Number Theory and Cryptography

Chapter 4

Chapter Motivation

- *Number theory* is the part of mathematics devoted to the study of the integers and their properties.
- Key ideas in number theory include divisibility and the primality of integers.
- Number theory has long been studied because of the beauty of its ideas, its accessibility, and its wealth of open questions.
- We'll use many ideas developed in Chapter 1 about proof methods and proof strategy in our exploration of number theory.
- Mathematicians have long considered number theory to be pure mathematics, but it has important applications to computer science and cryptography studied in Sections 4.5 and 4.6.

Chapter Summary

- Divisibility and Modular Arithmetic
- Primes and Greatest Common Divisors
- Solving Congruencies
- Applications of Congruencies
- Cryptography

Divisibility and Modular Arithmetic

Section 4.1

Section Summary

- Division
- Division Algorithm
- Modular Arithmetic

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 say that a is a *factor* or *divisor* of b and that b is a multiple of a.
- The notation $a \mid b$ denotes that a divides b.
- If $a \mid b$, then b/a is an integer.
- If a does not divide b, we write $a \nmid b$.

Example: Determine whether $3 \mid 7$ and whether $3 \mid 12$.

Properties of Divisibility

Theorem 1: Let a, b, and c be integers, where $a \neq 0$.

i. If $a \mid b$ and $a \mid c$, then $a \mid (b + c)$;
ii. If $a \mid b$, then $a \mid bc$ for all integers c;
iii. If $a \mid b$ and $b \mid c$, then $a \mid c$.

Proof: (i) Suppose $a \mid b$ and $a \mid c$, then it follows that there are integers s and t with $b = as$ and $c = at$. Hence,

$b + c = as + at = a(s + t)$. Hence, $a \mid (b + c)$

(Exercises 3 and 4 ask for proofs of parts (ii) and (iii).) ◄

Corollary: If a, b, and c be integers, where $a \neq 0$, such that $a \mid b$ and $a \mid c$, then $a \mid mb + nc$ whenever m and n are integers.

Can you show how it follows easily from from (ii) and (i) of Theorem 1?

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 is really a theorem.

Division Algorithm: 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 $a = dq + r$ (*proved in Section* 5.2).

- d is called the *divisor*.
- a is called the *dividend*.
- q is called the *quotient*.
- r is called the *remainder*.

Definitions of Functions **div** and **mod**

$q = a$ **div** d

$r = a$ **mod** d

Examples:

- What are the quotient and remainder when 101 is divided by 11?

  Solution: The quotient when 101 is divided by 11 is 9 = 101 **div** 11, and the remainder is 2 = 101 **mod** 11.
- What are the quotient and remainder when −11 is divided by 3?

  Solution: The quotient when −11 is divided by 3 is −4 = −11 **div** 3, and the remainder is 1 = −11 **mod** 3.

Congruence Relation

Definition: If a and b are integers and m is a positive integer, then a is *congruent* to b *modulo* m if m divides $a - b$.

- 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

$$a \not\equiv b \pmod{m}$$

Example: Determine whether 17 is congruent to 5 modulo 6 and whether 24 and 14 are congruent modulo 6.

Solution:

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

More on Congruences

Theorem 4: 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 $a \equiv b \pmod{m}$. ◀

The Relationship between (mod m) and **mod** m Notations

- The use of "mod" in $a \equiv b \pmod{m}$ and $a \textbf{ mod } m = b$ are different.
  - $a \equiv b \pmod{m}$ is a relation on the set of integers.
  - In $a \textbf{ mod } m = b$, the notation **mod** denotes a function.
- The relationship between these notations is made clear in this theorem.
- **Theorem 3**: Let a and b be integers, and let m be a positive integer. Then $a \equiv b \pmod{m}$ if and only if $a \textbf{ mod } m = b \textbf{ mod } m$. (*Proof in the exercises*)

Congruencies of Sums and Products

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

$$a + c \equiv b + d \pmod{m} \text{ and } ac \equiv bd \pmod{m}$$

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

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

$$77 = 7 \;\; 11 \equiv 2 + 1 = 3 \pmod{5}$$

Algebraic Manipulation of Congruencies

- Multiplying both sides of a valid congruence by an integer preserves validity.

  If $a \equiv b \pmod{m}$ holds then $c \cdot a \equiv c \cdot b \pmod{m}$, where c is any integer, holds by Theorem 5 with $d = c$.

- Adding an integer to both sides of a valid congruence preserves validity.

  If $a \equiv b \pmod{m}$ holds then $c + a \equiv c + b \pmod{m}$, where c is any integer, holds by Theorem 5 with $d = c$.

- Dividing a congruence by an integer does not always produce a valid congruence.

  Example: The congruence $14 \equiv 8 \pmod{6}$ holds. But dividing both sides by 2 does not produce a valid congruence since $14/2 = 7$ and $8/2 = 4$, but $7 \not\equiv 4 \pmod{6}$.

Computing the **mod** *m* Function of Products and Sums

- We use the following corollary to Theorem 5 to compute the remainder of the product or sum of two integers when divided by m from the remainders when each is divided by m.

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

$$(a + b)\ (\mathbf{mod}\ m) = ((a\ \mathbf{mod}\ m) + (b\ \mathbf{mod}\ m))\ \mathbf{mod}\ m$$

and

$$ab\ \mathbf{mod}\ m = ((a\ \mathbf{mod}\ m)\ (b\ \mathbf{mod}\ m))\ \mathbf{mod}\ m.$$

(*proof in text*)

Applications of Congruences

Section Summary

- Hashing Functions
- Pseudorandom Numbers
- Check Digits

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 $h(k) = k$ **mod** m, where m is the number of memory locations.
- Because this hashing function is onto, all memory locations are possible.

Example: Let $h(k) = k$ **mod** 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 **mod** 111 = 14

h(037149212) = 037149212 **mod** 111 = 65

h(107405723) = 107405723 **mod** 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*:

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

- 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

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

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

- If psudorandom 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 $x_{n+1} = (7x_n + 4)$ **mod** 9, with $x_0 = 3$.

  $x_1 = 7x_0 + 4 \textbf{ mod } 9 = 7 \cdot 3 + 4 \textbf{ mod } 9 = 25 \textbf{ mod } 9 = 7,$
  $x_2 = 7x_1 + 4 \textbf{ mod } 9 = 7 \cdot 7 + 4 \textbf{ mod } 9 = 53 \textbf{ mod } 9 = 8,$
  $x_3 = 7x_2 + 4 \textbf{ mod } 9 = 7 \cdot 8 + 4 \textbf{ mod } 9 = 60 \textbf{ mod } 9 = 6,$
  $x_4 = 7x_3 + 4 \textbf{ mod } 9 = 7 \cdot 6 + 4 \textbf{ mod } 9 = 46 \textbf{ mod } 9 = 1,$
  $x_5 = 7x_4 + 4 \textbf{ mod } 9 = 7 \cdot 1 + 4 \textbf{ mod } 9 = 11 \textbf{ mod } 9 = 2,$
  $x_6 = 7x_5 + 4 \textbf{ mod } 9 = 7 \cdot 2 + 4 \textbf{ mod } 9 = 18 \textbf{ mod } 9 = 0,$
  $x_7 = 7x_6 + 4 \textbf{ mod } 9 = 7 \cdot 0 + 4 \textbf{ mod } 9 = 4 \textbf{ mod } 9 = 4,$
  $x_8 = 7x_7 + 4 \textbf{ mod } 9 = 7 \cdot 4 + 4 \textbf{ mod } 9 = 32 \textbf{ mod } 9 = 5,$
  $x_9 = 7x_8 + 4 \textbf{ mod } 9 = 7 \cdot 5 + 4 \textbf{ mod } 9 = 39 \textbf{ mod } 9 = 3.$

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

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

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

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

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

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

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

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

The validity of an ISBN-10 number can be evaluated with the equivalent $\sum_{i=1}^{10} i x_i \equiv 0 \pmod{11}.$

a. Suppose that the first 9 digits of the ISBN-10 are 007288008. What is the check digit?

b. Is 084930149X a valid ISBN10?

Solution:

X is used for the digit 10.

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

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

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

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

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

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*

Definitions: Let $\mathbf{Z}_m$ be the set of nonnegative integers less than m: $\{0,1, \ldots., m-1\}$

- The operation $+_m$ is defined as $a +_m b = (a + b)$ **mod** m. This is *addition modulo m*.
- The operation $\cdot_m$ is defined as $a \cdot_m b = (a \,.\, b)$ **mod** 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:

- $7 +_{11} 9 = (7 + 9)$ **mod** $11 = 16$ **mod** $11 = 5$
- $7 \cdot_{11} 9 = (7 \cdot 9)$ **mod** $11 = 63$ **mod** $11 = 8$

Arithmetic Modulo *m*

- The operations $+_m$ and $\cdot_m$ satisfy many of the same properties as ordinary addition and multiplication.
  - *Closure*: If a and b belong to $\mathbf{Z}_m$, then $a +_m b$ and $a \cdot_m b$ belong to $\mathbf{Z}_m$.
  - *Associativity*: If a, b, and c belong to $\mathbf{Z}_m$, 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 and b belong to $\mathbf{Z}_m$, 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 belongs to $\mathbf{Z}_m$, then $a +_m 0 = a$ and $a \cdot_m 1 = a$.

continued →

Arithmetic Modulo *m*

- *Additive inverses*: If $a \neq 0$ belongs to $\mathbf{Z}_m$, then $m - a$ is the additive inverse of a modulo m and 0 is its own additive inverse.
  - $a +_m (m - a) = 0$ and $0 +_m 0 = 0$
- *Distributivity*: If *a*, *b*, and *c* belong to $\mathbf{Z}_m$, then
  - $a \cdot_m (b +_m c) = (a \cdot_m b) +_m (a \cdot_m c)$ and $(a +_m b) \cdot_m c = (a \cdot_m c) +_m (b \cdot_m c)$.
- Exercises 42-44 ask for proofs of these properties.
- Multiplicatative inverses have not been included since they do not always exist. For example, there is no multiplicative inverse of 2 modulo 6.
- (*optional*) Using the terminology of abstract algebra, $\mathbf{Z}_m$ with $+_m$ is a commutative group and $\mathbf{Z}_m$ with $+_m$ and $\cdot_m$ is a commutative ring.

Primes and Greatest Common Divisors

Section 4.3

Section Summary

- Prime Numbers and their Properties
- Greatest Common Divisors and Least Common Multiples
- The Euclidian Algorithm
- gcds as Linear Combinations

Primes

Definition: A positive integer p greater than 1 is called *prime* if the only positive factors of p are 1 and p. A positive integer that is greater than 1 and is not prime is called *composite*.

Example: The integer 7 is prime because its only positive factors are 1 and 7, but 9 is composite because it is divisible by 3.

The Fundamental Theorem of Arithmetic

Theorem: Every positive integer greater than 1 can be written uniquely as a prime or as the product of two or more primes where the prime factors are written in order of non decreasing size.

Examples:

- $100 = 2 \cdot 2 \cdot 5 \cdot 5 = 2^2 \cdot 5^2$
- $641 = 641$
- $999 = 3 \cdot 3 \cdot 3 \cdot 37 = 3^3 \cdot 37$
- $1024 = 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 = 2^{10}$

The Sieve of Erastosthenes

TABLE 1 **The Sieve of Eratosthenes.**

Integers divisible by 2 other than 2 receive an underline.

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

Integers divisible by 3 other than 3 receive an underline.

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

Integers divisible by 5 other than 5 receive an underline.

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

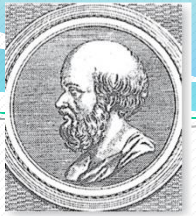
Integers divisible by 7 other than 7 receive an underline; integers in color are prime.

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

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

To see this, note that if $n = ab$, then $a \leq \sqrt{n}$ or $b \leq \sqrt{n}$.

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.



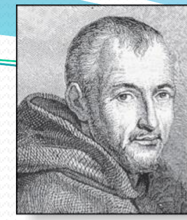
Erastothenes
(276-194 B.C.)

The Sieve of Erastosthenes

- The *Sieve of Erastosthenes* can be used to find all primes not exceeding a specified positive integer. For example, begin with the list of integers between 1 and 100.
  a. Delete all the integers, other than 2, divisible by 2.
  b. Delete all the integers, other than 3, divisible by 3.
  c. Next, delete all the integers, other than 5, divisible by 5.
  d. Next, delete all the integers, other than 7, divisible by 7.
  e. Since all the remaining integers are not divisible by any of the previous integers, other than 1, the primes 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}

continued →



Marin Mersenne
(1588-1648)

Mersenne Primes

Definition: Prime numbers of the form $2^p - 1$, where p is prime, are called *Mersenne primes*.

- $2^2 - 1 = 3$, $2^3 - 1 = 7$, $2^5 - 1 = 37$, and $2^7 - 1 = 127$ are Mersenne primes.
- $2^{11} - 1 = 2047$ is not a Mersenne prime since $2047 = 23 \cdot 89$.
- There is an efficient test for determining if $2^p - 1$ is prime.
- The largest known prime numbers are Mersenne primes.
- As of mid 2011, 47 Mersenne primes were known, the largest is $2^{43,112,609} - 1$, which has nearly 13 million decimal digits.
- The *Great Internet Mersenne Prime Search* (*GIMPS*) is a distributed computing project to search for new Mersenne Primes.

http://www.mersenne.org/

Greatest Common Divisor

Definition: Let a and b be integers, not both zero. The largest integer d such that $d \mid a$ and also $d \mid b$ is called the greatest common divisor of a and b. The greatest common divisor of a and b is denoted by gcd(a,b).

One can find greatest common divisors of small numbers by inspection.

Example: What is the greatest common divisor of 24 and 36?

Solution: gcd(24,36) = 12

Example: What is the greatest common divisor of 17 and 22?

Solution: gcd(17,22) = 1

Greatest Common Divisor

Definition: The integers a and b are *relatively prime* if their greatest common divisor is 1.

Example: 17 and 22

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

Example: Determine whether the integers 10, 17 and 21 are pairwise relatively prime.

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

Example: Determine whether the integers 10, 19, and 24 are pairwise relatively prime.

Solution: Because $\gcd(10,24) = 2$, 10, 19, and 24 are not pairwise relatively prime.

Finding the Greatest Common Divisor Using Prime Factorizations

- Suppose the prime factorizations of *a* and *b* are:

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

where each exponent is a nonnegative integer, and where all primes occurring in either prime factorization are included in both. Then:

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

- This formula is valid since the integer on the right (of the equals sign) divides both *a* and *b*. No larger integer can divide both *a* and *b*.

Example: $120 = 2^3 \cdot 3 \cdot 5 \quad 500 = 2^2 \cdot 5^3$

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

- Finding the gcd of two positive integers using their prime factorizations is not efficient because there is no efficient algorithm for finding the prime factorization of a positive integer.

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 lcm(*a*,*b*).

- The least common multiple can also be computed from the prime factorizations.

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

This number is divided by both *a* and *b* and no smaller number is divided by *a* and *b*.

Example: $\mathrm{lcm}(2^3 3^5 7^2,\ 2^4 3^3) = 2^{\max(3,4)}\ 3^{\max(5,3)}\ 7^{\max(2,0)} = 2^4\ 3^5\ 7^2$

- The greatest common divisor and the least common multiple of two integers are related by:

Theorem 5: Let a and b be positive integers. Then

$$ab = \gcd(a,b) \cdot \mathrm{lcm}(a,b)$$

(*proof is Exercise* 31)

Euclidean Algorithm

Euclid
(325 B.C.E. – 265 B.C.E.)

- The Euclidian algorithm is an efficient method for computing the greatest common divisor of two integers. It is based on the idea that gcd(a,b) is equal to gcd(a,c) when $a > b$ and c is the remainder when a is divided by b.

Example: Find gcd(91, 287):

- $287 = 91 \cdot 3 + 14$ Divide 287 by 91
- $91 = 14 \cdot 6 + 7$ Divide 91 by 14
- $14 = 7 \cdot 2 + 0$ Divide 14 by 7

Stopping condition

gcd(287, 91) = gcd(91, 14) = gcd(14, 7) = 7

continued →

Euclidean Algorithm

- The Euclidean algorithm expressed in pseudocode is:

```
procedure gcd(a, b: positive integers)
x := a
x := b
while  y ≠ 0
      r := x mod y
      x := y
      y := r
return x {gcd(a,b) is x}
```

- In Section 5.3, we'll see that the time complexity of the algorithm is $O(\log b)$, where $a > \mathrm{b}$.

Étienne Bézout (1730-1783)

gcds as Linear Combinations

Bézout's Theorem: If a and b are positive integers, then there exist integers s and t such that $\gcd(a,b) = sa + tb$.

Definition: If a and b are positive integers, then integers s and t such that $\gcd(a,b) = sa + tb$ are called *Bézout coefficients* of a and b. The equation $\gcd(a,b) = sa + tb$ is called *Bézout's identity*.

- By Bézout's Theorem, the gcd of integers a and b can be expressed in the form $sa + tb$ where s and t are integers. This is a *linear combination* with integer coefficients of a and b.
  - $\gcd(6,14) = (-2)\cdot 6 + 1\cdot 14$

Finding gcds as Linear Combinations

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

Solution: First use the Euclidean algorithm to show gcd(252,198) = 18

i. $252 = 1 \cdot 198 + 54$
ii. $198 = 3 \cdot 54 + 36$
iii. $54 = 1 \cdot 36 + 18$
iv. $36 = 2 \cdot 18$

- Now working backwards, from iii and i above
  - $18 = 54 - 1 \cdot 36$
  - $36 = 198 - 3 \cdot 54$
- Substituting the 2nd equation into the 1st yields:
  - $18 = 54 - 1 \cdot (198 - 3 \cdot 54) = 4 \cdot 54 - 1 \cdot 198$
- Substituting $54 = 252 - 1 \cdot 198$ (from i)) yields:
  - $18 = 4 \cdot (252 - 1 \cdot 198) - 1 \cdot 198 = 4 \cdot 252 - 5 \cdot 198$

- This method illustrated above is a two pass method. It first uses the Euclidian algorithm to find the gcd and then works backwards to express the gcd as a linear combination of the original two integers. A one pass method, called the *extended Euclidean algorithm*, is developed in the exercises.

Dividing Congruencies by an Integer

- Dividing both sides of a valid congruence by an integer does not always produce a valid congruence (see Section 4.1).
- But dividing by an integer relatively prime to the modulus does produce a valid congruence:

Theorem 7: Let m be a positive integer and let a, b, and c be integers. If $ac \equiv bc \pmod{m}$ and $\gcd(c,m) = 1$, then $a \equiv b \pmod{m}$.

Proof: Since $ac \equiv bc \pmod{m}$, $m \mid ac - bc = c(a - b)$ by Lemma 2 and the fact that $\gcd(c,m) = 1$, it follows that $m \mid a - b$. Hence, $a \equiv b \pmod{m}$. ◄

Solving Congruencies

Section 4.4

Section Summary

- Linear Congruencies
- The Chinese Remainder Theorem
- Fermat's Little Theorem
- Pseudo primes

Linear Congruencies

Definition: A congruence of the form

$$ax \equiv b(\bmod m),$$

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

- The solutions to a linear congruence $ax \equiv b(\bmod m)$ are all integers x that satisfy the congruence.

Definition: An integer $\bar{a}$ such that $\bar{a}a \equiv 1(\bmod m)$ is said to be an *inverse* of a modulo m.

Example: 5 is an inverse of 3 modulo 7 since $5 \cdot 3 = 15 \equiv 1(\bmod 7)$

- One method of solving linear congruencies makes use of an inverse $\bar{a}$, if it exists. Although we can not divide both sides of the congruence by a, we can multiply by $\bar{a}$ to solve for x.

Inverse of a modulo m

- The following theorem guarantees that an inverse of a modulo m exists whenever a and m are relatively prime. Two integers a and b are relatively prime when $\gcd(a,b) = 1$.

Theorem 1: If a and m are relatively prime integers and $m > 1$, then an inverse of a modulo m exists. Furthermore, this inverse is unique modulo m. (This means that there is a unique positive integer $\bar{a}$ less than m that is an inverse of a modulo m and every other inverse of a modulo m is congruent to $\bar{a}$ modulo m.)

Proof: Since $\gcd(a,m) = 1$, by Theorem 6 of Section 4.3, 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 an inverse of a modulo m.
- The uniqueness of the inverse is Exercise 7. ◀

Finding Inverses

- The Euclidean algorithm and Bézout coefficients gives us a systematic approaches to finding inverses.

  Example: Find an inverse of 3 modulo 7.

  Solution: Because $\gcd(3,7) = 1$, by Theorem 1, an inverse of 3 modulo 7 exists.

  - Using the Euclidian algorithm: $7 = 2\cdot 3 + 1$.
  - From this equation, we get $-2\cdot 3 + 1\cdot 7 = 1$, and see that -2 and 1 are Bézout coefficients of 3 and 7.
  - Hence, -2 is an inverse of 3 modulo 7.
  - Also every integer congruent to -2 modulo 7 is an inverse of 3 modulo 7, i.e., 5, -9, 12, etc.

Finding Inverses

Example: Find an inverse of 101 modulo 4620.

Solution: First use the Euclidian algorithm to show that gcd(101,4620) = 1.

$$4620 = 45 \cdot 101 + 75$$
$$101 = 1 \cdot 75 + 26$$
$$75 = 2 \cdot 26 + 23$$
$$26 = 1 \cdot 23 + 3$$
$$23 = 7 \cdot 3 + 2$$
$$3 = 1 \cdot 2 + 1$$
$$2 = 2 \cdot 1$$

Since the last nonzero remainder is 1, gcd(101,4260) = 1

Working Backwards:

$$1 = 3 - 1 \cdot 2$$
$$1 = 3 - 1 \cdot (23 - 7 \cdot 3) = -1 \cdot 23 + 8 \cdot 3$$
$$1 = -1 \cdot 23 + 8 \cdot (26 - 1 \cdot 23) = 8 \cdot 26 - 9 \cdot 23$$
$$1 = 8 \cdot 26 - 9 \cdot (75 - 2 \cdot 26) = 26 \cdot 26 - 9 \cdot 75$$
$$1 = 26 \cdot (101 - 1 \cdot 75) - 9 \cdot 75$$
$$= 26 \cdot 101 - 35 \cdot 75$$
$$1 = 26 \cdot 101 - 35 \cdot (42620 - 45 \cdot 101)$$
$$= -35 \cdot 42620 + 1601 \cdot 101$$

Bézout coefficients : − 35 and 1601

1601 is an inverse of 101 modulo 42620

Using Inverses to Solve Congruences

- We can solve the congruence $ax \equiv b(\bmod m)$ by multiplying both sides by $\bar{a}$.

  Example: What are the solutions of the congruence $3x \equiv 4(\bmod 7)$.

  Solution: We found that -2 is an inverse of 3 modulo 7 (two slides back). We multiply both sides of the congruence by -2 giving

  $$-2 \cdot 3x \equiv -2 \cdot 4 (\bmod 7).$$

  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 if every x with $x \equiv 6 \pmod 7$ is a solution. Assume that $x \equiv 6 \pmod 7$. By Theorem 5 of Section 4.1, it follows that $3x \equiv 3 \cdot 6 = 18 \equiv 4(\bmod 7)$ which shows that all such x satisfy the congruence.

  The solutions are the integers x such that $x \equiv 6 \pmod 7$, namely, $6, 13, 20 \ldots$ and $-1, -8, -15, \ldots$

The Chinese Remainder Theorem

Theorem 2: (*The Chinese Remainder Theorem*) Let $m_1, m_2, \ldots, m_n$ be pairwise relatively prime positive integers greater than one and $a_1, a_2, \ldots, a_n$ arbitrary integers. Then the system

$$
\begin{aligned}
x &\equiv a_1 \pmod{m_1} \\
x &\equiv a_2 \pmod{m_2} \\
&\cdot \\
&\cdot \\
&\cdot \\
x &\equiv a_n \pmod{m_n}
\end{aligned}
$$

has a unique solution modulo $m = m_1 m_2 \cdots m_n$.

(That is, there is a solution x with $0 \leq x < m$ and all other solutions are congruent modulo m to this solution.)

- **Proof**: We'll show that a solution exists by describing a way to construct the solution. Showing that the solution is unique modulo m is Exercise 30.

continued →

The Chinese Remainder Theorem

To construct a solution first let $M_k = m/m_k$ for $k = 1,2,\ldots,n$ and $m = m_1 m_2 \cdots m_n$.

Since $\gcd(m_k, M_k) = 1$, by Theorem 1, there is an integer y_k, an inverse of M_k modulo m_k, such that

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

Form the sum

$$x = a_1 M_1 y_1 + a_2 M_2 y_2 + \cdots + a_n M_n y_n \pmod{m}$$

Note that because $M_j \equiv 0 \pmod{m_k}$ whenever $j \neq k$, all terms except the kth term in this sum are congruent to 0 modulo m_k.

Because $M_k y_k \equiv 1 \pmod{m_k}$, we see that $x \equiv a_k M_k y_k \equiv a_k \pmod{m_k}$, for $k = 1,2,\ldots,n$.

Hence, x is a simultaneous solution to the n congruences.

$$x \equiv a_1 \pmod{m_1}$$
$$x \equiv a_2 \pmod{m_2}$$
$$\vdots$$
$$x \equiv a_n \pmod{m_n}$$

The Chinese Remainder Theorem

Example: Consider the 3 congruences from Sun-Tsu's problem:

$x \equiv 2 \pmod 3$, $x \equiv 3 \pmod 5$, $x \equiv 2 \pmod 7$.

- Let $m = 3 \cdot 5 \cdot 7 = 105$, $M_1 = m/3 = 35$, $M_3 = m/5 = 21$, $M_3 = m/7 = 15$.
- We see that
  - 2 is an inverse of $M_1 = 35$ modulo 3 since $35 \cdot 2 \equiv 2 \cdot 2 \equiv 1 \pmod 3$
  - 1 is an inverse of $M_2 = 21$ modulo 5 since $21 \equiv 1 \pmod 5$
  - 1 is an inverse of $M_3 = 15$ modulo 7 since $15 \equiv 1 \pmod 7$
- Hence,

$$x = a_1M_1y_1 + a_2M_2y_2 + a_3M_3y_3 \pmod m$$
$$= 2 \cdot 35 \cdot 2 + 3 \cdot 21 \cdot 1 + 2 \cdot 15 \cdot 1 = 233 \equiv 23 \pmod{105}$$

- We have shown that 23 is the smallest positive integer that is a simultaneous solution. Check it!

The Chinese Remainder Theorem

Word 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?

Solution:

We have the following congruences

$x \equiv 4 \pmod{5}$,

$x \equiv 6 \pmod{8}$,

$x \equiv 8 \pmod{9}$.

Cryptography

Section Summary

- Classical Cryptography
- Cryptosystems
- Public Key Cryptography
- RSA Cryptosystem
- Fermat's Little theorem

Caesar Cipher

Julius Caesar created secret messages by shifting each letter three letters forward in the alphabet (sending the last three letters to the first three letters.) For example, the letter B is replaced by E and the letter X is replaced by A. This process of making a message secret is an example of *encryption*.

Here is how the encryption process works:

- Replace each letter by an integer from $\mathbf{Z}_{26}$, that is an integer from 0 to 25 representing one less than its position in the alphabet.
- The encryption function is $f(p) = (p + 3)$ **mod** 26. It replaces each integer p in the set $\{0,1,2,\ldots,25\}$ by $f(p)$ in the set $\{0,1,2,\ldots,25\}$.
- Replace each integer p by the letter with the position $p + 1$ in the alphabet.

Example: Encrypt the message "MEET YOU IN THE PARK" using the Caesar cipher.

Solution: 12 4 4 19 24 14 20 8 13 19 7 4 15 0 17 10.

Now replace each of these numbers p by $f(p) = (p + 3)$ **mod** 26.

15 7 7 22 1 17 23 11 16 22 10 7 18 3 20 13.

Translating the numbers back to letters produces the encrypted message

"PHHW BRX LQ WKH SDUN."

Caesar Cipher

- To recover the original message, use $f^{-1}(p) = (p-3)$ **mod** 26. So, each letter in the coded message is shifted back three letters in the alphabet, with the first three letters sent to the last three letters. This process of recovering the original message from the encrypted message is called *decryption*.
- The Caesar cipher is one of a family of ciphers called *shift ciphers*. Letters can be shifted by an integer k, with 3 being just one possibility. The encryption function is

  $$f(p) = (p + k) \textbf{ mod } 26$$

  and the decryption function is

  $$f^{-1}(p) = (p-k) \textbf{ mod } 26$$

  The integer k is called a *key*.

Shift Cipher

Example 1: Encrypt the message "STOP GLOBAL WARMING" using the shift cipher with $k = 11$.

Solution: Replace each letter with the corresponding element of $\mathbf{Z}_{26}$.

18 19 14 15   6 11 14 1 0 11   22 0 17 12 8 13 6.

Apply the shift $f(p) = (p + 11)$ **mod** 26, yielding

3 4 25 0   17 22 25 12 11 22   7 11 2 23 19 24 17.

Translating the numbers back to letters produces the ciphertext

"DEZA RWZMLW HLCXTYR."

Shift Cipher

Example 2: Decrypt the message "LEWLYPLUJL PZ H NYLHA ALHJOLY" that was encrypted using the shift cipher with $k = 7$.

Solution: Replace each letter with the corresponding element of $\mathbf{Z}_{26}$.

11 4 22 11 24 15 11 20 9 11   15 25   7   13 24 11 7 0   0 11 7 9 14 11 24.

Shift each of the numbers by $-k = -7$ modulo 26, yielding

4 23 15 4 17 8 4 13 2 4   8 18   0   6 17 4 0 19   19 4 0 2 7 4 17.

Translating the numbers back to letters produces the decrypted message

"EXPERIENCE IS A GREAT TEACHER."

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.

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

$$c \equiv m^e \mod n$$

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

$$m \equiv c^d \mod n$$

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

Using RSA

Given pubKey = <e, n> and privKey = <d, n>

If Message = m

Then:

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

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

signature: s = md mod n, m < n

verification: m = se mod n

Example of RSA (1)

Choose p = 7 and q = 17.

Compute n = p*q= 119.

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

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

Compute d = _77_such that e*d=1 mod f(n).

- Public key: <5,119>
- Private key: <77,119>
- Message = 19
- Encryption: 19^5 mod 119 = 66
- Decryption: 66^{77} mod 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} mod 77 = 28; 04^{17} mod 77 = 16

– 11^{17} mod 77 = 44; – 11^{17} mod 77 = 44

– 14^{17} mod 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)

Fermat's Little Theorem

Pierre de Fermat
(1601-1665)

Theorem 3: (*Fermat's Little Theorem*) If p is prime and a is an integer not divisible by p, then $a^{p-1} \equiv 1 \pmod{p}$

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

(*proof outlined in Exercise 19*)

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

Example: Find 7^{222} **mod** 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} **mod** $11 = 5$.