and decryption. Use it to send a message to someone else in the class. (You may use smaller numbers than are usually used in implementing the RSA algorithm for the sake of efficiency. In other words, you may choose your numbers so that your computer can multiply them without overflow.)
- 14. For some non-zero $a \in Z_p$, where p is prime, consider the set

$$S = \{a^0 \bmod p, a^1 \bmod p, a^2 \bmod p, \dots, a^{p-2} \bmod p, a^{p-1} \bmod p\},\$$

and let s = |S|. Show that s is always a factor of p - 1.

15. Show that if $x^{n-1} \mod n = 1$ for all integers x that are not multiples of n, then n is prime. (The slightly weaker statement that $x^{n-1} \mod n = 1$ for all x relatively prime to n, does not imply that n is prime. There is a famous family of numbers called Carmichael numbers that are counterexamples.⁷)

⁷See, for example, Cormen, Leiserson, Rivest, and Stein, cited earlier.