1 Modern Cryptography

Modern cryptography generally begins with two, related, desires

- In all the ciphers we've seen so far, we assume Alice and Bob have some way of sharing a key. We haven't said much about *how* that happens, and we've also seen that the perfect secrecy requires very long keys. How can Alice and Bob share a long key with each other without Eve finding out about it?
- None of the ciphers we've seen so far is robust against chosen-plaintext attacks. If Eve has the ability to request the ciphertexts for any plaintexts she likes, she can gradually get more information about the key until she can break the code. Recall that the one-time pad is only safe if the key is only used once.

We would like a cryptosystem where the *decryption key* is only known to Bob. Alice and Eve and everyone else in the world has access to Bob's encryption key and can encrypt messages for Bob to see, but then only Bob can recover the plaintext. This allows us to avoid Alice and Bob having to share a common key, and would allow Eve to generate ciphertexts for any plaintext of her choosing while also making sure that the cryptosystem is safe against chosen-ciphertext attacks.

It turns out that there are a number of cryptosystems of this type, and they all fall under the heading of **public-key cryptography**, because the encryption key can be made public to the world. A recurring theme behind public-key cryptosystems is a "one-way function," which is a function that's easy to compute but hard to invert.

1.1 Interlude: Primes

Definition 1.1: Prime

A positive integer $p \geq 2$ is **prime** if its only positive divisors are 1 and itself. An integer $n \geq 2$ that is not prime is called **composite**.

For example, 5 is prime because 1 and 5 are its only divisors. 4, on the other hand, is not prime, since 2 is a divisor of 4 (in addition to 1 and 4).

(Exercise.) The "twin prime conjecture" is a famous open problem which says that there are infinitely many "twin primes," ie, pairs of primes that are 2 apart. For example, 3 and 5 are twin primes, as are 5 and 7, or 11 and 13. Give five more examples of twin primes.

The following pairs are five other examples:

$$(17, 19), (29, 31), (41, 43), (59, 61), (71, 73).$$

(Exercise.) There is a version of the twin prime conjecture which says that every even integer can be written as the difference of consecutive primes in infinitely many ways. For example, we have:

$$6 = 29 - 23 = 137 - 131 = 599 - 593 = \dots.$$

Express the integer 10 as the difference of two consecutive primes in five different ways.

Five other ways include:

$$10 = 13 - 3 = 17 - 7 = 29 - 19 = 41 - 31 = 71 - 61.$$

(Exercise.) Explain why, if an integer $n \geq 2$ is composite, it must be divisible by a prime p such that $p \leq \sqrt{n}$. Use this fact to determine whether or not 701 is prime.

Proof. Suppose n is composite. Then, we can write

$$n = ab$$

such that a, b are integers and $1 < a < n$, $1 < b < n$, and WLOG $a \leq b$.

Further suppose that $a > \sqrt{n}$. Then, $b \geq a > \sqrt{n}$. However, if $b \geq a > \sqrt{n}$, then

$$ab > \sqrt{n}\sqrt{n} = n,$$

a contradiction (since we wrote that $n = ab$). Therefore, $a \leq \sqrt{n}$. We know that there exists some prime p which divides a, so it follows that $p \leq a$ and $p \leq \sqrt{n}$. □

1.1.1 Ubiquity of Primes

Every positive integer $n \geq 2$ can be written as a product of primes. For example, $18 = 2 \cdot 3^2$ and both 2 and 3 are primes. An expression of an integer $n \geq 2$ as a product of primes is called a *prime factorization* of n.

Theorem 1.1: Fundamental Theorem of Arithmetic

Any positive integer $n \geq 2$ has a unique prime factorization. In other words, there exists an expression

$$n = p_1^{e_1} p_2^{e_2} \dots p_r^{e_r}$$

such that $p_1, \dots, p_r$ are primes and $e_1, \dots, e_r$ are positive integers, and the expression is unique up to reordering the indices.

For example, $60 = 2^2 \cdot 3^1 \cdot 5^1$ is a unique prime factorization of 60; the only other prime factorization involves reordering the factors around.

(Example.) Find the prime factorizations of 1231 and of 1232.

- For 1231, note that
$$1231 = 1231 \cdot 1.$$
- For 1232, note that
$$1232 = 616 \cdot 2 = 308 \cdot 2^2 = 154 \cdot 2^3 = 77 \cdot 2^4 = 11 \cdot 7 \cdot 2^4.$$

(Exercise.)

1. Find all prime factors of 50!. (Just a list of the prime factors is sufficient; you don't need to find the exponents of the prime factorization for this part.)

   Note that
   $$50! = 50 \cdot 49 \cdot 48 \cdot \ldots \cdot 3 \cdot 2 \cdot 1.$$
   In particular, any composite number less than or equal to 50 can be decomposed into primes that are also less than or equal to 50. Thus, we have
   $$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47.$$

2. Find the prime factorization of 10!.

We know that

$$10! = 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2,$$

where

$$\begin{aligned} 10 &= 5 \cdot 2, \\ 9 &= 3^2, \\ 8 &= 2^3, \\ 7 &= 7 \cdot 1, \\ 6 &= 3 \cdot 2, \\ 5 &= 5 \cdot 1, \\ 4 &= 2 \cdot 2, \\ 3 &= 3 \cdot 1, \\ 2 &= 2 \cdot 1. \end{aligned}$$

So,

$$10! = 2^8 \cdot 3^4 \cdot 5^2 \cdot 7.$$

3. Find the prime factorization of $11!/2^8$.

We know that

$$\begin{aligned} \frac{11!}{2^8} &= \frac{11 \cdot 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2}{2^8} \\ &= \frac{11 \cdot (5 \cdot 2) \cdot (3^2) \cdot (2^3) \cdot 7 \cdot (2 \cdot 3) \cdot 5 \cdot (2^2) \cdot 3 \cdot 2}{2^8} \\ &= 11 \cdot 5 \cdot (3^2) \cdot 7 \cdot 3 \cdot 5 \cdot 3 \\ &= (3^2) \cdot 3 \cdot 3 \cdot 5 \cdot 5 \cdot 7 \cdot 11 \\ &= 3^4 \cdot 5^2 \cdot 7 \cdot 11. \end{aligned}$$

Alternatively,

$$\begin{aligned} \frac{11!}{2^8} &= \frac{11 \cdot 10!}{2^8} \\ &= \frac{11 \cdot 2^8 \cdot 3^4 \cdot 5^2 \cdot 7}{2^8} \\ &= 11 \cdot 3^4 \cdot 5^2 \cdot 7. \end{aligned}$$

Lemma 1.1: Euclid's Lemma

Suppose a and b are integers and that d is a divisor of ab such that $\gcd(a, b) = 1$. Then, d is a divisor of b.

Proof. By Bezout's theorem, there exists x and y such that

$$1 = \gcd(a, d) = ax + dy.$$

Multiplying this by b gives us

$$b = abx + bdy.$$

Observe that d divides bdy, and it also divides ab. Thus, d divides the sum $abx + bdy = b$. □

1.1.2 Scarcity of Primes & Difficulty of Factoring

Having made the observation that primes are "ubiquitous," we should also note that primes are "scarce." The first sense is a literal sense: as a proportion of all integers, very few integers end up being prime. The second sense is that prime factorizations are difficult to actually calculate, even for moderately large numbers.

The naive way to find a prime factorization is to figure out what the factors are.

(Example.) Suppose we want to find the prime factorization of 75. First, we note that 2 does not divide 75. We then see that 3 does divide 75, and so we're left with 25 ($75/3 = 25$). Now, 3 no longer divides 25, and neither does 4. However, 5 does divide 25, so we're left with 5 ($25/5 = 5$). Since 5 divides itself, we're left with 1 ($5/5 = 1$), so the prime factorizations are

$$75 = 3^1 \cdot 5^2.$$

This process is extremely slow as the number gets larger or, more specifically, as the prime factors of the number get larger.

There are a number of deep, sophisticated, and clever techniques to speed this process up[1], but no one has yet found a factorization algorithm for classical computers that is substantially better than the naive method we just described. The difficulty of factoring can be leveraged to build modern cryptosystems that are in widespread use today.

Some things to further consider:

- Just because there are no substantially better algorithm for finding the prime factorizations doesn't mean one does exist. In other words, there could be an algorithm for finding prime factorizations in an efficient way, but we haven't found it.
- Efficient factorization algorithms exist for *quantum computers*, but quantum computing hardware has not yet caught up to our theoretical knowledge, so our modern cryptosystems are still safe for now. That being said, this won't last for much longer and so we will soon need new cryptosystems that are secure against quantum computers.

The difficulty of factoring suggests that the function μ, which takes as input two prime numbers p and q and outputs the product

$$\mu(p, q) = pq$$

is our first example of a **one-way** function. It's very easy to compute the product of two primes but, if the primes are large, it'll be very hard to invert this function (i.e., find the prime factors of some given number that has two large prime factors). This is the one-way function on which RSA is based, as we will soon see.

1.2 Euler's Phi Function

Definition 1.2: Euler's Phi Function

For a positive integer n, let $\phi(n)$ denote the number of integers r with $0 \leq r < n$ and $\gcd(n, r) = 1$. The function $n \mapsto \phi(n)$ is called *Euler's phi function* (or *Euler's totient function*).

For example, when we were counting that there are 312 affine encryption functions available in English, part of the process involved counting the number of numbers relatively prime to 26, and we found it was 12. Now, we can say this as

$$\phi(26) = 12.$$

[1] Which will not be covered in this course.

(Exercise.) Compute $\phi(12)$, $\phi(13)$, $\phi(14)$, and $\phi(15)$.

- For $\phi(12)$, we know that
$$\gcd(1,12) = \gcd(5,12) = \gcd(7,12) = \gcd(11,12) = 4,$$
so $\phi(12) = 4$.
- Because 13 is prime, $\phi(13) = 12$.
- For $\phi(14)$, we know that
$$\gcd(1,14) = \gcd(3,14) = \gcd(5,14) = \gcd(9,14) = \gcd(11,14) = \gcd(13,14) = 1,$$
so $\phi(14) = 6$.
- By the same reasoning, $\phi(15) = 8$.

(Exercise.) Explain why, in a language that uses an alphabet with n letters, the number of distinct affine encryption functions is $n\phi(n)$.

Recall that the affine encryption function is of the form
$$E(x) = (ax + b) \pmod{n}.$$
There are n possible options for b, and $\phi(n)$ possible options for a (otherwise, there's no inverse for a). This gives us $\phi(n) \cdot n$.

(Exercise.) Suppose p is prime.

1. Explain why $\phi(p) = p - 1$.

   If p is prime, then its only positive divisors are 1 and itself. So, in particular,
   $$\phi(1,p) = \phi(2,p) = \ldots = \phi(p-1,p) = 1$$
   and
   $$\phi(p,p) = p.$$

2. Explain why $\phi(p^e) = p^e - p^{e-1}$ for any $e \geq 1$.

   A number is relatively prime with p^e if and only if it is not divisible by p, so there are p^e numbers in total,
   $$\{1, 2, \ldots, p^e\}.$$
   From those, exactly p^{e-1} are divisible by p,
   $$\{1p, 2p, 3p, \ldots, p^{e-1}p\}.$$
   Therefore, there are $p^e - p^{e-1}$ numbers in the list not divisible by p.

Lemma 1.2

If $\gcd(a,b) = 1$, then $\phi(ab) = \phi(a)\phi(b)$.

With this in mind, the formula for $\phi(n)$ is

$$\phi(n) = p_1^{e_1-1}(p_1-1)\cdot\ldots\cdot p_r^{e_r-1}(p_r-1)$$

(Exercise.) Use the formula for Euler's phi function to calculate $\phi(n)$ for each of the following values of n.

(a) $n = 20 = 2^2 \cdot 5$.

$$\phi(20) = 2^{2-1}(2-1)\cdot 5^{1-1}(5-1) = 2\cdot 4 = 8.$$

(b) $n = 25 = 5^2$.

$$\phi(25) = 5^{2-1}(5-1) = 5\cdot 4 = 20.$$

(c) $n = 30 = 2\cdot 3\cdot 5$.

$$\phi(30) = 2^{1-1}(2-1)\cdot 3^{1-1}(3-1)\cdot 5^{1-1}(5-1) = 1\cdot 2\cdot 4 = 8.$$

(d) $n = 35 = 5\cdot 7$.

$$\phi(35) = 5^{1-1}(5-1)\cdot 7^{1-1}(7-1) = 4\cdot 6 = 24.$$

(e) $n = 40 = 2^3\cdot 5$.

$$\phi(40) = 2^{3-1}(2-1)\cdot 5^{1-1}(5-1) = 2^2\cdot 4 = 16.$$

We should also note that

$$\phi(n) = n\prod_{p|n}\left(1-\frac{1}{p}\right).$$

Theorem 1.2: Euler's Theorem

Suppose n is a positive integer and a is another integer with $\gcd(a,n) = 1$. Then,

$$a^{\phi(n)} \equiv 1 \pmod{n}.$$

(Example.) We know that $20 = 2^2\cdot 5$ so $\phi(20) = 8$. This means that

$$a^{\phi(20)} = a^8 \equiv 1 \pmod{20}$$

for any a that is relatively prime to 20. For example, when $a = 3$, we have $\gcd(3,20) = 1$ and, using the Modular Arithmetic Theorem, we have

$$\begin{aligned} 3^2 &= 9 \\ 3^4 &= (3^2)^2 \equiv 9^2 = 81 \equiv 1 \pmod{20} \\ 3^8 &= (3^4)^2 \equiv 1^2 = 1 \pmod{20}. \end{aligned}$$

We can use this theorem to compute very large powers of some number. For example, to compute

$7^{20232023} \pmod{20}$, we notice that 20232020 is divisible by 8 so

$$20232023 = 20232000 + 23 \equiv 23 \equiv 7 \pmod{8}.$$

So, $20232023 = 8q + 7$ for some $q \geq 0$, so

$$7^{20232023} = 7^{8q+7} = (7^8)^q 7^7 = 1^q 7^7 = 7^7 \pmod{20}.$$

Then, to compute $7^7 \pmod{20}$, we can do

$$\begin{aligned} 7^2 &= 49 \equiv 9 \pmod{20} \\ 7^4 &= (7^2)^2 \equiv 9^2 = 81 \equiv 1 \pmod{20} \\ 7^7 &= 7^4 \cdot 7^2 \cdot 7 \equiv 1 \cdot 9 \cdot 7 = 63 \equiv 3 \pmod{20}. \end{aligned}$$

So, $7^{20232023} \pmod{20} = 3$.

Notice how we could do this remainder calculation without having to compute $7^{20232023}$. This is important, since RSA will require us that to similar remainder calculations even when the numbers are so large that it is infeasible for computers to calculate the result of the exponentiation.

(Exercise.) Use Euler's Theorem to show that 51 divides $10^{32n+9} - 7$ for any integer $n \geq 0$.

Recall that

$$\phi(51) = 32,$$

so in particular

$$a^{32} \equiv 1 \pmod{51}$$

for any a that is relatively prime to 51. Then, it follows that

$$10^{32n+9} - 7 = 10^{32n} 10^9 - 7 = (10^{32})^n 10^9 - 7 \equiv 10^9 - 7 \pmod{51}.$$

Note here that 10^{32n} and 51 are relatively prime. In any case, it follows that

$$10^9 - 7 \pmod{51} = 0.$$

(Exercise.) Find the units digit of 3^{100} using Euler's theorem.

We have $\phi(10) = 4$ so

$$a^4 \equiv 1 \pmod{10}$$

for some a that is relatively prime to 10. Now, we note that

$$3^{100} = 3^{4 \cdot 25} = (3^4)^{25}$$

and so

$$(3^4)^{25} \equiv 1^{25} = 1 \pmod{10}.$$

Therefore, the unit digit of 3^{100} is 1.

(Exercise.) Fix a prime number p. There are two versions of "Fermat's little theorem."

1. If a is an integer that is not divisible by p, then $a^{p-1} \equiv 1 \pmod{p}$.

Recall that $\phi(p) = p - 1$, so

$$a^{\phi(p)} \equiv 1 \pmod{p} \implies a^{p-1} \equiv 1 \pmod{p}.$$

2. For any integer a, we have $a^p \equiv a \pmod{p}$.

Note that

$$a^{p-1} = a^p \frac{1}{a} \equiv 1 \pmod{p}$$

by part (1), so multiplying both sides by a gives us

$$a^p \equiv a \pmod{p}.$$