

Number Theory: Branch of mathematics concerned with properties of the positive integers (1, 2, 3, ...).


Discrete Mathematics, Chapter 4: Number Theory and Cryptography

Nouman M Durrani, FAST NUCES, Karachi.

Courtesy: Richard Mayr



Outline

- 1 Divisibility and Modular Arithmetic
 - 2 Primes and Greatest Common Divisors
 - 3 Solving Congruences
 - 4 Cryptography
- 

Terminology

- *Theorem*: A statement that has been proven to be true.
- *Axioms, postulates, hypotheses, premises*: Assumptions (often unproven) defining the structures about which we are reasoning.
- *Rules of inference*: Patterns of logically valid deductions from hypotheses to conclusions.
- *Lemma*: A minor theorem used as a stepping-stone to proving a major theorem.
- *Corollary*: A minor theorem proved as an easy consequence of a major theorem.
- *Conjecture*: A statement whose truth value has not been proven. (A conjecture may be widely believed to be true, regardless.)
- *Theory*: The set of all theorems that can be proven from a given set of axioms.

Division

Definition

If a and b are integers with $a \neq 0$, then a **divides** b if there exists an integer c such that $b = ac$.

- When a divides b we write $a|b$.
- We say that a is a **factor** or **divisor** of b and b is a **multiple** of a .
- If $a|b$ then b/a is an integer (namely the c above).
- If a does not divide b , we write $a \nmid b$.

$3 \mid (-12)$ $3 \mid 0$ $3 \nmid 7$ (where \nmid "not divides")

Division

Definition

If a and b are integers with $a \neq 0$, then a **divides** b if there exists an integer c such that $b = ac$.

Theorem

Let a, b, c be integers, where $a \neq 0$.

- ① If $a|b$ and $a|c$, then $a|(b + c)$. **Example:** $3 | 6$ and $3 | 9$, so $3 | 15$.
- ② If $a|b$, then $a|bc$ for all integers c . **Example:** $5 | 10$, so $5 | 20$, $5 | 30$, $5 | 40$.
- ③ If $a|b$ and $b|c$, then $a|c$. **Example:** $4 | 8$ and $8 | 24$, so $4 | 24$.

Proof.

We just prove the first; the others are similar. Assume $a|b$ and $a|c$. So, there exists integers d, e such that $b = da$ and $c = ea$. So $b + c = da + ea = (d + e)a$ and, therefore, $a|(b + c)$. □

Division Algorithm

When an integer is divided by a positive integer, there is a **quotient** and a **remainder**. This is traditionally called the “Division Algorithm”, but it is really a theorem.

Theorem

If a is an integer and d a positive integer, then there are unique integers q and r , with $0 \leq r < d$, such that $a = dq + r$

- a is called the dividend.
- d is called the divisor.
- q is called the quotient. $q = a \mathbf{div} d$
- r is called the remainder. $r = a \mathbf{mod} d$

q is quotient and r the remainder; $q = a \mathbf{div} d$ and $r = a \mathbf{mod} d$

$$a = 102 \text{ and } d = 12 \quad q = 8 \text{ and } r = 6 \quad 102 = 12 \cdot 8 + 6$$

$$a = -14 \text{ and } d = 6 \quad q = -3 \text{ and } r = 4 \quad -14 = 6 \cdot (-3) + 4$$

Congruence Relation

Definition

If a and b are integers and m is a positive integer, then a is **congruent** to b modulo m iff $m \mid (a - b)$.

- The notation $a \equiv b \pmod{m}$ says that a is congruent to b modulo m .
- We say that $a \equiv b \pmod{m}$ is a congruence and that m is its modulus.
- Two integers are congruent mod m if and only if they have the same remainder when divided by m .
- If a is not congruent to b modulo m , we write $a \not\equiv b \pmod{m}$.

Congruence: Examples

Example: Determine

- Whether 17 is congruent to 5 modulo 6, and
- Whether 24 and 14 are congruent modulo 6.

Clicker

- 1 No and No.
- 2 No and Yes.
- 3 Yes and No.
- 4 Yes and Yes.

Congruence: Examples

Example: Determine

- Whether 17 is congruent to 5 modulo 6, and
- Whether 24 and 14 are congruent modulo 6.

Clicker

- 1 No and No.
- 2 No and Yes.
- 3 Yes and No.
- 4 Yes and Yes.

Solution: $17 \equiv 5 \pmod{6}$ because 6 divides $17 - 5 = 12$.
 $24 \not\equiv 14 \pmod{6}$ since $24 - 14 = 10$ is not divisible by 6.

Theorem

*Let a and b be integers, and let m be a positive integer.
Then $a \equiv b \pmod{m}$ if and only if $a \bmod m = b \bmod m$*

The uses of “mod” in the following expressions are **different**.

- $a \equiv b \pmod{m}$, and
- $a \bmod m = b$

$a \equiv b \pmod{m}$ describes a **binary relation** on the set of integers.

In $a \bmod m = b$, the notation mod denotes a **function** (from integers to integers).

The relationship between these notations is made clear in this theorem.

Proof.

Assume $a \equiv b \pmod{m}$; so $m \mid (a - b)$. If $a = q_1m + r_1$ and $b = q_2m + r_2$ where $0 \leq r_1 < m$ and $0 \leq r_2 < m$ it follows that $r_1 = r_2$ and so $a \bmod m = b \bmod m$. If $a \bmod m = b \bmod m$ then a and b have the same remainder so $a = q_1m + r$ and $b = q_2m + r$; therefore $a - b = (q_1 - q_2)m$, and so $m \mid (a - b)$. □

A Theorem on Congruences

Theorem

Let m be a positive integer. The integers a and b are congruent modulo m if and only if there is an integer k such that $a = b + km$.

Proof.

- If $a \equiv b \pmod{m}$, then (by the definition of congruence) $m \mid (a - b)$. Hence, there is an integer k such that $a - b = km$ and equivalently $a = b + km$.
- Conversely, if there is an integer k such that $a = b + km$, then $km = a - b$. Hence, $m \mid (a - b)$ and $a \equiv b \pmod{m}$.




$$m \equiv 0 \pmod{m}$$

We always have $m \equiv 0 \pmod{m}$, and more generally $mk \equiv 0 \pmod{m}$ for any $k \in \mathbb{Z}$. In fact,

$$a \equiv 0 \pmod{m} \iff m|a,$$

so the congruence relation includes the divisibility relation as a special case: the multiples of m are exactly the numbers that “look like 0” modulo m . Because multiples of m are congruent to 0 modulo m , we will see that working with integers modulo m is tantamount to systematically ignoring additions and subtractions by multiples of m in algebraic calculations.

Congruences of Sums and Products

Theorem

Let m be a positive integer. If $a \equiv b \pmod{m}$ and $c \equiv d \pmod{m}$, then $a + c \equiv b + d \pmod{m}$ and $ac \equiv bd \pmod{m}$.

Proof.

Since $a \equiv b \pmod{m}$ and $c \equiv d \pmod{m}$, by the Theorem above there are integers s and t with $b = a + sm$ and $d = c + tm$. Therefore,

- $b + d = (a + sm) + (c + tm) = (a + c) + m(s + t)$, and
- $bd = (a + sm)(c + tm) = ac + m(at + cs + stm)$.

Hence, $a + c \equiv b + d \pmod{m}$ and $ac \equiv bd \pmod{m}$. □

Because $7 \equiv 2 \pmod{5}$ and $11 \equiv 1 \pmod{5}$, it follows from Theorem 5 that

$$18 = 7 + 11 \equiv 2 + 1 = 3 \pmod{5}$$

and that

$$77 = 7 \cdot 11 \equiv 2 \cdot 1 = 2 \pmod{5}.$$

Congruences of Sums and Products

Theorem

Let m be a positive integer. If $a \equiv b \pmod{m}$ and $c \equiv d \pmod{m}$, then $a + c \equiv b + d \pmod{m}$ and $ac \equiv bd \pmod{m}$.

Proof.

Since $a \equiv b \pmod{m}$ and $c \equiv d \pmod{m}$, by the Theorem above there are integers s and t with $b = a + sm$ and $d = c + tm$. Therefore,

- $b + d = (a + sm) + (c + tm) = (a + c) + m(s + t)$, and
- $bd = (a + sm)(c + tm) = ac + m(at + cs + stm)$.

Hence, $a + c \equiv b + d \pmod{m}$ and $ac \equiv bd \pmod{m}$. □

Corollary

Let m be a positive integer and let a and b be integers. Then

- $(a + b) \pmod{m} = ((a \pmod{m}) + (b \pmod{m})) \pmod{m}$
- $ab \pmod{m} = ((a \pmod{m})(b \pmod{m})) \pmod{m}$.

Hashing Functions

Definition: A hashing function h assigns memory location $h(k)$ to the record that has k as its key.

A common hashing function is $h(k) = k \bmod m$, where m is the number of memory locations.

➤ Because this hashing function is onto, all memory locations are possible.

Example: Let $h(k) = k \bmod 111$. This hashing function assigns the records of customers with social security numbers as keys to memory locations in the following manner:

$$h(064212848) = 064212848 \bmod 111 = 14$$

$$h(037149212) = 037149212 \bmod 111 = 65$$

$h(107405723) = 107405723 \bmod 111 = 14$, but since location 14 is already occupied, the record is assigned to the next available position, which is 15.

➤ The hashing function is not one-to-one as there are many more possible keys than memory locations. When more than one record is assigned to the same location, we say a *collision* occurs. Here a collision has been resolved by assigning the record to the first free location.

➤ For collision resolution, we can use a *linear probing function*:

$$h(k,i) = (h(k) + i) \bmod m, \text{ where } i \text{ runs from } 0 \text{ to } m - 1.$$

➤ There are many other methods of handling with collisions. You may cover these in a later CS course.

Pseudorandom Numbers

- Randomly chosen numbers are needed for many purposes, including computer simulations.
- Pseudorandom numbers* are not truly random since they are generated by systematic methods.
- The *linear congruential method* is one commonly used procedure for generating pseudorandom numbers.
- Four integers are needed: the *modulus* m , the *multiplier* a , the *increment* c , and seed x_0 , with $2 \leq a < m$, $0 \leq c < m$, $0 \leq x_0 < m$.
- We generate a sequence of pseudorandom numbers $\{x_n\}$, with $0 \leq x_n < m$ for all n , by successively using the recursively defined function

$$x_{n+1} = (ax_n + c) \bmod m.$$

(an example of a recursive definition, discussed in Section 5.3)

- If pseudorandom numbers between 0 and 1 are needed, then the generated numbers are divided by the modulus, x_n/m .

Pseudorandom Numbers

- ▶ **Example:** Find the sequence of pseudorandom numbers generated by the linear congruential method with modulus $m = 9$, multiplier $a = 7$, increment $c = 4$, and seed $x_0 = 3$.
- ▶ **Solution:** Compute the terms of the sequence by successively using the congruence

$$x_{n+1} = (7x_n + 4) \bmod 9, \text{ with } x_0 = 3.$$

$$x_1 = 7x_0 + 4 \bmod 9 = 7 \cdot 3 + 4 \bmod 9 = 25 \bmod 9 = 7,$$

$$x_2 = 7x_1 + 4 \bmod 9 = 7 \cdot 7 + 4 \bmod 9 = 53 \bmod 9 = 8,$$

$$x_3 = 7x_2 + 4 \bmod 9 = 7 \cdot 8 + 4 \bmod 9 = 60 \bmod 9 = 6,$$

$$x_4 = 7x_3 + 4 \bmod 9 = 7 \cdot 6 + 4 \bmod 9 = 46 \bmod 9 = 1,$$

$$x_5 = 7x_4 + 4 \bmod 9 = 7 \cdot 1 + 4 \bmod 9 = 11 \bmod 9 = 2,$$

$$x_6 = 7x_5 + 4 \bmod 9 = 7 \cdot 2 + 4 \bmod 9 = 18 \bmod 9 = 0,$$

$$x_7 = 7x_6 + 4 \bmod 9 = 7 \cdot 0 + 4 \bmod 9 = 4 \bmod 9 = 4,$$

$$x_8 = 7x_7 + 4 \bmod 9 = 7 \cdot 4 + 4 \bmod 9 = 32 \bmod 9 = 5,$$

$$x_9 = 7x_8 + 4 \bmod 9 = 7 \cdot 5 + 4 \bmod 9 = 39 \bmod 9 = 3.$$

The sequence generated is 3,7,8,6,1,2,0,4,5,3,7,8,6,1,2,0,4,5,3,...

It repeats after generating 9 terms.

- ▶ Commonly, computers use a linear congruential generator with increment $c = 0$. This is called a *pure multiplicative generator*. Such a generator with modulus $2^{31} - 1$ and multiplier $7^5 = 16,807$ generates $2^{31} - 2$ numbers before repeating.

Check Digits: UPCs

A common method of detecting errors in strings of digits is to add an extra digit at the end, which is evaluated using a function. If the final digit is not correct, then the string is assumed not to be correct.

Example: Retail products are identified by their *Universal Product Codes (UPCs)*. Usually these have 12 decimal digits, the last one being the check digit. The check digit is determined by the congruence:

$$3x_1 + x_2 + 3x_3 + x_4 + 3x_5 + x_6 + 3x_7 + x_8 + 3x_9 + x_{10} + 3x_{11} + x_{12} \equiv 0 \pmod{10}.$$

- a. Suppose that the first 11 digits of the UPC are 79357343104. What is the check digit?
- b. Is 041331021641 a valid UPC?

Solution:

a. $3 \cdot 7 + 9 + 3 \cdot 3 + 5 + 3 \cdot 7 + 3 + 3 \cdot 4 + 3 + 3 \cdot 1 + 0 + 3 \cdot 4 + x_{12} \equiv 0 \pmod{10}$

$$21 + 9 + 9 + 5 + 21 + 3 + 12 + 3 + 3 + 0 + 12 + x_{12} \equiv 0 \pmod{10}$$

$$98 + x_{12} \equiv 0 \pmod{10}$$

$$x_{12} \equiv 0 \pmod{10} \quad \text{So, the check digit is 2.}$$

b. $3 \cdot 0 + 4 + 3 \cdot 1 + 3 + 3 \cdot 3 + 1 + 3 \cdot 0 + 2 + 3 \cdot 1 + 6 + 3 \cdot 4 + 1 \equiv 0 \pmod{10}$

$$0 + 4 + 3 + 3 + 9 + 1 + 0 + 2 + 3 + 6 + 12 + 1 = 44 \equiv 4 \not\equiv 0 \pmod{10}$$

Hence, 041331021641 is not a valid UPC.

Check Digits:ISBNs

Books are identified by an *International Standard Book Number* (ISBN-10), a 10 digit code. The first 9 digits identify the language, the publisher, and the book. The tenth digit is a check digit, which is determined by the following congruence

$$x_{10} \equiv \sum_{i=1}^9 ix_i \pmod{11}.$$

The validity of an ISBN-10 number can be evaluated with the equivalent

$$\sum_{i=1}^{10} ix_i \equiv 0 \pmod{11}.$$

- a. Suppose that the first 9 digits of the ISBN-10 are 007288008. What is the check digit?
- b. Is 084930149X a valid ISBN10?

X is used
for the
digit 10.

Solution: a. $X_{10} \equiv 1 \cdot 0 + 2 \cdot 0 + 3 \cdot 7 + 4 \cdot 2 + 5 \cdot 8 + 6 \cdot 8 + 7 \cdot 0 + 8 \cdot 0 + 9 \cdot 8 \pmod{11}.$

$$X_{10} \equiv 0 + 0 + 21 + 8 + 40 + 48 + 0 + 0 + 72 \pmod{11}.$$

$$X_{10} \equiv 189 \equiv 2 \pmod{11}. \text{ Hence, } X_{10} = 2.$$

b. $1 \cdot 0 + 2 \cdot 8 + 3 \cdot 4 + 4 \cdot 9 + 5 \cdot 3 + 6 \cdot 0 + 7 \cdot 1 + 8 \cdot 4 + 9 \cdot 9 + X \cdot 10 =$

$$0 + 16 + 12 + 36 + 15 + 0 + 7 + 32 + 81 + 100 = 299 \equiv 2 \not\equiv 0 \pmod{11}$$

Hence, 084930149X is not a valid ISBN-10.

- A *single error* is an error in one digit of an identification number and a *transposition error* is the accidental interchanging of two digits. Both of these kinds of errors can be detected by the check digit for ISBN-10. (see text for more details)

Arithmetic modulo m

- Let $Z_m = \{0, 1, \dots, m-1\}$.
- The operation $+_m$ is defined as $a +_m b = (a + b) \bmod m$.
This is addition modulo m .
- The operation \cdot_m is defined as $a \cdot_m b = (a \cdot b) \bmod m$.
This is multiplication modulo m .
- Using these operations is said to be doing arithmetic modulo m .

Example: Find $7 +_{11} 9$ and $7 \cdot_{11} 9$.

Solution: Using the definitions above:

$$7 +_{11} 9 = (7 + 9) \bmod 11 = 16 \bmod 11 = 5$$

$$7 \cdot_{11} 9 = (7 \cdot 9) \bmod 11 = 63 \bmod 11 = 8$$

Arithmetic modulo m

The operations $+_m$ and \cdot_m satisfy many of the same properties as ordinary addition and multiplication.

Closure: If $a, b \in Z_m$, then $a +_m b$ and $a \cdot_m b$ belong to Z_m .

Associativity: If $a, b, c \in Z_m$, then $(a +_m b) +_m c = a +_m (b +_m c)$ and $(a \cdot_m b) \cdot_m c = a \cdot_m (b \cdot_m c)$.

Commutativity: If $a, b \in Z_m$, then $a +_m b = b +_m a$ and $a \cdot_m b = b \cdot_m a$.

Identity elements: The elements 0 and 1 are identity elements for addition and multiplication modulo m , respectively. If $a \in Z_m$ then $a +_m 0 = a$ and $a \cdot_m 1 = a$.

Additive inverses: If $0 \neq a \in Z_m$, then $m - a$ is the additive inverse of a modulo m . Moreover, 0 is its own additive inverse.
 $a +_m (m - a) = 0$ and $0 +_m 0 = 0$.

Distributivity: If $a, b, c \in Z_m$, then
 $a \cdot_m (b +_m c) = (a \cdot_m b) +_m (a \cdot_m c)$ and
 $(a +_m b) \cdot_m c = (a \cdot_m c) +_m (b \cdot_m c)$.

Primes

Definition

A positive integer $p > 1$ is called **prime** iff the only positive factors of p are 1 and p . Otherwise it is called **composite**.

Theorem (Fundamental Theorem of Arithmetic)

Every positive integer greater than 1 can be written uniquely as a prime or as the product of its prime factors, written in order of nondecreasing size.

Example: $765 = 3 \cdot 3 \cdot 5 \cdot 17 = 3^2 \cdot 5 \cdot 17$.

Theorem (Euclid (325-265 BCE))

There are infinitely many primes.

Proof by contradiction. If there were only finitely many primes then multiply them all and add 1. This would be a new prime. Contradiction.

The Sieve of Eratosthenes (276-194 BCE)

How to find all primes between 2 and n ?

- 1 Write the numbers $2, \dots, n$ into a list. Let $i := 2$.
- 2 Remove all strict multiples of i from the list.
- 3 Let k be the smallest number present in the list s.t. $k > i$.
Then let $i := k$.
- 4 If $i > \sqrt{n}$ then stop else goto step 2.

Trial division: A very inefficient method of determining if a number n is prime, is to try every integer $i \leq \sqrt{n}$ and see if n is divisible by i .

Testing if a number is prime can be done efficiently in polynomial time [Agrawal-Kayal-Saxena 2002], i.e., polynomial in the number of bits used to describe the input number.

Efficient randomized tests had been available previously.

TABLE 1 The Sieve of Eratosthenes.

Integers divisible by 2 other than 2 receive an underline.

1	2	3	<u>4</u>	5	<u>6</u>	7	<u>8</u>	9	<u>10</u>
11	<u>12</u>	13	<u>14</u>	15	<u>16</u>	17	<u>18</u>	19	<u>20</u>
21	<u>22</u>	23	<u>24</u>	25	<u>26</u>	27	<u>28</u>	29	<u>30</u>
31	<u>32</u>	33	<u>34</u>	35	<u>36</u>	37	<u>38</u>	39	<u>40</u>
41	<u>42</u>	43	<u>44</u>	45	<u>46</u>	47	<u>48</u>	49	<u>50</u>
51	<u>52</u>	53	<u>54</u>	55	<u>56</u>	57	<u>58</u>	59	<u>60</u>
61	<u>62</u>	63	<u>64</u>	65	<u>66</u>	67	<u>68</u>	69	<u>70</u>
71	<u>72</u>	73	<u>74</u>	75	<u>76</u>	77	<u>78</u>	79	<u>80</u>
81	<u>82</u>	83	<u>84</u>	85	<u>86</u>	87	<u>88</u>	89	<u>90</u>
91	<u>92</u>	93	<u>94</u>	95	<u>96</u>	97	<u>98</u>	99	<u>100</u>

Integers divisible by 3 other than 3 receive an underline.

1	2	3	<u>4</u>	5	<u>6</u>	7	8	<u>9</u>	<u>10</u>
11	<u>12</u>	13	<u>14</u>	<u>15</u>	<u>16</u>	17	<u>18</u>	19	<u>20</u>
<u>21</u>	<u>22</u>	23	<u>24</u>	25	<u>26</u>	<u>27</u>	<u>28</u>	29	<u>30</u>
31	<u>32</u>	<u>33</u>	<u>34</u>	35	<u>36</u>	37	<u>38</u>	<u>39</u>	<u>40</u>
41	<u>42</u>	43	<u>44</u>	<u>45</u>	<u>46</u>	47	<u>48</u>	49	<u>50</u>
<u>51</u>	<u>52</u>	53	<u>54</u>	55	<u>56</u>	<u>57</u>	<u>58</u>	59	<u>60</u>
61	<u>62</u>	<u>63</u>	<u>64</u>	65	<u>66</u>	67	<u>68</u>	<u>69</u>	<u>70</u>
71	<u>72</u>	73	<u>74</u>	<u>75</u>	<u>76</u>	77	<u>78</u>	79	<u>80</u>
<u>81</u>	<u>82</u>	83	<u>84</u>	85	<u>86</u>	<u>87</u>	<u>88</u>	89	<u>90</u>
91	<u>92</u>	<u>93</u>	<u>94</u>	95	<u>96</u>	97	<u>98</u>	<u>99</u>	<u>100</u>

Integers divisible by 5 other than 5 receive an underline.

1	2	3	<u>4</u>	5	<u>6</u>	7	<u>8</u>	<u>9</u>	<u>10</u>
11	<u>12</u>	13	<u>14</u>	<u>15</u>	<u>16</u>	17	<u>18</u>	19	<u>20</u>
<u>21</u>	<u>22</u>	23	<u>24</u>	<u>25</u>	<u>26</u>	<u>27</u>	<u>28</u>	29	<u>30</u>
31	<u>32</u>	<u>33</u>	<u>34</u>	<u>35</u>	<u>36</u>	37	<u>38</u>	<u>39</u>	<u>40</u>
41	<u>42</u>	43	<u>44</u>	<u>45</u>	<u>46</u>	47	<u>48</u>	49	<u>50</u>
<u>51</u>	<u>52</u>	53	<u>54</u>	<u>55</u>	<u>56</u>	<u>57</u>	<u>58</u>	59	<u>60</u>
61	<u>62</u>	<u>63</u>	<u>64</u>	<u>65</u>	<u>66</u>	67	<u>68</u>	<u>69</u>	<u>70</u>
71	<u>72</u>	73	<u>74</u>	<u>75</u>	<u>76</u>	77	<u>78</u>	79	<u>80</u>
<u>81</u>	<u>82</u>	83	<u>84</u>	<u>85</u>	<u>86</u>	<u>87</u>	<u>88</u>	89	<u>90</u>
91	<u>92</u>	<u>93</u>	<u>94</u>	<u>95</u>	<u>96</u>	97	<u>98</u>	<u>99</u>	<u>100</u>

Integers divisible by 7 other than 7 receive an underline; integers in color are prime.

1	2	3	<u>4</u>	5	<u>6</u>	7	<u>8</u>	<u>9</u>	<u>10</u>
11	<u>12</u>	13	<u>14</u>	<u>15</u>	<u>16</u>	17	<u>18</u>	19	<u>20</u>
<u>21</u>	<u>22</u>	23	<u>24</u>	<u>25</u>	<u>26</u>	<u>27</u>	<u>28</u>	29	<u>30</u>
31	<u>32</u>	33	<u>34</u>	<u>35</u>	<u>36</u>	37	<u>38</u>	39	<u>40</u>
41	<u>42</u>	43	<u>44</u>	<u>45</u>	<u>46</u>	47	<u>48</u>	49	<u>50</u>
51	52	53	<u>54</u>	<u>55</u>	<u>56</u>	57	<u>58</u>	59	<u>60</u>
61	62	63	<u>64</u>	<u>65</u>	<u>66</u>	67	<u>68</u>	69	<u>70</u>
71	72	73	<u>74</u>	<u>75</u>	<u>76</u>	77	<u>78</u>	79	<u>80</u>
81	82	83	<u>84</u>	<u>85</u>	<u>86</u>	87	<u>88</u>	89	<u>90</u>
91	92	93	<u>94</u>	<u>95</u>	<u>96</u>	97	<u>98</u>	99	<u>100</u>

Distribution of Primes

What part of the numbers are primes?

Do primes get scarce among the large numbers?

The prime number theorem gives an asymptotic estimate for the number of primes not exceeding x .

Theorem (Prime Number Theorem)

*The ratio of the number of primes not exceeding x and $x / \ln(x)$ approaches 1 as x grows without bound.
($\ln(x)$ is the natural logarithm of x .)*

- The theorem tells us that the number of primes not exceeding x , can be approximated by $x / \ln(x)$. For example, $100 / \ln(100) = 21.71$
- The odds that a randomly selected positive integer less than x is prime are approximately $(x / \ln(x)) / x = 1 / \ln(x)$.
- The k -th prime is approximately of size $k \cdot \ln(k)$.

For example, the odds that an integer near 10^{1000} is prime are approximately $1 / \ln 10^{1000}$, which is approximately $1/2300$.

Greatest Common Divisor

Definition

Let $a, b \in \mathbb{Z} - \{0\}$. The largest integer d such that $d|a$ and also $d|b$ is called the **greatest common divisor** of a and b . It is denoted by $\gcd(a, b)$.

Example: $\gcd(24, 36) = 12$.

Definition

The integers a and b are **relatively prime (coprime)** iff $\gcd(a, b) = 1$.

Example: 17 and 22. (Note that 22 is not a prime.)

Definition

The integers a_1, a_2, \dots, a_n are **pairwise relatively prime** iff $\gcd(a_i, a_j) = 1$ whenever $1 \leq i < j \leq n$.

Example: 10, 17 and 21 are pairwise relatively prime, since $\gcd(10, 17) = \gcd(10, 21) = \gcd(17, 21) = 1$.

Least Common Multiple

Definition

The **least common multiple** of the positive integers a and b is the smallest positive integer that is divisible by both a and b . It is denoted by $\text{lcm}(a, b)$.

Example: $\text{lcm}(45, 21) = 7 \cdot 45 = 15 \cdot 21 = 315$.

Gcd and Lcm by Prime Factorizations

Suppose that the prime factorizations of a and b are:

$$a = p_1^{a_1} p_2^{a_2} \cdots p_n^{a_n}, \quad b = p_1^{b_1} p_2^{b_2} \cdots p_n^{b_n}$$

where each exponent is a nonnegative integer (possibly zero). Then

$$\gcd(a, b) = p_1^{\min(a_1, b_1)} p_2^{\min(a_2, b_2)} \cdots p_n^{\min(a_n, b_n)}$$

This number clearly divides a and b . No larger number can divide both a and b . Proof by contradiction and the prime factorization of a postulated larger divisor.

$$\text{lcm}(a, b) = p_1^{\max(a_1, b_1)} p_2^{\max(a_2, b_2)} \cdots p_n^{\max(a_n, b_n)}$$

This number is clearly a multiple of a and b . No smaller number can be a multiple of both a and b . Proof by contradiction and the prime factorization of a postulated smaller multiple.

Factorization is a **very inefficient** method to compute \gcd and lcm .
The Euclidian algorithm is much better.

Because the prime factorizations of 120 and 500 are $120 = 2^3 \cdot 3 \cdot 5$ and $500 = 2^2 \cdot 5^3$, the greatest common divisor is

$$\gcd(120, 500) = 2^{\min(3, 2)} 3^{\min(1, 0)} 5^{\min(1, 3)} = 2^2 3^0 5^1 = 20.$$

What is the least common multiple of $2^3 3^5 7^2$ and $2^4 3^3$?

Solution: We have

$$\text{lcm}(2^3 3^5 7^2, 2^4 3^3) = 2^{\max(3, 4)} 3^{\max(5, 3)} 7^{\max(2, 0)} = 2^4 3^5 7^2.$$

Euclidian Algorithm

Lemma

*Let $a = bq + r$, where a, b, q , and r are integers.
Then $\gcd(a, b) = \gcd(b, r)$.*

Proof.

Suppose that d divides both a and b . Then d also divides $a - bq = r$. Hence, any common divisor of a and b must also be a common divisor of b and r .

For the opposite direction suppose that d divides both b and r . Then d also divides $bq + r = a$. Hence, any common divisor of b and r must also be a common divisor of a and b .

Therefore, $\gcd(a, b) = \gcd(b, r)$. □

This means that if $a > b$ then $\gcd(a, b) = \gcd(b, a \bmod b)$, which directly yields the algorithm.

(Note that both arguments have gotten smaller.) One can show that its complexity is $O(\log b)$.

Before describing the Euclidean algorithm, we will show how it is used to find $\gcd(91, 287)$. First, divide 287, the larger of the two integers, by 91, the smaller, to obtain

$$287 = 91 \cdot 3 + 14.$$

Any divisor of 91 and 287 must also be a divisor of $287 - 91 \cdot 3 = 14$. Also, any divisor of 91 and 14 must also be a divisor of $287 = 91 \cdot 3 + 14$. Hence, the greatest common divisor of 91 and 287 is the same as the greatest common divisor of 91 and 14. This means that the problem of finding $\gcd(91, 287)$ has been reduced to the problem of finding $\gcd(91, 14)$.

Next, divide 91 by 14 to obtain

$$91 = 14 \cdot 6 + 7.$$

Because any common divisor of 91 and 14 also divides $91 - 14 \cdot 6 = 7$ and any common divisor of 14 and 7 divides 91, it follows that $\gcd(91, 14) = \gcd(14, 7)$.

Continue by dividing 14 by 7, to obtain

$$14 = 7 \cdot 2.$$

Because 7 divides 14, it follows that $\gcd(14, 7) = 7$. Furthermore, because $\gcd(287, 91) = \gcd(91, 14) = \gcd(14, 7) = 7$, the original problem has been solved.

Find the greatest common divisor of 414 and 662 using the Euclidean algorithm.

Solution: Successive uses of the division algorithm give:

$$662 = 414 \cdot 1 + 248$$

$$414 = 248 \cdot 1 + 166$$

$$248 = 166 \cdot 1 + 82$$

$$166 = 82 \cdot 2 + 2$$

$$82 = 2 \cdot 41.$$

Hence, $\gcd(414, 662) = 2$, because 2 is the last nonzero remainder.



Gcd as a Linear Combination

Theorem (Bézout's Theorem)

If a and b are positive integers, then there exist integers s and t such that $\gcd(a, b) = sa + tb$

(Proof in exercises of Section 5.2).

The numbers s and t are called Bézout coefficients of a and b .

Example: $2 = \gcd(6, 14) = (-2) \cdot 6 + 1 \cdot 14$.

The Bézout coefficients can be computed as follows. First use the Euclidian algorithm to find the \gcd and then work backwards (by division and substitution) to express the \gcd as a linear combination of the original two integers.

Express $\gcd(252, 198) = 18$ as a linear combination of 252 and 198.

Solution: To show that $\gcd(252, 198) = 18$, the Euclidean algorithm uses these divisions:

$$252 = 1 \cdot 198 + 54$$

$$198 = 3 \cdot 54 + 36$$

$$54 = 1 \cdot 36 + 18$$

$$36 = 2 \cdot 18.$$

Using the next-to-last division (the third division), we can express $\gcd(252, 198) = 18$ as a linear combination of 54 and 36. We find that

$$18 = 54 - 1 \cdot 36.$$

The second division tells us that

$$36 = 198 - 3 \cdot 54.$$

Substituting this expression for 36 into the previous equation, we can express 18 as a linear combination of 54 and 198. We have

$$18 = 54 - 1 \cdot 36 = 54 - 1 \cdot (198 - 3 \cdot 54) = 4 \cdot 54 - 1 \cdot 198.$$

The first division tells us that

$$54 = 252 - 1 \cdot 198.$$

Substituting this expression for 54 into the previous equation, we can express 18 as a linear combination of 252 and 198. We conclude that

$$18 = 4 \cdot (252 - 1 \cdot 198) - 1 \cdot 198 = 4 \cdot 252 - 5 \cdot 198,$$

completing the solution.



Linear Congruences

Definition

A congruence of the form

$$ax \equiv b \pmod{m}$$

where m is a positive integer, a, b are integers and x is an integer variable is called a **linear congruence**.

The solution of the congruence are all the integers x that satisfy it.

Definition

An integer \bar{a} such that $\bar{a}a \equiv 1 \pmod{m}$ is called a **multiplicative inverse** of a modulo m .

Multiplicative inverses can be used to solve congruences. If $ax \equiv b \pmod{m}$ then $\bar{a}ax \equiv (\bar{a}b) \pmod{m}$ and thus $x \equiv (\bar{a}b) \pmod{m}$.


Find an inverse of 3 modulo 7 by first finding Bézout coefficients of 3 and 7. (Note that we have already shown that 5 is an inverse of 3 modulo 7 by inspection.)

Solution: Because $\gcd(3, 7) = 1$, Theorem 1 tells us that an inverse of 3 modulo 7 exists. The Euclidean algorithm ends quickly when used to find the greatest common divisor of 3 and 7:

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

From this equation we see that

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

This shows that -2 and 1 are Bézout coefficients of 3 and 7. We see that -2 is an inverse of 3 modulo 7. Note that every integer congruent to -2 modulo 7 is also an inverse of 3, such as 5, -9 , 12, and so on. 

Find an inverse of 101 modulo 4620.

Solution: For completeness, we present all steps used to compute an inverse of 101 modulo 4620. (Only the last step goes beyond methods developed in Section 4.3 and illustrated in Example 17 in that section.) First, we use the Euclidean algorithm to show that $\gcd(101, 4620) = 1$. Then we will reverse the steps to find Bézout coefficients a and b such that $101a + 4620b = 1$. It will then follow that a is an inverse of 101 modulo 4620. The steps used by the Euclidean algorithm to find $\gcd(101, 4620)$ are

$$4620 = 45 \cdot 101 + 75$$

$$101 = 1 \cdot 75 + 26$$

$$75 = 2 \cdot 26 + 23$$

$$26 = 1 \cdot 23 + 3$$


$$23 = 7 \cdot 3 + 2$$

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

$$2 = 2 \cdot 1.$$

Because the last nonzero remainder is 1, we know that $\gcd(101, 4620) = 1$. We can now find the Bézout coefficients for 101 and 4620 by working backwards through these steps, expressing $\gcd(101, 4620) = 1$ in terms of each successive pair of remainders. In each step we eliminate the remainder by expressing it as a linear combination of the divisor and the dividend. We obtain

$$\begin{aligned}1 &= 3 - 1 \cdot 2 \\&= 3 - 1 \cdot (23 - 7 \cdot 3) = -1 \cdot 23 + 8 \cdot 3 \\&= -1 \cdot 23 + 8 \cdot (26 - 1 \cdot 23) = 8 \cdot 26 - 9 \cdot 23 \\&= 8 \cdot 26 - 9 \cdot (75 - 2 \cdot 26) = -9 \cdot 75 + 26 \cdot 26 \\&= -9 \cdot 75 + 26 \cdot (101 - 1 \cdot 75) = 26 \cdot 101 - 35 \cdot 75 \\&= 26 \cdot 101 - 35 \cdot (4620 - 45 \cdot 101) = -35 \cdot 4620 + 1601 \cdot 101.\end{aligned}$$

That $-35 \cdot 4620 + 1601 \cdot 101 = 1$ tells us that -35 and 1601 are Bézout coefficients of 4620 and 101, and 1601 is an inverse of 101 modulo 4620. 

Example: Solving the linear congruence $5x \equiv 1 \pmod{67}$

Suppose we want to solve the equation $5x \equiv 1 \pmod{67}$. We first check to see if solutions exist. In this case, we know that $(5, 67) = 1$, and since $(5, 67) = 1 \mid 1$, we know there are solutions. In fact, we know that there is exactly 1 solution mod 67. To compute it, we first need to write $(5, 67)$ as a linear combination. We'll use the Euclidean Algorithm. This gives

$$\begin{aligned} 67 &= 13 \cdot 5 + 2 \\ 5 &= 2 \cdot 2 + 1 \\ 2 &= 2 \cdot 1 + 0 \end{aligned} \tag{1}$$

Now we can use these equations to express 1 as a combination of 5 and 67:

$$1 = 5 - 2 \cdot 2 = 5 - 2(67 - 13 \cdot 5) = 27 \cdot 5 - 2 \cdot 67. \tag{2}$$

Taking this equation modulo 67 shows that $1 \equiv 27 \cdot 5 \pmod{67}$, and so 27 is the multiplicative inverse of 5 modulo 67. \square

This example leads to the following

Definition: A solution to the linear congruence $ax \equiv 1 \pmod{m}$ is called a multiplicative inverse for a modulo m .

Example: Solving $5x \equiv 11 \pmod{67}$

Suppose we wish to solve $5x \equiv 11 \pmod{67}$. We could proceed as we have before — finding a gcd, writing that gcd as a linear combination, etc. Alternatively, we can use the fact that we've already computed the multiplicative inverse of 5 as 27. To take this latter route, notice that we have

$$5x \equiv 11 \pmod{67} \iff 27 \cdot 5x \equiv 27 \cdot 11 \pmod{67}. \quad (3)$$

(Notice: we're allowed to multiply by 27 on both sides of the expression without disturbing the solution set because $(27, 67) = 1$, and you'll recall our theorem which says that $ca \equiv cb \pmod{m}$ if and only if $a \equiv b \pmod{\frac{m}{(c,m)}}$).

Using the fact that $27 \cdot 5 \equiv 1 \pmod{67}$ by our previous example, this means that our solution is $x \equiv 27 \cdot 11 \pmod{67}$. \square

What are the solutions of the linear congruence $3x \equiv 4 \pmod{7}$?


Solution: By Example 1 we know that -2 is an inverse of 3 modulo 7 . Multiplying both sides of the congruence by -2 shows that

$$-2 \cdot 3x \equiv -2 \cdot 4 \pmod{7}.$$

Because $-6 \equiv 1 \pmod{7}$ and $-8 \equiv 6 \pmod{7}$, it follows that if x is a solution, then $x \equiv -8 \equiv 6 \pmod{7}$.

We need to determine whether every x with $x \equiv 6 \pmod{7}$ is a solution. Assume that $x \equiv 6 \pmod{7}$. Then, by Theorem 5 of Section 4.1, it follows that

$$3x \equiv 3 \cdot 6 = 18 \equiv 4 \pmod{7},$$

which shows that all such x satisfy the congruence. We conclude that the solutions to the congruence are the integers x such that $x \equiv 6 \pmod{7}$, namely, $6, 13, 20, \dots$ and $-1, -8, -15, \dots$ 

Multiplicative Inverses

Example: Let $m = 15$.

Find a multiplicative inverse of 8 modulo 15.

Clicker.

- 1 1
- 2 2
- 3 3
- 4 4
- 5 5
- 6 ≥ 6

$$2 \cdot 8 = 16 \equiv 1 \pmod{15}.$$

Thus 2 is a multiplicative inverse of 8 modulo 15.

Multiplicative Inverses

Find a multiplicative inverse of 7 modulo 15.

Clicker.

- 1 ≤ 3
- 2 between 4 and 8
- 3 between 9 and 11
- 4 between 12 and 14

Multiplicative Inverses

What is the multiplicative inverse of 5 modulo 15?

$$\begin{aligned}1 \cdot 5 &\equiv 5 \pmod{15} \\2 \cdot 5 &\equiv 10 \pmod{15} \\3 \cdot 5 &\equiv 0 \pmod{15} \\4 \cdot 5 &\equiv 5 \pmod{15} \\5 \cdot 5 &\equiv 10 \pmod{15} \\6 \cdot 5 &\equiv 0 \pmod{15} \\7 \cdot 5 &\equiv 5 \pmod{15} \\&\dots\end{aligned}$$

Where is the inverse??? 5 does not have any inverse modulo 15.

The multiplicative group Z_m^*

Theorem

If a and m are relatively prime integers and $m > 1$, then a multiplicative inverse of a modulo m exists. Furthermore, this inverse is unique modulo m .

Proof. Since $\gcd(a, m) = 1$, by Bézout's Theorem there are integers s and t such that $sa + tm = 1$.

Hence, $sa + tm \equiv 1 \pmod{m}$.

Since $tm \equiv 0 \pmod{m}$, it follows that $sa \equiv 1 \pmod{m}$.

Consequently, s is a multiplicative inverse of a modulo m .

Uniqueness: Exercise.

Definition

Let $Z_m^* = \{x \mid 1 \leq x < m \text{ and } \gcd(x, m) = 1\}$. Together with multiplication modulo m , this is called the multiplicative group modulo m . It is closed, associative, has a neutral element (namely 1) and every element has an inverse.

The Chinese Remainder Theorem

Let

$$x \equiv 2 \pmod{3}$$

$$x \equiv 3 \pmod{5}$$

$$x \equiv 2 \pmod{7}$$

What is x ? (Or rather, what is the smallest x that satisfies these?)

Theorem (Chinese Remainder Theorem)

Let m_1, m_2, \dots, m_n be pairwise relatively prime positive integers greater than one and a_1, a_2, \dots, a_n arbitrary integers. Then the system

$$x \equiv a_1 \pmod{m_1}$$

$$x \equiv a_2 \pmod{m_2}$$

\dots

$$x \equiv a_n \pmod{m_n}$$

has a unique solution modulo $m = m_1 m_2 \dots m_n$. (I.e., there is a solution x with $0 \leq x < m$ and all other solutions are congruent modulo m to this solution.)

The Chinese Remainder Theorem: Proof

We will construct a solution x .


First, let $M_k = m/m_k$ for $k = 1, 2, \dots, n$ and $m = m_1 m_2 \dots m_n$.

Since $\gcd(m_k, M_k) = 1$, the number M_k has a multiplicative inverse y_k modulo m_k . I.e.,

$$M_k y_k \equiv 1 \pmod{m_k}$$


Now we let

$$x = a_1 M_1 y_1 + a_2 M_2 y_2 + \dots + a_n M_n y_n$$


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

To solve the system of congruences in Example 4, first let $m = 3 \cdot 5 \cdot 7 = 105$, $M_1 = m/3 = 35$, $M_2 = m/5 = 21$, and $M_3 = m/7 = 15$. We see that 2 is an inverse of $M_1 = 35$ modulo 3, because $35 \cdot 2 \equiv 2 \cdot 2 \equiv 1 \pmod{3}$; 1 is an inverse of $M_2 = 21$ modulo 5, because $21 \equiv 1 \pmod{5}$; and 1 is an inverse of $M_3 = 15$ modulo 7, because $15 \equiv 1 \pmod{7}$. The solutions to this system are those x such that

$$\begin{aligned}x &\equiv a_1 M_1 y_1 + a_2 M_2 y_2 + a_3 M_3 y_3 = 2 \cdot 35 \cdot 2 + 3 \cdot 21 \cdot 1 + 2 \cdot 15 \cdot 1 \\&= 233 \equiv 23 \pmod{105}.\end{aligned}$$

It follows that 23 is the smallest positive integer that is a simultaneous solution. We conclude that 23 is the smallest positive integer that leaves a remainder of 2 when divided by 3, a remainder of 3 when divided by 5, and a remainder of 2 when divided by 7. 

Use the method of back substitution to find all integers x such that $x \equiv 1 \pmod{5}$, $x \equiv 2 \pmod{6}$, and $x \equiv 3 \pmod{7}$.

Solution: By Theorem 4 in Section 4.1, the first congruence can be rewritten as an equality, $x = 5t + 1$ where t is an integer. Substituting this expression for x into the second congruence tells us that

$$5t + 1 \equiv 2 \pmod{6},$$


which can be easily solved to show that $t \equiv 5 \pmod{6}$ (as the reader should verify). Using Theorem 4 in Section 4.1 again, we see that $t = 6u + 5$ where u is an integer. Substituting this expression for t back into the equation $x = 5t + 1$ tells us that $x = 5(6u + 5) + 1 = 30u + 26$. We insert this into the third equation to obtain

$$30u + 26 \equiv 3 \pmod{7}.$$


Solving this congruence tells us that $u \equiv 6 \pmod{7}$ (as the reader should verify). Hence, Theorem 4 in Section 4.1 tells us that $u = 7v + 6$ where v is an integer. Substituting this expression for u into the equation $x = 30u + 26$ tells us that $x = 30(7v + 6) + 26 = 210v + 206$. Translating this back into a congruence, we find the solution to the simultaneous congruences,

$$x \equiv 206 \pmod{210}.$$





Problem: Jessica breeds rabbits. She's not sure exactly how many she has today, but as she was moving them about this morning, she noticed some things. When she fed them, in groups of 5, she had 4 left over. When she bathed them, in groups of 8, she had a group of 6 left over. She took them outside to romp in groups of 9, but then the last group consisted of only 8. She's positive that there are fewer than 250 rabbits - but how many does she have?



Fermat's Little Theorem (Pierre de Fermat (1601-65))

Theorem

If p is prime and $p \nmid a$, then $a^{p-1} \equiv 1 \pmod{p}$.

Furthermore, for every integer a we have $a^p \equiv a \pmod{p}$.

Proof sketch: $a^x \pmod{p} = (a \pmod{p})^x \pmod{p}$. Also $p \nmid a$.

So without restriction we consider only $0 < a < p$.

Consider the powers of a^1, a^2, a^3, \dots modulo p .

These form a subgroup of Z_p^* which has some size k and we have $a^k \equiv 1 \pmod{p}$.

By Lagrange's theorem, k divides the size of Z_p^* which is $p - 1$, so $p - 1 = km$ for some positive integer m . Thus

$$a^{p-1} \equiv a^{km} \equiv (a^k)^m \equiv 1^m \equiv 1 \pmod{p}$$

This directly implies $a^p \equiv a \pmod{p}$.

In the other case where $p \mid a$ we trivially have $a^p \equiv a \equiv 0 \pmod{p}$.

Fermat's Little Theorem

Fermat's little theorem is useful in computing the remainders modulo p of large powers of integers.

Example: Find $7^{222} \bmod 11$.

By Fermat's little theorem, we know that $7^{10} \equiv 1 \pmod{11}$, and so $(7^{10})^k \equiv 1 \pmod{11}$ for every positive integer k .

Therefore, $7^{222} = 7^{22 \cdot 10 + 2} = (7^{10})^{22} 7^2 \equiv 1^{22} \cdot 49 \equiv 5 \pmod{11}$.

Hence, $7^{222} \bmod 11 = 5$.

Number Theory in Cryptography

Terminology: Two parties **Alice** and **Bob** want to communicate securely s.t. a third party **Eve** who intercepts messages cannot learn the content of the messages.

Symmetric Cryptosystems: Alice and Bob share a secret. Only they know a secret key K that is used to encrypt and decrypt messages. Given a message M , Alice encodes it (possibly with padding) into m , and then sends the ciphertext $encrypt(m, K)$ to Bob. Then Bob uses K to decrypt it and obtains $decrypt(encrypt(m, K), K) = m$.

Example: AES.

Public Key Cryptosystems: Alice and Bob do a-priori **not** share a secret. How can they establish a shared secret when others are listening to their messages?

Idea: Have a two-part key, i.e., a key pair. A public key that is used to encrypt messages, and a secret key to decrypt them. Alice uses Bob's public key to encrypt a message (everyone can do that). Only Bob can decrypt the message with his secret key.

RSA: an example of a Public Key Cryptosystem

- Named after Rivest, Shamir, Adelman (1976 at MIT). Discovered earlier by Clifford Cocks, working secretly for the UK government.
- Still widely used, e.g., in PGP and ssh.
- Described here because it is easy to explain with elementary number theory.

Cryptography: Caveats

- There do **not exist** any cryptosystems that are proven to be secure for complexity theoretic reasons (i.e., easy to encrypt, hard to decrypt).
- The only systems proven secure are so for information theoretic reasons. Random one-time pad: secret key longer than message and used only once (Vernam scheme). Message: $m_n \dots m_0$ bits. Secret key: $k_n \dots k_0$ bits. Ciphertext: $c_i = m_i \text{ xor } k_i$. Decryption: $m_i = c_i \text{ xor } k_i$.

Cryptography: More Caveats

- RSA could be broken with an efficient algorithm to factorize numbers, **but possibly also by other means**. It is an open question if an efficient method to break RSA would imply an efficient factorization method.
- A 768 bit RSA key has been broken, and experts believe 1024 bit could be broken with sufficient resources.
- Many experts increasingly doubt the security of RSA in general, and recommend to use Elliptic curve cryptography systems instead. (Also based on number theory, but harder to explain.)
- Key generation relies on strong random number generation. Vulnerabilities have been deliberately inserted by the NSA into some systems (e.g., Dual_EC_DRBG).
- **Closed source implementations** of cryptographic software are likely to contain more such backdoors, and can **not be considered secure**.

Description of RSA: Key generation

- Choose two distinct prime numbers p and q . Numbers p and q should be chosen at random, and be of similar bit-length. Prime integers can be efficiently found using a primality test.
- Let $n = pq$ and $k = (p - 1)(q - 1)$. (In particular, $k = |Z_n^*|$).
- Choose an integer e such that $1 < e < k$ and $\gcd(e, k) = 1$; i.e., e and k are coprime.
 e (for encryption) is released as the public key exponent.
(e must not be very small.)
- Let d be the multiplicative inverse of e modulo k , i.e., $de \equiv 1 \pmod{k}$. (Computed using the extended Euclidean algorithm.) d (for decryption) is the private key and kept secret.

The public key is (n, e) and the private key is (n, d) .

RSA: Encryption and Decryption

Alice transmits her public key (n, e) to Bob and keeps the private key secret.

Encryption: Bob then wishes to send message M to Alice. He first turns M into an integer m , such that $0 \leq m < n$ by using an agreed-upon reversible protocol known as a padding scheme. He then computes the ciphertext c corresponding to

$$c \equiv m^e \pmod{n}$$

This can be done quickly using the method of exponentiation by squaring. Bob then transmits c to Alice.

Decryption: Alice can recover m from c by using her private key exponent d via computing

$$m \equiv c^d \pmod{n}$$

Given m , she can recover the original message M by reversing the padding scheme.

Encrypt the message STOP using the RSA cryptosystem with key $(2537, 13)$. Note that $2537 = 43 \cdot 59$, $p = 43$ and $q = 59$ are primes, and


$$\gcd(e, (p - 1)(q - 1)) = \gcd(13, 42 \cdot 58) = 1.$$

Solution: To encrypt, we first translate the letters in STOP into their numerical equivalents. We then group these numbers into blocks of four digits (because $2525 < 2537 < 252525$), to obtain

1819 1415.

We encrypt each block using the mapping


$$C = M^{13} \bmod 2537.$$

Computations using fast modular multiplication show that $1819^{13} \bmod 2537 = 2081$ and $1415^{13} \bmod 2537 = 2182$. The encrypted message is 2081 2182. 

We receive the encrypted message 0981 0461. What is the decrypted message if it was encrypted using the RSA cipher from Example 8?

Solution: The message was encrypted using the RSA cryptosystem with $n = 43 \cdot 59$ and exponent 13. As Exercise 2 in Section 4.4 shows, $d = 937$ is an inverse of 13 modulo $42 \cdot 58 = 2436$. We use 937 as our decryption exponent. Consequently, to decrypt a block C , we compute

$$M = C^{937} \bmod 2537.$$

To decrypt the message, we use the fast modular exponentiation algorithm to compute $0981^{937} \bmod 2537 = 0704$ and $0461^{937} \bmod 2537 = 1115$. Consequently, the numerical version of the original message is 0704 1115. Translating this back to English letters, we see that the message is HELP. 

The RSA Algorithm

To generate a key pair:

- Pick large primes p and q (do not disclose them)
- Let $n = p \cdot q$
- For the public key, choose e that is relatively prime to $\phi(n) = (p-1)(q-1)$.
public key = $\langle e, n \rangle$
- For private key, find d that is the multiplicative inverse of e mod $\phi(n)$, i.e., $e \cdot d \equiv 1 \pmod{\phi(n)}$

Using RSA

Given $\text{pubKey} = \langle e, n \rangle$ and $\text{privKey} = \langle d, n \rangle$

If Message = m

Then:

encryption: $c = m^e \bmod n, m < n$

decryption: $m = c^d \bmod n$

signature: $s = md \bmod n, m < n$

verification: $m = se \bmod n$

Example of RSA (1)

Choose $p = 7$ and $q = 17$.

Compute $n = p \cdot q = 119$.

Compute $f(n) = (p-1)(q-1) = 96$.

Select $e = 5$, (a relatively prime to $f(n)$.)

Compute $d = \underline{\hspace{1cm}}_{77}$ such that $e \cdot d = 1 \pmod{f(n)}$.

- Public key: $\langle 5, 119 \rangle$
- Private key: $\langle 77, 119 \rangle$
- Message = 19
- Encryption: $19^5 \pmod{119} = 66$
- Decryption: $66^{77} \pmod{119} = 19$

Example of RSA (2)

$p = 7, q = 11, n = 77$

Alice chooses $e = 17$, making $d = 53$

Bob wants to send Alice secret message

HELLO (07 04 11 11 14)

– $07^{17} \bmod 77 = 28$; $04^{17} \bmod 77 = 16$

– $11^{17} \bmod 77 = 44$; – $11^{17} \bmod 77 = 44$

– $14^{17} \bmod 77 = 42$

• Bob sends **28 16 44 44 42**

Example of RSA (3)

Alice receives **28 16 44 44 42**

Alice uses private key, $d = 53$, to decrypt message:

- $28^{53} \bmod 77 = 07$; $16^{53} \bmod 77 = 04$
- $44^{53} \bmod 77 = 11$; $44^{53} \bmod 77 = 11$
- $42^{53} \bmod 77 = 14$

• Alice translates **07 04 11 11 14** to **HELLO**

No one else could read it, as only Alice knows her private key (needed for decryption)

RSA: Correctness of decryption

Given that $c \equiv m^e \pmod n$, why is $c^d \equiv m \pmod n$?

$$c^d \equiv (m^e)^d \equiv m^{ed} \pmod n.$$

By construction, d and e are each others multiplicative inverses modulo k , i.e., $ed \equiv 1 \pmod k$. Also $k = (p-1)(q-1)$.

Thus $ed - 1 = h(p-1)(q-1)$ for some integer h .

We consider m^{ed} modulo p . If $p \nmid m$ then

$$m^{ed} = m^{h(p-1)(q-1)} m = (m^{p-1})^{h(q-1)} m \equiv 1^{h(q-1)} m \equiv m \pmod p$$

by Fermat's little theorem. Otherwise $m^{ed} \equiv 0 \equiv m \pmod p$.

Symmetrically, $m^{ed} \equiv m \pmod q$.

Since p, q are distinct primes, we have $m^{ed} \equiv m \pmod{pq}$.

Since $n = pq$, we have $c^d \equiv m^{ed} \equiv m \pmod n$.