Module 5: Basic Number Theory

Theme 1: Division

Given two integers, say a and b, the quotient b/a may or may not be an integer (e.g., $16/4 = 4$ but $12/5 = 2.4$). Number theory concerns the former case, and discovers criteria upon which one can decide about divisibility of two integers.

More formally, for $a \neq 0$ we say that a **divides** b if there is another integer k such that

$$b = ka,$$

and we write $a|b$. In short:

$$a|b \quad \text{if and only if} \quad \exists_{k \in \mathbf{Z}} \;\; b = ka.$$

This simple definition leads to many properties of divisibility. For example, let us establish the following lemma.

Lemma 1 *If $a|b$ and $a|c$, then $a|(b+c)$.*

Proof. We give a direct proof. From the definition of divisibility and the hypotheses we know that there are integers t and s such that

$$b = ta, \quad c = sa.$$

Hence

$$b + c = a(s + t).$$

Since $s + t$ is an integer, we prove that $a|(b + c)$.

Exercise 5A: Prove the following two facts:
1. If $a|b$, then $a|bc$ for all integers c.
2. If $a|b$ and $b|c$, then $a|c$.

We already noted that an integer may be or not divisible by another integer. However, when dividing one number by another there is always a quotient and a remainder. More precisely, *if a and d are positive integers then there is a* **unique** *q and r such that*

$$a = dq + r$$

where $0 \leq r < d$ is a remainder. Observe that the remainder can take only d values $0, 1, \ldots, d - 1$.

Theme 2: Primes

Primes numbers occupy very prominent role in number theory. A **prime** number p is an integer greater than 1 that is divisible *only* by 1 and itself. A number that is not prime is called **composite**.

Example 1: The primes less than 100 are:

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97.$$

How many primes are there? We first prove that there are infinite number of primes.

Theorem 1. *There are infinite number of primes.*

Proof. We provide a proof by contradiction. Actually, it is due to Euclid and it is more than 2000 years old. Let us assume that there is a *finite* number of primes, say, $2, 3, 5, \ldots, p_k$ where p_k is the *largest* prime (there is the largest prime since we assumed there are only finitely many of them). Construct another number

$$M = 2 \cdot 3 \cdot 5 \cdots p_k + 1$$

which is a product of all primes plus one. First, observe that none of the primes $2, 3, 5, \ldots, p_k$ can divide M, since the remainder of dividing M by any of the primes is equal to 1. Since every number, including M, is divisible by at least two numbers, 1 and itself, there must be another prime, possible M itself, that is not among the primes $2, 3, 5, \ldots, p_k$. This contradicts the assumption that $2, 3, 5, \ldots, p_k$ are the only primes.

But how many primes are there smaller than n, where n is a fixed number. This is a very difficult problem that was solved only in the last century. Basically, there are approximately about $n/\log(n)$ primes smaller than n. For example, there are 25 primes smaller than 100, and $100/\log(100) \approx 22$.

Primes are important since every integer can be represented as a product of primes. This is known as the **Fundamental Theorem of Arithmetics** and we will prove it below.

Example 2: Observe that

$$\begin{aligned} 100 &= 2 \cdot 2 \cdot 5 \cdot 5 = 2^2 5^2, \\ 381 &= 3 \cdot 127, \\ 888 &= 2^3 \cdot 3 \cdot 37. \end{aligned}$$

Theorem 2. [**Fundamental Theorem of Arithmetics**] *Every positive integer can be written uniquely as the product of primes where the prime factors are written in order of increasing size, that is, if n is a natural numbers and $p_1 < p_2 < \cdots < p_m$ are distinct primes, then*

$$n = p_1^{e_1} \cdot p_2^{e_2} \cdots p_m^{e_m}$$

where e_i are exponents of p_i (i.e., the number of times p_i occurs in the factorization of n).

Proof. We give an indirect proof. Let us assume that there are two *different* prime factorizations of n, say

$$\begin{aligned} n &= p_1^{e_1} \cdot p_2^{e_2} \cdots p_m^{e_m} \\ n &= q_1^{d_1} \cdot q_2^{d_2} \cdots q_r^{d_r} \end{aligned}$$

where $q_1 < \cdots < q_r$ are primes. Since we factorize the same number n we must have

$$p_1^{e_1} \cdot p_2^{e_2} \cdots p_m^{e_m} = q_1^{d_1} \cdot q_2^{d_2} \cdots q_r^{d_r}.$$

We first prove that $p_1 = q_1$. If $p_1 \neq q_1$, then p_1 can not divide any of the primes $q_1, \ldots, q_r$ (we say that p_1 is *relatively prime* to all $q_1, \ldots, q_r$). Indeed, since p_1 and $q_1, \ldots, q_r$ are primes, none of them equal, then they must be relatively prime. But, then p_1 cannot divide $n = q_1^{d_1} \cdot q_2^{d_2} \cdots q_r^{d_r}$ which is nonsense since $n = p_1^{e_1} \cdot p_2^{e_2} \cdots p_m^{e_m}$. Thus, we must conclude that $p_1 = q_1$.

Now we prove that $e_1 = d_1$ provided $p_1 = q_1$ that we just established above. Again, assume contrary that $e_1 < d_1$, say $d_1 = e_1 + h$, $h > 0$. Then after dividing everything by $p_1^{e_1}$ we obtain

$$p_2^{e_2} \cdots p_m^{e_m} = q_1^{h} \cdot q_2^{d_2} \cdots q_r^{d_r}.$$

But then the right-hand side of the above is divisible by q_1 while the left-hand side is not, which is impossible since there is an equality sign between the left-hand side and the right-hand side of the above. This completes the proof.

How to find out whether an integer is a prime or not? Unfortunately, there is no fast way of doing it (i.e., there is no efficient algorithm), but one can use some properties of primes and composite numbers to speed up the process. Here is one useful result.

Lemma 2.*If n is a composite integer, then n has a prime divisor less than or equal to $\sqrt{n}$.*

Proof. Since n is a composite integer, it must have a factor d such that $1 < d < n$, that is, $n = d \cdot r$ where $r > 1$ is an integer. Let us now assume contrary that $d > \sqrt{n}$ and $r > \sqrt{n}$. But then

$$d \cdot r > \sqrt{n}\sqrt{n} = n$$

which is the desired contradiction since we assumed that $n = dr$. We must conclude that n has at least one divisor not exceeding $\sqrt{n}$. This divisor is prime or not. If it is not prime, it must have a prime divisor, which certainly must be smaller than $\sqrt{n}$.

We can use this lemma, in its contrapositive form, to decide whether n is a prime or not. Indeed. the above lemma is equivalent to: *if n has no prime divisor less than or equal to $\sqrt{n}$, then n is a prime number.*

Example 3: Let us show that 107 is a prime number. If 107 would be composite, then it has had prime divisor smaller than $\sqrt{107} \approx 10.34$. Primes smaller than 10 are $2, 3, 5$, and 7. None of it divides 107, thus it 107 must be a prime number.

There were several attempts to find a systematic way of computing prime numbers. Euclid suggested that $(k+1)$-st prime can be computed recursively as follows:

$$\begin{aligned} e_1 &= 2, \\ e_{k+1} &= e_1 e_2 \cdots e_k + 1. \end{aligned}$$

For example, the first few numbers are

$$\begin{aligned} e_2 &= 2 + 1 = 3, \\ e_3 &= 2 \cdot 3 + 1 = 7, \\ e_4 &= 2 \cdot 3 \cdot 7 + 1 = 43. \end{aligned}$$

This is an example of a recurrence that we already encountered in the previous module. All numbers computed so far are primes. But, unfortunately,

$$e_5 = 2 \cdot 3 \cdot 7 \cdot 43 + 1 = 1807 = 13 \cdot 139$$

is not a prime.

In the seventeenth century, a French mathematician Marin Marsenne suggested that $2^p - 1$ is prime provided p is prime. Unfortunately,

$$2^{11} - 1 = 2047 = 23 \cdot 89.$$

From now on we shall work under the assumption that there is no easy, simple and fast algorithm to compute prime numbers.

Theme 3: Greatest Common Divisor

The largest divisor that divides both m and n is called the **greatest common divisor** of n and m. It is denoted as $\gcd(m, n)$. Formally:

$$\gcd(m, n) := \max\{k : \ k|m \text{ and } k|n\}.$$

Example 4: What is the greatest common divisor of 24 and 36. One way of finding it is to list all divisors of 24 and 36 and pick up the largest common to both lists. For example,

$$\begin{aligned} \text{divisors of } 24 &= \{1, 2, 3, 4, 6, 8, 12, 24\}, \\ \text{divisors of } 36 &= \{1, 2, 3, 4, 6, 9, 12, 18, 36\}. \end{aligned}$$

Thus $\gcd(24,36) = 12$. Another, more systematic way is to do prime factorization of both numbers and pick up the largest common factors. In our case,

$$\begin{aligned} 24 &= 2^3 \cdot 3, \\ 36 &= 2^2 \cdot 3^2. \end{aligned}$$

Thus

$$\gcd(24,36) = 2^2 \cdot 3 = 12.$$

Generalizing the last example, let

$$\begin{aligned} m &= p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}, \\ n &= p_1^{b_1} p_2^{b_2} \cdots p_k^{b_k} \end{aligned}$$

be prime factorizations with possible zero exponents. Then

$$\gcd(m,n) = p_1^{\min\{a_1,b_1\}} p_2^{\min\{a_2,b_2\}} \cdots p_k^{\min\{a_k,b_k\}}$$

where $\min\{x,y\}$ is the minimum of x and y. Indeed, take the last example to see that

$$\gcd(24,36) = 2^{\min\{2,3\}} 3^{\min\{1,2\}}.$$

Exercise 5B: Let us define the **least common multiple** of m and n as the smallest positive integer that is divisible by both m and n. It is denoted as $\operatorname{lcm}(m,n)$ (e.g., $\operatorname{lcm}(5,8) = 40$). Prove that for any positive integers m and n

$$m \cdot n = \gcd(m,n) \cdot \operatorname{lcm}(m,n).$$

We need some more definitions. Two integers, say m and n, may be composite but the only common divisor of both is 1. In such a case we say that m and n are *relatively prime*. More generally:

Definition 1. *The integers $a_1, a_2, \ldots, a_k$ are* **pairwise relatively prime** *if*

$$\gcd(a_i, a_j) = 1, \qquad \forall\, 1 \le i < j \le k.$$

Unlike finding primes, there is an efficient **algorithm** (a procedure) that finds the greatest common divisor. We start with an example.

Example 5: Find $\gcd(91, 260)$. We first divide 260 by 91 to find

$$260 = 2 \cdot 91 + 78.$$

Observe that any divisor of 91 and 260 must also be a divisor of $260 - 2 \cdot 91 = 78$, and vice versa any divisor of 91 and 78 must be a divisor of $260 = 2 \cdot 91 + 78$. (Indeed, if d is a divisor of 260 and 91, then there are integers k and l such that $260 = d \cdot k$ and $91 = d \cdot l$, hence $260 - 2 \cdot 91 = d(k - 2l)$, so $260 - 2 \cdot 91 = 78$ is divisible by d.) Thus we concluded that

$$\gcd(91, 260) = \gcd(78, 91).$$

We now repeat this procedure: we divide 91 by 78 to get

$$91 = 1 \cdot 78 + 13.$$

Again any divisor of 78 and 91 must be a divisor of $91 - 78 = 13$, and vice versa. This means that

$$\gcd(91, 260) = \gcd(78, 91) = \gcd(13, 78).$$

But

$$78 = 13 \cdot 6,$$

hence finally

$$\gcd(91, 260) = \gcd(78, 91) = \gcd(13, 78) = \gcd(0, 13) = 13$$

and we conclude that $\gcd(91, 260) = 13$.

From the last example, we should conclude that the greatest common divisor of m and $n > m$ is the same as the greatest common divisor of m and the remainder of the division of n by m (i.e., $n = q \cdot m + r$, where q is an integer and $0 \leq r < m$). Indeed, if d is a divisor of m and n, then it must also divides $r = n - q \cdot m$, and vice versa if d divides m and r, then it divides $n = m \cdot q + r$. Therefore,

$$\gcd(m, n) = \gcd(r, m).$$

In previous modules we have used an abbreviation for a remainder. Indeed, we write

$$r := n \ \mathbf{mod}\ m$$

where $n = q \cdot m + r$. This is called **modular arithmetic** and we will be devoted the next section it. For now, we just use the fact that the remainder r can be also written as $n \ \mathbf{mod}\ m$. Then the last equation, can be expressed as

$$\gcd(m, n) = \gcd(n \ \mathbf{mod}\ m, m). \tag{1}$$

From the example above, we conclude that we can use (1) successively until we reach $\gcd(0, m) = m$.

In summary, we design the following algorithm that computes $\gcd(m, n)$:

ALGORITHM: **The Euclidean Algorithm**

```
$x := m$
$y := n$
while $y \neq 0$ do $r := x \,\mathbf{mod}\, y$
$x := y$
$y := r$
end
$\gcd(m,n) := x$.
```

Example 6: Find $\gcd(414, 662)$. According to the Euclidean algorithm we proceed as follows:

$$\gcd(414, 662) = \gcd(248, 414) = \gcd(166, 248) = \gcd(82, 166) = \gcd(82, 2) = \gcd(2, 0) = 2.$$

Theme 4: Modular Arithmetic

We have already seen in previous modules modular arithmetic. It is about the remainder of an integer when it is divided by another specific natural integer. It occurs in many applications (e.g., when counting time over a 24-hour clock since after 24:00 we have 1 am, 2 am, etc.).

We start with a definition.

Definition 2. (i) *Let n be an integer and M be a positive integer. We denote by*

$$r := n \,\mathbf{mod}\, M$$

the remainder r when n is divided by M, that is,

$$n = q \cdot M + r$$

where q is an integer and $0 \leq r < M$.

(ii) *Let n and m be integers and M a positive integer. We say that n* **is congruent to** *m* **modulo** *M if M divides $n - m$. We shall write*

$$n \equiv m \,\mathbf{mod}\, M \quad \textit{if and only if} \quad M|(n - m).$$

If n are m are not congruent modulo M, then we write $n \not\equiv m \,\mathbf{mod}\, M$.

Example 7: We have

$$18 \,\mathbf{mod}\, 5 = 3, \qquad -145 \,\mathbf{mod}\, 9 = 8.$$

We also have

$$17 \equiv 5 \,\mathbf{mod}\, 6, \qquad 24 \not\equiv 14 \,\mathbf{mod}\, 6.$$

Exercise 5C: Find $134 \,\mathbf{mod}\, 8 =?$. Is $-23 \,\mathbf{mod}\, 4 = 4$?

The following result is useful when computing congruences.

Theorem 3. *Let* $a \equiv b \;\mathbf{mod}\; m$ *and* $c \equiv d \;\mathbf{mod}\; m$. *Then*

$$\begin{aligned} a + c &\equiv b + d \;\mathbf{mod}\; m, &(2)\\ ac &\equiv bd \;\mathbf{mod}\; m. &(3)\end{aligned}$$

Proof. Since $a \equiv b \;\mathbf{mod}\; m$ and $c \equiv d \;\mathbf{mod}\; m$, hence there are integers s and t such that

$$\begin{aligned} b &= sm + a,\\ d &= c + tm.\end{aligned}$$

Therefore

$$\begin{aligned} b + d &= (a + c) + m(s + t),\\ bd &= (a + sm)(c + tm) = ac + m(at + cs + stm)\end{aligned}$$

which prove (2) and (3).

Example 8: Let $7 \equiv 2 \;\mathbf{mod}\; 5$ and $11 \equiv 1 \;\mathbf{mod}\; 5$. Then

$$18 = 7 + 11 \equiv 2 + 1 \;\mathbf{mod}\; 5$$

and

$$77 = 7 \cdot 11 \equiv 2 \cdot 1 \;\mathbf{mod}\; 5.$$

From Theorem 3 we conclude that

$$\begin{aligned} (a + b) \;\mathbf{mod}\; m &= [(a \;\mathbf{mod}\; m) + (b \;\mathbf{mod}\; m)] \;\mathbf{mod}\; m, &(4)\\ (a \cdot b) \;\mathbf{mod}\; m &= [(a \;\mathbf{mod}\; m) \cdot (b \;\mathbf{mod}\; m)] \;\mathbf{mod}\; m. &(5)\end{aligned}$$

Identities (4)–(5) are useful when one needs to compute modulo m over large numbers or products of large numbers. For example, let $a = 123$ and $b = 234$. Then

$$120 \cdot 234 \;\mathbf{mod}\; 5 = (123 \;\mathbf{mod}\; 5) \cdot (234 \;\mathbf{mod}\; 5) \;\mathbf{mod}\; 5 = 3 \cdot 4 \;\mathbf{mod}\; 5 = 2.$$

In fact, (5) is often used in the following form

$$a^s \;\mathbf{mod}\; m = (a \;\mathbf{mod}\; m)^s \;\mathbf{mod}\; m.$$

Let us compute $572^{29} \;\mathbf{mod}\; 713$. If one tries to estimate this directly on a computer, overflow will likely occur since 572^{29} is a huge number. But let us use (5). We first represent the exponent 29 as

$$29 = 2^4 + 2^3 + 2^2 + 2^0.$$

We now compute 572 to each of the powers $16, 8, 4$ and 1 modulo 713. Here is the calculation (observe how easy it is!):

$$\begin{aligned}
572^2 \ \mathbf{mod}\ 713 &= 327184 \ \mathbf{mod}\ 713 = 630, \\
572^4 \ \mathbf{mod}\ 713 &= (572^2 \ \mathbf{mod}\ 713)^2 \ \mathbf{mod}\ 713 = 630^2 \ \mathbf{mod}\ 713 = 472, \\
572^8 \ \mathbf{mod}\ 713 &= (572^4 \ \mathbf{mod}\ 713)^2 \ \mathbf{mod}\ 713 = 472^2 \ \mathbf{mod}\ 713 = 328, \\
572^{16} \ \mathbf{mod}\ 713 &= (572^8 \ \mathbf{mod}\ 713)^2 \ \mathbf{mod}\ 713 = 328^2 \ \mathbf{mod}\ 713 = 634, \\
572^{29} \ \mathbf{mod}\ 713 &= (572^{16} \ \mathbf{mod}\ 713) \cdot (572^8 \ \mathbf{mod}\ 713) \cdot (572^4 \ \mathbf{mod}\ 713) \cdot (572 \ \mathbf{mod}\ 713) \ \mathbf{mod}\ 713 \\
&= 634 \cdot 328 \cdot 472 \cdot 572 \ \mathbf{mod}\ 713 = 113.
\end{aligned}$$

Theme 5: Applications

We shall discuss here some applications of numbers theory, namely, hashing, pseudo random generators, and cryptosystems based on modular arithmetic.

Hashing

Often one needs a fast methods of locating a given record in a huge set of records. **Hashing** is a possible solution. It works as follows. Every record has a **key**, k, which uniquely identifies it. A **hashing function** $h(k)$ maps the set of keys into the available memory locations.

In practice, the most common hashing function is

$$h(k) = k \ \mathbf{mod}\ m$$

where m is the size of the memory.

Example 9: Let $m = 111$ and let keys be social security numbers of students. In particular,

$$\begin{aligned}
h(064212848) &= 064212848 \ \mathbf{mod}\ 111 = 14 \\
h(037149212) &= 037149212 \ \mathbf{mod}\ 111 = 65.
\end{aligned}$$

Observe that hashing is *not* one-to-one function, hence some records may be hashed into the same location. For example,

$$h(107405723) = 107405723 \ \mathbf{mod}\ 111 = 14.$$

Thus two records are mapped into the location 14. Since this location was already occupied by the previous record, the new collided record is moved to the next empty location modulo $m = 111$. In our case, it is at memory location 15.

Pseudo Random Number Generators

In many applications, including hashing, one needs to generate numbers that look randomly. For example, in hashing we want to spread out uniformly all records over the memory so to minimize the number of collisions. We should point out that most random generators compute *deterministically* numbers, therefore, we call them **pseudo** random generators. We require, however, that a statistical test applied to them will not distinguish these numbers from randomly generated numbers.

The most common procedure to generate pseudo random numbers is the **linear congruential method**. In this method we choose (very carefully) the **modulus** m, **multiplier** a, **increment** c, and **seed** x_0 with $2 \le a < m$, $0 \le c < m$, and $0 \le x_0 < m$. Then we generate recursively a sequence x_n as

$$x_{n+1} = (ax_n + c) \ \mathbf{mod} \ m$$

with x_0 given. Observe that $0 \le x_n < m$, hence at most after m generations a repetition occurs. Of course, this is not good for *random* generations, and one must select very carefully the parameters a, c and m (which should be large) to obtain a long sequence without a repetition.

The following result is known.

Theorem 4. [**T. Hull and A. Dobel, 1962**]*The linear congruential generator has a full period (i.e., there is no repetition in the first m generations) if and only if the following three conditions hold:*

(i) *Both m and c are relatively prime, that is,* $\gcd(m, c) = 1$.

(ii) *If q is a prime number that divides m, then q divides* $a - 1$.

(iii) *If 4 divides m, then 4 divides* $a - 1$.

Cryptology

One of the most important application of congruences is in **cryptology**, which is a study of secret messages. The first encryption algorithms were very simple. For example, Julius Caesar designed an encryption system by shifting each letter three letters in the alphabet. Mathematically speaking, in this case the encryption function $f(p)$ is defined as

$$f(p) = (p + 3) \ \mathbf{mod} \ 26.$$

Then decryption is merely finding the inverse function f^{-1}, which in this case is

$$F^{-1}(p) = (p - 3) \ \mathbf{mod} \ 26.$$

The above encryption system is too easy to break. Therefore, in mid-1970 the concept of **public key cryptosystem** was introduced. In such a system, every person can have a publicly known encryption key to send encrypted message, but only those who have secret key can decrypt the message. We describe below a system known as the **RSA encryption system** (RSA name is built from the initials of the inventors Rivest, Shamir and Adleman).

In the RSA system, the message M to be sent is first transformed into an integer representing it (with some abuse of notation we denote such an integer by M). The RSA is based on modular exponentiation modulo of the product of two large primes, say p and q. Define $n = pq$ and $\phi = (p-1)(q-1)$. In practice, p and q have 100 digits each, thus n has 200 digits. Define now an exponent e as

$$\gcd(e, \phi) = 1,$$

that is, e is relatively prime to $\phi = (p-1)(q-1)$. The cipher text C of the original message M is computed as follows

$$C = M^e \,\mathbf{mod}\, n \tag{6}$$

The RSA **decryption** works as follows: We first find a number d defined as

$$de = 1 \,\mathbf{mod}\, \phi = (p-1)(q-1).$$

The number d is called **inverse** of e modulo ϕ. It should be underlined that d can be found fast (based on the Euclidean algorithm) only if one knows **both** primes p and q, not the product pq. Then, it can be proved (see below) that

$$C^d \equiv M \,\mathbf{mod}\, n = pq. \tag{7}$$

Example 10: Let us encrypt the message $STOP$ using the RSA with $p = 43$ and $q = 59$. Thus $n = 43 \cdot 59 = 2537$, and one finds $e = 13$ since $\gcd(13, 42 \cdot 58) = 1$.

We now transform the message $STOP$ into its numerical equivalent (where $A = 00$, $B = 02, \ldots Z = 25$) and group them in pairs. We obtain

$$1819 \quad 1415.$$

We will encrypt each of the two blocks separately. We have

$$\begin{aligned} 1819^{13} \,\mathbf{mod}\, 2537 &= 2081, \\ 1415^{13} \,\mathbf{mod}\, 2537 &= 2182. \end{aligned}$$

Hence, the encrypted message is $C = 2081$ 2182.

Now, to decrypt it, we first find the inverse d. Using the Euclidean algorithm (and **knowing** $p = 43$ $q = 59$) we compute that $d = 937$. Then (with $n = 2537$)

$$2081^{937} \,\mathbf{mod}\, 2537 = 1819,$$

and

$$2182^{937} \,\mathbf{mod}\, 2537 = 1415.$$

hence, we recover the original message.

Mathematics behind RSA

In this subsection, we present in some details mathematical ideas used in the construction of the RSA algorithm. Our main goal is to justify mathematically the decoding procedure (7).

Let us start with introducing an **inverse modulo** m. We say that $\bar{a}$ is an inverse of a modulo m if

$$\bar{a}a \equiv 1 \;\mathbf{mod}\; m.$$

In order to compute the inverse, we must plunge into another aspect of number theory. We claim that *for any positive a and b there exist integers s and t such that*

$$\gcd(a, b) = sa + tb. \tag{8}$$

We explain how to construct these two numbers on an example.

Example 11: Let us use Euclidean algorithm to compute $\gcd(396, 504)$. We proceed according to the algorithm as follows:

$$\begin{aligned} 504 &= 396 + 108 \\ 396 &= 3 \cdot 108 + 72 \\ 108 &= 72 + 36 \\ 72 &= 2 \cdot 36, \end{aligned}$$

Thus $\gcd(396, 504) = 36$. To find the representation (8) we work backward the Euclidean algorithm starting from the next-to-last devision above, that is,

$$\begin{aligned} \gcd(396, 504) = 36 &= 108 - 72 \\ &= 108 - (396 - 3 \cdot 108) = 4 \cdot 108 - 396 \\ &= 4(504 - 396) - 396 = 4 \cdot 504 - 5 \cdot 396 \\ &= (-5)a + (4)b \end{aligned}$$

where $a = 396$ and $b = 504$. Thus $s = -5$ and $t = 4$ in the representation (8). It is not much harder to prove (8) in general terms.

Now we can go back to the inverse modulo m construction. Let us *assume* that $\gcd(a, m) = 1$. Then from the fact just proved we conclude that there must exist integers s and t such that

$$sa + tm = 1.$$

This certainly implies that

$$sa + tm \equiv 1 \;\mathbf{mod}\; m.$$

But since m divides tm we conclude that

$$sa \equiv 1 \textbf{ mod } m.$$

Consequently $s = \bar{a}$ is the inverse of a modulo m.

In summary, we've just established the following result.

Theorem 5. *If $m > 1$ and $\gcd(a, m) = 1$ (i.e., a and m are relatively prime), then an inverse of a modulo m exists and it is equal to s in the following representation of* $\gcd(a, m) = 1$

$$sa + tm = 1$$

which can be found efficiently by the Euclidean algorithm.

Example 12: Let's find the inverse of 3 modulo 7. Since $\gcd(3, 7) = 1$, the inverse exists, and the Euclidean algorithm gives:

$$7 = 2 \cdot 3 + 1$$

hence

$$-2 \cdot 3 + 1 \cdot 7 = 1$$

and the inverse of 3 modulo 7 is equal to -2.

We need two more results before we can explain the decryption algorithm of RSA. The first one goes back to ancient Chinese and Hindu mathematicians and it is known as the *Chinese Remainder Theorem*. Here is the problem: let $m_1, m_2, \ldots, m_n$ be pairwise relatively prime positive integers. Find a solution x modulo $m = m_1 m_2 \cdots m_n$ of the following system;

$$\begin{aligned} x &\equiv a_1 \textbf{ mod } m_1, \\ x &\equiv a_2 \textbf{ mod } m_2, \\ &\vdots \\ x &\equiv a_n \textbf{ mod } m_n, \end{aligned}$$

We now construct a solution to the above system of congruences. Let us define for $k = 1, \ldots, n$

$$M_k = \frac{m}{m_k} = m_1 \cdots m_{k-1} m_{k+1} \cdots m_n.$$

Observe that $\gcd(M_k, m_k) = 1$. Therefore, by Theorem 5 there exists inverse y_k of M_k modulo m_k, that is,

$$M_k y_k \equiv 1 \textbf{ mod } m_k.$$

Let us now define

$$x = a_1 M_1 y_1 + a_2 M_2 y_2 + \cdots + a_n M_n y_n \textbf{ mod } m. \tag{9}$$

We claim it is a simultaneous solution of the above system modulo m. Indeed, we first observe that $M_j \equiv 0 \mathbf{\ mod\ } m_k$ for $j \neq k$. But

$$x \equiv a_k M_k y_k \equiv a_k \mathbf{\ mod\ } m_k$$

since $M_k y_k \equiv 1 \mathbf{\ mod\ } m_k$. Thus we have shown that (9) is a simultaneous solution of the above n congruences. This is called the *Chinese Remainder Theorem*.

Example 13: Solve

$$\begin{aligned} x &\equiv 2 \mathbf{\ mod\ } 3, \\ x &\equiv 3 \mathbf{\ mod\ } 5, \\ x &\equiv 2 \mathbf{\ mod\ } 7. \end{aligned}$$

We have $m = 3 \cdot 5 \cdot 7 = 105$, and $M_1 = 35$, $M_2 = 21$ and $M_3 = 15$. We find that $y_1 = 2$ is inverse of $M_1 = 35$ modulo 3, $y_2 = 1$ is inverse of M_2 modulo 5, and $y_3 = 1$ is an inverse of M_3 modulo 7. Thus the solution of the above system of congruences

$$\begin{aligned} x &\equiv 2 \cdot 35 \cdot 2 + 3 \cdot 21 \cdot 1 + 2 \cdot 15 \cdot 1 \mathbf{\ mod\ } 105 \\ &= 233 \equiv 23 \mathbf{\ mod\ } 105, \end{aligned}$$

thus the solution $x = 23$.

Finally, we quote (without a proof) the *Fermat Little Theorem*.

Theorem 6. [**Fermat's Little Theorem**] *If p is a prime number and a is an integers not divisible by p, then*

$$a^{p-1} \equiv 1 \mathbf{\ mod\ } p$$

or equivalently

$$a^p \equiv a \mathbf{\ mod\ } p.$$

Now, we are ready to explain the decryption procedure (7) of the RSA algorithm. We recall that d is inverse of e modulo $\phi = (p-1)(q-1)$, that is,

$$de \equiv 1 \mathbf{\ mod\ } \phi.$$

This implies that there is an integer k such that

$$de = 1 + k\phi.$$

Therefore by the Fermat theorem

$$C^d = M^{de} = M \cdot (M^{p-1})^{k(q-1)} \equiv M \cdot 1 \equiv M \mathbf{\ mod\ } p,$$

and

$$C^d = M \cdot (M^{q-1})^{k(p-1)} \equiv M \cdot 1 \equiv M \ \mathbf{mod}\, q$$

since $M^{p-1} \equiv 1 \ \mathbf{mod} p$ and $M^{q-1} \equiv 1 \ \mathbf{mod} q$ by Fermat's theorem. But $\gcd(p, q) = 1$, hence it follows from the Chinese Remainder Theorem that

$$C^d \equiv M \ \mathbf{mod}\, pq$$

as desired.

Assignment 5.1: Basic Number Theory Problems

Each assignment is worth 10 points.

1. Show that if a, b and $c \neq 0$ are integers such that $(ac)|(bc)$, then $a|b$.

2. Find the prime factorization of $10!$.

3. Use the Euclidean algorithm to find

(a) $\gcd(1529, 14039)$,

(b) $\gcd(1111, 11111)$.

4. Find an inverse of 2 modulo 17.

5. Encrypt the message ATTACK using the RSA system with $n = 43 \cdot 59$ and $e = 13$, translating each letter into integers (where $A = 00$, $B = 01, \ldots Z = 25$) and grouping pairs of integers, as we did in our Example 10.

Solutions to Exercises

Solution to Exercise 5A

We first prove that *if* $a|b$*, then* $a|bc$ *for all integers* c*.* Indeed, since $a|b$ there must be an integer k such that $b = k \cdot a$. This implies $b \cdot c = k \cdot c \cdot a$, hence $a|bc$ for any integer c.

Now we prove *if* $a|b$ *and* $b|c$*, then* $a|c$. From the hypotheses we conclude that there are integers k and l such that $b = ka$ and $c = bl$. Therefore, $c = bl = kla$, hence $a|c$.