

## Chapter 7

# The Theorem of Euler-Fermat

In this chapter we will discuss the generalization of Fermat's Little Theorem to composite values of the modulus. We will also discuss applications in cryptography.

### 7.1 The Theorem of Euler-Fermat

Consider the unit group  $(\mathbb{Z}/15\mathbb{Z})^\times$  of  $\mathbb{Z}/15\mathbb{Z}$ . It consists of the eight residue classes  $[1], [2], [4], [7], [8], [11], [13], [14]$ . If we multiply each of these classes e.g. by  $[7]$  (or  $[8], [9]$ ), then we get

$$\begin{array}{lll} [1] \cdot [7] = [7] & [1] \cdot [8] = [8] & [1] \cdot [9] = [9] \\ [2] \cdot [7] = [14] & [2] \cdot [8] = [1] & [2] \cdot [9] = [3] \\ [4] \cdot [7] = [13] & [4] \cdot [8] = [2] & [4] \cdot [9] = [6] \\ [7] \cdot [7] = [4] & [7] \cdot [8] = [11] & [7] \cdot [9] = [3] \\ [8] \cdot [7] = [11] & [8] \cdot [8] = [4] & [8] \cdot [9] = [12] \\ [11] \cdot [7] = [2] & [11] \cdot [8] = [13] & [11] \cdot [9] = [9] \\ [13] \cdot [7] = [1] & [13] \cdot [8] = [14] & [13] \cdot [9] = [12] \\ [14] \cdot [7] = [8] & [14] \cdot [8] = [7] & [14] \cdot [9] = [6] \end{array}$$

As in our proof of Fermat's Little Theorem, the resulting residue classes (for multiplication by  $[7]$  and  $[8]$ ) are the classes we started with in a different order. Multiplying these equations we get

$$\prod_{(a,15)=1} [a] = \prod_{(a,15)=1} [7a] = [7]^8 \prod_{(a,15)=1} [a].$$

Since the  $a$  are coprime to 15, so is their product; thus we may cancel, and we find  $[7]^8 = [1]$ , or  $7^8 \equiv 1 \pmod{15}$ . Similarly, we find  $8^8 \equiv 1 \pmod{15}$ ; for multiplication by 9, however, the classes on the right hand side differ from

those on the left (they're all divisible by 3 since both 9 and 15 are), and we do *not* get  $9^8 \equiv 1 \pmod{15}$ .

The same idea works in general. Let  $m \geq 2$  be an integer, and let  $\phi(m)$  denote the number of residue classes coprime to  $m$ , that is,  $\phi(m) = \#(\mathbb{Z}/m\mathbb{Z})^\times$ . Then we have the following result, which is usually referred to as the Euler-Fermat Theorem: it is due to Euler, but contains Fermat's Little Theorem as a special case.

**Theorem 7.1.** *If  $a$  is an integer coprime to  $m \geq 2$ , then  $a^{\phi(m)} \equiv 1 \pmod{m}$ .*

For  $m = p$  prime, we have  $\phi(p) = p - 1$ , and Euler's Theorem becomes Fermat's Little Theorem.

*Proof.* Let  $[r_i]$ ,  $i = 1, \dots, t = \phi(m)$ , denote the residue classes in  $(\mathbb{Z}/m\mathbb{Z})^\times$ . Then we claim that  $[ar_1], \dots, [ar_t]$  are pairwise distinct. In fact, assume that  $[ar_i] = [ar_j]$  with  $i \neq j$ , that is,  $ar_i \equiv ar_j \pmod{m}$ . Since  $\gcd(a, m) = 1$ , we may cancel  $a$ , and get  $[r_i] = [r_j]$ : contradiction.

Since the classes  $[ar_1], \dots, [ar_t]$  are all in  $(\mathbb{Z}/m\mathbb{Z})^\times$  and different, and since there are only  $t$  different classes in  $(\mathbb{Z}/m\mathbb{Z})^\times$ , we must have  $(\mathbb{Z}/m\mathbb{Z})^\times = \{[ar_1], \dots, [ar_t]\}$ . But then  $\prod_{i=1}^t [r_i] = \prod_{i=1}^t [ar_i] = [a]^{\phi(m)} \prod_{i=1}^t [r_i]$ . Since the  $[r_i]$  are coprime to  $m$ , so is their product. Cancelling then gives  $[a]^{\phi(m)} = [1]$ , which proves the claim.  $\square$

## 7.2 Euler's Phi Function

For the application of Euler-Fermat we need a formula that allows us to compute  $\phi(n)$ . Let us first compute  $\phi(n)$  directly for some small  $n$ . For  $n = 6$ , there are 6 different residue classes modulo 6; the classes  $[0]$ ,  $[2]$ ,  $[3]$  and  $[4]$  are not coprime to 6 (or, in other words, do not have a multiplicative inverse), which leaves the classes  $[1]$  and  $[5]$  as the only ones that are coprime to 6: thus  $\phi(6) = 2$ . The classes mod 8 coprime to 8 are  $[1]$ ,  $[3]$ ,  $[5]$ ,  $[7]$ , hence  $\phi(8) = 4$ . If  $p$  is prime, then all the  $p - 1$  classes  $[1]$ ,  $[2]$ ,  $\dots$ ,  $[p - 1]$  are coprime to  $p$ , hence  $\phi(p) = p - 1$ .

$n$	3	4	5	6	7	8	9	10	12	15
$\phi(n)$	2	2	4	2	6	4	6	4	4	8

We can easily compute  $\phi(p^k)$  (Euler's phi function for prime powers): starting with all the nonzero classes  $[1], [2], \dots, [p^2 - 1]$  (there are  $p^2 - 1$  of them) we have to eliminate those that are not coprime to  $p^2$ , that is, exactly the multiples of  $p$  smaller than  $p^2$ : these are  $p, 2p, 3p, \dots, (p - 1)p$  (note that  $p \cdot p = p^2 > p^2 - 1$ ); since there are exactly  $p - 1$  of these multiples of  $p$ , there will be exactly  $p^2 - 1 - (p - 1) = p^2 - p = p(p - 1)$  classes left: thus  $\phi(p^2) = p(p - 1)$ .

The same method works for  $p^k$ : there are exactly  $p^k - 1$  nonzero classes, namely  $[1], [2], \dots, [p^k - 1]$ . The multiples of  $p$  among these classes are  $[p], [2p], \dots, p^k - p = (p^{k-1} - 1)p$ , and there are exactly  $p^{k-1} - 1$  of them. Thus  $\phi(p^k) = p^k - 1 - (p^{k-1} - 1) = p^k - p^{k-1} = p^{k-1}(p - 1)$ .

We have proved

**Proposition 7.2.** *For primes  $p$  and integers  $k \geq 1$ , we have*

$$\phi(p^k) = p^{k-1}(p-1).$$

Let us now compute  $\phi(pq)$  for a product of two different primes. We have  $pq - 1$  nonzero residue classes  $[1], [2], \dots, [pq - 1]$ . The classes that have a factor in common with  $pq$  are multiples of  $p$  and multiples of  $q$ , namely  $[p], [2p], \dots, [(q-1)p]$  and  $[q], [2q], \dots, [(p-1)q]$ . Since there are no multiples of  $p$  that are multiples of  $q$  (like  $[0], [pq]$ , etc) among these, there will be exactly  $pq - 1 - (p-1) - (q-1) = pq - p - q + 1 = (p-1)(q-1)$  classes left after eliminating multiples of  $p$  or  $q$ . Thus  $\phi(pq) = (p-1)(q-1) = \phi(p)\phi(q)$ .

The general result is

**Proposition 7.3.** *If  $m$  and  $n$  are coprime integers, then  $\phi(mn) = \phi(m)\phi(n)$ .*

Before we turn to the proof, let's see how it works in a specific example like  $m = 5$  and  $n = 3$ . What we'll do is take a residue class modulo 15 and coprime to 15, and map it to a pair of residue classes mod 3 and mod 5:

$a \bmod 15$	1	2	4	7	8	11	13	14
$a \bmod 3$	1	2	1	1	2	2	1	2
$a \bmod 5$	1	2	4	2	3	1	3	4

Thus we have the following pairs of residue classes modulo 3 and 5:  $(1, 1), (1, 2), (1, 3), (1, 4)$  and  $(2, 1), (2, 2), (2, 3), (2, 4)$ . In particular, there are  $\phi(5) = 4$  pairs with  $a \equiv 1 \pmod 3$  and 4 pairs with  $a \equiv 2 \pmod 3$ .

*Proof of Prop. 7.3.* We have to find a map sending a residue class modulo  $mn$  to two residue classes modulo  $m$  and  $n$ . Let's try

$$\psi : (\mathbb{Z}/mn\mathbb{Z})^\times \longrightarrow (\mathbb{Z}/m\mathbb{Z})^\times \times (\mathbb{Z}/n\mathbb{Z})^\times : [a]_{mn} \longmapsto ([a]_m, [a]_n).$$

All that's left to do is check that it works. First observe that  $\gcd(ab, n) = 1$  if and only if  $\gcd(a, n) = \gcd(b, n) = 1$ .

**Surjectivity:** We have to show that, given residue classes  $[r]_m$  and  $[s]_n$ , there exists a residue class  $[a]_{mn}$  such that  $[a]_m = [r]_m$  and  $[a]_n = [s]_n$ . At this point, Bezout comes in again: since  $\gcd(m, n) = 1$ , there exist  $x, y \in \mathbb{Z}$  such that  $1 = mx + ny$ . Now put  $a = rym + sxn$ : then  $a = rym + sxm \equiv rym \equiv 1 \pmod m$  since  $yn \equiv 1 \pmod m$  from the Bezout representation, and similarly  $a = rym + sxm \equiv sxm \equiv s \pmod n$ .

**Injectivity:** Assume that there are residue classes  $[a]_{mn}$  and  $[b]_{mn}$  such that  $[a]_m = [b]_m$  and  $[a]_n = [b]_n$ . Then  $m \mid (b - a)$  and  $n \mid (b - a)$ , and since  $\gcd(m, n) = 1$ , this implies that  $[a]_{mn} = [b]_{mn}$  and proves the injectivity of  $\phi$ .  $\square$

Here is how one could come up with the application of Bezout in the above proof. Given coprime residue classes  $r \pmod m$  and  $s \pmod n$ , we want a formula

for computing an integer  $a$  such that  $a \equiv r \pmod{m}$  and  $a \equiv s \pmod{n}$ . The first idea is to see whether  $a$  can be written as a linear combination of  $r$  and  $s$ , that is, to look for integers  $x, y$  such that  $a = xr + ys$ . Reduction modulo  $m$  gives

$$r \equiv a = xr + ys \pmod{m}. \quad (7.1)$$

The simplest way to achieve this is by taking  $x = 1$  and  $y = 0$ . But observe that we also need

$$s \equiv a \equiv xr + ys \pmod{n}. \quad (7.2)$$

Thus we need more leeway. The right idea is to observe that (7.1) will be satisfied if only  $x \equiv 1 \pmod{m}$ ,  $y \equiv 0 \pmod{m}$ . Similarly, (7.2) will be satisfied if  $x \equiv 0 \pmod{n}$  and  $y \equiv 1 \pmod{n}$ .

Is it possible to satisfy these four congruences simultaneously? Let's see:  $x \equiv 0 \pmod{n}$  and  $y \equiv 0 \pmod{m}$  mean  $x = an$  and  $y = bm$  for some  $a, b \in \mathbb{Z}$ . The two other congruences boil down to  $x = an \equiv 1 \pmod{m}$  and  $y = bm \equiv 1 \pmod{n}$ . But these are both solvable since  $\gcd(m, n) = 1$ , so  $n$  has an inverse  $a$  modulo  $m$ , and  $m$  has an inverse  $b$  modulo  $n$ . Inverses can be computed using Bezout, and collecting everything we now can see where the formulas in the above proof were coming from.

Combining the formulas for Euler's phi function for prime powers and for products of coprime integers, we now find that an integer

$$m = p_1^{a_1} \cdots p_r^{a_r}$$

has exactly

$$\begin{aligned} \phi(m) &= (p_1 - 1)p_1^{a_1 - 1} \cdots (p_r - 1)p_r^{a_r - 1} \\ &= p_1^{a_1} \cdots p_r^{a_r} \cdot \frac{p_1 - 1}{p_1} \cdots \frac{p_r - 1}{p_r} \\ &= m \left(1 - \frac{1}{p_1}\right) \cdots \left(1 - \frac{1}{p_r}\right) \end{aligned}$$

residue classes coprime to  $m$ .

## Chinese Remainder Theorem

In the proof of the multiplicativity of Euler's phi function we have shown that, given a system of congruences

$$\begin{aligned} x &\equiv a \pmod{m} \\ y &\equiv b \pmod{n} \end{aligned}$$

can always be solved if  $m$  and  $n$  are coprime. This result, or rather its generalization to system of arbitrarily many such congruences, is called the Chinese Remainder Theorem.

## The Abstract Version

There is more to the bijection

$$\psi : (\mathbb{Z}/mn\mathbb{Z})^\times \longrightarrow (\mathbb{Z}/m\mathbb{Z})^\times \times (\mathbb{Z}/n\mathbb{Z})^\times : [a]_{mn} \longmapsto ([a]_m, [a]_n)$$

constructed above than meets the eye: we claim that  $\psi$  induces an isomorphism  $(\mathbb{Z}/mn\mathbb{Z})^\times \longrightarrow (\mathbb{Z}/m\mathbb{Z})^\times \times (\mathbb{Z}/n\mathbb{Z})^\times$ .

A homomorphism between groups  $(G, \circ)$  and  $(H, *)$  is a map  $f : G \longrightarrow H$  that respects the group laws in the sense that we have  $f(g \circ g') = f(g) * f(g')$ . Here are some examples:

1. the exponential function is a homomorphism  $\exp : (\mathbb{R}, +) \longrightarrow (\mathbb{R}_{>0}, \cdot)$  because  $\exp(a+b) = \exp(a)\exp(b)$ .
2. the logarithm is a homomorphism  $\log : (\mathbb{R}_{>0}, \cdot) \longrightarrow (\mathbb{R}, +)$  because  $\log ab = \log a + \log b$ . Note that  $\exp$  and  $\log$  are inverse maps of each other.
3. The set  $C^\infty$  of all infinitely often differentiable functions  $(0, 1) \longrightarrow \mathbb{R}$  is an additive group, and  $\frac{d}{dx} : C^\infty \longrightarrow C^\infty$  is a homomorphism because  $(f+g)' = f' + g'$ .
4. If  $f : V \longrightarrow W$  is a linear map between  $K$ -vector spaces  $V$  and  $W$ , then  $f$  is also a homomorphism between the additive groups  $(V, +)$  and  $(W, +)$ .
5. The map  $\psi : (\mathbb{Z}/mn\mathbb{Z})^\times \longrightarrow (\mathbb{Z}/m\mathbb{Z})^\times \times (\mathbb{Z}/n\mathbb{Z})^\times$  is a homomorphism. In fact we have

$$\begin{aligned}\psi([ab]_{mn}) &= ([ab]_m, [ab]_n), \\ \psi([a]_{mn}) &= ([a]_m, [a]_n), \\ \psi([b]_{mn}) &= ([b]_m, [b]_n),\end{aligned}$$

and by the group law in direct products we see that

$$\psi([ab]_{mn}) = \psi([a]_{mn})\psi([b]_{mn}).$$

If  $(G, \circ)$  and  $(H, *)$  are groups, then the cartesian product  $G \times H$  can be given a group structure by defining  $(g, h)(g', h') = (g \circ g', h \circ h')$ . Checking the axioms is straightforward. Also, if  $G \times H$  is abelian if and only if  $G$  and  $H$  are.

Observe that if  $f : G \longrightarrow H$  is a homomorphism between additively written groups, then  $f(0) = 0$  and  $f(-g) = -f(g)$ . This follows easily from the axioms.

Since we have already seen that  $\psi$  is bijective, we can conclude that it is an isomorphism. Note that for any bijective homomorphism  $f : G \longrightarrow H$  there exists a homomorphism  $g : H \longrightarrow G$  such that  $f \circ g$  and  $g \circ f$  are the identity maps on  $H$  and  $G$ , respectively.

We can play this game also with rings: a map from a ring  $R$  to some ring  $S$  is called a ring homomorphism if  $f(r+r') = f(r) + f(r')$ ,  $f(rr') = f(r)f(r')$ , and  $f(1) = 1$ . It is then easy to show that  $\psi$  actually induces a ring isomorphism  $\mathbb{Z}/mn\mathbb{Z} \longrightarrow \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ : this is the abstract formulation of the Chinese Remainder Theorem.

### 7.3 The Order of Residue Classes

Assume that we are given an integer  $m$  and an integer  $a$  coprime to  $m$ . The smallest exponent  $n > 0$  such that  $a^n \equiv 1 \pmod{m}$  is called the order of  $a \pmod{m}$ ; we write  $n = \text{ord}_m(a)$ . Note that we always have  $\text{ord}_m(1) = 1$ . Here's a table for the orders of elements in  $(\mathbb{Z}/7\mathbb{Z})^\times$ :

$a \pmod{7}$	1	2	3	4	5	6
$\text{ord}_7(a)$	1	3	6	3	6	2

If  $m = p$  is prime, then Fermat's Little Theorem gives us  $a^{p-1} \equiv 1 \pmod{p}$ , i.e., the order of  $a \pmod{p}$  is at most  $p-1$ . In general, the order of  $a$  is not  $p-1$ ; it is, however, always a divisor of  $p-1$  (as the table above suggested):

**Proposition 7.4.** *Given a prime  $p$  and an integer  $a$  coprime to  $p$ , let  $n$  denote the order of  $a$  modulo  $p$ . If  $m$  is any integer such that  $a^m \equiv 1 \pmod{p}$ , then  $n \mid m$ . In particular,  $n$  divides  $p-1$ .*

*Proof.* Write  $d = \text{gcd}(n, m)$  and  $d = nx + my$ ; then  $a^d = a^{nx+my} \equiv 1 \pmod{p}$  since  $a^n \equiv a^m \equiv 1 \pmod{p}$ . The minimality of  $n$  implies that  $n \leq d$ , but then  $d \mid n$  shows that we must have  $d = n$ , hence  $n \mid m$ .  $\square$

Here comes a pretty application to prime divisors of Mersenne and Fermat numbers.

**Corollary 7.5.** *If  $p$  is an odd prime and if  $q \mid M_p$ , then  $q \equiv 1 \pmod{2p}$ .*

*Proof.* It suffices to prove this for prime values of  $q$  (why?). So assume that  $q \mid 2^p - 1$ ; then  $2^p \equiv 1 \pmod{q}$ . By Proposition 7.4, the order of  $2 \pmod{q}$  divides  $p$ , and since  $p$  is prime, we find that  $p = \text{ord}_q(2)$ .

On the other hand, we also have  $2^{q-1} \equiv 1 \pmod{q}$  by Fermat's little theorem, so Proposition 7.4 gives  $p \mid (q-1)$ , and this proves the claim because we clearly have  $q \equiv 1 \pmod{2}$ .  $\square$

Example:  $M_{11} = 2047 = 23 \cdot 89$ .

Fermat numbers are integers  $F_n = 2^{2^n} + 1$  (thus  $F_1 = 5$ ,  $F_2 = 17$ ,  $F_3 = 257$ ,  $F_4 = 65537$ ,  $\dots$ ), and Fermat conjectured (and once even seemed to claim he had a proof) that these integers are all primes. These integers became much more interesting when Gauss succeeded in proving that a regular  $p$ -gon,  $p$  an odd prime, can be constructed with ruler and compass if  $p$  is a Fermat prime. Gauss also stated that he had proved the converse, namely that if a regular  $p$ -gon can be constructed by ruler and compass, then  $p$  is a Fermat prime, but the first (almost) complete proof was given by Pierre Wantzel.<sup>1</sup>

**Corollary 7.6.** *If  $q$  divides  $F_n$ , then  $q \equiv 1 \pmod{2^{n+1}}$ .*

<sup>1</sup>Pierre Wantzel, 1814 (Paris) – 1848 (Paris).

*Proof.* It is sufficient to prove this for prime divisors  $q$ . Assume that  $q \mid F_n$ ; then  $2^{2^n} + 1 \equiv 1 \pmod{q}$ , hence  $2^{2^n} \equiv -1 \pmod{q}$  and  $2^{2^{n+1}} \equiv 1 \pmod{q}$ . We claim that actually  $2^{n+1} = \text{ord}_q(2)$ : in fact, Proposition 7.4 says that the order divides  $2^{n+1}$ , hence is a power of 2. But  $2^{n+1}$  is clearly the smallest power of 2 that does it.

On the other hand,  $2^{q-1} \equiv 1 \pmod{q}$  by Fermat's Little Theorem, and Proposition 7.4 gives  $2^{n+1} \mid (q-1)$ , which proves the claim.  $\square$

In particular, the possible prime divisors of  $F_5 = 4294967297$  are of the form  $q = 64m + 1$ . After a few trial divisions one finds  $F_5 = 641 \cdot 6700417$ . This is how Euler disproved Fermat's conjecture. Today we know the prime factorization of  $F_n$  for all  $n \leq 11$ , we know that  $F_n$  is composite for  $5 \leq n \leq 30$  (and several larger values up to  $n = 382447$ ), and we don't know any factors for  $n = 14, 20, 22$  and  $24$ . See

<http://www.prothsearch.net/fermat.html>  
for more.

## 7.4 RSA

Cryptography deals with methods that allow us to transmit information safely, that is, in such a way that eavesdroppers have no chance of reading it. Simple methods for encrypting messages were known and widely used in military circles for several millenia; basically all of these codes are easy to break with computers.

An example of such a classical code is Caesar's cipher: permute the letters of the alphabet by sending  $X \mapsto A$ ,  $Y \mapsto B$ ,  $Z \mapsto C$ ,  $A \mapsto D$  etc; the text "ET TU, BRUTE" would be encrypted as "BQ QR, YORQB". For longer texts, analyzing the frequency of letters (for given languages) makes breaking this and similar codes a breeze, in particular if you are equipped with a computer.

Another common feature of these ancient methods of encrypting messages is the following: anyone who knows the key, that is, the method with which messages are encrypted, can easily break the code by inverting the encryption. In 1976, Diffie and Hellman suggested the existence of public key cryptography: these are methods for encrypting messages that do not allow you to read encrypted messages even if you know the key. The most famous of all public key cryptosystems is called RSA after its discoverers Ramir, Shamir and Adleman (1978).

Here's the simple idea: assume that Bob wants to receive secure messages; he selects two (large) primes  $p$  and  $q$  and forms their product  $n = pq$ . Bob also chooses an integer  $E < n$  coprime to  $(p-1)(q-1)$ . The integers  $n$  and  $E$  are made public and constitute the key, so everybody can encrypt messages. For decrypting messages, however, one needs to know the prime factors  $p$  and  $q$ , and if  $p$  and  $q$  are large enough (say about 150 digits each) then known factorization methods cannot factor  $n$  in any reasonable amount of time (say 100 years).

How does the encryption work? It is a simple matter to transform any text into a sequence of numbers, for example by using  $a \mapsto 01$ ,  $b \mapsto 02$ ,  $\dots$ , with a

couple of extra numbers for blanks, commas, etc. We may therefore assume that our message is a sequence of integers  $T < n$  (if the text is longer, break it up into smaller pieces). Alice encrypts each integer  $T$  as  $C \equiv T^E \pmod n$  and sends the sequence of  $C$ 's to Bob (by email, say). Now Bob can decrypt the message as follows: since he knows  $p$  and  $q$ , he can form the product  $m = (p - 1)(q - 1)$  and run the Euclidean algorithm on the pair  $(E, m)$  to find an integer  $D$  such that  $DE \equiv 1 \pmod m$ . Now he takes the message  $C$  and computes  $C^D \pmod n$ . The result is  $C^D \equiv (T^E)^D = T^{DE} \pmod n$ , but since  $DE \equiv 1 \pmod m = \phi(n)$ , the theorem of Euler-Fermat shows that  $C^D \equiv T \pmod n$ , and Bob has got the original text that Alice sent him.

Now assume that Celia is eavesdropping. Of course she knows the pair  $(n, E)$  (which is public anyway), and she also knows the message  $C$  that Alice sent to Bob. That does not suffice for decrypting the message, however, since one seems to need an inverse  $D$  of  $E \pmod{(p - 1)(q - 1)}$  to do that; it is likely that one needs to know the factors of  $n$  in order to compute  $D$ .

**Baby Example.** The following choice of  $n = 1073$  with  $p = 29$  and  $q = 37$  is not realistic because this number can be factored easily; its only purpose is to illustrate the method.

So assume that Bob picks the key  $(n, E) = (1073, 25)$ . Alice wants to send the message "miss piggy" to Bob. She starts by transforming the message into a string of integers as follows:

	m	i	s	s		p	i	g	g	y
T	13	9	19	19	27	16	9	7	7	25

Next she encrypts this sequence by computing  $C \equiv T^{25} \pmod n$  for each of these  $T$ : starting with  $13^{25} \equiv 671 \pmod{1073}$ , she finds

T	13	9	19	19	27	16	9	7	7	25
C	671	312	901	901	656	1011	312	922	922	546

Alice sends this string of  $C$ 's to Bob. Knowing the prime factorization of  $n$ , Bob is able to compute the inverse of  $25 \pmod{(p - 1)(q - 1)}$  as follows: he multiplies  $p - 1 = 28$  and  $q - 1 = 36$  to get  $(p - 1)(q - 1) = 28 \cdot 36 = 1008$ . Then he applies the extended Euclidean algorithm to  $(25, 1008)$  and finds  $1 = 25 \cdot 121 - 1008 \cdot 3$ , and this shows that  $D = 121$ .

Now Bob takes the string of  $C$ 's he got from Alice and decrypts them: starting with  $671^{121} \equiv 13 \pmod n$  he can get back the string of  $T$ 's, and hence the original message.

**Remark.** There is a big problem with this baby example: if we encrypt the message letter for letter, then equal letters will have equal code, and the cryptosystem can be broken (if the message is long enough) by analyzing the frequency with which each letter occurs (say in English). This problem vanishes into thin air when we use (realistic) key sizes of about 200 digits: there we encrypt the message in blocks of about 100 letters, and since the chance that any two blocks of 100 letters inside a message coincide is practically 0, an attack based on the frequency of letters will not be successful for keys of this size.



RSA can also be applied to the signature problem. Assume that Alice receives an email from someone claiming to be Bob. How can Alice verify that this is true? Here's the simple trick in a nutshell: both Bob and Alice choose public keys, say  $(n_A, E_A)$  for Alice and  $(n_B, E_B)$  for Bob. Moreover, Alice knows  $D_A$  with  $D_A E_A \equiv 1 \pmod{\phi(n_A)}$ , while Bob knows  $D_B$  with  $D_B E_B \equiv 1 \pmod{\phi(n_B)}$ . Now Bob encrypts his message as above, but instead of sending the T's to Alice, he computes  $U = T^{D_B} \pmod{n_B}$  and sends the U's. In order to decrypt the message, Alice computes first  $T \equiv U^{E_D} \pmod{n_B}$  and then decrypts the T's as in the original version of RSA using her  $D_A$ . If this works, then Alice can be sure that the message came from Bob because in order to encrypt the message this way, the sender has to know  $D_B$ .

## 7.5 Pollard's $p - 1$ -Factorization Method

Pollard is definitely the world champion in inventing new methods for factoring integers. One of his earliest contributions were the  $p - 1$ -method (ca. 1974), his  $\rho$ -method followed shortly after, and his latest invention is the number field sieve (which is based on ideas from algebraic number theory).

The idea behind Pollard's  $p - 1$ -method is incredibly simple. Assume that we are given an integer  $N$  that we want to factor. Fix an integer  $a > 1$  and check that  $\gcd(a, N) = 1$  (should  $d = \gcd(a, N)$  be not trivial, then we have already found a factor  $d$  and continue with  $N$  replaced by  $N/d$ ).

Let  $p$  be a factor of  $N$ ; by Fermat's Little Theorem we know that  $a^{p-1} \equiv 1 \pmod{p}$ , hence  $D := \gcd(a^{p-1} - 1, N)$  has the properties  $p \mid D$  and  $D \mid N$ . Thus  $D$  is a nontrivial factor of  $N$  unless  $D = N$  (which should not happen too often).

The procedure above is not much of a factorization algorithm as long as we have to know the prime factor  $p$  beforehand. The prime  $p$  occurs at two places in the method above: first, as the modulus when computing  $a^{p-1} \pmod{p}$ . But this problem is easily taken care of because we may simply compute  $a^{p-1} \pmod{N}$ . It is more difficult to get rid of the  $p$  in the exponent: the fundamental observation is that we can replace the exponent  $p - 1$  above by any multiple, and  $D$  still will be divisible by  $p$  (note though that the chance that  $D = N$  has become slightly larger). Does this help us? Not always; assume, however, that  $p - 1$  is the product of *small* primes (say of primes below a bound  $B$  that in practice can be taken to be  $B = 10^5$  or  $B = 10^6$ , depending on the computing power of your hardware). Then it is not too hard to come up with good candidates for multiples of  $p - 1$ : we might simply pick  $k = B!$ , or, in a similar vein,

$$k = \prod_i p_i^{a_i}, \quad \text{where } p_i^{a_i} \leq B < p_i^{a_i+1}. \quad (7.3)$$

If we  $(p - 1) \mid k$ , then  $a^k \equiv 1 \pmod{p}$ , hence  $p \mid D = \gcd(a^k - 1, N)$ .

Thus the following algorithm has a good chance of finding those factors  $p$  of  $N$  for which  $p - 1$  has only small prime factors:

1. Pick  $a > 1$  and check that  $\gcd(a, N) = 1$
2. Choose a bound  $B$ , say  $B = 10^4, 10^5, 10^6, \dots$
3. Pick  $k$  as in (7.3) and compute  $D = \gcd(a^k - 1, N)$ .

Note that the computation of  $a^k$  can be done modulo  $N$ ; if  $p \mid N$  and  $(p - 1) \mid k$ , then  $a^k \equiv 1 \pmod{p}$ , hence  $p \mid D$ .

If  $D = 1$ , we may increase  $k$ ; if  $D = N$ , we can reduce  $k$  and repeat the computation.

Among the record factors found by the  $p - 1$ -method is the 37-digit factor  $p = 6902861817667290192729108442204980121$  of  $71^{77} - 1$  with  $p - 1 = 2^3 \cdot 3^3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 \cdot 401 \cdot 409 \cdot 3167 \cdot 83243 \cdot 83983 \cdot 800221 \cdot 2197387$  discovered by Dubner.

A list of record factors can be found at

<http://www.users.globalnet.co.uk/~aads/Pminus1.html>

Here's a baby example: take  $N = 1769$ ,  $a = 2$  and  $B = 6$ . Then we compute  $k = 2^2 \cdot 3 \cdot 5$  and we find  $2^{60} \equiv 306 \pmod{1769}$ ,  $\gcd(305, 1769) = 61$  and  $N = 29 \cdot 61$ . Note that  $61 - 1 = 2^2 \cdot 3 \cdot 5$ , so the factor 61 was found, while  $29 - 1 = 2^2 \cdot 7$  explains why 29 wasn't (although  $29 < 61$ ).

Another large class of factorization algorithms is based on an algorithm invented by Fermat: the idea is to write an integer  $n$  as a difference of squares. If  $n = x^2 - y^2$ , then  $n = (x - y)(x + y)$ , and unless this is the trivial factorization  $n = 1 \cdot n$ , we have found a factor.

Another baby example: take  $n = 1073$ ; then  $\sqrt{n} = 32.756\dots$ , so we start by trying to write  $n = 33^2 - y^2$ . Since  $33^2 - 1073 = 16$ , we find  $n = 33^2 - 4^2 = (33 - 4)(33 + 4) = 29 \cdot 37$ . If the first attempt would have been unsuccessful, we would have tried  $n = 34^2 - y^2$ , etc.

In modern algorithms (continued fractions, quadratic sieve, number field sieve) the equation  $N = x^2 - y^2$  is replaced by a congruence  $x^2 \equiv y^2 \pmod{N}$ : if we have such a thing, then  $\gcd(x - y, N)$  has a good chance of being a nontrivial factor of  $N$ . The first algorithm above constructed such pairs  $(x, y)$  by computing the continued fraction expansion of  $\sqrt{n}$  (which we have not discussed), the number field sieve produces such pairs by factoring certain elements in algebraic number fields.

## Exercises

- 7.1 Compute the addition and multiplication tables for the ring  $\mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}$ , and compare the result to those for  $\mathbb{Z}/4\mathbb{Z}$ .
- 7.2 Do the same exercise for the rings  $\mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/3\mathbb{Z}$  and  $\mathbb{Z}/6\mathbb{Z}$ .
- 7.3 Find all integers with  $\phi(m) = 6$ .
- 7.4 Show that  $m$  is prime if and only if  $\phi(m) = m - 1$ .

7.5 Solve the system of congruences

$$x \equiv 12 \pmod{13},$$

$$x \equiv 7 \pmod{19}.$$