LECTURES ON ELEMENTARY NUMBER THEORY

SHAOFANG HONG

Mathematical College
Sichuan University
Chengdu 610064
The People's Republic of China

Prof. Dr. Shaofang Hong, Mathematical College, Sichuan University, Chengdu 610064, China.

.

Contents

Chapter 1. Divisibility and the Fundamental Theorem of Arithmetic

§1.1. Division algorithm and m-adic representation

Divisibility is a basic and significant concept in number theory. Let us first introduce this concept.

Definition. Let a and d be integers. If there exists an integer q such that $a = dq$, then we say that d *divides* a, or that a is *divisible* by d, or that d is a *divisor* of a, or that a is a *multiple* of d. If there does not exist any integer q, we say that d does not divide a, or that a is not divisible by d, or that d is not a divisor of a, or that a is not a multiple of d.

If d divides a, then we write $d|a$. Otherwise we write $d \not| a$.

The following properties of divisibility are clear to be true and so we omit their proofs.

Proposition 1.1.1. *Assume that a, b, c are integers. Then each of the following is true:*

(i). *We have $a|0$, $1|b$ and $a|a$;*

(ii). *If $a|b, b|c$, then $a|c$;*

(iii). *If $a|b$, then $ac|bc$;*

(iv). *If $a|b, a|c$, then for any integers m, n we have $a|mb + nc$;*

(v). *If $a|b$ and $b \neq 0$, then $|a| \leq |b|$;*

(vi). *If $ac|bc$ and $c \neq 0$, then $a|b$;*

(vii). *If $a|b$ and $a \neq 0$, then $\frac{b}{a}|b$.*

Proof. This is clear true. □

The *minimum principle* states that every nonempty set of integers bounded below contains a smallest element. For example, a nonempty set of nonnegative integers must contain a smallest element. We can see the necessity of the condition that the nonempty set be bounded below by considering the example of the set $\mathbf{Z}$ of all integers, positive, negative and zero.

The minimum principle is all we need to prove the following important result.

Theorem 1.1.2. (Division algorithm) *Let a and d be integers with $d \geq 1$. There exist unique integers q and r such that*

$$a = dq + r \tag{1.1}$$

and

$$0 \leq r \leq d - 1. \tag{1.2}$$

The integer q is called the *quotient* and the integer r is called the *remainder* in the division of a by d.

Proof. Let

$$S := \{n \geq 0 | n = a - dx, x \in \mathbf{Z}\}.$$

Then S is a nonempty set of nonnegative integers. Actually, if $a > 0$, then $a = a - d \times 0 \in S$. If $a < 0$, then let $x = -y$, where y is a positive integer. Since d is positive, we have $a - dx = a + dy \in S$ if y is sufficiently large. Thus S is nonempty.

Now by the minimum principle, S contains the smallest element r, and $r = a - dq \geq 0$ for some $q \in \mathbf{Z}$. Assume that $r \geq d$. Then

$$0 \leq r - d = a - d(q + 1) < r$$

and $r-d \in S$, which contradicts with the minimality of r. Therefore q and r satisfy (1.1) and (1.2).

Let q_1, r_1, q_2, r_2 be integers such that $a = dq_1+r_1 = dq_2+r_2$ and $0 \le r_1, r_2 \le d-1$. Then $|r_1 - r_2| \le d-1$ and $d(q_1-q_2) = r_2 - r_1$. If $q_1 \ne q_2$, then $|q_1 - q_2| \ge 1$ and

$$d \le d|q_1 - q_2| = |r_2 - r_1| \le d-1,$$

which is impossible. Thus we must have $q_1 = q_2$ and $r_1 = r_2$. This proves that the quotient and remainder are unique. □

Example. Division of 17 by 6 gives the quotient 2 and remainder 5, i.e.

$$17 = 6 \times 2 + 5.$$

Division of -17 by 6 gives quotient -3 and the remainder 1, namely,

$$-17 = 6 \times (-3) + 1.$$

A simple geometric way to picture the division algorithm is to imagine the real number line with dots at the positive integers. Let q be a positive integer, and put a large dot on each multiple of q. The integer a either lies on one of these large dots, in which case a is a multiple of q, or a lies on a dot strictly between two large dots, that is, between two successive multiple of q, and the distance r between a and largest multiple of q that is less than a is a positive integer no greater than $q-1$. For example, if $q=7$ and $a = \pm 16$, we have the following picture.

The *principle of mathematical induction* states that if $S(k)$ is some statement about integers $k \ge k_0$ such that $S(k_0)$ is true and such that the truth of $S(k-1)$ implies the truth of $S(k)$, then $S(k)$ holds for all integers $k \ge k_0$. Another form of the principle of mathematical induction states that if $S(k_0)$ is true and if the truth of $S(k_0), S(k_0+1), ..., S(k-1)$ implies the truth of $S(k)$, then $S(k)$ holds for all integers $k \ge k_0$. Mathematical induction is equivalent to the minimum principle (Exercise to you!).

Using mathematical induction and the division algorithm, we can prove the existence and uniqueness of *m-adic representations* of integers.

Theorem 1.1.3. *Let m be an integer, $m \ge 2$. Every positive integer n can be represented uniquely in the form*

$$n = a_0 + a_1 m + a_2 m^2 + ... + a_k m^k, \tag{1.3}$$

where k is the nonnegative integer such that $m^k \le n < m^{k+1}$ and $a_0, a_1, ..., a_k$ are integers such that $1 \le a_k \le m-1$ and $0 \le a_i \le m-1$ for $0 \le i \le k-1$.

This is called the *m-adic representation* of n. The integers a_i are called the *digits* of n to base m (or is called the *digits* of m-adic representation of n). Equivalently, we can write $n = \sum\limits_{i=0}^{\infty} a_i m^i$, where $\le a_i \le m-1$ for all i, and $a_i = 0$ for all sufficiently large integers i.

Proof. For $k \ge 0$, let $S(k)$ be the statement that every integer in the interval $m^k \le n < m^{k+1}$ has a unique m-adic representation. Evidently, to show Theorem 1.1.3, it is sufficient to show that $S(k)$ is true for all nonnegative integers k. We use induction on k to prove it. The statement $S(0)$ is true because if $1 \le n < m$, then $n = a_0$ is the unique m-adic representation.

Let $k \geq 1$, and assume that the statements $S(0), S(1), ..., S(k-1)$ are true. We shall prove $S(k)$. Let $m^k \leq n < m^{k+1}$. By the division algorithm, we can divide n by m^k and obtain $n = a_k m^k + r$, where $0 \leq r < m^k$. Then

$$0 < m^k - r \leq n - r = a_k m^k \leq n < m^{k+1}.$$

Dividing this inequality by m^k, we obtain $0 < a_k < m$. Since m and a_k are integers, it follows that $1 \leq a_k \leq m-1$. If $r = 0$, then $n = a_k m^k$ is an m-adic representation. If $r \geq 1$, then $m^{k'} \leq r < m^{k'+1}$ for some nonnegative integer $k' \leq k-1$. By the induction assumption, $S(k')$ is true and r has a unique m-adic representation of the form

$$r = a_0 + a_1 m + ... + a_{k-1} m^{k-1}$$

with $0 \leq a_i \leq m-1$ for $i = 0, 1, ..., k-1$. It follows that n has the m-adic representation

$$n = a_0 + a_1 m + ... + a_{k-1} m^{k-1} + a_k m^k.$$

In the following we show that this representation is unique. Let

$$n = b_0 + b_1 m + ... + b_l m^l$$

be another m-adic representation of n, where $0 \leq b_j \leq m-1$ for all $j = 0, 1, ..., l$ and $b_l \geq 1$. First we show that $k = l$. If $l \geq k+1$, then

$$n < m^{k+1} \leq b_l m^l \leq n,$$

which is impossible. Thus $l \leq k$. Then by the symmetry of k and l, one gets that $k \leq l$. Therefore $k = l$. Consequently, we show that $a_k = b_k$. If $a_k < b_k$, then

$$\begin{aligned} n &= a_0 + a_1 m + ... + a_{k-1} m^{k_1} + a_k m^k \\ &\leq (m-1) + (m-1)m + ... + (m-1)m^{k-1} + a_k m^k \\ &= (m^k - 1) + a_k m^k \\ &< (a_k + 1)m^k \leq b_k m^k \leq n, \end{aligned}$$

which again is impossible. Therefore, $b_k \leq a_k$. By symmetry, we have $a_k \leq b_k$ and so $a_k = b_k$. Then

$$\begin{aligned} n - a_k m^k &= a_0 + a_1 m + a_2 m^2 + ... + a_{k-1} m^{k-1} \\ &= b_0 + b_1 m + b_2 m^2 + ... + b_{k-1} m^{k-1}. \end{aligned}$$

Since $n - a_k m^k < m^k$, by the induction assumption, $a_i = b_i$ for $i = 0, 1, ..., k-1$. Thus, the m-adic representation of n exists and is unique, and $S(k)$ is true. By mathematical induction, $S(k)$ holds for all $k \geq 0$. □

Example. (i). The 2-adic representation of 100 is $100 = 1 \times 2^2 + 1 \times 2^5 + 1 \times 2^6$,
3-adic representation of 100 is $100 = 1 + 2 \times 3^2 + 1 \times 3^4$,
4-adic representation of 100 is $100 = 1 \times 4 + 2 \times 4^2 + 1 \times 4^3$,
and 5-adic representation of 100 is $100 = 4 \times 5^2$.
(ii). The 10-adic representation of 2014 is $2014 = 4 + 1 \times 10^1 + 2 \times 10^3$.
The 10-adic representation of 2020 is $2020 = 2 \times 10^1 + 2 \times 10^3$.

§1.2. Greatest common divisors, Euclidean algorithm, Bezout identity, least common multiple and continued fractions

Definition. Let A be a finite nonempty set of integers, not all 0. If an integer d divides a for all elements $a \in A$, then d is called a *common divisor* of A.

For example, 1 is a common divisor of every finite nonempty set of integers. Evidently, the set of all the common divisors of a finite nonempty set of integers is nonempty and finite.

Definition. A positive integer d is called *greatest common divisor* of the set A, denoted by $d = \gcd(A)$, if d is a common divisor of A and every common divisor of A is no more than d.

Obviously, the greatest common divisor $d = \gcd(A)$ of every finite nonempty set A of integers exists.

For integers a and b with $b > 1$, there is a simple and efficient method which is called *Euclidean algorithm* to compute the greatest common divisor (a, b) of a and b, and to express (a, b) explicitly in the form $ax + by$. This is done in what follows.

Define $r_0 = a$ and $r_1 = b$. By the division algorithm, there exist integers q_0 and r_2 such that

$$r_0 = r_1 q_0 + r_2 \text{ with } 0 \leq r_2 < r_1.$$

If an integer d divides r_0 and r_1, then d also divides r_1 and r_2. Similarly, if an integer d divides r_1 and r_2, the d also divides r_0 and r_1. Therefore *the set of common divisors of* r_0 *and* r_1 *is the same as the set of common divisors of* r_1 *and* r_2 and so

$$(a, b) = (r_0, r_1) = (r_1, r_2).$$

If $r_2 = 0$, then $a = bq_0$ and $(a, b) = b = r_1$. If $r_2 > 0$, then we divide r_2 into r_1 and obtain integers q_1 and r_3 such that

$$r_1 = r_2 q_1 + r_3, \text{ where } 0 \leq r_3 < r_2 < r_1$$

and $(a, b) = (r_1, r_2) = (r_2, r_3)$. Moreover, $q_1 \geq 1$ since $r_2 < r_1$. If $r_3 = 0$, then $(a, b) = r_2$. If $r_3 > 0$, then there exist integers q_2 and r_4 such that

$$r_2 = r_3 q_2 + r_4, \text{ where } r_2 \geq 1 \text{ and } 0 \leq r_4 < r_3 < r_2 < r_1$$

and $(a, b) = (r_2, r_3) = (r_3, r_4)$. If $r_4 = 0$, then $(a, b) = r_3$.

Iterating this process k times, we obtain an integer q_0, a sequence of positive integers $q_1, q_2, ..., q_{k-1}$, and a strictly decreasing sequence of nonnegative integers $r_1, r_2, ..., r_{k+1}$ such that

$$r_{i-1} = r_i q_{i-1} + r_{i+1}$$

for $i = 1, 2, ..., k$ and

$$(a, b) = (r_0, r_1) = (r_1, r_2) = ... = (r_k, r_{k+1}).$$

If $r_{k+1} > 0$, then we can divide r_k by r_{k+1} and obtain

$$r_k = r_{k+1} q_k + r_{k+2},$$

where $0 \leq r_{k+2} < r_{k+1}$. Since a strictly decreasing sequence of nonnegative integers must be finite, it follows that there exists an integer $n \geq 1$ such that $r_{n+1} = 0$. Then we have an integer q_0, a sequence of positive integers $q_1, q_2, ..., q_{n-1}$ and a strictly decreasing sequence of positive integers $r_1, r_2, ..., r_n$ with

$$(a, b) = (r_n, r_{n+1}) = r_n.$$

The n applications of the division algorithm produce n equations

$$r_0 = r_1 q_0 + r_2$$

$$r_1 = r_2 q_1 + r_3$$

$$r_2 = r_3 q_2 + r_4$$
$$\vdots$$
$$r_{n-2} = r_{n-1} q_{n-2} + r_n$$
$$r_{n-1} = r_n q_{n-1}.$$

Since $r_n < r_{n-1}$, it follows that $q_{n-1} \geq 2$.

This procedure is called the *Euclidean algorithm.* We call n the *length* of the Euclidean algorithm for a and b. This is the number of divisions required to find the greatest common divisor. The sequence $q_0, q_1, ..., q_{n-1}$ is called the *sequence of partial quotients.* The sequence $r_2, r_3, ..., r_n$ is called the *sequence of remainders.* In what follows we express r_k as the integral linear combination of a and b for $1 \leq k \leq n$.

Proposition 1.2.1. *For $k = 1, ..., n$, we have $aX_{k-1} - bY_{k-1} = (-1)^k r_k$, where $X_0 = 0, Y_0 = 1, X_1 = 1, Y_1 = q_0, X_l = q_{l-1}X_{l-1} + X_{l-2}, Y_l = q_{l-1}Y_{l-1} + Y_{l-2}$ for $2 \leq l \leq n$.*

Proof. We use induction on k to prove the Proposition. Clearly the result is true if $k = 1$ and 2. Let's now consider the $k = 3$ case, that is, X_2 and Y_2. Since

$$r_3 = b - r_2 q_1 = b - q_1(a - bq_0) = -q_1 a + (1 + q_0 q_1)b,$$

we have

$$-r_3 = q_1 a - (1 + q_0 q_1)b.$$

So one obtains that $X_2 = q_1 = q_1 X_1 + X_0$ and $Y_2 = 1 + q_0 q_1 = q_1 Y_1 + Y_0$. This concludes the result for the $k = 3$ case. In the following let $k \geq 3$. Assume that the result holds for all the no less than $k \geq 3$ case. In what follows we show the result is true for the $k + 1$ case.

Since $r_{k-1} = r_k q_{k-1} + r_{k+1}$, we have

$$(-1)^{k+1} r_{k+1} = (-1)^{k+1}(r_{k-1} - r_k q_k) = (-1)^{k+1} r_{k-1} + q_{k-1} \cdot (-1)^k r_k.$$

Then using inductive hypothesis, we get that

$$(-1)^{k+1} r_{k+1} = (aX_{k-2} - bY_{k-2}) + q_{k-1}(aX_{k-1} - bY_{k-1})$$
$$= a(q_{k-1}X_{k-1} + X_{k-2}) - b(q_{k-1}Y_{k-1} + Y_{k-2}).$$

Thus we derive that $X_k = q_{k-1}X_{k-1} + X_{k-2}$ and $Y_k = q_{k-1}Y_{k-1} + Y_{k-2}$ as desired. So the result is true for the $k + 1$ case.

This completes the proof of Proposition 1.2.1. □

In particular, we have the following result which expresses (a, b) as an integral linear combination of a and b.

Theorem 1.2.1. *There are integers x and y such that $ax + by = (a, b)$.*

Proof. This theorem follows immediately from Proposition 1.2.1. □

Proposition 1.2.2. *Let a, b, c be integers. Then each of the following is true.*

(i). *Every common divisor of a and b is a divisor of (a, b).*

(ii). *Commutative law*: $(a, b) = (b, a)$.

(iii).*Associative law*: $((a, b), c) = (a, (b, c)) := (a, b, c)$.

(iv). $(a, b)|c| = (ac, bc)$.

(v). $(a, b, c) = (|a|, |b|, |c|)$.

Proof. We omit the details of the proof. □

Definition. The integers $a_1, ..., a_k$ are called *relatively prime* if their greatest common divisor is 1, that is, $(a_1, ...a_k) = 1$. The integers $a_1, ..., a_k$ are called *pairwise relatively prime* if $(a_i, a_j) = 1$ for $i \neq j$.

For example, the three integers 6,10,15 are relatively prime but not pairwise relatively prime, since $(6, 10, 15) = 1$ but $(6, 10) = 2, (6, 15) = 3$, and $(10, 15) = 5$.

Theorem 1.2.2. (Bezout's identity) *Let $a_1, ..., a_n$ be any positive integers. Then there are n integers $x_1, ..., x_n$ such that $a_1x_1 + ... + a_nx_n = (a_1, ..., a_n)$.*

Proof. Using induction on n, the result follows immediately from Theorem 1.2.2. We here omit the detail of the proof. □

Example. Let us use the Euclidean algorithm to find (574, 252) and express it as a linear combination of 574 and 252. We have

$$574 = 252 \times 2 + 70,$$
$$252 = 70 \times 3 + 42,$$
$$70 = 42 \times 1 + 28,$$
$$42 = 28 \times 1 + 14,$$
$$28 = 14 \times 2,$$

and so $(574, 252) = 14$.

The sequence of partial quotients is $(2, 3, 1, 1, 2)$ and the sequence of partial remainders is (70, 42, 28, 14). The Euclidean algorithm for 574 and 252 has length 5. Note that $574 = 14 \times 41$ and $252 = 16 \times 28$ and that 41 and 28 are relatively prime. Working backwards through the Euclidean algorithm to express 14 as a linear combination of 574 and 252, we obtain

$$\begin{aligned} 14 &= 42 - 28 \times 1 = 42 - 70 - 42 \times 1 = 42 \times 2 - 70 \times 1 \\ &= (252 - 70 \times 3) \times 2 - 70 \times 1 = 252 \times 2 - 70 \times 7 \\ &= 252 \times 2 - (574 - 252 \times 2) \times 7 = 252 \times 16 - 574 \times 7. \end{aligned}$$

If $A = \{a_1, ..., a_k\}$ is nonempty, finite set of integers, not all 0, we write $\gcd(A) = (a_1, ..., a_k)$. For example, since $35 \times (-5) + 91 \times 2 = 7$, we have $(35, 91) = \gcd(\{35, 91\}) = 7$.

Theorem 1.2.3. *Let $a_1, ..., a_k$ be integers, not all zero. Then $(a_1, ..., a_k) = 1$ if and only if (iff) there exist integers $x_1, ..., x_k$ such that $a_1x_1 + ... + a_kx_k = 1$.*

Proof. This follows immediately from Theorem 1.2.3. □

Definition. Let $a_1, ..., a_k$ be nonzero integers. An integer m' is called a *common multiple* of $a_1, ..., a_k$ if it is a multiple of a_i for all $i = 1, ..., k$, that is, every integer a_i divides m'. The *least common multiple* of $a_1, ..., a_k$ is a positive integer m such that m is a common multiple of $a_1, ..., a_k$ and m is no more than every positive common multiple of $a_1, ..., a_k$.

Clearly, for arbitrary nonzero integers $a_1, ..., a_k$, the product $a_1...a_kl$ is a common multiple of $a_1, ..., a_k$ for any integer l. Further, the least common multiple of $a_1, ..., a_k$ exists and is unique. If $A = \{a_1, ..., a_k\}$ is a nonempty, finite set of integers, not all 0, then we write $\text{lcm}(A) = [a_1, ..., a_k]$.

Example. 910 is a common multiple of 35 and 91 and 455 is the least common multiple, namely $[35, 91] = 455$. We have $[12, 16] = 48, [1, 2, 3, 4, 5, 6, 7, 8] = 840$.

Proposition 1.2.3. *Let a, b, c be integers. Then:*

(i). $[a, b] = [b, a]$.

(ii). $[[a, b], c] = [a, [b, c]] := [a, b, c]$.

(iii). $[a, b, c] = [|a|, |b|, |c|]$.
(iv). $[a, b]|c| = [ac, bc]$.

Proof. (iv). Let $m_1 = [a, b]|c|$ and $m_2 = [ac, bc]$. One may let $c \neq 0$. On the one hand, since $m_1 = [a, b]|c|$, one has $\frac{m_1}{|c|} = [a, b]$ that infers that $a|\frac{m_1}{|c|}$ and $b|\frac{m_1}{|c|}$. Thus $a|c||m_1$ and $b|c||m_1$, and so $ac|m_1$ and $bc|m_1$. Namely, m_1 is a common multiple of ac and bc. Then by the definition, one knows that $m_2 \leq m_1$.

On the other hand, since $m_2 = [ac, bc]$, one has $a|c||ac|m_2$ and $b|c||bc|m_2$. Hence $a|\frac{m_2}{|c|}$ and $b|\frac{m_2}{|c|}$. That is, $\frac{m_2}{|c|}$ is a common multiple of a and b. But $m_1 = [a, b]|c|$ tells us that $[a, b] = \frac{m_1}{|c|}$. Thus $\frac{m_1}{|c|} \leq \frac{m_2}{|c|}$, and $m_1 \leq m_2$. This concludes that $m_1 = m_2$ as required. Part (iv) is proved.

We omit the details of the proofs of parts (i), (ii) and (v). □

One has a basic relation between the gcd and the lcm of any two integers. To prove it, we need to show two lemmas.

Lemma 1.2.1 (Euclid's lemma). *Let a, b, c be integers. If a divides bc and $(a, b) = 1$, then a divides c.*

Proof. Since a divides bc, we have $bc = aq$ for some integer q. Since a and b are relatively prime, Theorem 1.2.1 implies that there exist integers x and y such that $1 = ax + by$. Multiplying by c, we obtain $c = acx + bcy = acx + aqy = a(cx + qy)$, and so a divides c. This completes the proof. □

Lemma 1.2.2. *Let a and b be not all zero integers. Then every common multiple of a and b is a multiple of $\frac{ab}{(a,b)}$.*

Proof. Let l be any common multiple of a and b. Then $a|l$ and $b|l$, and so there are integers m and n such that $l = am = bn$. Let $d = (a, b)$ and $a = da'$ and $b = db'$, where $(a', b') = 1$. Then it follows from $am = bn$ that $a'm = b'n$ which implies that $a'|b'n$. But $(a', b') = 1$. So by Euclid's lemma (Lemma 1.2.1), one gets that $a'|n$. Then writing $n = a'z$ with $z \in \mathbf{Z}$. Hence we have $l = bn = ba'z = \frac{ab}{d}z$. Thus $\frac{ab}{(a,b)}|l$ as one desires. □

Now we can state the relation between the gcd and the lcm of any two integers.

Theorem 1.2.4. *Let a and b be positive integers. Each of the following is true:*
(i). One has $(a, b)[a, b] = ab$.
(ii). *Every common multiple of a and b is a multiple of $[a, b]$.*

Proof. (i). First we consider the case $(a, b) = 1$. Then we need only to show that $[a, b] = ab$. Clearly ab is a common multiple of a and b. Now let l be any positive common multiple of a and b. Then Lemma 1.2.2 tells us that $ab|l$ which infers that $ab \leq l$, and so $[a, b] = ab$ as required.

Now we consider the general case. Note that $(a, b)[a, b] = ab$ is equivalent to $[a, b] = \frac{ab}{(a,b)}$. Let's now show the latter one. Let $d = (a, b)$. Then $(a/d, b/d) = 1$ and so by the result for the case $(a, b) = 1$, we have $[a/d, b/d] = ab/d^2$. Thus $d[a/d, b/d] = ab/d$. But by Proposition 1.2.3 (iv), we have $d[a/d, b/d] = [a, b]$. Hence we arrive at $[a, b] = ab/d$ as desired. Part (i) is proved.

(ii). It follows immediately from Lemma 1.2.2 and part (i).

The proof of Theorem 1.2.4 is complete. □

From Theorem 1.2.4 and using induction method, one derives that for arbitrary nonzero integers $a_1, ..., a_k$, the least common multiple $[a_1, ..., a_k]$ divides any common multiple of $a_1, ..., a_k$. In other words, every common multiple of $a_1, ..., a_k$ is a multiple of $[a_1, ..., a_k]$.

Finally, we describe the relation between the continued fractions and Euclidean algorithm. Let us first give the concept of continued fraction.

Definition. Let $a_0, a_1, ..., a_N$ be real numbers with $a_i > 0$ for $i = 1, ..., N$. We define the *finite simple continued fraction*, denoted by $\langle a_0, a_1, ..., a_N\rangle$, as follows:

$$\langle a_0, a_1, ..., a_N\rangle = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots \cfrac{1}{a_{N-1} + \frac{1}{a_N}}}}}.$$

Another notation for a continued fraction is

$$\langle a_0, a_1, ..., a_N\rangle = a_0 + \frac{1}{a_1+}\frac{1}{a_2+}\cdots\frac{1}{a_N}.$$

The numbers $a_0, a_1, ...a_N$ are called the *partial quotients* of the continued fraction.

Example. One has

$$\langle 2, 1, 1, 2\rangle = 2 + \cfrac{1}{1 + \frac{1}{1+\frac{1}{2}}} = \frac{13}{5}.$$

We can write a finite simple continued fraction as a rational function in the variables $a_0, a_1, ..., a_N$. For example,

$$\langle a_0\rangle = a_0, \langle a_0, a_1\rangle = \frac{a_0a_1 + 1}{a_1} \text{ and } \langle a_0, a_1, a_2\rangle = \frac{a_0a_1a_2 + a_0 + a_2}{a_1a_2 + 1}.$$

Proposition 1.2.5. *If $N \geq 1$, then $\langle a_0, a_1, ..., a_N\rangle = a_0 + \frac{1}{\langle a_1,...,a_N\rangle}$.*

Proof. By definition, it is clear. □

We can use the Euclidean Algorithm to write a rational number as a finite simple continued fraction with integral partial quotients. For example, to represent 574/252, we have

$$\begin{aligned}\frac{574}{252} &= 2 + \frac{70}{252}\\ &= 2 + \cfrac{1}{3 + \frac{42}{70}}\\ &= 2 + \cfrac{1}{3 + \frac{1}{1+\frac{28}{42}}}\\ &= 2 + \cfrac{1}{3 + \frac{1}{1+\frac{1}{1+\frac{14}{28}}}}\\ &= 2 + \cfrac{1}{3 + \frac{1}{1+\frac{1}{1+\frac{1}{2}}}}\\ &= \langle 2, 3, 1, 1, 2\rangle.\end{aligned}$$

Notice that *the partial quotients in the Euclidean Algorithm are the partial quotients in the continued fraction.*

Theorem 1.2.5. *Let a and b be integers with $b \geq 1$. If the Euclidean Algorithm for a and b has length n with sequence of partial quotients $q_0, q_1, ..., q_{n-1}$, then*

$$\frac{a}{b} = \langle q_0, q_1, ..., q_{n-1} \rangle.$$

Proof. Let $r_0 = a$ and $r_1 = b$. The proof is by induction on n. If $n = 1$, then $r_0 = r_1 q_0$ and $\frac{a}{b} = \frac{r_0}{r_1} = q_0 = \langle q_0 \rangle$. If $n = 2$, then $r_0 = r_1 q_0 + r_2, r_1 = r_2 q_1$, and

$$\frac{a}{b} = \frac{r_0}{r_1} = q_0 + \frac{r_2}{r_1} = q_0 + \frac{1}{\frac{r_1}{r_2}} = q_0 + \frac{1}{q_1} = \langle q_0, q_1 \rangle.$$

Let $n \geq 2$ and assume that the theorem is true for any integers a and $b \geq 1$ whose Euclidean Algorithm has length n. Let a and $b \geq 1$ be integers whose Euclidean Algorithm has length $n+1$ and whose sequence of partial quotients is $\langle q_0, q_1, ..., q_n \rangle$. Let

$$\begin{aligned} r_0 &= r_1 q_0 + r_2 \\ r_1 &= r_2 q_1 + r_3 \\ &\vdots \\ r_{n-1} &= r_n q_{n-1} + r_{n+1} \\ r_n &= r_{n+1} q_n. \end{aligned}$$

be the $n+1$ equations in the Euclidean algorithm for $a = r_0$ and $b = r_1$. The Euclidean Algorithm for positive integers r_1 and r_2 has length n with sequence of partial quotients $q_1, ..., q_n$. It follows from the induction hypothesis that $\frac{r_1}{r_2} = \langle q_1, ..., q_n \rangle$ and so by Proposition 1.2.3, we have

$$\frac{a}{b} = \frac{r_0}{r_1} = q_0 + \frac{1}{\frac{r_1}{r_2}} = q_0 + \frac{1}{\langle q_1, ..., q_n \rangle} = \langle q_0, q_1, ..., q_n \rangle.$$

This completes the proof. □

It is also true that the representation of a rational number as a finite simple continues fraction is essentially unique (Exercise to U!).

§1.3. Fundamental theorem of arithmetic and p-adic valuation of factorials

Definition. A *prime number* is an integer p greater than 1 whose only positive divisors are 1 and p. A positive integer greater than 1 that is not prime is called *composite.* If n is a composite, then it has a divisor d such that $1 < d < n$, and so $n = dd'$, where also $1 < d' < n$. The primes less than 100 are the following:

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.

If d is positive divisor of n, then $d' = n/d$ is called the *conjugate divisor* to d. Clearly if $n = dd'$ and $d \leq d'$, then $d \leq\sqrt{n}$.

We shall prove that every positive integer can be written as the product of prime numbers (with the convention that the empty product is equal to 1) and that this representation is unique except for the order in which the prime factors are written. This result is called the *fundamental theorem of arithmetic.*

Lemma 1.3.1. *Let $k \geq 2$ and $a, b_1, b_2, ..., b_k$ be integers.*

(i). *If $(a, b_i) = 1$ for all $i = 1, 2, ..., k$, then we have $(a, b_1b_2...b_k) = 1$.*

(ii). *If a prime number p divides a product of some integers, then p divides at least one of these integers.*

Proof. (i). The proof is by induction on k. Let $k = 2$ and $d = (a, b_1b_2)$. we must show that $d = 1$. Since d divides a and $(a, b_1) = 1$, it follows that $(d, b_1) = 1$. Since d divides b_1b_2, Euclid's lemma implies that d divides b_2. Therefore d is a common divisor of a and b_2, but $(a, b_2) = 1$ and so $d = 1$.

Let $k \geq 3$ and assume that the result holds for $k - 1$. Let $a, b_1, ..., b_k$ be integers such that $(a, b_i) = 1$ for $i = 1, ..., k$. The induction assumption implies that $(a, b_1...b_{k-1}) = 1$. Since we also have $(a, b_k) = 1$, it follows from the case $k = 2$ that $(a, b_1...b_{k-1}b_k) = 1$. This completes the proof of part (i).

(ii). Let $b_1, b_2, ..., b_k$ be integers such that p divides $b_1...b_k$. By part (i) of this theorem, we have $(p, b_i) > 1$ for some i. Since p is prime, it follows that p divides b_i. □

Theorem 1.3.1. (Fundamental theorem of arithmetic) *Every positive integer can be written uniquely (up to order) as the product of prime numbers.*

Proof. First we prove that every positive integer can be written as a product of primes. Since an empty product is equal to 1, we can write 1 as the empty product of primes. Let $n \geq 2$. Suppose that every positive integer less than n is a product of primes. If n is prime, we are done. If n is composite, then $n = dd'$, where $1 < d \leq d' < n$. By the induction hypothesis, d and d' are both products of primes, and so $n = dd'$ is a product of primes.

Next we use induction to prove this representation is unique. The representation of 1 as the product of the empty set of primes is unique. Let $n \geq 2$ and assume that the statement is true for all positive integers less than n. We must show that if $n = p_1...p_k = p'_1...p'_l$, where $p_1, ..., p_k, p'_1, ..., p'_l$ are primes, then $k = l$ and there is a permutation σ of $1, ..., k$ such that $p_i = p'_{\sigma(i)}$ for $i = 1, ..., k$. Since p_k divides $p'_1...p'_l$, by Theorem 1.4.3 there exists an integer $j_0 \in \{1, ..., l\}$ such that p_k divides p'_{j_0}, and so $p_k = p'_{j_0}$ since p'_{j_0} is prime. Therefore,

$$\frac{n}{p_k} = p_1...p_{k-1} = \prod_{j=1, j\neq j_0}^{l} p'_j < n.$$

It follows from the induction hypothesis that $k-1 = l-1$, and there is a one-to-one map σ from $\{1, ..., k-1\}$ into $\{1, ..., k\} \setminus \{j_0\}$ such that $p_i = p'_{\sigma(i)}$ for $i = 1, ..., k-1$. Let $\sigma(k) = j_0$. This defines the permutation σ and the proof is complete. □

Definition. For any nonzero integer n and prime number p, we define $\nu_p(n)$ as the greatest integer r such that p^r divides n. Then $\nu_p(n)$ is a nonnegative integer, and $\nu_p(n) \geq 1$ if and only if p divides n. If $\nu_p(n) = r$, then we say that the prime power p^r *exactly divides* n, and write $p^r \parallel n$. The *standard factorization* of n is $n = \prod_{p|n} p^{\nu_p(n)}$.

Since every positive integer is divisible by only a finite number of primes, we can also write $n = \prod_p p^{\nu_p(n)}$, where the product is an infinite product over the set of all prime numbers, and $\nu_p(n) = 0$ for all but finitely many primes p.

Definition. The function $\nu_p(n)$ is called the *p-adic value* of n. It is *completely additive* in the sense that $\nu_p(mn) = \nu_p(m) + \nu_p(n)$ for all positive integers m and n (Exercise).

For example, since $n! = 1 \times 2 \times 3 \times ... \times n$, we have $\nu_p(n!) = \sum_{m=1}^{n} \nu_p(m)$.

Example. The standard factorizations of the first 100 positive integers are

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

Let $a \in \mathbf{Z}^+$ and $d|a|m$. It follows from the fundamental theorem of arithmetic that if $a = \prod_{i=1}^{k} p_i^{v_{p_i}(a)}$, where $p_1, ..., p_k$ are distinct primes, then we must have

$$d = \prod_{i=1}^{k} p_i^{s_i} \text{ and } m = m' \prod_{i=1}^{k} p_i^{t_i},$$

where $0 \leq s_i \leq v_{p_i}(a)$, $t_i \geq v_{p_i}(a)$ and m' is an integer coprime to $p_1...p_k$. The following results give expressions for the standard factorizations of the greatest common divisor and the least common multiple.

Theorem 1.3.2. *Let $a_1, ..., a_k$ be positive integers. Then*

$$(a_1, ..., a_k) = \prod_p p^{\min(\nu_p(a_1), ..., \nu_p(a_k))}$$

and

$$[a_1, ..., a_k] = \prod_p p^{\max(\nu_p(a_1), ..., \nu_p(a_k))}.$$

Proof. Let $d = \prod_p p^{\min(\nu_p(a_1),...,\nu_p(a_k))}$ and $m = \prod_p p^{\max(\nu_p(a_1),...,\nu_p(a_k))}$.

First of all, since for all integers i with $1 \leq i \leq k$, one has

$$\min(\nu_p(a_1), ..., \nu_p(a_k)) \leq \nu_p(a_i) \leq \max(\nu_p(a_1), ..., \nu_p(a_k))$$

that implies that for any prime p, we have

$$p^{\min(\nu_p(a_1),...,\nu_p(a_k))} | p^{\nu_p(a_i)} | p^{\max(\nu_p(a_1),...,\nu_p(a_k))}.$$

It follows that for all integers i with $1 \le i \le k$, one has

$$\prod_p p^{\min(\nu_p(a_1),...,\nu_p(a_k))}\Big|\prod_p p^{\nu_p(a_i)}\Big|\prod_p p^{\max(\nu_p(a_1),...,\nu_p(a_k))}.$$

That is, $d|a_i|m$ for all integers i with $1 \le i \le k$. Thus d and m are a common divisor and a common multiple of $a_1, ..., a_{k-1}$ and a_k, respectively.

Consequently, let d' and m' be any positive common divisor and any positive common multiple of $a_1, ..., a_{k-1}$ and a_k, respectively. Then $d'|a_i|m'$ for all integers i with $1 \le i \le k$. So for any prime p, we have that $\nu_p(d') \le \nu_p(a_i) \le \nu_p(m')$ for all integers i with $1 \le i \le k$. It infers that for any prime p,

$$\nu_p(d') \le \min(\nu_p(a_1), ..., \nu_p(a_k))$$

and

$$\max(\nu_p(a_1), ..., \nu_p(a_k)) \le \nu_p(m').$$

Therefore

$$p^{\nu_p(d')} \le p^{\min(\nu_p(a_1),...,\nu_p(a_k))}$$

and

$$p^{\max(\nu_p(a_1),...,\nu_p(a_k))} \le p^{\nu_p(m')}$$

which tell us that

$$\prod_p p^{\nu_p(d')} \le \prod_p p^{\min(\nu_p(a_1),...,\nu_p(a_k))}$$

and

$$\prod_p p^{\max(\nu_p(a_1),...,\nu_p(a_k))} \le \prod_p p^{\nu_p(m')}.$$

In other words, one has $d' \le d$ and $m \le m'$. This concludes that d and m are the greatest common divisor and the least common multiple of $a_1, ..., a_{k-1}$ and a_k, respectively. Namely, $d = (a_1, ..., a_k)$ and $m = [a_1, ..., a_k]$ as desired.

This finishes the proof of Theorem 1.3.2. □

Let x be a real number. Recall that the integer part of x is the greatest integer not exceeding x, that is, the unique integer n such than $n \le x < n + 1$. We denote the integer part of x by $[x]$. For example, $[\frac{4}{3}] = 1$, $[\sqrt{7}] = 2$ and $[-\frac{4}{3}] = -2$. The *fractional part* of x is the real number $\{x\} = x - [x]$. Then $\{x\} \in [0, 1)$. Thus $\{\frac{4}{3}\} = \frac{1}{3}$ and $\{-\frac{4}{3}\} = \frac{2}{3}$. We can use the greatest integer function to compute the standard factorization of factorials.

Theorem 1.3.3. *For every positive integer n and prime p, we have*

$$\boxed{\nu_p(n!) = \sum_{r=1}^{[\frac{\log n}{\log p}]} \left[\frac{n}{p^r}\right].}$$

Proof. Let $1 \le m \le n$. If p^r divides m, then $p^r \le m \le n$ and $r \le \log n/\log p$. Since r is an integer, we have $r \le [\log n/\log p]$ and

$$\nu_p(m) = \sum_{r=1, p^r|m}^{[\frac{\log n}{\log p}]} 1.$$

Since the number of positive integers not exceeding n that are divisible by p^r is exactly $[n/p^r]$, we have

$$\nu_p(n!) = \sum_{m=1}^{n} \nu_p(m) = \sum_{m=1}^{n} \sum_{r=1,p^r|m}^{[\frac{\log n}{\log p}]} 1 = \sum_{r=1}^{[\frac{\log n}{\log p}]} \sum_{m=1,p^r|m}^{n} 1 = \sum_{r=1}^{[\frac{\log n}{\log p}]} \left[\frac{n}{p^r}\right].$$

This completes the proof of Theorem 1.3.3. □

Alternative proof of Theorem 1.3.3. Clearly the number of integers k with fixed $v = \nu_p(k)$ that appear in the product $n!$ is equal to the number of multiples of p^v that are not multiples of p^{v+1} (and are less than or equal to n), namely $\left[\frac{n}{p^v}\right] - \left[\frac{n}{p^{v+1}}\right]$, where $[x]$ denotes the integral of the real number x. Thus we have

$$\nu_p(n!) = \sum_{v\geq 1}\left(\left[\frac{n}{p^v}\right] - \left[\frac{n}{p^{v+1}}\right]\right)v = \left[\frac{n}{p}\right] + \left[\frac{n}{p^2}\right] + \left[\frac{n}{p^3}\right] + \cdots = \sum_{r\geq 1}\left[\frac{n}{p^r}\right].$$

as expected. □

Theorem 1.3.4. *For every positive integer n and prime p, we have*

$$\boxed{\nu_p(n!) = \frac{n - \sigma_p(n)}{p-1},}$$

where $\sigma_p(n)$ means the sum of the digits of p-adic representation of n.

Proof. We need to compute $\nu_p(n!) = \sum_{1\leq k\leq n} \nu_p(k)$. Fix an integer $k \leq n$ with order $\nu_p(k) = v$ and write the p-adic representation of k as follows:

$$k = k_v p^v + \cdots + k_l p^l \quad (v \leq l, k_v \neq 0).$$

Then

$$k - 1 = (p-1) + \cdots + (p-1)p^{v-1} + (k_v - 1)p^v + \cdots k_l p^l,$$

and hence $\sigma_p(k-1) = (p-1)v + \sigma_p(k) - 1$. Equivalently,

$$v = \nu_p(k) = \frac{1}{p-1}(1 + \sigma_p(k-1) - \sigma_p(k)).$$

Summing over all values of $k \leq n$ we obtain a telescoping sum

$$\nu_p(n!) = \frac{1}{p-1}\sum_{1\leq k\leq n}(1 + \sigma_p(k-1) - \sigma_p(k)) = \frac{1}{p-1}(n - \sigma_p(n)).$$

as required. So Theorem 1.3.4 is proved. □

Alternative proof of Theorem 1.3.4: Let $n = n_0 + n_1 p + n_2 p^2 + \cdots$ (a finite sum) be the p-adic representation of n. Then

$$\begin{aligned}
\left[\frac{n}{p}\right] &= n_1 + n_2 p + n_3 p^2 + \cdots, \\
\left[\frac{n}{p^2}\right] &= n_2 + n_3 p + n_4 p^2 + \cdots, \\
&\cdots\cdots, \\
\left[\frac{n}{p^j}\right] &= n_j + n_{j+1} p + n_{j+2} p^2 + \cdots, \\
&\cdots\cdots.
\end{aligned}$$

Therefore

$$n = n_0 + p\left[\frac{n}{p}\right], \left[\frac{n}{p}\right] = n_1 + p\left[\frac{n}{p^2}\right], \cdots, \left[\frac{n}{p^j}\right] = n_j + p\left[\frac{n}{p^{j+1}}\right], \cdots.$$

Key

Summing all these and by Theorem 1.3.3, we obtain that

$$n + \sum_{j\geq 1}\left[\frac{n}{p^j}\right] = \sum_{j\geq 0} n_j + p\sum_{j\geq 1}\left[\frac{n}{p^j}\right].$$

That is,

$$n + \nu_p(n!) = \sigma_p(n) + p\ \nu_p(n!).$$

So $n - \sigma_p(n) = (p-1)\nu_p(n!)$ and the desired result follows immediately. □

Example. We shall use Theorems 1.3.5 and 1.3.6 to compute the standard factorization of 10!. The primes not exceeding 10 are 2, 3, 5 and 7, and

$$\nu_2(10!) = [\frac{10}{2}] + [\frac{10}{4}] + [\frac{10}{8}] = 5 + 2 + 1 = 8, \sigma_2(10) = 2 \Rightarrow \nu_2(10!) = \frac{10-2}{2-1} = 8,$$

$$\nu_3(10!) = [\frac{10}{3}] + [\frac{10}{9}] = 4, \sigma_3(10) = 2 \Rightarrow \nu_3(10!) = \frac{10-2}{3-1} = 4,$$

$$\nu_5(10!) = [\frac{10}{5}] = 2, \sigma_5(10) = 2 \Rightarrow \nu_5(10!) = \frac{10-2}{5-1} = 2,$$

$$\nu_7(10!) = [\frac{10}{7}] = 1, \sigma_7(10) = 4 \Rightarrow \nu_7(10!) = \frac{10-4}{7-1} = 1, .$$

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

Theorem 1.3.5. *For any positive integer m and n with $m \leq n$ and prime p, we have*

$$\nu_p(\begin{pmatrix} n \\ m \end{pmatrix}) = \frac{\sigma_p(m) + \sigma_p(n-m) - \sigma_p(n)}{p-1}.$$

Proof. Since $\begin{pmatrix} p \\ k \end{pmatrix} = \frac{n!}{m!(n-m)!}$, applying Theorem 1.3.4 one obtains that

$$\begin{aligned}\nu_P(\begin{pmatrix} p \\ k \end{pmatrix}) =& \nu_P(n!) - \nu_p(m!) - \nu_P((n-m)!) \\ =& \frac{n-\sigma_p(n)}{p-1} - \frac{m-\sigma_p(m)}{p-1} - \frac{n-m-\sigma_p(n-m)}{p-1} \\ =& \frac{\sigma_p(m) + \sigma_p(n-m) - \sigma_p(n)}{p-1}\end{aligned}$$

as desired. This concludes the proof of Theorem 1.3.5. □

§1.4. Euclid's theorem and the sieve of Eratosthenes

How many primes are there? The fundamental theorem of arithmetic tells us that every number is uniquely the product of primes, but it does not give us the number of primes. Euclid proved that the number of primes is infinite. The following proof is also due to Euclid. It has retained its power for more than two thousand years.

Theorem 1.4.1 (Euclid's theorem). *There are infinitely many primes.*

Proof. Let $p_1, ..., p_n$ be any finite set of primes numbers. Consider the integer

$$N = p_1 ... p_n + 1.$$

Since $N > 1$, it follows from the fundamental theorem of arithmetic that N is divisible by some prime p. If $p = p_i$ for some $i = 1, 2, ..., n$, then p divides $N -$

$p_1...p_n = 1$, which is absurd. Therefore, $p \neq p_i$ for $i = 1, ..., n$. This means that, for any finite set of primes, there always exists a prime that does not belong to the set, and so the number of primes is infinite. □

Let $\pi(x)$ denote the number of primes not exceeding x. Then $\pi(x) = 0$ for $x < 2$, $\pi(x) = 1$ for $2 \leq x < 3$, $\pi(x) = 2$ for $3 \leq x < 5$, and so on. Euclid's theorem says that there are infinitely many prime numbers, that is,

$$\lim_{x \to \infty} \pi(x) = \infty,$$

but it does not tell us how to determine them. We can compute all the prime number up to x by using a beautiful and efficient method called the *sieve of Eratosthenes.* The sieve is based on a simple observation. If the positive integer n in composite, then n can be written in the form $n = dd'$, where $1 < d \leq d' < n$. If $d > \sqrt{n}$, then

$$n = dd' > \sqrt{n}\sqrt{n} = n,$$

which is absurd. Therefore *if n is composite, then n has a divisor d such that $1 < d \leq \sqrt{n}$.* In particular, *every composite number $n \leq x$ is divisible by a prime $p \leq \sqrt{x}$.*

To find all the primes up to x, we write down the integers between 1 and x, and eliminate numbers from the list according to the following rule: Cross out 1. The first number in the list which is not eliminated is 2; cross out all multiples of 2 that are greater than 2. The iterative procedure is as follows: *Let d be the smallest number on the list whose multiples have not already been eliminated . If $d \leq \sqrt{x}$, then cross out all multiples of d that are greater than d. If $d > \sqrt{x}$, stop.* This algorithm must terminate after at most $\sqrt{x}$ steps. *The prime numbers up to x are the numbers that have not been crossed out.*

We shall demonstrate this method to find the prime numbers up to 60. We must sieve out by the prime numbers less than $\sqrt{60}$, that is, by 2,3,5,7. Here is the list of numbers up to 60:

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

We cross out 1 and all multiples of 2 beginning with 4:

/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

Next we cross out all multiples of 3 beginning with 6:

/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

Next we cross out all multiples of 5 beginning with 10:

/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

Finally, we cross out all multiples of 7 beginning with 14:

/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

The numbers that have not been crossed out are:

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59.$$

These are the prime numbers up to 60.

§1.5. A linear Diophantine equation and Frobenius problem

A *diophantine equation* is an equation of the form $f(x_1, ...x_k) = b$ that we want to solve in rational numbers,integers,or nonnegative integers. This means that the values of the variables $x_1, ..., x_k$ will be rational numbers, integers or nonnegative integers. Usually the function $f(x_1, ..., x_k)$ is a polynomial with rational or integer coefficients.

In this section we consider the linear diophantine equation

$$a_1x_1 + ... + a_kx_k = b.$$

We would like to know when the above equation has a solution in integers, and when it has a solution in nonnegative integers. For example, the equation

$$3x_1 + 5x_2 = b$$

has a solution in integers for every integer b, and a solution in nonnegative integers for $b = 0, 3, 5, 6$ and all $b \geq 8$ (Exercise to you!).

Theorem 1.5.1. *Let $a_1, ..., a_k$ be integers, not all zero. For any integer b, there exist integers $x_1, ..., x_k$ such that*

$$a_1x_1 + ... + a_kx_k = b \tag{1.4}$$

if and only if b is a multiple of $(a_1, ..., a_k)$. In particular, the linear equation (1.4) has a solution for every integer b if and only if the numbers $a_1, .., a_k$ are relatively prime.

Proof. Let $d = (a_1, ..., a_k)$. If equation (1.4) is solvable in integers x_i, then d divides b since d divides each integer a_i. Conversely, if d divides b, then $b = dq$ for some integer q. By Theorem 1.2.3, there exist integers $y_1, ..., y_k$ such that

$$a_1y_1 + ... + a_ky_k = d.$$

Let $x_i = y_iq$ for $i = 1, ..., k$. Then

$$a_1x_1 + ... + a_kx_k = a_1(y_1q) + ... + a_k(y_kq) = dq = b$$

is a solution of (1.4). It follows that (1.4) is solvable in integers for every b if and only if $(a_1, ..., a_k) = 1$.

Lemma 1.5.1. *Let $k \geq 2$ be an integer and $a_1, ..., a_{k-1}$ and a_k be integers such that $a_k \geq 1$ and $(a_1, ..., a_k) = 1$. Then for any integer b, there are integers $x_1, ..., x_k$ with $0 \leq x_i \leq a_k - 1$ for $1 \leq i \leq k - 1$ such that $a_1x_1 + ... + a_kx_k = b$.*

Proof. By Theorem 1.5.1, there exist integers $z_1, ..., z_k$ such that $a_1z_1 + ... + a_kz_k = b$. Using the division algorithm, we can divide each of the integers $z_1, ..., z_{k-1}$ by a_k so that $z_i = a_kq_i + x_i$ and $0 \leq x_i \leq a_k - 1$ for $i = 1, ..., k-1$. Let $x_k = z_k + \sum_{i=1}^{k-1} a_iq_i$. Then

$$\begin{aligned} b &= a_1z_1 + ... + a_{k-1}z_{k-1} + a_kz_k \\ &= a_1(a_kq_1 + x_1) + ... + a_{k-1}(a_kq_{k-1} + x_{k-1}) + a_kz_k \\ &= a_1x_1 + ... + a_{k-1}x_{k-1} + a_k\Big(z_k + \sum_{i=1}^{k-1} a_iq_i\Big) \\ &= a_1x_1 + ... + a_{k-1}x_{k-1} + a_kx_k \end{aligned}$$

as required. Thus Lemma 1.5.1 is proved. □

Theorem 1.5.2. *Let $a_1, ..., a_k$ be positive integers such that $(a_1, ..., a_k) = 1$. If*

$$b \geq (a_k - 1)\Big(\sum_{i=1}^{k-1} a_i - 1\Big),$$

then there exist nonnegative integers $x_1, ..., x_k$ such that $a_1x_1 + ... + a_kx_k = b$.

Proof. By Lemma 1.5.1 we know that there are integers $x_1, ..., x_k$ with $0 \leq x_i \leq a_k - 1$ for $1 \leq i \leq k - 1$ such that $a_1x_1 + ... + a_kx_k = b$. It follows that

$$b \leq (a_k - 1)\sum_{i=1}^{k-1} a_i + a_kx_k,$$

where x_k is an integer, possibly negative. Since $b \geq (a_k - 1)(\sum_{i=1}^{k-1} a_i - 1)$, one then derives that

$$a_kx_k \geq b - (a_k - 1)\sum_{i=1}^{k-1} a_i \geq -(a_k - 1) > -a_k.$$

Hence $x_k > -1$. But x_k is an integer. So $x_k \geq 0$. This completes the proof of Theorem 1.5.2. □

Let $a_1, ..., a_k$ be relatively prime positive integers. Since *every sufficiently large integer can be written as a nonnegative integral linear combination of* $a_1, ..., a_k$, it follows that there exists a smallest integer $G(a_1, ..., a_k)$ such that every integer $b \geq G(a_1, ..., a_k)$ can be represented in the form (1.4), where the variables $x_1, ..., x_k$ are nonnegative integers. The example above shows that $G(3, 5) = 8$.

The *linear Diophantine problem of Frobenius* is to determine $G(a_1, ..., a_k)$ for all finite sets of relatively prime positive integers $a_1, ..., a_k$. This is a difficult open problem, but there are some special cases where the solution is known. The following theorem solves the Frobenius problem in the case $k = 2$. We need a lemma.

Lemma 1.5.2. *Let a_1 and a_2 be coprime positive integers. For every integer b, there exist a unique pair (x_1, x_2) of integers with $0 \leq x_1 \leq a_2 - 1$ such that*

$$b = a_1x_1 + a_2x_2. \tag{1.5}$$

Proof. We saw in the proof of Theorem 1.5.2 that for every integer b there exist integers x_1 and x_2 such that (1.5) holds. If we have another representation $b = a_1x'_1 + a_2x'_2$ and $0 \leq x'_1 \leq a_2 - 1$, then

$$a_1(x_1 - x'_1) = a_2(x'_2 - x_2).$$

since a_2 divides $a_1(x_1 - x'_1)$ and $(a_1, a_2) = 1$, Euclid's Lemma (Lemma 1.3.1) implies that a_2 divides $x_1 - x'_1$. Then $x_1 = x'_1$ since $|x_1 - x'_1| \leq a_2 - 1$. It follows that $x_2 = x'_2$, and so the representation (1.5) is unique. Lemma 1.5.2 is proved. □

Theorem 1.5.3. *Let a_1 and a_2 be relatively prime positive integers. Then $G(a_1, a_2) = (a_1 - 1)(a_2 - 1)$.*

Proof. First, by Theorem 1.5.2 with $k = 2$, we have

$$G(a_1, a_2) \leq (a_1 - 1)(a_2 - 1).$$

So, to finish the proof of Theorem 1.6.3, it suffices to show that $G(a_1, a_2) \geq (a_1 - 1)(a_2-1)$. To prove the latter one, in what follows we show that $(a_1-1)(a_2-1)-1 = a_1a_2 - a_1 - a_2$ cannot be expressed as nonnegative integral linear combinations of $a_1, ..., a_{k-1}$ and a_k.

Since $a_1(a_2 - 1) + a_2(-1) = a_1a_2 - a_1 - a_2 < a_1a_2$, it follows that if

$$a_1a_2 - a_1 - a_2 = a_1x_1 + a_2x_2$$

for any nonnegative integers x_1 and x_2, then $0 \leq x_1 \leq a_2 - 1$. By the uniqueness of the representation (1.5), we must have $x_1 = a_2 - 1$ and $x_2 = -1$.Therefore the integer $a_1a_2 - a_1 - a_2$ cannot be represented as a nonnegative integral linear combination of a_1 and a_2, and so

$$G(a_1, a_2) \geq a_1a_2 - a_1 - a_2 + 1 = (a_1 - 1)(a_2 - 1).$$

Therefore we arrive at the desired $G(a_1, a_2) = (a_1 - 1)(a_2 - 1)$. Theorem 1.5.3 is proved. □

Frobenius problem is kept widely open when $k \geq 3$. Even for the first nontrivial case $k = 3$, this question still remains unknown and attracted many authors' attention.

§1.6. Dirichlet drawer principle and Pell equations

Dirichlet Drawer Principle, also called *Pigeon Hole Principle*, is basic and important in number theory and combinatorics et al. Let's now state it.

Dirichlet Drawer Principle. *Suppose that* $a = \sum_{i=1}^{b} l_i - b + 1$ *objects are sorted into* b *boxes* $B_1, ..., B_b$. *Then there must be some* $1 \leq i \leq b$ *such that box* B_i *contains at least* l_i *objects.*

Theorem 1.6.1. *Let* $1 \leq a_1 < a_2 < ... < a_{n+1} \leq 2n$. *Then there exist integers* $1 \leq i < j \leq n+1$ *such that* $a_i | a_j$.

Proof. Let $a_i = 2^{e_i} b_i$ with $2 \nmid b_i$ for $1 \leq i \leq n+1$. Then $1 \leq b_i \leq 2n$. We take the n odd numbers in the interval $[1, n]$ as the n boxes. Since each of $b_1, ..., b_{n+1}$ is odd, we hold $n+1$ objects $b_1, ..., b_{n+1}$. Thus by the Dirichlet Drawer Principle, there exist $1 \leq i \neq j \leq n+1$ such that $b_i = b_j$. Since either $e_i \leq e_j$ or $e_i \geq e_j$, we have either $a_i | a_j$ or $a_j | a_i$. Theorem 1.6.1 is proved. □

Actually, a more general result is true as the following theorem shows.

Theorem 1.6.2. *Let* p *be a prime and* $n \geq 1$ *be an integer. If* $a_i \in \mathbf{Z}$ *for* $1 \leq i \leq (p-1)n+1$ *and* $1 \leq a_1 < a_2 < ... < a_{(p-1)n+1} \leq pn$, *then there exist integers* $1 \leq i < j \leq (p-1)n+1$ *such that* $a_i | a_j$.

Proof. Let $a_i = p^{e_i} b_i$ with $p \nmid b_i$ for $1 \leq i \leq (p-1)n+1$. Then $1 \leq b_i \leq pn$. We take the $(p-1)n$ numbers in the interval $[1, pn]$ which is not divisible by p as the $(p-1)n$ boxes.

On the other hand, since each of $b_1, ..., b_{(p-1)n+1}$ is not divisible by p, we have $(p-1)n+1$ objects $b_1, ..., b_{(p-1)n+1}$. It follows immediately from the Dirichlet Drawer Principle that there exist $1 \leq i \neq j \leq (p-1)n+1$ such that $b_i = b_j$. Since either $e_i \leq e_j$ or $e_i \geq e_j$, we have either $a_i | a_j$ or $a_j | a_i$. So Theorem 1.6.2 is proved. □

Theorem 1.6.3. (Dirichlet) *Let* ξ *be an irrational number. Then there are infinitely many rational numbers* $\frac{x}{y}$ *with* x *and* y *being integers and* $(x, y) = 1$ *such that* $|\frac{x}{y} - \xi| < \frac{1}{y^2}$.

Proof. First of all, we claim that for any given integer $n \geq 1$, there are integers x, y with $(x, y) = 1$ and $0 < y \leq n$ such that

$$|\frac{x}{y} - \xi| < \frac{1}{ny}. \tag{1.6}$$

To prove (1.6), we partition the half-open interval $[0, 1)$ into the union of the n distinct half-open intervals $[\frac{i}{n}, \frac{i+1}{n})$ with $0 \leq i \leq n-1$. Take the n distinct half-open intervals

$$\Big[0, \frac{1}{n}\Big), \Big[\frac{1}{n}, \frac{2}{n}\Big), ..., \Big[\frac{n-1}{n}, \frac{n}{n}\Big) \tag{1.7}$$

as the n boxes. Note that for any real number x, its fractional part can be written as $x - \lfloor x \rfloor$, where $\lfloor x \rfloor$ means the largest integer no more than x. Then by the Dirichlet Drawer Principle, we know that the fractional parts

$$0, \{\xi\} = \xi - \lfloor \xi \rfloor, \{2\xi\} = 2\xi - \lfloor 2\xi \rfloor, ..., \{n\xi\} = n\xi - \lfloor n\xi \rfloor$$

of $n+1$ distinct real numbers $0, \xi, 2\xi, ..., n\xi$ must lie in the same subinterval of the n distinct half-open intervals in (1.7). It then follows that there are integers i and j with $0 \leq i < j \leq n$ such that

$$|j\xi - \lfloor j\xi \rfloor - (i\xi - \lfloor i\xi \rfloor)| < \frac{1}{n}. \tag{1.8}$$

Now picking $x = \lfloor j\xi \rfloor - \lfloor i\xi \rfloor$ and $y = j - i$, then from (1.8) one derives that $|x - y\xi| < \frac{1}{n}$, and so (1.6) follows immediately. It also follows from (1.6) and $0 < y \le n$ that $|\frac{x}{y} - \xi| < \frac{1}{y^2}$.

Since ξ is irrational, one has $\frac{x}{y} - \xi \neq 0$. Then take $n_1 \in \mathbf{Z}$, such that $n_1 > \frac{1}{|\frac{x}{y} - \xi|}$. The claim applied to n_1 gives us that there are relatively prime integers x_1, y_1 $0 < y_1 \le n_1$ such that

$$|\frac{x_1}{y_1} - \xi| < \frac{1}{n_1 y_1}.$$

Since $0 < y_1 \le n_1$, we have

$$|\frac{x_1}{y_1} - \xi| < \frac{1}{y_1^2}.$$

Now take an integer n_2 such that $n_2 > \frac{1}{|\frac{x_1}{y_1} - \xi|}$. Then the above claim applied to n_2 gives us the existence of the pair (x_2, y_2) of integers with $0 < y_2 \le n_2$ such that

$$|\frac{x_2}{y_2} - \xi| < \frac{1}{n_2 y_2}. \tag{1.9}$$

Since $0 < y_2 \le n_2$, we have

$$|\frac{x_2}{y_2} - \xi| < \frac{1}{y_2^2}.$$

Further, since $n_2 > \frac{1}{|\frac{x_1}{y_1} - \xi|}$ and $y_2 \ge 1$, one has

$$\frac{1}{n_2 y_2} < \frac{|\frac{x_1}{y_1} - \xi|}{y_2} \le |\frac{x_1}{y_1} - \xi|. \tag{1.10}$$

Thus from (1.9) and (1.10), we infer that

$$|\frac{x_2}{y_2} - \xi| < \frac{1}{n_2 y_2} < |\frac{x_1}{y_1} - \xi|.$$

The above procedure can continue and produces infinitely many pairs (x_i, y_i) of integers with $y_i > 0$ such that $|\frac{x_i}{y_i} - \xi| < \frac{1}{y_i^2}$. Thus Theorem 1.6.3 is proved. □

Proposition 1.6.1. *Let d be a positive square-free integer. Then there exist infinitely many pairs $(x, y) \in \mathbf{Z}^2$ such that $|x^2 - dy^2| < 1 + 2\sqrt{d}$.*

Proof. First, Theorem 1.6.3 applied to $\xi = \sqrt{d}$ gives us the existence of infinitely many pairs (x, y) of integers with $y > 0$ such that $|x - \sqrt{d}y| < \frac{1}{y}$. So we deduce that

$$|x + \sqrt{d}y| \le |x - \sqrt{d}y| + 2\sqrt{d}|y| < \frac{1}{y} + 2\sqrt{d}y$$

It then follows that

$$|x^2 - dy^2| = |x + \sqrt{d}y||x - \sqrt{d}y)| < |x + \sqrt{d}y|\frac{1}{y} \le 2\sqrt{d} + \frac{1}{y^2} \le 1 + 2\sqrt{d}$$

as desired. Thus Proposition 1.6.1 is proved. □

We can then use Proposition 1.6.1 to show the following important theorem about Pell's equation.

Theorem 1.6.4. *Let d be a positive square-free integer. Then $x^2 - dy^2 = 1$ has infinitely many integral solutions. Furthermore, there is a solution (x_1, y_1) such that each solution is of the form (x_n, y_n), where $x_n + \sqrt{d}y_n = (x_1 + \sqrt{d}y_1)^n, n \in \mathbf{Z}$.*

Exercises for Chapter 1

(1). Find all divisors of 180, 270 and 520.

(2). Find then quotient and reminder for a divided by d when

(i). $a = 521$ and $d = 12$;

(ii). $a = 271$ and $d = 23$;

(iii). $a = 2006$ and $d = 18$;

(iii). $a = 2008$ and $d = 37$.

(3). Compute the m-adic representation of 687 for $m = 2, 3, 7$ and 9.

(4). Show that for any integer n, we have $6|n(n+1)(2n+1)$.

(5). Let a, x, y, z and w be integers. Show that if $a|(2006x-y)$ and $a|(2006z-w)$, then $a|(xw-yz)$.

(6). Let $m_0 = 1$ and $\{m_i\}_{i=0}^{\infty}$ be a strictly increasing divisor chain of positive integers, that is $m_i|m_{i+1}$ for all $i \geq 0$. Show that every positive integer n can be represented uniquely in the form $n = \sum_{i=0}^{\infty} a_i m_i$, where $0 \leq a_i \leq \frac{m_{i+1}}{m_i} - 1$ for all $i \geq 0$ and $m_i = 0$ for all but finitely many integers i.

(7). Compute $(206, 208)$ and $[168, 252, 294]$.

(8). Show that $5n+2$ and $7n+3$ are relatively prime for every integer n.

(9). Show that $(n, n+2) = 1$ if n is odd and $(n, n+2) = 2$ if n is even.

(10). Let a and b be positive integers. Show that $(a, b) = a$ iff a divides b. Show also that $[a, b] = a$ iff b divides a.

(11). Use the Euclidean algorithm to calculate the greatest common divisor of 14 and 91, and express (14, 91) as a linear combination of 14 and 91. Compute the simple continued fraction for 91/14.

(12). We define a sequence $\{f_n\}$ by: $f_0 = 0, f_1 = 1$ and $f_n = f_{n-1} + f_{n-2}$ for $n \geq 2$. The integer f_n is called the nth *Fibonacci number*. Compute f_n for $n = 2, 3, ..., 20$. Show that $(f_n, f_{n+1}) = 1$ for all nonnegative integers n.

(13). Find the gcd and lcm of $a = 2^8 5^5 7^3 11^7$ and $b = 3^7 5^9 11^3 13^8$.

(14). Suppose that a, b, m are integers such that $a \neq 0, b \neq 0$ and $a|m, b|m$. Show that $\left(\frac{m}{a}, \frac{m}{b}\right) = \frac{m}{[a,b]}$ and $\left[\frac{m}{a}, \frac{m}{b}\right] = \frac{m}{(a,b)}$.

(15). Let $n \geq 2$. Show that $(n+1)!+k$ is a composite number for $k = 2, ..., n+1$.

(16). Let $a, b \in \mathbf{Z}^*$. Show that $a|b$ iff one has $\nu_p(a) \leq \nu_p(b)$ for all primes p.

(17). Prove that rad(mn)=rad(m)rad(n) iff $(m, n) = 1$.

(18). For an integer $n \geq 1$, let $s_n := 1 + \frac{1}{2} + ... + \frac{1}{n}$. Show that s_n is NOT an integer for any $n \geq 2$.

(19). Let S be a set of positive integers. We say that S is *gcd-closed* (resp. *lcm-closed*) if $(x, y) \in S$ (resp. $[x, y] \in S$) for any $x, y \in S$. For a real number e, we define the e-th power set S^e by $S^e := \{x^e | x \in S\}$. Prove that if e is a positive integer, then the set S is gcd-closed (resp. lcm-closed) if and only if the power set S^e is gcd-closed (resp. lcm-closed).

(20). Use the sieve of Eratosthenes to find the prime numbers up to 300. Compute $\varphi(300)$.

(21). Let $2 = p_1 < p_2 < ...$ be the sequence of primes in increasing order. Show that $p_n \leq 2^{2^{n-1}}$ for all $n \geq 1$.

(22). Let $k \in \mathbf{Z}^+$. (i). Show that if $2^k + 1$ is prime, then $k = 2^n$. Primes of the form $F_n := 2^{2^n} + 1$ are called *Fermat primes*.

(ii). Show that F_n is prime if $n = 1, 2, 3, 4$ but not prime when $n = 5$.

(23). Let $a, n \in \mathbf{Z}^+$.

(i). Show that $a^n - 1$ is a prime only if $a = 2$ and $n = p$ is a prime. Primes of the form $M_p := 2^p - 1$ are called *Mersenne primes*.

(ii). Compute the first five Mersenne primes.

(24). Show that the equation $3x_1 + 5x_2 = b$ has a solution in integers for every integer b, and a solution in nonnegative integers for $b = 0, 3, 5, 6$ and all $b \geq 8$.

(25). Find all nonnegative integers that cannot be represented by the form $3x_1 + 10x_2 + 14x_3$ with x_1, x_2, x_3 nonnegative integers. Compute $G(3, 10, 14)$.

(26). Let p be a prime. Show that each of the following is true:

(i). For any $1 \leq k \leq p - 1$, the binomial number $p\begin{pmatrix} p \\ k \end{pmatrix}$ is divisible by p.

(ii). For any $e \geq 1$, the binomial number $\begin{pmatrix} p^e \\ p^{e-1} \end{pmatrix}$ is divisible by p.

(27). Let $n, k \in \mathbf{Z}$ and $k > 0$. Show that:

(i). $(n-1)^2 | (n^k - 1)$ if and only if $(n-1) | k$;

(ii). $(n-1)^3 | (n^k - 1)$ if and only if $2(n-1)^2 | k(2 + (k-1)(n-1))$.

(28). Show that $(a, b) = (a + b, [a, b])$.

(29). Show that for any positive integers m and n, $(2^m - 1, 2^n - 1) = 2^{(m,n)} - 1$.

(30). Show that if $a > 1$ is an integer, then for any positive integers m and n, we have $(a^m - 1, a^n - 1) = a^{(m,n)} - 1$.

(31). Let m, n, a and b be positive integers such that $a > b$ and $(a, b) = 1$. Show that $(a^m - b^m, a^n - b^n) = a^{(m,n)} - b^{(m,n)}$.

(32). Show the following multiplicative property of the gcd:

$$(am, bn) = (a, b)(m, n)\Big(\frac{a}{(a, b)}, \frac{n}{(m, n)}\Big)\Big(\frac{b}{(a, b)}, \frac{m}{(m, n)}\Big).$$

Particularly this shows that $(am, bn) = (a, n)(b, m)$ whenever $(a, b) = (m, n) = 1$.

(33). Show that $n^{4l^4} + 4l^4$ is composite if either $n > 1$ or $l > 1$.

(34). Let $1 \leq a_1 < a_2 < ... < a_{2n+1} \leq 3n$. Show that there exist $1 \leq i < j \leq 2n + 1$ such that $a_i | a_j$.

(35). Let $1 \leq a_1 < a_2 < ... < a_{2010n+1} \leq 2011n$. Show that there exist $1 \leq i < j \leq 2010n + 1$ such that $a_i | a_j$.

(36). Let $k \in \mathbf{Z}$. Let $a = k + \frac{1}{2} + \sqrt{k^2 + \frac{1}{4}}$ and $b = 2k + \frac{1}{2} + \sqrt{4k^2 + \frac{1}{4}}$. Show that for any positive integer n, we have $k | [a^n]$ and $2k | [b^n]$.

(37). Let n be any positive integers. Let p be a prime such that $n < p < 2n$. Then p divides $\begin{pmatrix} 2n \\ n \end{pmatrix}$.

(38). Let a, b, c be positive integers. Then we have:

(i). $a + b + c + \min(a, b, c) = \max(a, b, c) + \min(a, b) + \min(b, c) + \min(c, a)$;

(ii). $[a, b, c] = \frac{abc(a,b,c)}{(a,b)(b,c)(c,a)}$.

(39). Let $a \in \mathbf{Z}$ and $b \in \mathbf{Z}$ be coprime. Show that if $a + b \equiv 1 \pmod 2$, then $(a + b, a - b) = 1$.

(40). Let $m, n \in \mathbf{Z}$ and m odd. Show that $(2^m - 1, 2^n + 1) = 1$ and $(4^m - 1, 4^n + 1) = 1$.

(41). Let $m, n \in \mathbf{Z}$ and m odd. Show that $(a^m - 1, a^n + 1)$ equals 1 if a is even, and equals 2 if a is odd.

(42). Let a, b and c be positive integers. Show that $[(a, b), c] = ([a, c], [b, c])$ and $([a, b], c) = [(a, c), (b, c)]$.

(43). Let p be an odd prime and $(a, b) = 1$ and $a + b \neq 0$. Show that $(a + b, \frac{a^p + b^p}{a+b}) = 1$ or p.

(44). Let n and k be positive integers such that $k > [\frac{n+1}{2}]$. Show that if $1 \leq a_1 < a_2 < ... < a_k \leq n$ are integers, then there exist indices $1 \leq i < j \leq k$ such that $a_i + a_1 = a_j$.

(45). Let $n \geq 1$ be an integer. Compute $\gcd(\binom{2n}{1}, \binom{2n}{3}, ..., \binom{2n}{2n-1})$.

(46). If $(a, b) = 1$, then there exist $x > 0$ and $y > 0$ such that $ax - by = 1$.

(47). (i). If $(a, b) = 1$ then for every $n > ab$ there exist positive x and y such that $n = ax + by$.

(ii). If $(a, b) = 1$ then there are no positive x and y such that $ab = ax + by$.

(48). Given x and y, let $m = ax + by, n = cx + dy$, where $ad - bc = \pm 1$. Prove that $(m, n) = (x, y)$.

(49). If $m \neq n$, compute the $\gcd(a^{2^m} + 1, a^{2^n} + 1)$ in terms of a.

(50). Let $n \geq 1$ be an integer. Show that $2^{2^n} - 1$ has at least n distinct prime divisors.

(51). Let $m > 0$ and a, b be integers such that $(a, b) = 1$. Show that there are infinitely many integers in the arithmetic progression $\{a + bi\}_{i=0}^{\infty}$ which are coprime to m.

(52). Let a and b be integers. Show that a and b are coprime if and only if there are integers s and t such that $as + bt = 1$.

(53). Let $n \geq 1$ be an integer and $S_n = \sum_{k=1}^{n}(k^5 + k^7)$. Find the greatest common divisor of S_n and S_{3n}.

(54). Assume that $a_1, ..., a_k$ are k integers such that $1 < a_1 < ... < a_k \leq n$ and $\text{lcm}(a_i, a_j) > n$ if $i \neq j$. Show that $\sum_{i=1}^{k} \frac{1}{a_i} < \frac{3}{2}$.

(55). Show that any sequence of $mn + 1$ distinct integers contains either an increasing subsequence of length greater than m or a decreasing subsequence of length greater than n.

(56). For every nonzero integer m, the *radical* of m, denoted by $\text{rad}(m)$, is defined to be the product of the distinct primes that divides m. Let a and b be nonzero integers. Show that there exists a positive integer k such that b divides a^k if and only if $\text{rad}(b)$ divides $\text{rad}(a)$.

(57). Let $d_1, ..., d_n$ be positive integers, and let $u_i = (d_i, \frac{d_1 ... d_n}{d_i})$ for $1 \leq i \leq n$. Show that for any $1 \leq i \leq n$, one has $u_i = (u_i, \frac{u_1 ... u_n}{u_i})$.

(58). Let $n \geq 1$ be an integer. Show that the following identities hold:

$$\lfloor\sqrt{n} + \sqrt{n+1}\rfloor = \lfloor\sqrt{4n+1}\rfloor, \lfloor\sqrt{n} + \sqrt{n+1} + \sqrt{n+2}\rfloor = \lfloor\sqrt{9n+8}\rfloor,$$

$$\lfloor\sqrt{n} + \sqrt{n+1} + \sqrt{n+2} + \sqrt{n+3}\rfloor = \lfloor\sqrt{16n+20}\rfloor,$$

$$\lfloor\sqrt{n} + \sqrt{n+1} + \sqrt{n+2} + \sqrt{n+3} + \sqrt{n+4}\rfloor = \lfloor\sqrt{25n+49}\rfloor,$$

$$\lfloor\sqrt{n} + \sqrt{n+1} + \sqrt{n+2} + \sqrt{n+3} + \sqrt{n+4} + \sqrt{n+5}\rfloor = \lfloor\sqrt{36n+89}\rfloor.$$

(59). For an integer $n \geq 1$, let $t_n := 1 + \frac{1}{3} + ... + \frac{1}{2n-1}$. Show that t_n is NOT an integer for any $n \geq 2$.

(60). Let $a \geq 1$ be an integer such that $2^a + 1$ is a prime. Show that $2^a - 1$ contains at least $\log_2 a$ distinct prime factors.

(61). Let n be a positive integer. Show that n is divisible by 3 (resp. 9) if and only if the digital sum $\sigma_{10}(n)$ of 10-adic representation of n is divisible by 3 (resp. 9).

(62). Let n be a positive integer and $n = \sum_{k=0}^{e} a_k \times 10^k$ with all a_k satisfying that $0 \leq a_k \leq 9$ and $a_e \neq 0$. Show that n is divisible by 11 if and only if the difference

$$\sum_{\substack{k=0 \\ k \text{ even}}}^{e} a_k - \sum_{\substack{k=0 \\ k \text{ odd}}}^{e} a_k$$

is divisible by 11.

Chapter 2. Congruences

§2.1. Ring of congruence classes

Let m be a positive integer. If a and b are integers such that $a-b$ is divisible by m, then we say that a and b are *congruent modulo* m, and write

$$a \equiv b \pmod{m}.$$

Integers a and b are called *incongruent modulo* m if they are not congruent modulo m. For example, $-12 \equiv 43 \pmod 5$ and $-12 \equiv 43 \pmod{11}$, but $-12 \not\equiv 43 \pmod 7$. Every even integer is congruent to 0 modulo 2, if x is not divisible by 3, then $x^2 \equiv 1 \pmod 3$. We have the following properties.

Proposition 2.1.1. *Let $m \in \mathbf{Z}^+$ and $a_i, b_i \in \mathbf{Z}$ for $i = 1, 2$ be such that $a_i \equiv b_i \pmod m$. Then each of the following is true:*

(i). $a_1 + a_2 \equiv b_1 + b_2 \pmod m$.

(ii). $a_1 - a_2 \equiv b_1 - b_2 \pmod m$.

(iii). $a_1 a_2 \equiv b_1 b_2 \pmod m$.

(iv). *For any integer c and d, $ca_1 + da_2 \equiv cb_1 + db_2 \pmod m$.*

(v). *For any positive integer n, $a_1^n \equiv b_1^n \pmod m$.*

(vi). *For $f(x) \in \mathbf{Z}[x]$, one has $f(a_1) \equiv f(b_1) \pmod m$.*

Proof. □

Congruence modulo m is an equivalence relation since for all integers a, b and c we have

(i). Reflexivity: $a \equiv a \pmod m$.

(ii). Symmetry: If $a \equiv b \pmod m$, then $b \equiv a \pmod m$.

(iii). Transitivity: If $a \equiv b \pmod m$ and $b \equiv c \pmod m$, then $a \equiv c \pmod m$.

Properties (i) and (ii) follow immediately from the definition of congruence. To prove (iii), we observe that if $a \equiv b \pmod m$ and $b \equiv c \pmod m$, then there exist integers x and y such that $a - b = mx$ and $b - c = my$. Since

$$a - c = (a - b) + (b - c) = mx + my = m(x + y),$$

it follows that $a \equiv c \pmod m$. The equivalence class of an integer a under this relation is called the *congruence class* of a modulo m and written $a + m\mathbf{Z}$. Thus $a + m\mathbf{Z}$ is the set of all integers b such that $b \equiv a \pmod m$, that is, the set of all integers of the form $a + mx$ for some integer x. If $(a + m\mathbf{Z}) \cap (b + m\mathbf{Z}) \neq \phi$, then $a + m\mathbf{Z} = b + m\mathbf{Z}$. We denote by $\mathbf{Z}/m\mathbf{Z}$ the set of all congruence classes modulo m.

A congruence class modulo m is also called a *residue class modulo m.*

By the division algorithm, we can write every integer a in the form $a = mq + r$, where q and r are integers and $0 \leq r \leq m - 1$. Then $a \equiv r \pmod m$, and r is called the *least nonnegative residue* of a modulo m.

If $a \equiv 0 \pmod m$ and $|a| < m$, then $a = 0$ since 0 is the only integral multiple of m in the open interval $(-m, m)$. This implies that if $a \equiv b \pmod m$ and $|a-b| < m$, then $a = b$. In particular, if $r_1, r_2 \in \{0, 1, \ldots, m-1\}$ and if $a \equiv r_1 \pmod m$ and $a \equiv r_2 \pmod m$, then $r_1 = r_2$. Thus every integer belongs to a unique congruence class of the form $r + m\mathbf{Z}$, where $0 \leq r \leq m - 1$, and so

$$\mathbf{Z}/m\mathbf{Z} = \{m\mathbf{Z}, 1 + m\mathbf{Z}, \ldots, (m-1) + m\mathbf{Z}\}.$$

The integers $0, 1, \ldots, m-1$ are pairwise incongruent modulo m.

A set of integers $R = \{r_1, \dots, r_m\}$ is called a *complete set of residues* modulo m if $r_1, \dots, r_m$ are pairwise incongruent modulo m and every integer x is congruent modulo m to some integer $r_i \in R$. For example, the set $\{0, 2, 4, 6, 8, 10, 12\}$ is a complete set of residues modulo 7. The set $\{0, 3, 6, 9, 12, 15, 18, 21\}$ is a complete set of residues modulo 8. The set $\{0, 1, \dots, m-1\}$ is a complete set of residues modulo m for every positive integer m.

There is a natural way to define addition, subtraction and multiplication of congruence classes. If

$$a_1 \equiv a_2 \pmod{m}$$

and

$$b_1 \equiv b_2 \pmod{m},$$

then

$$a_1 \pm b_1 \equiv a_2 \pm b_2 \pmod{m},$$

and

$$a_1 b_1 \equiv a_2 b_2 \pmod{m}.$$

There statements are consequences of the identities

$$(a_1 + b_1) - (a_2 + b_2) = (a_1 - a_2) + (b_1 - b_2) \equiv 0 \pmod{m},$$
$$(a_1 - b_1) - (a_2 - b_2) = (a_1 - a_2) - (b_1 - b_2) \equiv 0 \pmod{m}$$

and

$$a_1 b_1 - a_2 b_2 = a_1(b_1 - b_2) + (a_1 - a_2) b_2 \equiv 0 \pmod{m}.$$

Addition, subtraction and multiplication in $\mathbf{Z}/m\mathbf{Z}$ are well-defined if we define the sum, difference and product of congruence classes modulo m by

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

and

$$(a + m\mathbf{Z}) \cdot (b + m\mathbf{Z}) = ab + m\mathbf{Z}.$$

Addition of congruence classes is associative and commutative since

$$\begin{aligned}
&((a + m\mathbf{Z}) + (b + m\mathbf{Z})) + (c + m\mathbf{Z}) \\
&= ((a + b) + m\mathbf{Z}) + (c + m\mathbf{Z}) \\
&= ((a + b) + c) + m\mathbf{Z} \\
&= (a + (b + c)) + m\mathbf{Z} \\
&= (a + m\mathbf{Z}) + ((b + c) + m\mathbf{Z}) \\
&= (a + m\mathbf{Z}) + ((b + m\mathbf{Z}) + (c + m\mathbf{Z}))
\end{aligned}$$

and

$$\begin{aligned}
&(a + m\mathbf{Z}) + (b + m\mathbf{Z}) = (a + b) + m\mathbf{Z} \\
&= (b + a) + m\mathbf{Z} \\
&= (b + m\mathbf{Z}) + (a + m\mathbf{Z}).
\end{aligned}$$

The congruence class $m\mathbf{Z}$ is a zero element for addition since $m\mathbf{Z} + (a + m\mathbf{Z}) = a + m\mathbf{Z}$ for all $a + m\mathbf{Z} \in \mathbf{Z}/m\mathbf{Z}$, and the additive inverse of the congruence class $a + m\mathbf{Z}$ is $-a + m\mathbf{Z}$ since

$$(a + m\mathbf{Z}) + (-a + m\mathbf{Z}) = (a - a) + m\mathbf{Z} = m\mathbf{Z}.$$

From these identities we see that the set of congruence classes modulo m is an abelian group under addition.

We have also defined multiplication in $\mathbf{Z}/m\mathbf{Z}$. Multiplication is associative and commutative since

$$((a+m\mathbf{Z})(b+m\mathbf{Z}))(c+m\mathbf{Z}) = (ab)c+m\mathbf{Z} = a(bc)+m\mathbf{Z} = (a+m\mathbf{Z})((b+m\mathbf{Z})(c+m\mathbf{Z}))$$

and

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

The congruence class $1+m\mathbf{Z}$ is an identity for multiplication since $(1+m\mathbf{Z})(a+m\mathbf{Z}) = a+m\mathbf{Z}$ for all $a+m\mathbf{Z} \in \mathbf{Z}/m\mathbf{Z}$. Finally, multiplication of congruence classes is *distributive with respect to addition* in the sense that

$$\begin{aligned}&(a+m\mathbf{Z})((b+m\mathbf{Z})+(c+m\mathbf{Z}))\\ &\quad = a(b+c)+m\mathbf{Z}\\ &\quad = (ab+m\mathbf{Z})+(ac+m\mathbf{Z})\\ &\quad = (a+m\mathbf{Z})(b+m\mathbf{Z})+(a+m\mathbf{Z})(c+m\mathbf{Z})\end{aligned}$$

for all $a+m\mathbf{Z}, b+m\mathbf{Z}, c+m\mathbf{Z} \in \mathbf{Z}/m\mathbf{Z}$.

Definition. A *ring* is a set R with two binary operations, addition and multiplication such that R is an abelian group under addition with additive identity 0 and multiplication satisfies the following axioms:

(i) Associativity: For all $x, y, z \in R, (xy)z = x(yz)$.

(ii) Identity element: There exists an element $1 \in R$ such that for all $x \in R$, $1 \cdot x = x \cdot 1 = x$. The element 1 is called the *multiplicative identity* of the ring.

(iii) Distributivity: For all $x, y, z \in R,\ x(y+z) = xy+xz$. The ring R is *commutative* if multiplication also satisfies the axiom:

(iv) Commutativity: For all $x, y \in R, xy = yx$.

The integers, rational numbers, real numbers and complex numbers are example of commutative rings. The set $M_2(\mathbf{C})$ of 2×2 matrices with complex coefficients and the usual matrix addition and multiplication is a noncommutative ring.

Definition. Let R and S be rings with multiplicative identities 1_R and 1_S, respectively. A map $f : R \longrightarrow S$ is called a *ring homomorphism* if $f(x+y) = f(x)+f(y)$ and $f(xy) = f(x)f(y)$ for all $x, y \in R$ and $f(1_R) = 1_S$. An element a in the ring R is called a *unit* if there exists an element $x \in R$ such that $ax = xa = 1$. If a is a unit in R and $x, y \in R$ are both inverses of a, then $x = x(ay) = (xa)y = y$ and so the inverse of a is unique. We denote the inverse of a by a^{-1}. The set $R^{\times}$ of all units in R is a multiplicative group, called the *group of units* in the ring R.

Definition. A *field* is a commutative ring in which every nonzero element is a unit.

For example, the rational, real and complex numbers are fields. The integers form a ring but not a field and the only units in the ring of integers are ± 1.

The various properties of sums and products of congruence classes that we proved in this section are equivalent to the following statement.

Theorem 2.1.1. *For every integer $m \geq 2$, the set $\mathbf{Z}/m\mathbf{Z}$ of congruence classes modulo m is a commutative ring.*

§2.2. Linear congruences and Wilson's theorem

The following theorem is one of the most useful and important tools in elementary number theory.

Theorem 2.2.1. *Let m,a,b be integers with $m \geq 1$ and $d = (a, m)$ be the greatest common divisor of a and m. The congruence*

$$ax \equiv b \pmod{m} \tag{2.1}$$

has a solution if and only if

$$b \equiv 0 \pmod{d}.$$

If $b \equiv 0 \pmod{d}$, then the congruence (2.1) has exactly d solutions in integers that are pairwise incongruence modulo m. In particular, if $(a, m) = 1$, then for every integer b the congruence (2.1) has a unique solution modulo m.

Proof. Let $d = (a, m)$. Congruence (2.1) has a solution if and only if there exist integers x, y such that $ax - b = my$, or equivalently, $b = ax - my$. By Theorem 1.6.1, this is possible if and only if $b \equiv 0 \pmod{d}$.

If x, x_1 are solutions of (2.1), then

$$a(x_1 - x) \equiv ax_1 - ax \equiv b - b \equiv 0 \pmod{m},$$

and so $a(x_1 - x) = mz$ for some integer z. If d is the greatest common divisor of a and m, then $(a/d, m/d) = 1$ and $\frac{a}{d}(x - x_1) = \frac{m}{d}z$. By Euclid's Lemma (Theorem 1.4.1), m/d divides $x_1 - x$ and so $x_1 = x + \frac{im}{d}$ for some integer i, that is $x_1 \equiv x \pmod{\frac{m}{d}}$. Moreover, every integer x_1 of this form is a solution of (2.1). An integer x_1 congruent to x modulo m/d is congruent to $x + im/d$ modulo m for some integer $i = 0, 1, \ldots, d - 1$ and the d integers $x + im/d$ with $i = 0, 1, \ldots, d - 1$ are pairwise incongruent modulo m. Thus the congruence (2.1) has exactly d pairwise incongruent solutions. This completes the proof. □

Theorem 2.2.2. *If p is a prime, then $\mathbf{Z}/p\mathbf{Z}$ is a field.*

Proof. If $a + p\mathbf{Z}$ and $a + p\mathbf{Z} \neq p\mathbf{Z}$, then a is an integer not divisible by p. By Theorem 2.2.1, there exists an integer x such that $ax \equiv 1 \pmod{p}$. This implies that $(a + p\mathbf{Z})(x + p\mathbf{Z}) = 1 + p\mathbf{Z}$ and so $a + p\mathbf{Z}$ is invertible. Thus every nonzero congruence class in $\mathbf{Z}/p\mathbf{Z}$ is a unit and $\mathbf{Z}/p\mathbf{Z}$ is a field. □

Here are some examples of linear congruences. The congruence $7x \equiv 3 \pmod{5}$ has a unique solution modulo 5 since $(7, 5) = 1$. The solution is $x \equiv 4 \pmod{5}$. The congruence

$$35x \equiv -14 \pmod{91} \tag{2.2}$$

is solvable since $(35, 91) = 7$ and $-14 \equiv 0 \pmod{7}$.

Congruence(2.2) is equivalent to the congruence

$$5x \equiv -2 \pmod{13} \tag{2.3}$$

which has the unique solution $x \equiv 10 \pmod{13}$. Every solution of (2.2) satisfies $x \equiv 10 \pmod{13}$ and so a complete set of solutions that are pairwise incongruent modulo 91 is $\{10, 23, 36, 49, 62, 75, 88\}$.

Lemma 2.2.3. *Let p be a prime number. Then $x^2 \equiv 1 \pmod{p}$ if and only if $x \equiv \pm 1 \pmod{p}$.*

Proof. If $x \equiv \pm 1 \pmod{p}$, then $x^2 \equiv 1 \pmod{p}$. Conversely, if $x^2 \equiv 1 \pmod{p}$, then p divides $x^2 - 1 = (x - 1)(x + 1)$ and so p must divide $x - 1$ or $x + 1$. □

Theorem 2.2.4. (Wilson) *If p is a prime number, then $(p-1)! \equiv -1 \pmod{p}$.*

Proof. This is true for $p = 2$ and $p = 3$. Let $p \geq 5$. By Theorem 2.2.2, to each integer $a \in \{1, 2, \ldots, p-1\}$ there is a unique integer $a^{-1} \in \{1, 2, \ldots, p-1\}$ such that $aa^{-1} \equiv 1 \pmod{p}$. By Lemma 2.2.3, $a = a^{-1}$ if and only if $a = 1$ or $a = p-1$. Therefore we can partition the $p-3$ numbers in the set $\{2, 3, \ldots, p-2\}$ into $(p-3)/2$ pairs of integers $\{a_i, a_i^{-1}\}$ such that $a_i a_i^{-1} \equiv 1 \pmod{p}$ for $i = 1, \ldots, (p-3)/2$. Then

$$\begin{aligned}(p-1)! &\equiv 1 \cdot 2 \cdots (p-2)(p-1) \\ &\equiv (p-1) \prod_{i=1}^{(p-3)/2} a_i a_i^{-1} \\ &\equiv p - 1 \equiv -1 \pmod{p}.\end{aligned}$$

This completes the proof. □

For example, $4! \equiv 24 \equiv -1 \pmod{5}$. The converse of Wilson's theorem is also true (Exercise).

If $a \equiv b \pmod{m}$, then $a = b + mx$ for some integer x. An integer d is a common divisor of a and m if and only if d is a common divisor of b and m and so $(a, m) = (b, m)$. In particular, if a is relatively prime to m, then every integer in the congruence class of $a + m\mathbf{Z}$ is relatively prime to m. A congruence class modulo m is called *relatively prime to* m if some (and, consequently, every) integer in the class is relatively prime to m.

We denote by $\varphi(m)$ the number of congruence class in $\mathbf{Z}/m\mathbf{Z}$ that are relatively prime to m. The function $\varphi(m)$ is called the *Euler phi function*. Equivalently, $\varphi(m)$ is the number of integers in the set $\{0, 1, 2, ..., m-1\}$ that are relatively prime to m. The Euler phi function is also called the *totient function*.

Definition. A set of integers $\{r_1, ..., r_{\varphi(m)}\}$ is called a *reduced set of residues* modulo m if every integer x such that $(x, m) = 1$ is congruent modulo m to some integer r_i. An integer a is called *invertible modulo* m or a *unit modulo* m if there exists an integer x such that $ax \equiv 1 \pmod{m}$.

For example, the sets $\{1, 2, 3, 4, 5, 6\}$ and $\{2, 4, 6, 8, 10, 12\}$ are reduced sets of residues modulo 7. The set $\{1, 3, 5, 7\}$ and $\{3, 9, 15, 21\}$ are reduced sets of residues modulo 8.

By Theorem 2.2.1, a is invertible modulo m if and only if a is relatively prime to m. Moreover, if a is invertible and $ax \equiv 1 \pmod{m}$, then x is unique modulo m. The congruence class $a + m\mathbf{Z}$ is called *invertible* if there exists a congruence class $x + m\mathbf{Z}$ such that $(a + m\mathbf{Z})(x + m\mathbf{Z}) = 1 + m\mathbf{Z}$.

We denote the inverse of the congruence class $a + m\mathbf{Z}$ by $(a + m\mathbf{Z})^{-1} = a^{-1} + m\mathbf{Z}$. The invertible congruence classes are the units in the ring $\mathbf{Z}/m\mathbf{Z}$. We denote the group of units in $\mathbf{Z}/m\mathbf{Z}$ by $(\mathbf{Z}/m\mathbf{Z})^{\times}$. If $R = \{r_1, \ldots, r_{\varphi(m)}\}$ is a reduced set of residues modulo m, then $(\mathbf{Z}/m\mathbf{Z})^{\times} = \{r + m\mathbf{Z} : r \in R\}$. and $|(\mathbf{Z}/m\mathbf{Z})^{\times}| = \varphi(m)$. For example, $(\mathbf{Z}/6\mathbf{Z})^{\times} = \{1 + 6\mathbf{Z}, 5 + 6\mathbf{Z}\}$.

If $a + m\mathbf{Z}$ is a unit in $\mathbf{Z}/m\mathbf{Z}$, then $(a, m) = 1$ and we can apply the Euclidean algorithm to compute $(a + m\mathbf{Z})^{-1}$. If we can find integers x and y such that $ax + my = 1$, then $(a + m\mathbf{Z})(x + m\mathbf{Z}) = 1 + m\mathbf{Z}$, and $x + m\mathbf{Z} = (a + m\mathbf{Z})^{-1}$.

For example, to find the inverse of $13 + 17\mathbf{Z}$, we use the Euclidean algorithm to obtain $17 = 13 \cdot 1 + 4$, $13 = 4 \cdot 3 + 1$, $4 = 1 \cdot 4$. This gives $1 = 13 - 4 \cdot 3 = 13 - (17 - 13 \cdot 1)3 = 13 \cdot 4 - 17 \cdot 3$, and so $13 \cdot 4 \equiv 1 \pmod{17}$. Therefore, $(13 + 17\mathbf{Z})^{-1} = 4 + 17\mathbf{Z}$.

§2.3. Euler Phi Function

An *arithmetical function* is a function defined on the positive integers. The Euler phi function $\varphi(m)$ is the arithmetic function that counts the number of integers in the set $\{0, 1, 2, \ldots, m-1\}$ that are relatively prime to m. We have

$$\begin{aligned} &\varphi(1)=1, \quad \varphi(6)=2,\\ &\varphi(2)=1, \quad \varphi(7)=6\\ &\varphi(3)=2, \quad \varphi(8)=4,\\ &\varphi(4)=2, \quad \varphi(9)=6,\\ &\varphi(5)=4, \quad \varphi(10)=4. \end{aligned}$$

If p is a prime number, then $(a,p)=1$ for $a=1,\ldots,p-1$ and $\varphi(p)=p-1$. If p^r is a prime power and $0 \le a \le p^r-1$, then $(a,p^r)>1$ if and only if a is a multiple of p. The integral multiples of p in the interval $[0,p^r-1]$ are the p^{r-1} numbers $0,p,2p,3p,\ldots,(p^{r-1}-1)p$ and so $\varphi(p^r)=p^r-p^{r-1}=p^r\left(1-\frac{1}{p}\right)$. In this section we shall obtain some important properties of the Euler phi function.

Theorem 2.3.1. *Let m and n be relatively prime positive integers. For every integer c there exist unique integers a and b such that $0 \le a \le n-1$, $0 \le b \le m-1$, and*

$$c \equiv ma+nb \pmod{mn}. \tag{2.4}$$

Moreover, $(c,mn)=1$ if and only if $(a,n)=(b,m)=1$ in the representation (2.4).

Proof. Let a_1, a_2, b_1, b_2 be integers such that

$$ma_1+nb_1 \equiv ma_2+nb_2 \pmod{mn}.$$

Since $(m,n)=1$, it follows that $a_1 \equiv a_2 \pmod{n}$ and so $a_1=a_2$. Similarly, $b_1=b_2$. It follows that the mn integers $ma+nb$ are pairwise incongruent modulo mn. Since there are exactly mn distinct congruence classes modulo mn, the congruence (2.4) has a unique solution for every integer c.

Let $c \equiv ma+nb \pmod{mn}$. Since $(m,n)=1$, we have

$$(c,m)=(ma+nb,m)=(nb,m)=(b,m)$$

and

$$(c,n)=(ma+nb,n)=(ma,n)=(a,n).$$

It follows that $(c,mn)=1$ if and only if $(c,m)=(c,n)=1$ if and only if $(b,m)=(a,n)=1$. This completes the proof. □

For example, we can represent the congruence classes modulo 6 as linear combinations of 2 and 3 as follows:

$$\begin{aligned} 0 &\equiv 0\cdot 2+0\cdot 3 \pmod{6}\\ 1 &\equiv 2\cdot 2+1\cdot 3 \pmod{6}\\ 2 &\equiv 1\cdot 2+0\cdot 3 \pmod{6}\\ 3 &\equiv 0\cdot 2+1\cdot 3 \pmod{6}\\ 4 &\equiv 2\cdot 2+0\cdot 3 \pmod{6}\\ 5 &\equiv 1\cdot 2+1\cdot 3 \pmod{6} \end{aligned}$$

A *multiplicative function* is an arithmetic function $f(m)$ such that $f(mn)=f(m)f(n)$ for all pairs of relatively prime positive integers m and n. If $f(m)$ is multiplicative, then it is easy to prove by induction on k that if $m_1,\ldots,m_k$ are pairwise relatively prime positive integers, then $f(m_1\ldots m_k)=f(m_1)\ldots f(m_k)$.

Theorem 2.3.2. *The Euler phi function is multiplicative. Moreover we have*

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

Proof. Let $(m,n) = 1$. There are $\varphi(mn)$ congruence classes in the ring $\mathbf{Z}/mn\mathbf{Z}$ that are relatively prime to mn. By Theorem 2.3.1, every congruence class modulo mn can be written uniquely in the form $ma+nb+mn\mathbf{Z}$, where a and b are integers such that $0 \le a \le n-1$ and $0 \le b \le m-1$. Moreover, the congruence class $ma+nb+mn\mathbf{Z}$ is prime to mn if and only if $(b,m) = (a,n) = 1$. Since there are $\varphi(n)$ integers $a \in [0,n-1]$ that are relatively prime to n, and $\varphi(m)$ integers $b \in [0,m-1]$ that are relatively prime to m, it follows that $\varphi(mn) = \varphi(m)\varphi(n)$ and so the Euler phi function is multiplicative. If $m_1,\ldots,m_k$ are pairwise relatively prime positive integers, then $\varphi(m_1\cdots m_k) = \varphi(m_1)\cdots\varphi(m_k)$. In particular, if $m = p_1^{r_1}\cdots p_k^{r_k}$ is the standard factorization of m, where $p_1,\ldots,p_k$ are distinct primes and $r_1,\ldots,r_k$ are positive integers, then

$$\varphi(m) = \prod_{i=1}^{k}\varphi(p_i^{r_i}) = \prod_{i=1}^{k}p_i^{r_i}\left(1-\frac{1}{p_i}\right) = m\prod_{p|m}\left(1-\frac{1}{p}\right)$$

This completes the proof. □

For example, we have $7875 = 3^2 5^3 7$ and

$$\varphi(7875) = \varphi(3^2)\varphi(5^3)\varphi(7) = (9-3)(125-25)(7-1) = 3600.$$

Theorem 2.3.3. (Gauss Theorem) *For every positive integer m, we have*

$$\sum_{d|m}\varphi(d) = m.$$

Proof. Let $m = p_1^{t_1}p_2^{t_2}\cdots p_k^{t_k}$, where $p_1,\ldots,p_k$ are distinct prime numbers and $t_1,\ldots,t_k$ are positive integers. Clearly every divisor d of m is of the form $d = p_1^{r_1}p_2^{r_2}\cdots p_k^{r_k}$, where $0 \le r_i \le t_i$ for $i = 1,\ldots,k$. By Theorem 2.3.2, $\varphi(d)$ is multiplicative. Hence $\varphi(d) = \varphi(p_1^{r_1})\varphi(p_2^{r_2})\cdots\varphi(p_k^{r_k})$. Thus

$$\begin{aligned}\sum_{d|m}\varphi(d) &= \sum_{r_1=0}^{t_1}\cdots\sum_{r_k=0}^{t_k}\varphi(p_1^{r_1}\cdots p_k^{r_k}) = \sum_{r_1=0}^{t_1}\cdots\sum_{r_k=0}^{t_k}\varphi(p_1^{r_1})\cdots\varphi(p_k^{r_k})\\ &= \prod_{i=1}^{k}\sum_{r_i=0}^{t_i}\varphi(p_i^{r_i}) = \prod_{i=1}^{k}\left(1+\sum_{r_i=1}^{t_i}(p_i^{r_i}-p_i^{r_i-1})\right) = \prod_{i=1}^{k}p_i^{t_i} = m\end{aligned}$$

as required. This completes the proof. □

Another proof of Gauss Theorem: Consider the set of fractions $M := \{\frac{1}{m},\ldots,\frac{m}{m}\}$. Reduce each element of M to a reduced fraction and get a new set M'. Then we have $M' = \bigcup_{d|m}M_d$, where $M_d = \left\{\frac{l}{d}|1 \le l \le d, (l,d) = 1\right\}$. Note that $M_{d_1}\bigcap M_{d_2} = \phi$ if $d_1 \ne d_2$. Since $|M_d| = \varphi(d)$, we have $|M'| = \sum_{d|m}|M_d| = \sum_{d|m}\varphi(d)$. But $|M'| = |M| = m$. Therefore we obtain $\sum_{d|m}\varphi(d) = m$ as desired. □

§2.4. Chinese remainder theorem and its extension

Theorem 2.4.1. *Let m and n be positive integers. For any integers a and b there exists an integer x such that*

$$x \equiv a \pmod{m} \tag{2.5}$$

and

$$x \equiv b \pmod{n} \tag{2.6}$$

if and only if

$$a \equiv b \pmod{(m,n)}$$

If x is a solution of congruences (2.5) and (2.6), then the integer y is also a solution if and only if

$$x \equiv y \pmod{[m,n]}.$$

Proof. If x is a solution of congruence (2.5), then $x = a + mu$ for some integer u. If x is also a solution of congruence (2.6), then $x = a + mu \equiv b \pmod{n}$, that is, $a + mu = b + nv$ for some integer v. It follows that

$$a - b = nv - mu \equiv 0 \pmod{(m,n)}.$$

Conversely, if $a - b \equiv 0 \pmod{(m,n)}$, then by the Bezout's Theorem there exist integer u and v such that $a-b = nv - mu$. Then $x = a + mu = b + nv$ is a solution of the two congruences.

An integer y is another solution of the congruences if and only if $y \equiv a \equiv x \pmod{m}$ and $y \equiv b \equiv x \pmod{n}$, that is, if and only if $x - y$ is a common multiple of m and n, or equivalently, $x - y$ is divisible by the least common multiple $[m,n]$. This completes the proof. □

Example. The system of congruences

$$\begin{aligned} x &\equiv 5 \pmod{21} \\ x &\equiv 19 \pmod{56} \end{aligned}$$

has a solution, since $(56, 21) = 7$ and $19 \equiv 5 \pmod{7}$.

The integer x is a solution if there exists an integer u such that $x = 5+21u \equiv 19 \pmod{56}$, that is $21u \equiv 14 \pmod{56}$, $3u \equiv 2 \pmod{8}$, or $u \equiv 6 \pmod{8}$. Then

$$x = 5 + 21u = 5 + 21(6 + 8v) = 131 + 168v$$

is a solution of the system of congruence for any integer v and so the set of all solutions is the congruence class $131 + 168\mathbf{Z}$.

Theorem 2.4.2. (Chinese remainder theorem) *Let $k \geq 2$. If $a_1, \ldots, a_k$ are integers and $m_1, \ldots, m_k$ are pairwise relatively prime positive integers, then there exists an integers x such that $x \equiv a_i \pmod{m_i}$ for all $i = 1, \ldots, k$. If x is any solution of this set of congruences, then the integer y is also a solution if and only if $x \equiv y \pmod{m_1 \cdots m_k}$.*

Proof. We prove the theorem by induction on k. If $k = 2$, then $[m_1, m_2] = m_1m_2$ and this is a special case of Theorem 2.4.1.

Let $k \geq 3$ and assume that the statement is true for $k - 1$ congruences. Then there exists an integer z such that $z \equiv a_i \pmod{m_i}$ for $i = 1, \ldots, k-1$. Since

$m_1, \ldots, m_k$ are pairwise relatively prime integers, we have $(m_1 \cdots m_{k-1}, m_k) = 1$, and so, by the case $k = 2$, there exists an integer x such that

$$\begin{aligned} x &\equiv z \pmod{m_1 \cdots m_{k-1}}, \\ x &\equiv a_k \pmod{m_k}. \end{aligned}$$

Then

$$x \equiv z \equiv a_i \pmod{m_i}$$

for $i = 1, \ldots, k-1$.

If y is another solution of the system of k congruences, then $x - y$ is divisible by m_i for all $i = 1, \ldots, k$. Since $m_1, \ldots, m_k$ are pairwise relatively prime, it follows that $x - y$ is divisible by $m_1 \cdots m_k$. This completes the proof. □

Example. The system of congruences

$$\begin{aligned} x &\equiv 2 \pmod{3}, \\ x &\equiv 3 \pmod{5}, \\ x &\equiv 5 \pmod{7}, \\ x &\equiv 7 \pmod{11} \end{aligned}$$

has a solution, since the moduli are pairwise relatively prime. The solution to the first two congruences is the congruence class $x \equiv 8 \pmod{15}$. The solution to the first three congruences is the congruence class $x \equiv 68 \pmod{105}$. The solution to the four congruences is the congruence class $x \equiv 1118 \pmod{1155}$.

Consequently, we give an extension of Chinese remainder theorem.

Theorem 2.4.3. *Let $m_1, ..., m_k$ be positive integers and $a_1, ..., a_k$ be any integers. Then*

(i). *The simultaneous congruences*

$$x \equiv a_1 \pmod{m_1}, ..., x \equiv a_k \pmod{m_k} \tag{2.10}$$

have a solution if and only if $a_i \equiv a_j \pmod{\gcd(m_i, m_j)}$ whenever $i \neq j$;

(ii). *When this condition is satisfied, the general solution forms a single congruence class $\pmod{m}$ where $m := \operatorname{lcm}(m_1, ..., m_k)$.*

Proof. (i). ($\Longrightarrow$) Suppose that (2.10) has a solution. For $i \neq j$, since $x \equiv a_i \pmod{m_i}$ and $x \equiv a_j \pmod{m_j}$, we have $x \equiv a_i \pmod{\gcd(m_i, m_j)}$ and $x \equiv a_j \pmod{\gcd(m_i, m_j)}$. So we deduce that $a_i \equiv a_j \pmod{\gcd(m_i, m_j)}$ as desired.

($\Longleftarrow$) Assume that $a_i \equiv a_j \pmod{\gcd(m_i, m_j)}$ whenever $i \neq j$. In what follows we show that (2.10) has a solution. Let $m = \prod_{i=1}^{l} p_i^{e_i}$. WLOG we may let $p_i^{e_i} \parallel m_{j_i}$ for some $1 \leq j_i \leq k$ $(1 \leq i \leq l)$. We claim that (2.10) has a solution is equivalent to the following system of linear congruences has a solution:

$$x \equiv a_{j_1} \pmod{p_1^{e_1}}, ..., x \equiv a_{j_l} \pmod{p_l^{e_l}}, \tag{2.11}$$

where $1 \leq j_1, ..., j_l \leq k$ are not necessarily distinct l integers such that $p_i^{e_i} \parallel m_{j_i}$ for $1 \leq i \leq l$. But by the Chinese remainder theorem (Theorem 2.4.2) we know that (2.11) clearly has a solution since all the modulus $p_1^{e_1}, ..., p_l^{e_l}$ are coprime. It remains to show the equivalence of (2.10) and (2.11). Suppose that (2.10) is solvable, then by Theorem 2.4.3, (2.11) is also solvable. Conversely, assume that (2.11) is solvable and x_0 is such a solution. Then

$$x_0 \equiv a_{j_1} \pmod{p_1^{e_1}}, ..., x_0 \equiv a_{j_l} \pmod{p_l^{e_l}}. \tag{2.12}$$

Let $m_1 = \prod_{t\in T} p_t^{f_t}$ with $1 \le f_t \le e_t$ for $t \in T \subseteq \{1, ..., l\}$. Now let $t \in T$. By (2.12) we have $x_0 \equiv a_{j_t} \pmod{p_t^{f_t}}$. Since $a_1 \equiv a_{j_t} \pmod{\gcd(m_1, m_{j_t})}$ and note that $p_t^{e_t} \parallel m_{j_t}$ implying that $p_t^{f_t} | m_{j_t}$, we have $a_1 \equiv a_{j_t} \pmod{p_t^{f_t}}$. So $x_0 \equiv a_1 \pmod{p_t^{f_t}}$ and thus we have $x_0 \equiv a_1 \pmod{m_1}$. Similarly one can prove that $x_0 \equiv a_2 \pmod{m_2}, ..., x_0 \equiv a_k \pmod{m_k}$. That is: x_0 is a solution of (2.10). This completes the proof of part (i).

(ii). Let x_1 be a given solution and x_2 be any solution of (2.10). Then we have $x_1 \equiv x_2 \pmod{m_1}, ..., x_1 \equiv x_2 \pmod{m_k}$. It follows immediately that $x_1 \equiv x_2 \pmod{\operatorname{lcm}(m_1, ..., m_k)}$. That is x_1 and x_2 are in the same congruence class $(\bmod \operatorname{lcm}(m_1, ..., m_k))$. Part (ii) is proved. □

Example. Find all solutions of the system of congruences:

$$x \equiv 11 \pmod{36},\ x \equiv 7 \pmod{40},\ x \equiv 32 \pmod{75}.$$

Then one can check that the conditions $m_{ij} | (a_i - a_j)$ are all satisfied, and so there are solutions. These three congruences can be simplified to:

$$x \equiv 7 \pmod 8,\ x \equiv 11 \equiv 2 \pmod 9,\ x \equiv 32 \equiv 7 \pmod{25}.$$

Finally we find the general solution $x \equiv 407 \pmod{1800}$.

Now consider the problem of solving diophantine equations of the form

$$f(x_1, ..., x_k) \equiv 0 \pmod m, \tag{2.7}$$

where $f(x_1, \ldots, x_k)$ is a polynomial with integer coefficients in one or several variables. This equation is *solvable modulo* m if there exist integers $a_1, \ldots, a_k$ such that

$$f(a_1, \ldots, a_k) \equiv 0 \pmod m.$$

There is an important application of the Chinese remainder theorem for solving (2.7). By the Chinese remainder theorem we can reduce the question of the solvability of the congruence modulo m to the special case of prime power moduli p^r. For simplicity, we consider polynomials in only one variable.

Theorem 2.4.6. *Let $m = p_1^{r_1} \cdots p_k^{r_k}$ be the standard factorization of the positive integer m. Let $f(x)$ be a polynomial with integral coefficients. The congruence*

$$f(x) \equiv 0 \pmod m$$

is solvable if and only if the congruences

$$f(x) \equiv 0 \pmod{p_i^{r_i}}$$

are solvable for all $i = i, \ldots, k$.

Proof. If $f(x) \equiv 0 \pmod m$ has a solution in integers, then there exists an integer a such that m divides $f(a)$. Since $p_i^{r_i}$ divides m, it follows that $p_i^{r_i}$ divides $f(a)$ and so the congruences $f(x) \equiv 0 \pmod{p_i^{r_i}}$ are solvable for $i = i, \ldots, k$.

Conversely, suppose that the congruences $f(x) \equiv 0 \pmod{p_i^{r_i}}$ are solvable for $i = i, \ldots, k$. Then for each i there exists an integer a_i such that

$$f(a_i) \equiv 0 \pmod{p_i^{r_i}}.$$

Since the prime power $p_1^{r_1}, \ldots, p_k^{r_k}$ are pairwise relatively prime, the Chinese remainder theorem tells us that there exists an integer a such that

$$a \equiv a_i \pmod{p_i^{r_i}}$$

for all i. Then

$$f(a) \equiv f(a_i) \equiv 0 \pmod{p_i^{r_i}}$$

for all i. Since $f(a)$ is divisible by each of the prime powers $p_i^{r_i}$, it is also divisible by their product m and so $f(a) \equiv 0 \pmod{m}$. This completes the proof. □

Example. Consider the congruence

$$f(x) = x^2 - 34 \equiv 0 \pmod{495}.$$

Since $495 = 3^2 \cdot 5 \cdot 11$, it suffices to solve the congruences

$$\begin{aligned} f(x) = x^2 - 34 \equiv x^2 + 2 \equiv 0 \pmod{9} \\ f(x) = x^2 - 34 \equiv x^2 + 1 \equiv 0 \pmod{5} \end{aligned}$$

and

$$f(x) = x^2 - 34 \equiv x^2 - 1 \equiv 0 \pmod{11}$$

These congruences have solutions

$$\begin{aligned} f(5) \equiv 0 \pmod{9} \\ f(2) \equiv 0 \pmod{5} \end{aligned}$$

and

$$f(1) \equiv 0 \pmod{11}.$$

By the Chinese remainder theorem, there exists an integer a such that

$$\begin{aligned} a \equiv 5 \pmod{9} \\ a \equiv 2 \pmod{5} \\ a \equiv 1 \pmod{11} \end{aligned}$$

Solving these congruences, we obtain

$$a \equiv 122 \pmod{495}.$$

We can check that

$$f(122) = 122^2 - 34 = 14850 = 30 \cdot 495,$$

and so

$$f(122) \equiv 0 \pmod{495}.$$

Let $N(f(x_1, ..., x_k) \equiv 0 \pmod{m})$ denotes the number of such k-tuples $(a_1, ..., a_k)$ such that $f(a_1, ..., a_k) \equiv 0 \pmod{m}$ and $0 \leq a_i \leq m - 1$ for all $1 \leq i \leq k$. Let $i^2 = -1$ and define $\exp(i\theta) := \cos\theta + i\sin\theta$ for any real number θ. Then we have the following result.

Theorem 2.4.4. *We have*

$$N(f(x_1, ..., x_k) \equiv 0 \pmod{m}) = \frac{1}{m} \sum_{l=0}^{m-1} \sum_{x_1=0}^{m-1} ... \sum_{x_k=0}^{m-1} \exp\left(\frac{2\pi i l f(x_1, ..., x_k)}{m}\right).$$

Proof. For any $(x_1, ..., x_k)$ such that $f(x_1, ..., x_k) \not\equiv 0 \pmod{m}$, we have

$$\exp\left(\frac{2\pi i f(x_1, ..., x_k)}{m}\right) \neq 1$$

and thus

$$\frac{1}{m} \sum_{l=0}^{m-1} \exp\left(\frac{2\pi i l f(x_1, ..., x_k)}{m}\right) = \frac{1}{m} \sum_{l=0}^{m-1} \left(\exp\left(\frac{2\pi i f(x_1, ..., x_k)}{m}\right)\right)^l$$

$$=\frac{1}{m}\frac{\left(\exp\left(\frac{2\pi i f(x_1,...,x_k)}{m}\right)\right)^m-1}{\exp\left(\frac{2\pi i f(x_1,...,x_k)}{m}\right)-1}=\frac{1}{m}\frac{\exp\left(m\cdot\frac{2\pi i f(x_1,...,x_k)}{m}\right)-1}{\exp\left(\frac{2\pi i f(x_1,...,x_k)}{m}\right)-1}=0. \qquad (2.8)$$

For any $(x_1, ..., x_k)$ such that $f(x_1, ..., x_k) \equiv 0 \pmod m$, we have $\exp\left(\frac{2\pi i l f(x_1,...,x_k)}{m}\right) = 1$ for all $0 \le l \le m-1$ and so

$$\frac{1}{m}\sum_{l=0}^{m-1}\exp\left(\frac{2\pi i l f(x_1, ..., x_k)}{m}\right)=\frac{1}{m}\sum_{l=0}^{m-1}1=1.$$

Hence we obtain that

$$\frac{1}{m}\sum_{l=0}^{m-1}\sum_{f(x_1,...,x_k)\equiv 0 \pmod m}\exp\left(\frac{2\pi i l f(x_1, ..., x_k)}{m}\right)=N(f(x_1, ..., x_k)\equiv 0 \pmod m). \qquad (2.9)$$

Collaborating (2.8) and (2.9) gives us the desired result. This completes the proof of Theorem 2.4.4. □

As a special case of Theorem 2.4.4, we have the following result.

Theorem 2.4.5. *We have*

$$N(f(x)\equiv 0 \pmod m)=\frac{1}{m}\sum_{l=0}^{m-1}\sum_{x=0}^{m-1}\exp\left(\frac{2\pi i l f(x)}{m}\right).$$

§2.5. Systems of linear congruences with the same modulo

We will consider systems of more than one congruence that involve the same number of unknowns as congruences, where all congruences have the same modulus. Readers unfamiliar with linear algebra may wish to skip this section.

Systems of n linear congruences involving n unknowns will arise in our subsequent cryptographic studies. To study such systems when n is large, it is helpful to use the language of matrices. We will use some of the basic notions of matrix arithmetic, which are discussed in most linear algebra texts. Before proceeding, we need to define congruences of matrices.

Definition. Let A and B be $n \times k$ matrices with integer entries, with (i, j)-th entries a_{ij} and b_{ij}, respectively. We say that A is *congruent to B modulo m* if $a_{ij} \equiv b_{ij} \pmod m$ for all pairs (i, j) with $1 \le i \le n$ and $1 \le j \le k$. We write $A \equiv B \pmod m$ if A is congruent to B modulo m.

The matrix congruence $A \equiv B \pmod m$ provides a succinct way of expressing the nk congruences $a_{ij} \equiv b_{ij} \pmod m$ for $1 \le i \le n$ and $1 \le j \le k$.

The following proposition will be needed.

Proposition 2.5.1. *If A and B are $n \times k$ matrices with $A \equiv B \pmod m$, C is a $k \times p$ matrix and D is a $p \times n$ matrix, all with integer entries, then $AC \equiv BC \pmod m$ and $DA \equiv DB \pmod m$.*

Proof. Let the entries of A and B be a_{ij} and b_{ij}, respectively, for $1 \le i \le n$ and $1 \le j \le k$, and let the entries of C be c_{ij} for $1 \le i \le k$ and $1 \le j \le p$. The (i, j)-th entries of AC and BC are $\sum_{t=1}^{k} a_{it}c_{tj}$ and $\sum_{t=1}^{k} b_{it}c_{tj} \pmod m$. Consequently, $AC \equiv BC \pmod m$.

The proof of $DA \equiv DB \pmod m$ is similar and so we omit the details. □

Definition. If A and $\bar{A}$ are $n \times n$ matrices of integers and $\bar{A}A \equiv A\bar{A} \equiv I$ (mod m), where $I = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \dots & 1 \end{pmatrix}$ is the identity matrix of order n, then $\bar{A}$ is said to be an *inverse of A modulo m.*

If $\bar{A}$ is an inverse of A and $B \equiv \bar{A} \pmod{m}$, then B is also an inverse of A. This follows from Theorem 2.5.2, because $BA \equiv \bar{A}A \equiv I \pmod{m}$. Conversely, if B_1 and B_2 are both inverses of A, then $B_1 \equiv B_2 \pmod{m}$. To see this, using Theorem 2.5.2 and the congruence $B_1A \equiv B_2A \equiv I \pmod{m}$, we have $B_1AB_1 \equiv B_2AB_2 \pmod{m}$. Because $AB_1 \equiv I \pmod{m}$, we conclude that $B_1 \equiv B_2 \pmod{m}$. The following proposition gives an easy method for finding inverses for 2×2 matrices.

Example. Let $A = \begin{pmatrix} 3 & 4 \\ 2 & 5 \end{pmatrix}$. Because 2 is an inverse of $\det A = 7$ modulo 13, we have

$$\bar{A} \equiv 2\begin{pmatrix} 5 & -4 \\ -2 & 3 \end{pmatrix} \equiv \begin{pmatrix} 10 & -8 \\ -4 & 6 \end{pmatrix} \equiv \begin{pmatrix} 10 & 5 \\ 9 & 6 \end{pmatrix} \pmod{13}.$$

To provide a formula for an inverse of an $n \times n$ matrix, where n is a positive integer greater than 2, we need a result from linear algebra. It involves the notion of the adjoint of a matrix, which is defined as follows.

Definition. The *adjoint* of an $n \times n$ matrix A is the $n \times n$ matrix with (i, j)th entry C_{ij}, where C_{ij} is $(-1)^{i+j}$ times the determinant of the matrix obtained by deleting the jth row and ith column from A. The adjoint of A is denoted by $\mathrm{adj}(A)$, or simply $\mathrm{adj}A$.

Proposition 2.5.2. *If A is an $n \times n$ matrix, then $A(\mathrm{adj}A) = (\det A)I$, where $\mathrm{adj}A$ is the adjoint of A.*

Using this theorem, the following theorem follows readily.

Theorem 2.5.1. *Let A be an $n \times n$ matrix with integer entries and m be a positive integer. Then A is invertible modulo m if and only if $(\det A, m) = 1$. Moreover, if A is invertible modulo m, then the matrix $\bar{A} = \bar{\Delta}(\mathrm{adj}A)$ is an inverse of A modulo m with $\bar{\Delta}$ being an inverse of $\Delta = \det A$ modulo m.*

Proof. First, we let $(\det A, m) = 1$. Then we know that $\det A \neq 0$. Hence, by Proposition 2.5.3, we have

$$A \cdot \mathrm{adj}A = (\det A)I = \Delta I.$$

Since $(\det A, m) = 1$, there is an inverse $\bar{\Delta}$ of $\Delta = \det A$ modulo m. Hence,

$$A(\bar{\Delta}\ \mathrm{adj}A) \equiv A \cdot (\mathrm{adj}A)\bar{\Delta} \equiv \Delta\bar{\Delta}I \equiv I \pmod{m}.$$

and

$$\bar{\Delta}(\ \mathrm{adj}A)A \equiv \bar{\Delta}(\mathrm{adj}A \cdot A) \equiv \bar{\Delta}\Delta I \equiv I \pmod{m}.$$

This shows that $\bar{A} = \bar{\Delta} \cdot (\mathrm{adj}A)$ is an inverse of A modulo m.

Conversely, we let A be invertible modulo m. Then there is a matrix B of order n such that $AB \equiv I_n \pmod{m}$. So $AB = I_n + mC$ for some matrix C of order n. It follows that

$$\det(A)\det(B) = \det(AB) = \det(I_n + mC) = 1 + mt$$

for some integer t. Hence $\det(A)\det(B) \equiv 1 \pmod{m}$. This implies that $(\det A, m) = 1$ as required. The proof of Theorem 2.5.1 is complete. □

Proposition 2.5.3. *Let* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *be a matrix of integers such that* $\Delta = \det A = ad - bc$ *is relatively prime to the positive integer* m. *Then the matrix*

$$\bar{A} = \bar{\Delta} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

where $\bar{\Delta}$ *is the inverse of* Δ *modulo* m, *is an inverse of* A *modulo* m.

Example. Let $A = \begin{pmatrix} 2 & 5 & 6 \\ 2 & 0 & 1 \\ 1 & 2 & 3 \end{pmatrix}$. Then $\det A = -5$. Furthermore, we have $(\det A, 7) = 1$ and we see that 4 is an inverse of $\det A = -5 \pmod 7$. Consequently, we find that

$$\bar{A} = 4(\text{adj}A) = 4 \begin{pmatrix} -2 & -3 & 5 \\ -5 & 0 & 10 \\ 4 & 1 & -10 \end{pmatrix}$$

$$= \begin{pmatrix} -8 & -12 & 20 \\ -20 & 0 & 40 \\ 16 & 4 & -40 \end{pmatrix} \equiv \begin{pmatrix} 6 & 2 & 6 \\ 1 & 0 & 5 \\ 2 & 4 & 2 \end{pmatrix} \pmod 7.$$

Now let us consider the system of congruences:

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \equiv b_1 \pmod m$$

$$\cdots\cdots$$

$$a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n \equiv b_n \pmod m.$$

Clearly, this system of n congruences is equivalent to the matrix congruence $AX \equiv B \pmod m$, where

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & & \ddots & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \text{and} \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

Example. We can rewrite the system

$$13x + 4y \equiv 15 \pmod{23}$$

$$22x - 5y \equiv 17 \pmod{23}$$

as

$$\begin{pmatrix} 13 & 4 \\ 22 & -5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \equiv \begin{pmatrix} 15 \\ 17 \end{pmatrix} \pmod{23}.$$

We now develop a method for solving congruences of the form $AX \equiv B \pmod m$. This method is based on finding a matrix $\bar{A}$ such that $\bar{A}A \equiv I \pmod m$, where I is the identity matrix. We can use an inverse of A modulo m to solve the system

$$AX \equiv B \pmod m,$$

where $(\det A, m) = 1$. By Theorem 2.5.2, when we multiply both sides of this congruence by an inverse $\bar{A}$ of A, we obtain

$$\begin{aligned} \bar{A}(AX) &\equiv \bar{A}B \pmod m \\ (\bar{A}A)X &\equiv \bar{A}B \pmod m \\ X &\equiv \bar{A}B \pmod m. \end{aligned}$$

Hence we find the solution X by forming $\bar{A}B \pmod{m}$. Therefore we have the following result.

Proposition 2.5.4. *Let A be an $n \times n$ matrix with integer entries and m be a positive integer such that $(\det A, m) = 1$. Then the system*

$$AX \equiv B \pmod{m}$$

has a unique solution modulo m given by

$$X \equiv \bar{A}B \pmod{m},$$

where $\bar{A}$ is an inverse of A modulo m.

Corollary 2.5.6. *Let a, b, c, d, e, f and m be integers, $m > 0$, such that $(\Delta, m) = 1$, where $\Delta = ad - bc$. Then the system of congruences*

$$ax + by \equiv e \pmod{m},$$

$$cx + dy \equiv f \pmod{m}$$

has a unique solution modulo m given by:

$$x \equiv \bar{\Delta}(de - bf) \pmod{m},$$

$$y \equiv \bar{\Delta}(af - ce) \pmod{m},$$

where $\bar{\Delta}$ is an inverse of Δ modulo m.

Proof. Let $AX \equiv B$, where $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $X = \begin{pmatrix} x \\ y \end{pmatrix}$ and $B = \begin{pmatrix} e \\ f \end{pmatrix}$. If $\Delta = \det A = ad - bc$ is relatively prime to m, then

$$\begin{pmatrix} x \\ y \end{pmatrix} = X \equiv \bar{A}B \equiv \bar{\Delta}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}\begin{pmatrix} e \\ f \end{pmatrix} = \bar{\Delta}\begin{pmatrix} de - bf \\ af - ce \end{pmatrix} \pmod{m}.$$

This demonstrates that (x, y) is a solution if and only if

$$x \equiv \bar{\Delta}(de - bf) \pmod{m}, \qquad y \equiv \bar{\Delta}(af - ce) \pmod{m}.$$

Next, we give an example of the solution of a system of three congruences in three unknowns using matrices.

Example. We consider the system of three congruences

$$\begin{aligned} 2x_1 + 5x_2 + 6x_3 &\equiv 3 \pmod{7} \\ 2x_1 + x_3 &\equiv 4 \pmod{7} \\ x_1 + 2x_2 + 3x_3 &\equiv 1 \pmod{7} \end{aligned}$$

This is equivalent to the matrix congruence

$$\begin{pmatrix} 2 & 5 & 6 \\ 2 & 0 & 1 \\ 1 & 2 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \equiv \begin{pmatrix} 3 \\ 4 \\ 1 \end{pmatrix} \pmod{7}.$$

We have previously shown that the matrix $\begin{pmatrix} 6 & 2 & 6 \\ 1 & 0 & 5 \\ 2 & 4 & 2 \end{pmatrix}$ is an inverse of $\begin{pmatrix} 2 & 5 & 6 \\ 2 & 0 & 1 \\ 1 & 2 & 3 \end{pmatrix}$ (mod 7). Hence we have

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6 & 2 & 6 \\ 1 & 0 & 5 \\ 2 & 4 & 2 \end{pmatrix}\begin{pmatrix} 3 \\ 4 \\ 1 \end{pmatrix} = \begin{pmatrix} 32 \\ 8 \\ 24 \end{pmatrix} \equiv \begin{pmatrix} 4 \\ 1 \\ 3 \end{pmatrix} \pmod{7}.$$

Before leaving this subject, we should mention that many methods for solving systems of linear equations may be adapted to solve systems of congruences. For instance, Gaussian elimination may be adapted to solve systems of congruences, where division is always replaced by multiplication by inverses modulo m. In closing this section, we list a result for solving systems of congruences analogous to Cramer's rule.

Theorem 2.5.2. (Analogue of Cramer's rule) *Let A be an $n \times n$ matrix with integer entries and m be a positive integer such that $(\text{det}A, m) = 1$. Then the system*

$$AX \equiv B \pmod m$$

has a unique solution modulo m given by

$$X \equiv (\text{det}A)^{-1}(\text{adj}A)B \pmod m,$$

where $(\text{det}A)^{-1}$ denotes an inverse of $\text{det}A$ modulo m.

§2.6. Some special congruences

Euler's theorem and its corollary, Fermat's theorem, are fundamental results in number theory with many applications in mathematics and computer science. In the following sections we shall see how the Euler and Fermat theorems can be used to determine whether an integer is prime or composite and how they are applied in cryptography.

Theorem 2.6.1. (Euler) *Let m be a positive integer and let a be an integer relatively prime to m. Then $a^{\varphi(m)} \equiv 1 \pmod m$.*

Proof. Let $\{r_1, \ldots, r_{\varphi(m)}\}$ be a reduced set of residues modulo m. Since $(a, m) = 1$ we have $(ar_i, m) = 1$ for $i = 1, \ldots, \varphi(m)$. Consequently, for every $i \in \{1, \ldots, \varphi(m)\}$ there exists $\sigma(i) \in \{1, \ldots, \varphi(m)\}$ such that $ar_i \equiv r_{\sigma(i)} \pmod m$. Moreover, $ar_i \equiv ar_j \pmod m$ if and only if $i = j$ and so σ is a permutation of the set $\{1, \ldots, \varphi(m)\}$ and $\{ar_1, \ldots, ar_{\varphi(m)}\}$ is also a reduced set of residues modulo m. It follows that

$$\begin{aligned} a^{\varphi(m)} r_1 r_2 \cdots r_{\varphi(m)} &\equiv (ar_1)(ar_2) \cdots (ar\varphi(m)) \pmod m \\ &\equiv r_{\sigma(1)} r_{\sigma(2)} \cdots r_{\sigma(\varphi(m))} \pmod m \\ &\equiv r_1 r_2 \cdots r_{\varphi(m)} \pmod m. \end{aligned}$$

Dividing by $r_1 r_2 \cdots r_{\varphi(m)}$, we obtain the desired result $a^{\varphi(m)} \equiv 1 \pmod m$.

This completes the proof of Theorem 2.6.1. □

The following corollary is sometimes called *Fermat's little Theorem.*

Theorem 2.6.2. (Fermat) *Let p be a prime number. If the integer a is not divisible by p, then $a^{p-1} \equiv 1 \pmod p$. Moreover, for every integer a, one has $a^p \equiv a \pmod p$.*

Proof. If p is prime and does not divide a, then $(a, p) = 1, \varphi(p) = p - 1$, and

$$a^{p-1} = a^{\varphi(p)} \equiv 1 \pmod p.$$

by Euler's theorem. Multiplying this congruence by a, we obtain

$$a^p \equiv a \pmod p$$

If p divides a, then this congruence also holds for a. □

Let m be a positive integer and let a be an integer that is relatively prime to m. By Euler's theorem, $a^{\varphi(m)} \equiv 1 \pmod{m}$. The order of a with respect to the modulus m is the smallest positive integer d such that $a^d \equiv 1 \pmod{m}$. Then $1 \leq d \leq \varphi(m)$. We denote the order of a modulo m by $\nu_m(a)$. We shall prove that $\nu_m(a)$ *divides* $\varphi(m)$ *for every integer* a *relatively prime to* m.

Proposition 2.6.1. *Let* m *be a positive integer and* a *an integer relatively prime to* m*. If* d *is the order of* a *modulo* m*, then* $a^k \equiv a^l \pmod{m}$ *if and only if* $k \equiv l \pmod{d}$*. In particular,* $a^n \equiv 1 \pmod{m}$ *if and only if* d *divides* n *and so* d *divides* $\varphi(m)$*.*

Proof. Since a has order d modulo m, we have $a^d \equiv 1 \pmod{m}$. If $k \equiv l \pmod{d}$, then $k = l + dq$ and so

$$a^k = a^{l+dq} = a^l(a^d)^q \equiv a^l \pmod{m}.$$

Conversely, suppose that $a^k \equiv a^l \pmod{m}$. By the division algorithm, there exist integers q and r such that

$$k - l = dq + r \text{ and } 0 \leq r \leq d - 1$$

Then

$$a^k = a^{l+dq+r} = a^l(a^d)^q a^r \equiv a^k a^r \pmod{m}.$$

Since $(a^k, m) = 1$, we can divide this congruence by a^k and obtain $a^r \equiv 1 \pmod{m}$. Since $0 \leq r \leq d - 1$ and d is the order of a modulo m, it follows that $r = 0$.and so $k \equiv l \pmod{d}$.

If $a^n \equiv 1 \equiv a^0 \pmod{m}$, then d divides n. In particular, d divides $\varphi(m)$, since $a^{\varphi(m)} \equiv 1 \pmod{m}$ by Euler's theorem. □

For example, let $m = 15$ and $a = 7$. Since $\varphi(15) = 8$. Euler's theorem tells us that $7^8 \equiv 1 \pmod{15}$. Moreover, the order of 7 with respect to 15 is a divisor of 8. We can compute the order as follows

$$\begin{aligned} 7^1 \equiv&\quad 7 \pmod{15}, \\ 7^2 \equiv&\quad 49 \equiv 4 \pmod{15}, \\ 7^4 \equiv&\quad 91 \equiv 1 \pmod{15}, \end{aligned}$$

and so the order of 7 is 4.

Now let's give the well-known Wolstenholme theorem as follows. In order to do so, we need to show a series of results.

Proposition 2.6.2. (Lagrange) *Let* p *be a prime and let* $f(x) = a_0 + a_1x + ... + a_nx^n \in \mathbf{Z}[x]$ *satisfy* $a_n \not\equiv 0 \pmod{p}$*. Then the polynomial congruence* $f(x) \equiv 0 \pmod{p}$ *has at most* n *solutions.*

Proof. Use the induction on n, the degree of f. If $n = 1$, the result is clear true.

Assume now that the theorem is true for all the polynomials of degree $n - 1$. Assume also that the congruence $f(x) \equiv 0 \pmod{p}$ has $n+1$ incongruent solutions modulo p, say $x_0, x_1, ..., x_n$, where $f(x_k) \equiv 0 \pmod{p}$ for each $k = 0, 1, ..., n$. We shall obtain a contradiction. We have

$$f(x) - f(x_0) = \sum_{r=1}^{n} a_i(x^i - x_0^i) = (x - x_0)g(x)$$

where $g(x) \in \mathbf{Z}[x]$ is of degree $n-1$ with leading coefficient c_n. Since $f(x_i) - f(x_0) \equiv 0 \pmod{p}$, we have $f(x) - f(x_0) = (x_i - x_0)g(x) \equiv 0 \pmod{p}$. But $x_i - x_0 \not\equiv 0 \pmod{p}$ if $i \neq 0$. So we have $g(x_i) \equiv 0 \pmod{p}$ for each $i \neq 0$. It follows that the congruence $g(x) \equiv 0 \pmod p$ has n incongruent solutions modulo p, contradicting our induction hypothesis. The proof is complete. □

Proposition 2.6.3. *Let p be a prime and $f(x) = a_0 + a_1x + ... + a_nx^n \in \mathbf{Z}[x]$ be a polynomial with integer coefficients. If the congruence has more than n solutions, then every coefficient of f is divisible by p.*

Proof. Assume that there is some coefficient not divisible by p. Let a_l be the one with largest index. Then $l \leq n$ and the congruence

$$a_0 + a_1x + ... + a_lx^l \equiv 0 \pmod p$$

has more than l solutions. On the other hand, since $a_l \not\equiv 0 \pmod p$, then by Lagrange's theorem (Proposition 2.6.2), the congruence has at most l solutions. This is a contradiction. So this Proposition is proved. □

Proposition 2.6.4. *Let p be any given prime. Then all the coefficients of the polynomial $f(x) := (x-1)(x-2)...(x-p+1) - x^{p-1} + 1$ are divisible by p.*

Proof. Let $g(x) = f(x) = (x-1)(x-2)...(x-p+1)$ and $h(x) = x^{p-1} - 1$. Then the roots of $g(x)$ are the numbers $1, ..., p-1$. So these numbers satisfy $g(x) \equiv 0 \pmod p$. But by the Fermat Little Theorem (Theorem 2.6.2) they also satisfy the congruence $h(x) \equiv 0 \pmod p$. Thus the difference $f(x) = g(x) - h(x)$ has degree $p-2$ but the congruence $f(x) \equiv 0 \pmod p$ has $p-1$ solutions, $1, 2, ..., p-1$. It then follows from Proposition 2.6.3 that each coefficient of $f(x)$ is divisible by p. □

Theorem 2.6.3. (Wolstenholme Theorem) *Let $p \geq 5$ be a prime. Then*

$$\sum_{k=1}^{p-1} \frac{(p-1)!}{k} \equiv 0 \pmod{p^2}.$$

Equivalently, we have

$$\binom{2p-1}{p-1} \equiv 1 \pmod{p^3}.$$

Or equivalently we have

$$\sum_{k=1}^{p-1} k^* \equiv 0 \pmod{p^2},$$

where for $1 \leq k \leq p-1$, k^* is an integer such that $kk^* \equiv 1 \pmod{p^2}$.

Proof. The proof of the equivalence of three statements is left to the reader. We here just prove the first statement.

The sum in question is equal to the coefficient of $-x$ in the polynomial

$$g(x) = f(x) = (x-1)(x-2)...(x-p+1).$$

We can expand and get that

$$g(x) = x^{p-1} - S_1x^{p-2} + S_2x^{p-3} - ... + S_{p-3}x^2 - S_{p-2}x + (p-1)!,$$

where the coefficient S_i is the ith elementary symmetric function of the roots $1, 2, ..., p-1$, that is, the sum of the products of the numbers $1, 2, ..., p-1$, taken i at a time. $1, 2, ..., p-1$. Proposition 2.6.4 shows that each of the numbers $S_1, ..., S_{p-2}$ is divisible by p. In the following we show that S_{p-2} is divisible by p^2.

Evidently we have $g(p) = (p-1)!$. Thus

$$(p-1)! = p^{p-1} - S_1p^{p-2} + S_2p^{p-3} - ... + S_{p-3}p^2 - S_{p-2}p + (p-1)!.$$

Canceling $(p-1)!$ and reducing the equation mod p^3 we find, since $p \geq 5$, we obtain

$$pS_{p-2} \equiv 0 \pmod{p^3},$$

and hence $S_{p-2} \equiv 0 \pmod{p^2}$, as desired. Therefore Wolstenholme Theorem is proved. □

In concluding this section, we give another famous congruence which has many applications in p-adic analysis and coding theory. We begin with a special case of Locus congruence.

Lemma 2.6.1. *Let a and b be positive integers with $a \geq b$. Then for any integer r and s with $0 \leq r, s \leq p-1$, one has*

$$\left(\begin{array}{c} ap+r \\ bp+s \end{array} \right) \equiv \left(\begin{array}{c} a \\ b \end{array} \right) \left(\begin{array}{c} r \\ s \end{array} \right) \pmod{p}.$$

Proof. First we show that for $1 \leq i \leq p-1$, $p | \left(\begin{array}{c} p \\ i \end{array} \right)$ (exercise 26 in Chapter 1). In fact we have $i! \left(\begin{array}{c} p \\ i \end{array} \right) = p \cdot \frac{(p-1)!}{(p-i)!}$. So $p | i! \left(\begin{array}{c} p \\ i \end{array} \right)$. But $(p, i!) = 1$ since $1 \leq i \leq p-1$. Thus we have $p | \left(\begin{array}{c} p \\ i \end{array} \right)$ as desired. Hence we derive that $(1+x)^p \equiv 1 + x^p \pmod{p}$. It then follows that for any integers $a \geq 1$ and $0 \leq r \leq p-1$, we have

$$(1+x)^{ap+r} \equiv (1+x^p)^a (1+x)^r \pmod{p}.$$

Comparing the coefficients of x^{bp+s} (where $0 \leq s \leq p-1$) on both sides yields that for any integers $b \geq 1$ and $0 \leq s \leq p-1$, we have

$$\left(\begin{array}{c} ap+r \\ bp+s \end{array} \right) \equiv \left(\begin{array}{c} a \\ b \end{array} \right) \left(\begin{array}{c} r \\ s \end{array} \right) \pmod{p}$$

as required. This completes the proof of Lemma 2.6.1. □

Theorem 2.6.4. (Lucas Congruence) *Let p be a prime. Let $m = m_0 + m_1p + ... + m_lp^l$ and $n = n_0 + n_1p + ... + n_lp^l$ be the p-adic representations of positive integers m and n, i.e. m_i and n_i are nonnegative integers less than p for $1 \leq i \leq l$. If $m \geq n$, then we have*

$$\left(\begin{array}{c} m \\ n \end{array} \right) \equiv \prod_{i=0}^{l} \left(\begin{array}{c} m_i \\ n_i \end{array} \right) \pmod{p}.$$

Proof. First of all, applying Lemma 2.6.1 gives us that

$$\left(\begin{array}{c} m_0 + m_1p \\ n_0 + n_1p \end{array} \right) \equiv \left(\begin{array}{c} m_0 \\ n_0 \end{array} \right) \left(\begin{array}{c} m_1 \\ n_1 \end{array} \right) \pmod{p}.$$

So the theorem for the case $l = 1$ is proved.

Assume that the theorem holds for the case $l \geq 1$. Now consider the case $l+1$. Let $m = m_0 + m_1p + ... + m_lp^l + m_{l+1}p^{l+1}$ and $n = n_0 + n_1p + ... + n_lp^l + n_{l+1}p^{l+1}$. By the induction hypothesis we have

$$\left(\begin{array}{c} m_1 + m_2p + ... + m_{l+1}p^l \\ n_1 + n_2p + ... + n_{l+1}p^l \end{array} \right) \equiv \prod_{i=1}^{l+1} \left(\begin{array}{c} m_i \\ n_i \end{array} \right) \pmod{p}.$$

But by () we get

$$\binom{m}{n} \equiv \binom{m_0}{n_0}\binom{m_1+m_2p+...+m_{l+1}p^l}{n_1+n_2p+...+n_{l+1}p^l} \pmod{p}.$$

It then follows from () and () that

$$\binom{m}{n} \equiv \prod_{i=0}^{l+1}\binom{m_i}{n_i} \pmod{p}$$

as required. Thus the theorem is true for the case $l+1$. The proof of Theorem 2.6.4 is complete. □

§2.7. Pseudoprimes and Carmichael numbers

Suppose we are given an odd integer $n \geq 3$ and we want to determine whether n is prime or composite. If n is ”small”, we can simply divide n by all odd integers d such that $3 \leq d \leq \sqrt{n}$. If some d divides n, then n is composite; otherwise, n is prime. If n is ”big”, however, this method is time-consuming and impractical. We need to find other primality tests.

Fermat's Little theorem can be applied to this problem. By Fermat's Little theorem, if n is an odd prime, then $2^{n-1} \equiv 1 \pmod{n}$. Therefore if n is odd and $2^{n-1} \not\equiv 1 \pmod{n}$, then n must be composite. In general, we can choose any integer b that is relatively prime to n. By Fermat's theorem, if n is prime, then $b^{n-1} \equiv 1 \pmod{n}$. It follows that

Theorem 2.7.1. *Let* $(b, n) = 1$*. If* $b^{n-1} \not\equiv 1 \pmod{n}$*, then* n *must be composite.*

Thus for every base b, Fermat's theorem gives a *primality test*, that is, a necessary condition for an integer to be prime.

Example. Suppose that we want to know whether $n = 851$ is prime or composite. We shall compute $2^{850} \pmod{851}$. An efficient method is to use the 2-adic representation of 850: $850 = 2 + 2^4 + 2^6 + 2^8 + 2^9$. Since $2^{2^n} = \left(2^{2^{n-1}}\right)^2$, we have

$$\begin{aligned}
2^2 &\equiv 4 \pmod{851},\\
2^{2^2} &\equiv 16 \pmod{851},\\
2^{2^3} &\equiv 256 \pmod{851},\\
2^{2^4} &\equiv 9 \pmod{851},\\
2^{2^5} &\equiv 81 \pmod{851},\\
2^{2^6} &\equiv 604 \pmod{851},\\
2^{2^7} &\equiv 588 \pmod{851},\\
2^{2^8} &\equiv 238 \pmod{851},\\
2^{2^9} &\equiv 478 \pmod{851}.
\end{aligned}$$

Then

$$\begin{aligned}
2^{850} &\equiv 2^2 2^{2^4} 2^{2^6} 2^{2^8} 2^{2^9} \pmod{851}\\
&\equiv 4 \cdot 9 \cdot 604 \cdot 238 \cdot 478 \pmod{851}\\
&\equiv 169 \not\equiv 1 \pmod{851}
\end{aligned}$$

and so 851 is composite. In fact we have

$$851 = 900 - 49 = 30^2 - 7^2 = (30-7)(30+7) = 23 \cdot 37.$$

This test can prove that an integer is composite, but it cannot prove that an integer is prime.

Example. Consider the composite number $n = 341 = 11 \cdot 31$. Choosing base $b = 2$, we have $2^{10} \equiv 1 \pmod{11}$, and so $2^{340} \equiv (2^{10})^{34} \equiv 1 \pmod{11}$. Similarly, $2^5 \equiv 1 \pmod{31}$. and so $2^{340} \equiv (2^5)^{68} \equiv 1 \pmod{31}$. Since $2^{340} - 1$is divisible by both 11 and 31, it is divisible by their product, that is, $2^{340} \equiv 1 \pmod{341}$.

Definition. A composite number n is called a *pseudoprime to the base* b if $(b, n) = 1$ and $b^{n-1} \equiv 1 \pmod{n}$.

Thus 341 is a pseudoprime to base 2.

We can show that 341 is composite by choosing the base $b = 7$. Since $7^3 = 343 \equiv 2 \pmod{341}$ and $2^{10} = 1024 \equiv 1 \pmod{341}$. It follows that

$$\begin{aligned} 7^{340} &= 7(7^3)^{113} \\ &\equiv 7 \cdot 2^{113} \pmod{341} \\ &\equiv 7 \cdot 2^3 \cdot (2^{10})^{11} \pmod{341} \\ &\equiv 56 \pmod{341} \\ &\not\equiv 1 \pmod{341}. \end{aligned}$$

Question. Can every composite number be *proved* composite by some primality test based on Fermat's theorem?

It is a surprising fact that the answer is "no". There exist composite numbers n that cannot be proved composite by any congruence of the form $b^{n-1} \pmod{n}$ with $(b, n) = 1$.

Example. $561 = 3 \cdot 11 \cdot 17$ is composite. Let b be an arbitrary integer relatively prime to 561. Then $b^2 \equiv 1 \pmod{3}$ and so $b^{560} = (b^2)^{280} \equiv 1 \pmod{3}$. Similarly, $b^{10} \equiv 1 \pmod{11}$, and so $b^{560} = (b^{10})^{56} \equiv 1 \pmod{11}$. Finally, $b^{16} \equiv 1 \pmod{17}$, and so $b^{560} = (b^{16})^{35} \equiv 1 \pmod{17}$. Since $b^{560} - 1$ is divisible by 3,11 and 17, it is also divisible by their product, hence $b^{560} \equiv 1 \pmod{561}$. This proves that 561 is a pseudoprime to base b for every b such that $(b, n) = 1$.

A *Carmichael number* (also called an *absolute pseudoprime*) is a positive integer n such that n is composite but $b^{n-1} \equiv 1 \pmod{n}$ for every integer b relatively prime to n. Thus 561 is a Carmichael number. But 341 is NOT a Carmichael number, just a pseudoprime.

Theorem 2.7.2. *A composite number n is a Carmichael number if and only if $n = q_1 ... q_k$, where $k \geq 3$ and $q_1, ..., q_k$ are distinct primes satisfying $(q_j - 1)|(n - 1)$ for all j.*

Proof. First we show the sufficiency. Let $n = q_1 ... q_k$, where $k \geq 3$ and $q_1, ..., q_k$ are distinct primes satisfying $(q_j - 1)|(n - 1)$ for all j. Then for any integer b coprime to n and for any $1 \leq j \leq k$, Fermat's little theorem tells us that $b^{q_j - 1} \equiv 1 \pmod{q_j}$ since $(b, n) = 1$ implying that $(b, q_j) = 1$. But $(q_j - 1)|(n - 1)$. Hence we deduce that $b^{n-1} \equiv 1 \pmod{q_j}$. Since $q_1, ..., q_k$ are distinct primes, it follows that $b^{n-1} \equiv 1 \pmod{n}$. So n is a Carmichael number as desired. The sufficiency part is proved.

Now let's show the necessity part. Let n be a composite number. Then .

□

In 1912, Carmichael conjectured that there are infinitely many Carmichael numbers. In 1994, following some ideas of Mingzhi Zhang at Sichuan University who searched large Carmichael numbers, three mathematicians named Alford, Granville and Pomerance confirmed Carmichael conjecture by showing the following interesting result.

Theorem 2.7.3. *There are infinitely many Carmichael numbers.*

We here don't give the detail of its proof since the proof is beyond the scope of this book. See Annals of Math. 139 (1994), 703-722) for the details of the proof.

Exercises for Chapter 2

(1). Show that every integer is congruent modulo 11 to a unique integer r such that $-5 \le r \le 5$.

(2). Show that $a^4 \equiv 1 \pmod 5$ if $a \in \mathbf{Z}$ and $5 \not| a$.

(3). Show that if $x_1, ..., x_m$ is a sequence of m not necessarily distinct integers, then there exist integers $1 \le k \le l \le m$ such that $\sum_{i=k}^{l} x_i \equiv 0 \pmod m$.

(4). Show that any positive integer n such that $n \equiv 3 \pmod 4$ cannot be written as the sum of two squares.

(5). Let p be prime, $m \ge 1$ and $0 \le k \le p-1$. Show that

$$\left(\begin{array}{c} mp+k \\ p \end{array} \right) \equiv m \pmod p.$$

(6). Find all solutions of the following congruences:

(i). $5x \equiv 9 \pmod{11}$; (ii). $18x \equiv 3 \pmod{51}$; (iii). $28x \equiv 45 \pmod{70}$.

(7). Let m be a composite number. Show that each of the following is true:

(i). If $m \ne 4$, then $(m-1)! \equiv 0 \pmod m$.

(ii). One has $m \not| ((m-1)!+1)$. This implies that the converse of Wilson's theorem is true.

(8). Show that if $p \ge 5$ is an odd prime, then $6(p-4)! \equiv 1 \pmod p$.

(9). Let m and a be integers such that $m \ge 1$ and $(a, m) = 1$. Show that if $\{r_1, ..., r_{\varphi(m)}\}$ is a reduced set of residues modulo m, then $\{ar_1, ..., ar_{\varphi(m)}\}$ is also a reduced set of residues modulo m.

(10). We say that an integer a is nilpotent modulo m if there exists a positive integer k such that $a^k \equiv 0 \pmod m$. Show that a is nilpotent modulo m if and only if $a \equiv 0 \pmod{\mathrm{rad}(m)}$.

(11). For $n \ge 1$, let $h_n = \sum_{k=1}^{n} \frac{1}{k} = \frac{u_n}{v_n}$, where u_n and v_n are positive integers. Show that if p is an odd prime, then the numerator u_{p-1} of h_{p-1} is divisible by p.

(12). Compute $\varphi(2006)$ and $\varphi(6993)$.

(13). Show that $\varphi(m)$ is even for all $m \ge 3$.

(14). Show that $\varphi(m^k) = m^{k-1}\varphi(m)$ for all positive integers m and k.

(15). Show that $\varphi(m)|\varphi(n)$ if m divides n.

(16). Show that $\varphi(m) = \varphi(2m)$ if and only if m is odd.

(17). Show that $\varphi(m) = m-1$ if and only if m is prime.

(18). Find all positive integers n such that $\varphi(5n) = 5\varphi(n)$.

(19). Find all solutions of the system of congruences:

$$x \equiv 5 \pmod{12}, \ x \equiv 8 \pmod{10}.$$

(20). Find all solutions of the congruences:

(i). $6x^3 + 27x^2 + 17x + 24 \equiv 0 \pmod{30}$;

(ii). $5x^3 - 93 \equiv 0 \pmod{231}$.

(21). Find all solutions of the system of congruences:

$$x \equiv 23 \pmod{36}, \ x \equiv 27 \pmod{40}, \ x \equiv 17 \pmod{75}.$$

(22). Find the solutions of the following system of linear congruences:

$$2x + 3y \equiv 5 \pmod 7, \ x + 5y \equiv 6 \pmod 7.$$

(23). Find the matrix C such that

$$C \equiv \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 2 & 1 \end{pmatrix} \pmod 5$$

and all entries of C are nonnegative integers less than 5.

(24). How many incongruent solutions does the following system of congruences have?

$$2x + y + z \equiv 1 \pmod 5, \ x + 2y + z \equiv 1 \pmod 5, \ x + y + 2z \equiv 1 \pmod 5.$$

(25). Find the reminder when 7^{51} is divided by 144.

(26). Show that if $n \geq 2$, then $2^n - 1$ is not divisible by n.

(27). Show that if m and n are relatively prime positive integers, then

$$m^{\varphi(n)} + n^{\varphi(m)} \equiv 1 \pmod{mn}.$$

(28). Let p be an odd prime. By Fermat Little Theorem, if $(a, p) = 1$, then $f_p(a) := \frac{a^{p-1}-1}{p} \in \mathbf{Z}$. Show that if $(ab, p) = 1$, then $f_p(ab) \equiv f_p(a) + f_p(b) \pmod p$.

(29). Show that 1729 is a pseudoprime to bases 2, 3 and 5.

(30). Show that 6601 is a Carmichael number.

(31). Let p be a prime. Show that for any integer $1 \leq k \leq p - 1$, we have

$$\begin{pmatrix} p \\ k \end{pmatrix} \equiv 0 \pmod p.$$

(32). Find the solutions of the following systems of linear congruences:

1). $x + 2y - 23 \equiv 0 \pmod{209}$, $4x - 7y + 88 \equiv 0 \pmod{209}$;

2). $x + 4y - 29 \equiv 0 \pmod{143}$, $2x - 9y + 84 \equiv 0 \pmod{143}$.

(33). Show that if $m > 2$ is an integer and $\{a_1, ..., a_{\varphi(m)}\}$ is any reduced set of residues modulo m, then $\sum_{i=1}^{\varphi(m)} a_i \equiv 0 \pmod m$.

(34). Show that for any integers x and y, $3x + 2y \equiv 0 \pmod{17}$ if and only if $10x + y \equiv 0 \pmod{17}$.

(35). Let a and b be integers and p an odd prime. Show that if $a^p + b^p \equiv 0 \pmod p$, then $a^p + b^p \equiv 0 \pmod{p^2}$.

(36). Show that $(m - 1)! \equiv 0 \pmod m$ for any composite m.

(37). Show that $n^9 - n^3 \equiv 0 \pmod{504}$ for any integer m.

(38). Show that $61! + 1 \equiv 0 \pmod{71}$.

(39). Let m be any integer. Show that the congruence

$$6xy - 2x - 3y + 1 \equiv 0 \pmod m$$

has integer solution, but the Diophantine equation $6xy - 2x - 3y + 1 = 0$ has no integer solution.

(40). Let p be an odd prime and $l = \frac{p-1}{2}$. Show that $(l!)^2 + (-1)^l \equiv 0 \pmod p$.

(41). Let $p > 3$ be a prime and $P = \{1, ..., p-1\}$. For $k \in T$, define $t_k \in T$ to be such that $kt_k \equiv 1 \pmod p$. Let $s_k := \frac{kt_k - 1}{p}$. Show that $\sum_{k=1}^{p-1} ks_k \equiv \frac{p-1}{2} \pmod p$.

(42). Let p be a prime and $m \geq 1$ an integer. Show that each of the following is true:

(i). $\begin{pmatrix} m \\ p \end{pmatrix} \equiv \left[\frac{m}{p}\right] \pmod{p^2}$;

(ii). If $l > 0$ is an integer and $p^l \big| \left[\frac{m}{p}\right]$, then $p^l \Big| \begin{pmatrix} p \\ m \end{pmatrix}$.

(43). Let $f(x)$ be a cubic polynomial with integer coefficients such that its constant is not congruent to 0 modulo 5. Show that if there is an integer m such that $f(m) \equiv 0 \pmod 5$, then there is an integer n such that $n^3 f(\frac{1}{n}) \equiv 0 \pmod 5$.

(44). Let $f(x) \in \mathbf{Z}[x]$ and $m > 0$, s and t are integers satisfying $s \equiv t \pmod m$. Show that $f(s) \equiv f(t) \pmod m$.

(45). Show that for any integer x, $\frac{7}{15}x + \frac{1}{3}x^3 + \frac{1}{5}x^5$ is an integer.

(46). Find the smallest positive integer x, such that $\frac{1}{2}x$ is a square, $\frac{1}{3}x$ is a cube and $\frac{1}{5}x$ is the 5-th power of an integer.

(47). Let p be a odd prime. Show that each of the following is true:

(i). $1^2 \cdot 3^2 \cdot ... \cdot (p-2)^2 \equiv (-1)^{\frac{p+1}{2}} \pmod p$;

(ii). $2^2 \cdot 4^2 \cdot ... \cdot (p-1)^2 \equiv (-1)^{\frac{p+1}{2}} \pmod p$.

(48). Let p be a prime, $p \geq 5$, and write $1 + \frac{1}{2} + \frac{1}{3} + ... + \frac{1}{p} = \frac{r}{ps}$. Show that $p^3 | (r - s)$.

(49). Let a, b and x_0 be positive integers and define $x_n = ax_{n-1} + b$ for $n = 1, 2, ...$ Prove that not all the x_n can be primes.

(50). Let a, b and n be positive integers such that n divides $a^n - b^n$. Prove that n also divides $\frac{a^n - b^n}{a - b}$.

(51). Let a be an odd integer. Define $a!! := \prod_{1 \leq k \leq a, (k,2)=1} k$. Show that if $k \geq 1$ and $e \geq 3$ are integers, then $(k \cdot 2^e - 3)!! + 1 \equiv 0 \pmod{2^{e+1}}$.

(52). Let $n > 1$ be an odd number. Show that there is one element in the set $\{2 - 1, 2^2 - 1, 2^3 - 1, ..., 2^{n-1} - 1\}$ which is divisible by n.

(53). Show the following generalization of Wilson's theorem: For any integer k with $1 \leq k \leq p - 1$, one has $(p - k)!(k - 1)! \equiv (-1)^k \pmod p$.

(54). Find the smallest integer n such that $n/2$ is a square, $n/3$ is a cube, $n/5$ is the 5-th power of an integer.

(55). Find the smallest integer n such that $n/3$ is a cube, $n/5$ is the 5-th power of an integer and $n/7$ is the 7-th power of an integer.

(56). Let p be a prime and k be an integer such that $1 \leq k \leq p - 1$. Show that the following congruence holds: $\frac{(kp)!}{k!p^k} \equiv (-1)^k \pmod p$.

(57). Let m and d be positive integers such that d divides m. Show that if a is an integer relatively prime to d, then there exists an integer a' such that $a' \equiv a \pmod d$ and a' is relatively prime to m.

(58). Let p be a prime factor of the Fermat number $F_m = 2^{2^m} + 1$. Show that $p^2 | F_m$ if and only if $2^{p-1} \equiv 1 \pmod{p^2}$.

(59). Show that there are infinitely many integers $k \geq 1$ such that all the elements in the sequence $\{k2^n + 1\}_{n=1}^{\infty}$ are composite numbers.

(60). Let $k \geq 1$ be any given integer. Show the existence of infinitely many pairs (m, n) of integers with $m > 1$ and $n > 1$ such that $n | m$ and $n^k | \varphi(m)$.

(61). Let p be an odd number. Show that p and $p + 2$ are both primes if and only if $4(p - 1)! + 4 \equiv -p \pmod{p(p + 2)}$.

(62). Show that $m^2 \equiv 1 \pmod 3$ if m is an integer coprime to 3.

(63). Show that the square of each odd number is congruent to 1 modulo 8.

(64). Find all solutions of the following systems of congruences:

$$(1).\ x \equiv 1 \pmod 5,\ x \equiv 2 \pmod 6,\ x \equiv 3 \pmod 7,\ x \equiv 4 \pmod 8.$$

$$(2).\ x \equiv 1 \pmod{15},\ x \equiv 2 \pmod{16},\ x \equiv 3 \pmod{17},\ x \equiv 4 \pmod{18}.$$

(65). Let p be a prime and $f(x)$ be an integer polynomial of degree d with its constant not being congruent to 0 modulo p. Show that if there is an integer m such that $f(m) \equiv 0 \pmod{p}$, then there is an integer n such that $n^d f(\frac{1}{n}) \equiv 0 \pmod{p}$.

Chapter 3. Arithmetical Functions and Mean Values

§3.1 Ring of arithmetic functions and Dirichlet inverse

An *arithmetical function* is a complex-valued function whose domain is the set of positive integers. For instance, the divisor function $d(n)$ and the Euler phi function $\varphi(n)$ are arithmetical functions.

The *pointwise sum* $f+g$ of the arithmetic functions f and g is defined by

$$(f+g)(n)=f(n)+g(n). \tag{3.1}$$

There are two natural ways to multiply arithmetic functions f and g. The first is the *pointwise product* $f\cdot g$, defined by $(f\cdot g)(n)=f(n)\cdot g(n)$. The second is the *Dirichlet convolution* $f*g$, defined by

$$(f*g)(n)=\sum_{d|n}f(d)g(n/d)=\sum_{dd'=n}f(d)g(d') \tag{3.2}$$

where the sum is over all positive divisor d of n. Dirichlet convolution occurs frequently in multiplicative problems in elementary number theory.

We define the arithmetic functions $\delta, 0, 1$ and id_e by $\delta(n)=\{^{1\ \text{ if } n=1,}_{0\ \text{ if } n\geq 2,}, 0(n)=0, 1(n)=1, \mathrm{id_e}(n)=n^e$ for all n. We define the negative function $-f$ by $(-f)(n)=-f(n)$. We denote by $\mathcal{F}$ the set of all arithmetic functions.

Proposition 3.1.1. *Let $f,g,h\in\mathcal{F}$. Then each of the following is true:*
(i). $f+g=g+f, f*g=g*f$.
(ii). $f+(g+h)=f+(g+h), (f*g)*h=f*(g*h)$.
(iii). $f+0=f, f*\delta=f$.
(iv). $f+(-f)=0$.
(v). $f*(g+h)=f*g+f*h$.

Proof. These are straightforward calculations. We have

$$(f*g)(n)=\sum_{d|n}f(d)g(n/d)=\sum_{d|n}g(n/d)f(d)=\sum_{d|n}g(d)f(n/d)=g*f(n)$$

and

$$\begin{aligned}((f*g)*h)(n)&=\sum_{d|n}(f*g)(d)h(\frac{n}{d})=\sum_{dm=n}(f*g)(d)h(m)\\ &=\sum_{dm=n}\sum_{kl=d}f(k)g(l)h(m)=\sum_{klm=n}f(k)g(l)h(m)\\ &=\sum_{k|n}f(k)\sum_{lm=n/k}g(l)h(m)=\sum_{k|n}f(k)\sum_{l|(n/k)}g(l)h(\frac{n}{kl})\\ &=\sum_{k|n}f(k)(g*h)(\frac{n}{k})=(f*(g*h))(n).\end{aligned}$$

Similarly,

$$(f*(g+h))(n)=\sum_{d|n}f(d)(g(n/d)+h(n/d))$$

$$=\sum_{d|n}f(d)g(n/d)+\sum_{d|n}f(d)h(n/d)=(f*g)(n)+(f*h)(n)=(f*g+f*h)(n).$$

Finally, we observe that $(\delta * f)(n) = \sum_{d|n} \delta(d) f(n/d)$ for every arithmetic function f, and so the arithmetic functions form a commutative ring with multiplicative identity $\delta(n)$. □

Theorem 3.1.1. *The set of all complex-valued arithmetical functions, with addition defined by pointwise sum and multiplication defined by Dirichlet convolution, is a commutative ring with additive identity 0(n) and multiplicative identity $\delta(n)$.*

Proof. By Proposition 3.1.1, one knows that the set $\mathcal{F}$ of arithmetic functions is an additive abelian group with the zero function 0 as the additive identity.

Consequently, Proposition 3.1.1 tells us that Dirichlet convolution is commutative, associative, and distributes over addition.

This completes the proof of Theorem 3.1.1. □

Recall that a *derivation* on a ring R is an additive homomorphism $D : R \to R$ such that $D(xy) = D(x)y + xD(y)$ for all $x, y \in R$.

Proposition 3.1.2. *Consider the arithmetic function L(n) defined by*

$$L(n) = \log n \; for \; all \; n \geq 1.$$

Pointwise multiplication by L(n) is a derivation on the ring of arithmetic functions.

Proof. Observe that if d is a positive divisor of n, then $[L(n) = L(d) + L(n/d)$. We must prove that $L \cdot (f * g) = (L \cdot f) * g + f * (L \cdot g)$ for all arithmetic functions f and g. We have

$$\begin{aligned}
(L \cdot (f * g))(n) &= L(n) \sum_{d|n} f(d) g(n/d) \\
&= \sum_{d|n} L(n) f(d) g(n/d) \\
&= \sum_{d|n} (L(d) + L(n/d)) f(d) g(n/d) \\
&= \sum_{d|n} L(d) f(d) g(n/d) + \sum_{d|n} f(d) L(n/d) g(n/d) \\
&= ((L \cdot f) * g + f * (L \cdot g))(n).
\end{aligned}$$

This completes the proof. □

Since $(f * g)(1) \neq 0$ if $f(1) \neq 0$ and $g(1) \neq 0$, the set of all arithmetical functions f with $f(1) \neq 0$ forms an abelian group respect to the operation $*$.

Theorem 3.3.2. *Let f be an arithmetical function. Then f has Dirichlet inverse f^{-1} if and only if $f(1) \neq 0$. Furthermore, if $f(1) \neq 0$, then the Dirichlet inverse f^{-1} is determined by: $f^{-1}(1) = \frac{1}{f(1)}$ and*

$$f^{-1}(n) = \frac{-1}{f(1)} \sum_{d|n, d<n} f(\frac{n}{d}) f^{-1}(d) \quad for \quad n > 1.$$

Proof. First, since $(f * f^{-1})(1) = 1$, we deduce $f(1)f^{-1}(1) = 1$. But $f(1) \neq 0$. Then $f^{-1}(1) = \frac{1}{f(1)}$. Now let $n > 1$ be an integer and assume that the function values $f^{-1}(k)$ have been uniquely determined for all $k < n$. Then we solve the equation

$$\sum_{d|n} f(\frac{n}{d})f^{-1}(d) = 0.$$

That is,

$$f(1)f^{-1}(n) + \sum_{d|n,d<n} f(\frac{n}{d})f^{-1}(d) = 0.$$

It follows that

$$f^{-1}(n) = \frac{-1}{f(1)} \sum_{d|n,d<n} f(\frac{n}{d})f^{-1}(d),$$

since $f(1) \neq 0$. Thus by induction, the existence and uniqueness of f^{-1} are established. The proof is complete. □

§3.2 Möbius function and Möbius inversion formula

The $M\ddot{o}bius function$ $\mu(n)$ is defined as follows:

$$\mu(n) = \begin{cases} 1, & \text{if } n = 1, \\ (-1)^k, & \text{if } n \text{ is the product of } k \text{ distinct primes }, \\ 0, & \text{if } n \text{ is divisible by the square of a prime.} \end{cases}$$

For example, we have $\mu(1) = 1, \mu(2) = -1, \mu(3) = 1, \mu(4) = 0, \mu(5) = -1, \mu(6) = 1, \mu(7) = -1, \mu(8) = 0, \mu(9) = 0, \mu(10) = 1$. An integer is called *square-free* if it is not divisible by the square of a prime. Thus, $\mu(n) \neq 0$ if and only if n is square-free. Recall that an arithmetic function $f(n)$ is *multiplicative* if $f(mn) = f(m)f(n)$ whenever $(m, n) = 1$.

Theorem 3.2.1. *Möbius function $\mu(n)$ is multiplicative and*

$$\sum_{d|n} \mu(d) = \begin{cases} 1, & \textit{if } n = 1, \\ 0, & \textit{if } n > 1. \end{cases}$$

Proof. The multiplicativity follows immediately from the definition of $M\ddot{o}bius$ function, since if m and n are relatively prime square-free integers with k and l prime factors, respectively, then mn is square-free with $k + l$ factors, and

$$\mu(m)\mu(n) = (-1)^k(-1)^l = (-1)^{k+l} = \mu(mn).$$

Next we prove the convolution formula. If $n = 1$, then $\sum_{d|n} \mu(d) = \mu(1) = 1$. For $n \geq 2$ let $n = p_1^{r_1} \dots p_k^{r_k}$ be the standard factorization of the integer n. Then $r \geq 1$. Recall that the *radical* of n is the largest square-free divisor of n, that is, $\text{rad}(n) = p_1 \dots p_r$ is the product of the distinct primes dividing n. Let $m = rad(n)$. If d divides n and $\mu(d) \neq 0$, then d is square-free, and so d divides m. Since m is the product of k primes, it follows that there are exactly $\binom{k}{i}$ divisors of m that can be written as the product of i distinct primes, that is, the number of divisor d of m such that $\omega(d) = i$ is $\binom{k}{i}$. Therefore,

$$\sum_{d|n} \mu(d) = \sum_{d|m} \mu(d) = \sum_{i=0}^{k} \sum_{\substack{d|m \\ \omega(d)=i}} \mu(d) = \sum_{i=0}^{k} \sum_{\substack{d|m \\ \omega(d)=i}} (-1)^i = \sum_{i=0}^{k} \binom{k}{i}(-1)^i = (1-1)^k = 0.$$

This completes the proof. □

We defined the arithmetic function $1(n)$ by $1(n) = 1$ for all n. Using the Dirichlet convolution, we can restate Theorem 3.. as follows

$$\mu * 1 = \delta,$$

and so the Möbius function μ is a unit with inverse 1.

Theorem 3.2.2. (Möbius inversion formula) *If f is any arithmetic function, and g is the arithmetic function defined by $g(n) = \sum_{d|n} f(d)$, then $f(n) = \sum_{d|n} \mu(\frac{n}{d})g(d)$. Conversely, if g is any arithmetic function, and f is the arithmetic function defined by $f(n) = \sum_{d|n} \mu(\frac{n}{d})g(d)$, then $g(n) = \sum_{d|n} f(d)$.*

Proof. We use Theorem 3.2.1 and the commutativity and associativity of Dirichlet convolution. The definition $g(n) = \sum_{d|n} f(d)$ is equivalent to $g = f * 1$. Then

$$g * \mu = (f * 1) * \mu = f * (1 * \mu) = f * \delta = f.$$

Conversely, if $f = g * \mu$, then $f * 1 = (g * \mu) = g * (\mu * 1) = g * \delta = g$.

This concludes the proof. □

Theorem 3.2.3. *Let f be any arithmetic function and n a positive integer. Then $\sum_{d|n}(f * \mu)(d) = f(n)$.*

Proof. Since $\mu * 1 = \delta$, one has $f = f * \delta = f * (\mu * 1) = (f * \mu) * 1$. It follows that for any positive integer n, we have

$$f(n) = (f * \mu) * 1(n) = \sum_{d|n}(f * \mu)(d)1\big(\frac{n}{d}\big) = \sum_{d|n}(f * \mu)(d)$$

as required. This finishes the proof of Theorem 3.2.3. □

Remark. Evidently Theorems 2.3.3 (Gauss Theorem) and 3.2.1 are special cases of Theorem 3.2.3.

The following result gives a useful identity for sum functions of arithmetical functions. The proof can be described geometrically as a sum over the lattice points (m, d) under the hyperbola $v = x/u$ in the uv-plane.

Theorem 3.2.4. *Let f(n) be an arithmetic function and $F(x) = \sum_{n \le x} f(n)$. Then $\sum_{m \le x} F(\frac{x}{m}) = \sum_{d \le x} f(d)[\frac{x}{d}] = \sum_{n \le x} \sum_{d|n} f(d)$.*

Proof. We have

$$\sum_{m \le x} F(\frac{x}{m}) = \sum_{m \le x} \sum_{d \le x/m} f(d) = \sum_{dm \le x} f(d) = \sum_{d \le x} f(d) \sum_{m \le x/d} 1 = \sum_{d \le x} f(d)[\frac{x}{d}].$$

On the other hand, we have

$$\sum_{m \le x} F(\frac{x}{m}) = \sum_{dm \le x} f(d) = \sum_{n \le x} \sum_{d|n} f(d).$$

This completes the proof. □

Theorem 3.2.5. *One has $\sum_{n \le x} \frac{\mu(n)}{n} = O(1)$.*

Proof. Applying Theorem 3.2.4 with $f(n) = \mu(n)$ and $F(x) = \sum_{n \le x} \mu(n)$, we obtain

$$\sum_{m \le x} F(\frac{x}{m}) = \sum_{d \le x} \mu(d)[\frac{x}{d}] = \sum_{n \le x} \sum_{d|n} f(d) = 1,$$

by Theorem 3.2.1. Since

$$\sum_{d\leq x}\mu(d)[\frac{x}{d}]=x\sum_{d\leq x}\frac{\mu(d)}{d}-\sum_{d\leq x}\mu(d)\{\frac{x}{d}\}=x\sum_{d\leq x}\frac{\mu(d)}{d}+O(x),$$

it follows that

$$x\sum_{d\leq x}\frac{\mu(d)}{d}+O(x)=1.$$

Therefore,

$$x\sum_{d\leq x}\frac{\mu(d)}{d}=O(x),$$

and so

$$\sum_{d\leq x}\frac{\mu(d)}{d}=O(1).$$

This completes the proof. □

Theorem 3.2.6.

$$\sum_{n\leq x}\frac{\mu(n)}{n^2}=\frac{6}{\pi^2}+O(\frac{1}{x}).$$

Proof. The *Riemann zeta function* $\zeta(s)=\sum_{n=1}^{\infty}\frac{1}{n^s}$ converges absolutely for $s>1$. Similarly, the function $G(s)=\sum_{n=1}^{\infty}\frac{\mu(n)}{n^s}$ converges absolutely for $s>1$. Therefore by Theorem 3.2.1 we have

$$\begin{aligned}\zeta(s)G(s)&=\sum_{k=1}^{\infty}\frac{1}{k^s}\sum_{d=1}^{\infty}\frac{\mu(d)}{d^s}\\&=\sum_{k=1}^{\infty}\sum_{d=1}^{\infty}\frac{\mu(d)}{(kd)^s}\\&=\sum_{n=1}^{\infty}\frac{1}{n^s}\sum_{d|n}f(d)=1,\end{aligned}$$

and so $\frac{1}{\zeta(s)}=\sum_{n=1}^{\infty}\frac{\mu(n)}{n^s}$ for $s>1$. Since $\zeta(2)=\sum_{n=1}^{\infty}\frac{1}{n^2}=\frac{\pi^2}{6}$, it follows that $\frac{1}{\zeta(2)}=\sum_{n=1}^{\infty}\frac{\mu(n)}{n^2}=\frac{6}{\pi^2}$, and so

$$\Big|\sum_{n\leq x}\frac{\mu(n)}{n^2}-\frac{6}{\pi^2}\Big|=\Big|\sum_{n>x}\frac{\mu(n)}{n^2}\Big|<\sum_{n>x}\frac{1}{n^2}\ll\frac{1}{x}.$$

This completes the proof. □

§3.3 Multiplicative functions and completely multiplicative functions

In this section we prove some general properties about multiplicative functions.

Theorem 3.3.1. *Let f and g be arithmetical functions. Then that two of f, g and $f*g$ are multiplicative implies that the third one is also multiplicative.*

Proof. It is easy to check that the multiplicativity of f and g implies the multiplicativity of $f*g$.

Now assume that f and $f*g$ are multiplicative. We want to show that g is also multiplicative. Let $h=f*g$. Suppose that g is not multiplicative. We shall derive that h is not multiplicative. Since g is not multiplicative, there exist positive integers m and n with $(a,b)=1$ such that $g(ab)\neq g(a)g(b)$. Choose such a pair a and b for which the product ab is as small as possible.

If $ab=1$, then $g(1)\neq g(1)g(1)$ thus $g(1)\neq 1$. Since $h(1)=f(1)g(1)=g(1)\neq 1$. So h is not multiplicative.

If $ab>1$, then for all positive integers c and d with $(c,d)=1$ and $cd<ab$, we have $g(cd)=g(c)g(d)$. We then have

$$h(ab)=\sum_{c|a,d|b,cd<ab} g(cd)f(\frac{ab}{cd})+f(1)g(ab)=\sum_{c|a,d|b,cd<ab} g(a)g(b)f(\frac{a}{c})g(\frac{b}{d})+g(ab)$$

$$=\big(\sum_{c|a} f(\frac{a}{c})g(c)\big)\big(\sum_{d|b} f(\frac{b}{d})g(d)\big)-g(a)g(b)+g(ab)=h(a)h(b)-g(a)g(b)+g(ab).$$

Since $g(ab)\neq g(a)g(b)$, then by the above identity we get $h(ab)\neq h(a)h(b)$. This implies that h is not multiplicative. It is a contradiction. Therefore g is multiplicative. □

Theorem 3.3.2. *Let f hold Dirichlet inverse. Then f is multiplicative if and only if f^{-1} is multiplicative.*

Proof. □

Theorem 3.3.3. *If f is a multiplicative function, then*

$$f([m,n])f((m,n))=f(m)f(n)$$

for all positive integers m and n.

Proof. Let $p_1,\ldots,p_r$ be the prime numbers that divide m or n. Then

$$n=\prod_{i=1}^{r} p_i^{k_i} \quad\text{and}\quad m=\prod_{i=1}^{r} p_i^{l_i},$$

where $k_1,\ldots,k_r,l_1,\ldots,l_r$ are nonnegative integers. Then

$$[m,n]=\prod_{i=1}^{r} p_i^{\max(k_i,l_i)} \quad\text{and}\quad (m,n)=\prod_{i=1}^{r} p_i^{\min(k_i,l_i)}.$$

Since f is multiplicative and

$$\{\max(k_i,l_i),\min(k_i,l_i)\}=\{k_i,l_i\}$$

it follows that

$$\begin{aligned} f([m,n])f((m,n)) &= \prod_{i=1}^{r} f(p_i^{\max(k_i,l_i)})\prod_{i=1}^{r} f(p_i^{\min(k_i,l_i)}) \\ &= \prod_{i=1}^{r} f(p_i^{k_i})\prod_{i=1}^{r} f(p_i^{l_i})=f(m)f(n). \end{aligned}$$

This completes the proof. □

Theorem 3.3.4. *Let f be a multiplicative function with $f(1)=1$. Then*

$$\sum_{d|n}\mu(d)f(d)=\prod_{p|n}(1-f(p)).$$

Proof. The identity holds for $n = 1$. For $n \geq 2$, let $m = \mathrm{rad}(n)$ be the product of the distinct primes dividing n. Since $\mu(d) = 0$ if d is not square-free, it follows that

$$\sum_{d|n} \mu(d) f(d) = \sum_{d|m} \mu(d) f(d) = \prod_{p|m} (1 - f(p)) = \prod_{p|n} (1 - f(p)).$$

This ends the proof. □

The sequences of prime powers is the sequence

$$2, 3, 4, 5, 7, 8, 9, 11, 13, 16, 17, 19, 23, 25, 27, \ldots.$$

The smallest power that is not a prime power is 36. We can now state a significant theorem.

Theorem 3.3.5. *Let $f(n)$ be a multiplicative function. If*

$$\lim_{p^k \to \infty} f(p^k) = 0$$

as p^k runs through the sequences of all prime powers, then $\lim_{n \to \infty} f(n) = 0$.

Proof. Since $\lim_{p^k \to \infty} f(p^k) = 0$, it follows that there exist only finitely many prime powers p^k such that $|f(p^k)| \geq 1$, and so we can define

$$A = \prod_{|f(p^k)| \geq 1} |f(p^k)|.$$

Then $A \geq 1$.

Let $0 < \varepsilon < 1$. There exist only finitely many prime powers p^k such that $|f(p^k)| \geq \varepsilon / A$, and so there are only finitely many integer n such that $|f(p^k)| \geq \frac{\varepsilon}{A}$ for every prime power p^k that exactly divides n. Therefore, if n is sufficiently large, then n is divisible by at least one prime power p^k such that $|f(p^k)| < \varepsilon / A$, and so n can be written in the form

$$n = \prod_{i=1}^{r} p_i^{k_i} \prod_{i=r+1}^{r+s} p_i^{k_i} \prod_{i=r+s+1}^{r+s+t} p_i^{k_i},$$

where $p_1, \ldots, p_{r+s+t}$ are distinct prime numbers such that

$$|f(p_i^{k_i})| \geq 1 \text{ for } i = 1, \ldots, r,$$

$$\frac{\varepsilon}{A} \leq |f(p_i^{k_i})| < 1 \text{ for } i = r+1, \ldots, r+s,$$

$$|f(p_i^{k_i})| < \frac{\varepsilon}{A} \text{ for } i = r+s+1, \ldots, r+s+t,$$

and

$$r \geq 0, s \geq 0, t \geq 1.$$

Since f is multiplicative,

$$|f(n)| = \prod_{i=1}^{r} |f(p_i^{k_i})| \prod_{i=r+1}^{r+s} |f(p_i^{k_i})| \prod_{i=r+s+1}^{r+s+t} |f(p_i^{k_i})| < A(\varepsilon/A)^t \leq \varepsilon.$$

This completes the proof. □

Definition. An arithmetical function f is called *completely multiplicative function* if $f(mn) = f(m)f(n)$ for all positive integers m and n.

The functions $f(n) = n^e, \delta(n)$ and the Liouville's function $\lambda(n) := (-1)^{\Omega(n)}$, where $\Omega(p_1^{k_1}...p_r^{k_r}) := k_1 + ... + k_r$, are completely multiplicative.

The following result gives us a formula of Dirichlet inverse of a completely multiplicative function.

Theorem 3.3.6. *Let f be a multiplicative function with $f(1) = 1$. Then f is completely multiplicative iff we have $f^{-1}(n) = \mu(n)f(n)$ for all $n \geq 1$.*

Proof. Let $g = \mu \cdot f$ and assume that f is completely multiplicative. Then

$$(g * f)(n) = \sum_{d|n} \mu(d)f(d)f(\frac{n}{d}) = f(n)\sum_{d|n} \mu(d) = \delta(n).$$

Thus $g = f^{-1}$.

Conversely, suppose that $f^{-1}(n) = \mu(n)f(n)$ for all $n \geq 1$. Clearly, to prove that f is completely multiplicative, it suffices to show that $f(p^a) = f(p)^a$ for any prime power p^a. Since $f^{-1}(n) = \mu(n)f(n)$ implying that $\sum_{d|n} \mu(d)f(d)f(\frac{n}{d}) = 0$ for all $n > 1$, then letting $n = p^a$ gives us $\mu(1)f(1)f(p^a) + \mu(p)f(p)f(p^{a-1}) = 0$. From this we derive that $f(p^a) = f(p)f(p^{a-1})$. So using induction on a we get $f(p^a) = f(p)^a$. Hence f is completely multiplicative. This completes the proof of Theorem 3.3.6. □

Theorem 3.3.7. *Let f be a multiplicative function with $f(1) = 1$. Then f is completely multiplicative iff we have $f^{-1}(p^a) = 0$ for any prime p and any integer $a \geq 2$.*

Proof. ($\Rightarrow$): Since f is completely multiplicative, by Theorem 3.3.6 we have $f^{-1}(p^a) = \mu(p^a)f(p^a) = 0$ for any prime p and any integer $a \geq 2$.

($\Leftarrow$): Evidently it is true that $f^{-1}(p^a) = \mu(p^a)f(p^a)$ for any prime p and any integer $a \geq 2$ since $f^{-1}(p^a) = 0$ and $\mu(p^a) = 0$. On the other hand, $0 = (f * f^{-1})(p) = f(1)f^{-1}(p) + f(p)f^{-1}(1) = f^{-1}(p) + f(p)$. Thus $f^{-1}(p) = -f(p) = \mu(p)f(p)$. That is, we have $f^{-1}(p^a) = \mu(p^a)f(p^a)$ for any prime p and any integer $a \geq 1$. Finally, by the multiplicativity of f and f^{-1}, we arrive at $f^{-1}(n) = \mu(n)f(n)$ for all $n \geq 1$. So Theorem 3.3.7 is proved. □

§3.4 Mean values of arithmetic functions

We define the *mean value* $F(x)$ of an arithmetic function $f(n)$ by

$$F(x) = \sum_{n \leq x} f(n),$$

where the sum is over all positive integers $n \leq x$. In particular, $F(x) = 0$ for $x < 1$. The function $F(x)$ is also called the *sum function* of f. We shall describe two simple but powerful tools for estimating sun functions in number theory. The first is integration and the second is partial summation.

The *integer part* of the real number x, denoted by $[x]$, is the unique integer n such that $n \leq x < n+1$. The *fractional part* of x is the real number $\{x\} = x - [x] \in [0, 1)$. For example, $[-\frac{5}{3}] = -2$ and $\{-\frac{5}{3}\} = \frac{1}{3}$. Every real number x can be written uniquely in the form $x = [x] + x$.

A function $f(t)$ is called *unimodal* on an interval I if there exists a number $t_0 \in I$ such that $f(t)$ is increasing for $t \leq t_0$ and decreasing for $t \geq t_0$. For example, the function $f(t) = \log^k t/t$ is unimodal on the interval $[1, \infty)$ with $t_0 = e^k$.

It is proved in the real analysis that every function that is monotonic or unimodal on a closed interval $[a, b]$ is integrable.

Theorem 3.4.1. (i). *Let a and b be integers with $a < b$, and let f(t) be a function that is monotonic on the interval [a,b]. Then*

$$\min(f(a), f(b)) \leq \sum_{n=a}^{b} f(n) - \int\limits_{a}^{b} f(t)dt \leq \max(f(a), f(b)); \tag{3.3}$$

(ii). *Let x and y be real numbers with $y < [x]$, and let $f(t)$ be a nonnegative monotonic function on [y,x]. Then*

$$\Big| \sum_{y<n\leq x} f(n) - \int\limits_{y}^{x} f(t)dt \Big| \leq \max(f(y), f(x)); \tag{3.4}$$

(iii). *If $f(t)$ is a nonnegative unimodal function on [1,∞), then*

$$F(x) = \sum_{n\leq x} f(n) = \int\limits_{1}^{x} f(t)dt + O(1). \tag{3.5}$$

Proof. If $f(t)$ is increasing on $[n, n+1]$, then $f(n) \leq \int\limits_{n}^{n+1} f(t)dt \leq f(n+1)$. If $f(t)$ is increasing on the interval $[a, b]$, then

$$f(a) + \int\limits_{a}^{b} f(t)dt \leq \sum_{n=a}^{b} f(n) \leq f(b) + \int\limits_{a}^{b} f(t)dt.$$

Similarly, if $f(t)$ is decreasing on the interval $[n, n+1]$, then $f(n+1) \leq \int\limits_{n}^{n+1} f(t)dt \leq f(n)$. If $f(t)$ is decreasing on the interval $[a, b]$, then

$$f(b) + \int\limits_{a}^{b} f(t)dt \leq \sum_{n=a}^{b} f(n) \leq f(a) + \int\limits_{a}^{b} f(t)dt.$$

This proves (3.3).

Let $f(t)$ be nonnegative and monotonic on the interval $[y, x]$. Let $a = [y] + 1$ and $b = [x]$. We have $y < a \leq b \leq x$. If $f(t)$ is increasing, then

$$\sum_{y<n\leq x} f(n) = \sum_{a\leq n\leq b} f(n) \leq \int\limits_{a}^{b} f(t)dt + f(b) \leq \int\limits_{y}^{x} f(t)dt + f(x).$$

Since $f(a) \geq \int\limits_{y}^{a} f(t)dt$ and $f(x) \geq \int\limits_{b}^{x} f(t)dt$, it follows that

$$\sum_{y<n\leq x} f(n) \geq \int\limits_{a}^{b} f(t)dt + f(a) = \int\limits_{y}^{x} f(t)dt - \int\limits_{b}^{x} f(t)dt + f(a) - \int\limits_{y}^{a} f(t)dt \geq \int\limits_{y}^{x} f(t)dt - f(x).$$

Therefore

$$\Big|\sum_{y<n\leq x} f(n) - \int\limits_{y}^{x} f(t)dt\Big| \leq f(x).$$

If $f(t)$ is decreasing, then

$$\sum_{y<n\leq x f(n)} = \sum_{a\leq n\leq b} f(n) \leq \int\limits_{a}^{b} f(t)dt + f(a) \leq \int\limits_{y}^{x} f(t)dt + f(y).$$

Since $f(b) \geq \int\limits_{b}^{x} f(t)dt$ and $f(y) \geq \int_y^a f(t)dt$, it follows that

$$\sum_{y<n\leq x} f(n) \geq \int_a^b f(t)dt + f(b) \geq \int_y^x f(t)dt + f(b) - \int_b^x f(t)dt - \int_y^a f(t)dt \geq \int_y^x f(t)dt - f(y)$$

and

$$\Big|\sum_{y<n\leq x} f(n) - \int_y^x f(t)dt\Big| \leq f(y).$$

This proves (3.4).

If function $f(t)$ is nonnegative and unimodal on $[1,\infty)$, then $f(t)$ is bounded and (3.5) follows from (3.4). □

Theorem 3.4.2. *For $x \geq 2$, $\sum_{n\leq x} \log n = x\log x - x + O(\log x)$.*

Proof. The function $f(t) = \log t$ is increasing on $[1, x]$. By Theorem 3.4.1,

$$\int_1^x \mathrm{log}t\mathrm{d}t \leq \sum_{n\leq x} \mathrm{log}n \leq \int_1^x \mathrm{log}t\mathrm{d}t + \mathrm{log}x,$$

and so

$$\sum_{n\leq x} \log n = x\log x - x + O(\log x).$$

This completes the proof. □

Theorem 3.4.3. *Let r be a nonnegative integer. For $x \geq 1$,*

$$\sum_{n\leq x} \frac{\log^r n}{n} = \frac{1}{r+1}\log^{r+1} x + O(1),$$

where the implied constant depends only on r.

Proof. The function $f(t) = \log^r t/t$ is nonnegative and unimodal on $[1,\infty)$ with maximum value $(r/e)^r$ at $t_0 = e^r$. By Theorem 3.4.1,

$$\sum_{n\leq x} \frac{\log^r n}{n} = \int_1^x \frac{\log^r tdt}{t} + O(1) = \frac{1}{r+1}\log^{r+1} x + O(1).$$

This completes the proof. □

Theorem 3.4.4. *Let k be a nonnegative integer.For $x \geq 1$,*

$$\sum_{n\leq x} \frac{\log^k (x/n)}{n} = \frac{1}{k+1}\log^{k+1} x + O(\log^k x),$$

where the implied constant depends only on k.

Proof. The idea is to expand $\log^k(x/n)$ by the binomial theorem and apply Theorem 3.4.3. We have

$$\begin{aligned}\sum_{n\le x}\frac{\log^k(x/n)}{n} &= \sum_{n\le x}\frac{(\log x-\log n)^k}{n}\\ &= \sum_{n\le x}\frac{1}{n}\sum_{r=0}^{k}\binom{k}{r}(-1)^r\log^{k-r}x\log^r n\\ &= \sum_{r=0}^{k}\binom{k}{r}(-1)^r\log^{k-r}x\sum_{n\le x}\frac{\log^r n}{n}\\ &= \sum_{r=0}^{k}\binom{k}{r}(-1)^r\log^{k-r}x\left(\frac{1}{r+1}\log^{r+1}x+O(1)\right)\\ &= \sum_{r=0}^{k}\binom{k}{r}\frac{(-1)^r}{r+1}\log^{k+1}x+O\left(\sum_{r=0}^{k}\binom{k}{r}\log^{k-r}x\right)\\ &= \frac{1}{k+1}\log^{k+1}x+O(\log^k x),\end{aligned}$$

since

$$\sum_{r=0}^{k}\frac{(-1)^r}{r+1}\binom{k}{r}=\frac{1}{k+1}.$$

(Exercise to you) □

Theorem 3.4.5. *Let k be a positive integer. Then*

$$\sum_{n_1\ldots n_k\le x}\frac{1}{n_1\ldots n_k}=\frac{1}{k!}\log^k x+O(\log^{k-1}x),$$

where $\sum_{n_1\ldots n_k\le x}$ denotes the sum over all k-tuples of positive integers $(n_1,\ldots,n_k)$ such that $n_1\ldots n_k\le x$.

Proof. By induction on k. For $k=1$, we set $r=0$ in the Theorem 3.2.3 and obtain $\sum_{n_1\le x}\frac{1}{n_1}=\log x+O(1)$. Assume that the result holds for the positive integer k. Then

$$\begin{aligned}\sum_{n_1\ldots n_kn_{k+1}\le x}\frac{1}{n_1\ldots n_kn_{k+1}} &= \sum_{n_{k+1}\le x}\frac{1}{n_{k+1}}\sum_{n_1\ldots n_k\le x/n_{k+1}}\frac{1}{n_1\ldots n_k}\\ &= \sum_{n_{k+1}\le x}\frac{1}{n_{k+1}}(\frac{1}{k!}\log^k(x/n_{k+1})+O(\log^{k-1}(x/n_{k+1})))\\ &= \sum_{n_{k+1}\le x}\frac{1}{k!n_{k+1}}(\log x-\log n_{k+1})^k+O\left(\log^{k-1}x\sum_{n_{k+1}\le x}\frac{1}{n_{k+1}}\right)\\ &= \sum_{n\le x}\frac{1}{k!n}(\log x-\log n)^k+O(\log^k x).\end{aligned}$$

We use the binomial theorem and Theorem 3.4.3 to compute the main term as follows:

$$\begin{aligned}\sum_{n\le x}\frac{1}{k!n}(\log x-\log n)^k &= \sum_{n\le x}\frac{1}{k!n}\sum_{r=0}^{k}(-1)^r\binom{k}{r}\log^{k-r}x\log^r n\\ &=\sum_{r=0}^{k}\frac{(-1)^r}{k!}\binom{k}{r}\log^{k-r}x\sum_{n\le x}\frac{\log^r n}{n}\\ &=\sum_{r=0}^{k}\frac{(-1)^r}{k!}\binom{k}{r}\log^{k-r}x\left(\frac{1}{r+1}\log^{r+1}x+O(1)\right)\\ &=\frac{1}{k!}\log^{k+1}x\sum_{r=0}^{k}\frac{(-1)^r}{r+1}\binom{k}{r}+O(\log^k x).\end{aligned}$$

□

Theorem 3.4.6. (Partial summation) *Let $f(n)$ and $g(n)$ be arithmetic functions. Consider the sum function*

$$F(x)=\sum_{n\le x}f(n);$$

(i). *Let a and b be nonnegative integers with $a<b$. Then*

$$\sum_{n=a+1}^{b}f(n)g(n)=F(b)g(b)-F(a)g(a+1)-\sum_{n=a+1}^{b-1}F(n)(g(n+1)-g(n)); \quad (3.6)$$

(ii). *Let x and y be nonnegative real numbers with $[y]<[x]$, and let $g(t)$ be a function with a continuous derivative on the interval $[y,x]$. Then*

$$\sum_{y<n\le x}f(n)g(n)=F(x)g(x)-F(y)g(y)-\int_y^x F(t)g'(t)dt; \quad (3.7)$$

(iii). *In particular , if $x\ge 2$ and $g(t)$ is continuously differentiable on $[1,x]$, then*

$$\sum_{n\le x}f(n)g(n)=F(x)g(x)-\int_1^x F(t)g'(t)dt. \quad (3.8)$$

Proof. Identity (3.6) is a straightforward calculation:

$$\begin{aligned}\sum_{n=a+1}^{b}f(n)g(n) &= \sum_{n=a+1}^{b}(F(n)v-F(n-1))g(n)\\ &=\sum_{n=a+1}^{b}F(n)g(n)-\sum_{n=a}^{b-1}F(n)g(n+1)\\ &=F(b)g(b)-F(a)g(a+1)-\sum_{n=a+1}^{b-1}F(n)(g(n+1)-g(n)).\end{aligned}$$

If the function $g(t)$ is continuously differentiable on $[y, x]$, then $g(n+1)-g(n) = \int_n^{n+1} g'(t)dt$. Since $F(t) = F(n)$ for $n \le t < n+1$, it follows that

$$F(n)(g(n+1) - g(n)) = \int_n^{n+1} F(t)g'(t)dt.$$

Let $a = [y]$ and $b = [x]$. Since $a \le y < a+1 \le b \le x < b+1$, we have

$$\begin{aligned}\sum_{y<n\le x} f(n)g(n) &= \sum_{n=a+1}^{b} f(n)g(n)\\ &= F(b)g(b) - F(a)g(a+1) - \sum_{n=a+1}^{b-1} F(n)(g(n+1) - g(n))\\ &= F(x)g(b) - F(y)g(a+1) - \sum_{n=a+1}^{b-1} \int_n^{n+1} F(t)g'(t)dt\\ &= F(x)g(x) - F(y)g(y) - F(x)(g(x) - g(b)) - F(y)(g(a+1) - g(y))\end{aligned}$$
$$- \int_{a+1}^{b} F(t)g'(t)dt$$
$$= F(x)g(x) - F(y)g(y) - \int_y^x F(t)g'(t)dt.$$

This proves (3.7).

IF $x \ge 2$ and $g(t)$ is continuously differentiable on $[1, x]$, then

$$\begin{aligned}\sum_{n\le x} f(n)g(n) &= f(1)g(1) + \sum_{1<n\le x} f(n)g(n)\\ &= f(1)g(1) + F(x)g(x) - F(1)g(1) - \int_1^x F(t)g'(t)dt\\ &= F(x)g(x) - \int_1^x F(t)g'(t)dt.\end{aligned}$$

This proves (3.8). □

Letting $r = 0$ in Theorem 3.4.3, we obtain $\sum_{n\le x} 1/n = \log x + O(1)$. Using partial summation, we can obtain a more precise result.

Theorem 3.4.7. *For* $x \ge 1$, *one has* $\sum_{n\le x} \frac{1}{n} = \log x + \gamma + r(x)$, *where* $0 < \gamma = 1 - \int_1^\infty \frac{\{t\}}{t^2} dt < 1$ *and* $|r(x)| < \frac{1}{x}$.

The number $\gamma = 0.577\ldots$ is called *Euler's constant*. A famous unsolved problem in number theory is to determine whether γ is rational of irrational.

Proof. Since $0 \le \{t\} < 1$ for all t, we have $0 < \int_1^\infty \frac{\{t\}}{t^2} dt < \int_1^\infty \frac{1}{t^2} dt = 1$, and so $\gamma \in (0, 1)$. We apply partial summation to the functions $f(n) = 1$ and $g(t) = 1/t$.

Then $F(t) = \sum_{n \leq t} 1 = [t]$ and

$$\begin{aligned}\sum_{n \leq x} \frac{1}{n} &= \sum_{n \leq x} f(n)g(n) \\ &= \frac{[x]}{x} + \int_1^x \frac{[t]}{t^2} dt \\ &= 1 - \frac{\{x\}}{x} + \int_1^x \frac{1}{t} dt - \int_1^x \frac{\{t\}}{t^2} dt \\ &= \log x + (1 - \int_1^\infty \frac{\{t\}}{t^2} dt) + \int_x^\infty \frac{\{t\}}{t^2} dt - \frac{\{x\}}{x} \\ &= \log x + \gamma + r(x),\end{aligned}$$

where

$$r(x) = \int_x^\infty \frac{\{t\}}{t^2} dt - \frac{\{x\}}{x}.$$

Moreover, $|r(x)| < 1/x$ since $0 \leq \{x\}/x < 1$ and

$$0 < \int_x^\infty \frac{\{t\}}{t^2} dt < \int_x^\infty \frac{1}{t^2} dt = \frac{1}{x}.$$

□

Theorem 3.4.8. *Let $A = \{a_i\}_{i=1}^\infty$ be an infinite set of positive integer with $a_1 < a_2 < a_3 < \ldots$. If $A(x) = \sum_{a_i \leq x} 1 = O(\frac{x}{\log^2 x})$ for $x \geq 2$, then the series $\sum_{i=1}^\infty \frac{1}{a_i}$ converges.*

Proof. Let $\chi_A(n)$ be the characteristic function of A, that is, $\chi_A(n) = \{\begin{smallmatrix} 1 & \text{if } n \in A, \\ 0 & \text{if } n \notin A. \end{smallmatrix}$ There exists a number c such that

$$A(x) = \sum_{n \leq x} \chi_A(n) \leq \frac{cx}{\log^2 x}$$

for all $x \geq 2$ and $A(x) \leq 1$ for $1 \leq x \leq 2$. Applying partial summation, we obtain

$$\begin{aligned}\sum_{a_i \leq x} \frac{1}{a_i} &= \sum_{n \leq x} \frac{\chi_A(n)}{n} \\ &= \frac{A(x)}{x} + \int_1^x \frac{A(t)dt}{t^2} \\ &\leq \frac{c}{\log^2 x} + \frac{1}{2} + c \int_{\log 2}^{\log x} \frac{du}{u^2} < \infty.\end{aligned}$$

This completes the proof. □

Theorem 3.4.9. *For $x \geq 2$, $\sum_{n \leq x} \log^2 n = x \log^2 x - 2x \log x + 2x + O(\log^2 x)$.*

Proof. We use partial summation with $f(n) = 1$ and $g(t) = \log^2 t$. Then $F(t) = [t]$ and $g'(t) = 2 \log t / t$. Hence

$$\begin{aligned}\sum_{n\le x}\log^2 n &= [x]\log^2 x - 2\int_1^x \frac{[t]\log t}{t}dt\\ &= (x-\{x\})\log^2 x - 2\int_1^x \frac{(t-\{t\})\log t}{t}dt\\ &= x\log^2 x + O(\log^2 x) - 2\int_1^x \log t dt + 2\int_1^x \frac{\{t\}\log t}{t}dt\\ &= x\log^2 x - 2x\log x + 2x + O(\log^2 x).\end{aligned}$$

This completes the proof. □

Theorem 3.4.10. *For $x \ge 2$, one has $\sum_{n\le x}\log^2\frac{x}{n} = 2x + O(\log^2 x)$.*

Proof. From Theorem 3.4.2 and Theorem 3.4.9, we obtain

$$\begin{aligned}\sum_{n\le x}\log^2\frac{x}{n} &= \sum_{n\le x}(\log x - \log n)^2\\ &= \sum_{n\le x}(\log^2 x - 2\log x\log n + \log^2 n)\\ &= [x]\log^2 x - 2\log x\sum_{n\le x}\log n + \sum_{n\le x}\log^2 n\\ &= x\log^2 x - 2\log x(x\log x - x) + x\log^2 x - 2x\log x + 2x + O(\log^2 x)\\ &= 2x + O(\log^2 x).\end{aligned}$$

This completes the proof. □

The Euler phi function is

$$\varphi(n) = n\prod_{p|n}\left(1-\frac{1}{p}\right) = n\sum_{d|n}\frac{\mu(d)}{d} = \sum_{d'd=n} d'\mu(d). \qquad (3.10)$$

We shall find an asymptotic formula for the mean value of the Euler phi function.

Theorem 3.4.11. *For $x \ge 1$, we have $\Phi(x) = \sum_{n\le x}\varphi(n) = \frac{3x^2}{\pi^2} + O(x\log x)$.*

Proof. We have

$$\begin{aligned}\Phi(x) &= \sum_{n\le x}\varphi(n) = \sum_{n\le x}\sum_{d'd=n} d'\mu(d) = \sum_{d\le x}\mu(d)\sum_{d'\le x/d} d'\\ &= \frac{1}{2}\sum_{d\le x}\mu(d)\left[\frac{x}{d}\right]\left(\left[\frac{x}{d}\right]+1\right) = \frac{1}{2}\sum_{d\le x}\mu(d)\left(\left(\frac{x}{d}\right)^2 + O\left(\frac{x}{d}\right)\right)\\ &= \frac{x^2}{2}\sum_{d\le x}\frac{\mu(d)}{d^2} + O\left(x\sum_{d\le x}\frac{1}{d}\right) = \frac{3x^2}{\pi^2} + O(x\log x).\end{aligned}$$

This completes the proof. □

Theorem 3.4.12. *The probability that two positive integers are relatively prime is $6/\pi^2$.*

Proof. Let $N \geq 1$. The number of ordered pairs of positive integers (m, n) such that $1 \leq m \leq n \leq N$ is equal to $N(N+1)/2$. The number of positive integers $m \leq n$ that are relatively prime is $\varphi(n)$, and so the number of pairs of positive integers (m, n) such that $1 \leq m \leq n \leq N$ and m and n are relatively prime is

$$\sum_{n \leq N} \varphi(n) = \frac{3N^2}{\pi^2} + O(N \log N).$$

Therefore, the frequency of relatively prime pairs of positive integers not exceeding N is

$$\frac{\frac{3N^2}{\pi^2} + O(N \log N)}{N(N+1)/2} = \frac{6}{\pi^2} + O\left(\frac{\log N}{N}\right) \to \frac{6}{\pi^2}$$

as $N \to \infty$. This completes the proof. □

§3.5. Inequalities for $\pi(n)$ and p_n

Theorem 3.5.1. *For every integer $n \geq 2$, we have*

$$\frac{1}{6}\frac{n}{\log n} < \pi(n) < 6\frac{n}{\log n}.$$

Proof. Omitted.

□

Using Theorem 3.5.1 we can give upper and lower bounds on the size of the nth prime as follows.

Theorem 3.5.2. *For $n \geq 1$ the nth prime p_n satisfies the inequalities*

$$\frac{1}{6} n \log n < p_n < 12(n \log n + n \log \frac{12}{e}).$$

Proof. □

§3.6 Principle of cross-classification and its generalization

First we state the principle of cross-classification. Since it is a special case of our extension, we omit the details of the proof.

Theorem 3.6.1. *Let R be any given finite set. For a subset T of R, we denote by $\bar{T}$ the set of those elements of R which are not in T, i.e., $\bar{T} = R \backslash T$. If $R_1, ..., R_m$ are given m distinct subsets of R, then we have*

$$|\cap_{i=1}^m \bar{R}_i| = |R| + \sum_{t=1}^m (-1)^t \sum_{1 \leq i_1 < ... < i_t \leq m} |\cap_{j=1}^t R_{i_j}|.$$

Consequently we give our extension of the cross-classification principle. In the following, for a subset T of R, we denote by $\bar{T}$ the set $R \setminus T$ of the elements of R that are not in T. Moreover $\cap_{j \in \phi} R_j$ is meant to be R.

Theorem 3.6.2. (Hong, 1996) *Let R be any given finite set and f any complex-valued function defined on R. For a subset T of R, we denote by $\bar{T}$ the set of those elements of R which are not in T, i.e., $\bar{T} = R \backslash T$. If $R_1, ..., R_m$ are given m distinct subsets of R, then*

$$\sum_{x \in \cap_{i=1}^m \bar{R}_i} f(x) = \sum_{x \in R} f(x) + \sum_{t=1}^m (-1)^t \sum_{1 \leq i_1 < ... < i_t \leq m} \sum_{x \in \cap_{j=1}^t R_{i_j}} f(x). \tag{3.11}$$

Proof. We adapte the proof given in [S. Hong, Linear Algebra Appli. 345 (2002) 225-233]. We use induction on m to prove Lemma 2.1. Let M denote the left-hand side of Equation (3.11), and let N denote the right-hand side of Equation (3.11).

If $m=1$, then the result is clear. Now let $m=2$. Then

$$\begin{aligned} M &= \sum_{x\in\bar R_1\cap\bar R_2} f(x) = \sum_{x\in\overline{R_1\cup R_2}} f(x) \\ &= \sum_{x\in R} f(x) - \sum_{x\in R_1\cup R_2} f(x) \\ &= \sum_{x\in R} f(x) - \sum_{x\in R_1} f(x) - \sum_{x\in R_2} f(x) + \sum_{x\in R_1\cap R_2} f(x) = N, \end{aligned}$$

as desired. In what follows let $m\ge 3$. Assume that Lemma 2.1 is true for the case $m-1$. Now consider the case m. We have

$$M = \sum_{x\in(\cap_{i=1}^{m-1}\bar R_i)\cap\bar R_m} f(x) = \sum_{x\in\overline{\cup_{i=1}^{m-1}R_i}\cap\bar R_m} f(x).$$

It then follows from the result for the case $m=2$ that

$$\begin{aligned} M &= \sum_{x\in R} f(x) - \sum_{x\in\cup_{i=1}^{m-1}R_i} f(x) - \sum_{x\in R_m} f(x) + \sum_{x\in(\cup_{i=1}^{m-1}R_i)\cap R_m} f(x) \\ &= \sum_{x\in\overline{\cup_{i=1}^{m-1}R_i}} f(x) - \sum_{x\in R_m} f(x) + \sum_{x\in\cup_{i=1}^{m-1}(R_i\cap R_m)} f(x) \\ &= \sum_{x\in\overline{\cup_{i=1}^{m-1}R_i}} f(x) - \sum_{x\in R_m} f(x) + \sum_{x\in R} f(x) - \sum_{x\in\overline{\cup_{i=1}^{m-1}R_i\cap R_m}} f(x) \\ &= \sum_{x\in\cap_{i=1}^{m-1}\bar R_i} f(x) - \sum_{x\in R_m} f(x) + \sum_{x\in R} f(x) - \sum_{x\in\cap_{i=1}^{m-1}\overline{R_i\cap R_m}} f(x). \end{aligned} \tag{3.12}$$

By the inductive hypothesis, we have

$$\sum_{x\in\cap_{i=1}^{m-1}\bar R_i} f(x) = \sum_{x\in R} f(x) + \sum_{t=1}^{m-1}(-1)^t \sum_{1\le i_1<\ldots<i_t\le m-1}\ \sum_{x\in\cap_{j=1}^{t}R_{i_j}} f(x) \tag{3.13}$$

and

$$\sum_{x\in\cap_{i=1}^{m-1}\overline{R_i\cap R_m}} f(x) = \sum_{x\in R} f(x) + \sum_{t=1}^{m-1}(-1)^t \sum_{1\le i_1<\ldots<i_t\le m-1}\ \sum_{x\in\cap_{j=1}^{t}R_{i_j}\cap R_m} f(x). \tag{3.14}$$

It follows from Equations (3.12)-(3.14) that $M=N$. Thus the result for the case m is also true. The proof of Theorem 3.6.2 is complete. □

Finally, we give some applications of the principle of cross-classification.

Theorem 3.6.3. *Let $r_1,\ldots,r_m$ be any m given positive integers. Then we have*

$$\max(r_1,\ldots,r_m) = \sum_{i=1}^m r_i + \sum_{t=2}^m (-1)^{t-1} \sum_{1\le i_1<\ldots<i_t\le m} \min(r_{i_1},\ldots,r_{i_t}).$$

Proof. Pick any fixed integer $r>\max(r_1,\ldots,r_m)$. Let R be the set of all positive integers no more than r. Then $|R|=r$. For all $1\le i\le m$, let R_i be the set of all positive integers no more than r_i. Then $|\cap_{j=1}^t R_{i_j}| = \min(r_{i_1},\ldots,r_{i_t})$ and

$|\cap_{i=1}^m \bar{R}_i| = r - \max(r_1, ..., r_m)$. So by the cross-classification principle (Theorem 3.6.1), we have

$$r - \max(r_1, ..., r_m) = r - \sum_{i=1}^m r_i + \sum_{t=2}^m (-1)^t \sum_{1 \le i_1 < ... < i_t \le m} \min(r_{i_1}, ..., r_{i_t}).$$

The desired result then follows immediately. □

Theorem 3.6.4. *Let $a_1, a_2, \ldots, a_m$ be any m positive integers. Then we have*

$$\operatorname{lcm}(a_1, a_2, \ldots, a_m) = a_1 a_2 \ldots a_m \cdot \prod_{t=2}^m \prod_{1 \le i_1 < ... < i_t \le m} (\gcd(a_{i_1}, ..., a_{i_t}))^{(-1)^{t-1}}.$$

Proof. For any prime p and for $1 \le i \le m$, let $r_i = v_p(a_i)$. Then noting the following identities

$$\operatorname{lcm}(a_1, a_2, \ldots, a_m) = \prod_{p \text{ prime}} p^{\max(v_p(a_1), \ldots, v_p(a_m))}$$

and

$$\gcd(a_{i_1}, ..., a_{i_t}) = \prod_{p \text{ prime}} p^{\min(v_p(a_{i_1}), \ldots, v_p(a_{i_t}))},$$

Theorem 3.6.4 follows from Theorem 3.6.3. □

Similarly, we have the following dual results.

Theorem 3.6.5. *Let $r_1, ..., r_m$ be any m given positive integers. Then we have*

$$\min(r_1, ..., r_m) = \sum_{i=1}^m r_i + \sum_{t=2}^m (-1)^{t-1} \sum_{1 \le i_1 < ... < i_t \le m} \max(r_{i_1}, ..., r_{i_t}).$$

Proof. Pick any fixed integer $r > \max(r_1, ..., r_m)$. Let R be the set of all positive integers no more than r. Then $|R| = r$. For all $1 \le i \le m$, let R_i be the set of all positive integers no less than r_i. Then $|\cap_{j=1}^t R_{i_j}| = r - \max(r_{i_1}, ..., r_{i_t})$ and $|\cap_{i=1}^m \bar{R}_i| = \min(r_1, ..., r_m)$. So by the cross-classification principle (Theorem 3.6.1), we obtain that

$$\min(r_1, ..., r_m) = r + \sum_{t=1}^m (-1)^t \sum_{1 \le i_1 < ... < i_t \le m} (r - \max(r_{i_1}, ..., r_{i_t}))$$

$$= r + \sum_{t=1}^m (-1)^t \sum_{1 \le i_1 < ... < i_t \le m} r + \sum_{t=1}^m (-1)^{t-1} \sum_{1 \le i_1 < ... < i_t \le m} \max(r_{i_1}, ..., r_{i_t})$$

$$= r(1 + \sum_{t=1}^m (-1)^t \sum_{1 \le i_1 < ... < i_t \le m} 1) + \sum_{t=1}^m (-1)^{t-1} \sum_{1 \le i_1 < ... < i_t \le m} \max(r_{i_1}, ..., r_{i_t})$$

$$= r(1 + \sum_{t=1}^m (-1)^t \cdot \begin{pmatrix} m \\ t \end{pmatrix}) + \sum_{t=1}^m (-1)^{t-1} \sum_{1 \le i_1 < ... < i_t \le m} \max(r_{i_1}, ..., r_{i_t})$$

$$= r \cdot (1-1)^m + \sum_{t=1}^m (-1)^{t-1} \sum_{1 \le i_1 < ... < i_t \le m} \max(r_{i_1}, ..., r_{i_t})$$

$$= \sum_{t=1}^m (-1)^{t-1} \sum_{1 \le i_1 < ... < i_t \le m} \max(r_{i_1}, ..., r_{i_t}).$$

Theorem 3.6.5 is proved. □

Theorem 3.6.6. *Let $a_1, a_2, \dots, a_m$ be any m positive integers. Then we have*

$$\gcd(a_1, a_2, \dots, a_m) = a_1 a_2 \dots a_m \cdot \prod_{t=2}^{m} \prod_{1\le i_1<\dots<i_t\le m} (\mathrm{lcm}(a_{i_1}, ..., a_{i_t}))^{(-1)^{t-1}}.$$

Proof. For any prime p and for $1 \le i \le m$, let $r_i = v_p(a_i)$. Since

$$\gcd(a_1, a_2, \dots, a_m) = \prod_{p \text{ prime}} p^{\min(v_p(a_1),\dots,v_p(a_m))}$$

and

$$\mathrm{lcm}(a_{i_1}, ..., a_{i_t}) = \prod_{p \text{ prime}} p^{\max(v_p(a_{i_1}),\dots,v_p(a_{i_t}))},$$

Theorem 3.6.6 then follows from Theorem 3.6.5. □

Exercises for Chapter 3

(1). Define the arithmetical function $1(n)$ by $1(n) = 1$ for all n. Show that $(1 * 1)(n) = d(n)$.

(2). Let f and g be arithmetical functions. Show that $f * g = 0$ if and only if $f = 0$ or $g = 0$. It follows that the ring of arithmetical functions is an integral domain.

(3). Let $\mathcal{A}$ be the ring of complex-valued arithmetical functions. An arithmetical function f is called a *unit* in $\mathcal{A}$ if there exists an arithmetical function g such that $f * g = \delta$. Prove that $f \in \mathcal{A}$ is a unit if and only if $f(1) \neq 0$.

(4). Let f and g be arithmetical functions. Show that

$$L^n(f * g) = \sum_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} L^{n-k} f * L^k g.$$

(5). For arithmetical functions f and g, define the product $f \star g$ by

$$(f \star g)(n) = \sum_{k=1}^{n-1} f(k) g(n-k).$$

Is this product commutative? Is it associative? What is $f \star \delta$?

(6). Show that $e\left(\frac{n}{e}\right)^n < n! < en\left(\frac{n}{e}\right)^n$.

(7). For $0 < a < 1$, let $\gamma(a) = \frac{a}{1-a} + a\int_1^\infty \frac{\{t\}}{t^{a+1}}dt$. Show that

$$\sum_{n\le x} \frac{1}{n^a} = \frac{x^{1-a}}{1-a} - \gamma(a) + O(x^{-a}).$$

(8). Show that $\sum_{n\le x} \log\frac{x}{n} = x + O(\log x)$.

(9). Let $1 \le a < b$ and $k \ge 2$ be integers. Show that

$$\sum_{n=a}^{b} \frac{1}{n^k} = \frac{1}{k-1}\left(\frac{1}{b^{k-1}} - \frac{1}{a^{k-1}}\right) + O\left(\frac{1}{a^k}\right).$$

(10). Compute $\mu(k)$ for $11 \le k \le 60$.

(11). Let $d(n)$ be the divisor function. Show that for every integer n, we have $\sum_{k|n} d(k)\mu(\frac{n}{k}) = 1$.

(12). Define the von Mangoldt function

$$\Lambda(n)=\begin{cases} \log p & \text{if } n=p^k \text{ is a prime power,}\\ 0 & \text{otherwise.}\end{cases}$$

Let $L(n)=\log n$. Show that $L=1*\Lambda$ and $\Lambda(n)=-\sum_{d|n}\mu(d)\log d$.

(13). Let f be a multiplicative function. Show that if $f(1)=0$, then f is identically equal to 0, that is $f(n)=0$ for all n. Show that if f is not identically equal to 0, then $f(1)=1$.

(14). Show that Liouville's function $\lambda(n)=(-1)^{\Omega(n)}$ is completely multiplicative. Show that

$$\sum_{d|n}\lambda(d)=\begin{cases} 1 & \text{if } n \text{ is a square,}\\ 0 & \text{otherwise.}\end{cases}$$

(15). Let $\sigma(n)$ denote the sum of the positive divisors of n, that is, $\sigma(n)=\sum_{d|n}d$. Show that for every positive integer n, we have $\sum d|n\sigma(d)\mu(n/d)=n$

(16). Show that for every $\delta>0$, we have $\lim_{n\to\infty}\frac{\varphi(n)}{n^{1-\delta}}=\infty$.

(17). Show that

$$\prod_{p|n}\left(1-\frac{1}{p^2}\right)\geq\prod_{k=2}^{n}\left(1-\frac{1}{k^2}\right)>\frac{1}{2}.$$

(18). Show that $\frac{1}{2}<\frac{\varphi(n)\sigma(n)}{n^2}<1$.

(19). Show that for every $\delta>0$ we have $n<\sigma(n)\ll n^{1+\delta}$.

(20). Use Möbius inversion to prove identity: $\varphi(n)=n\sum_{d|n}\frac{\mu(d)}{d}$.

(21). Show that $\limsup_{n\to\infty}\frac{\varphi(n)}{n}=1$.

(22). (Smith's determinant, 1875) Let f be an arithmetical function and $(f(\gcd(i,j)))$ denote the $n\times n$ matrix having f evaluated at the greatest common divisor $\gcd(i,j)$ of i and j as its (i,j)-entry, where $1\leq i,j\leq n$. Show that $\det(f\,(\gcd(i,j)))=\prod_{k=1}^{n}(f*\mu)(k)$.

(23). (Smith's determinant, 1875) Let f be an arithmetical function and $(f(\mathrm{lcm}(i,j)))$ denote the $n\times n$ matrix having f evaluated at the least common multiple $\mathrm{lcm}(i,j)$ of i and j as its (i,j)-entry, where $1\leq i,j\leq n$. Show that $\det(f\,(\mathrm{lcm}(i,j)))=\prod_{k=1}^{n}f^2(k)(\frac{1}{f}*\mu)(k)$.

(24). Let m be any given integer. Define the arithmetical function f for any positive integer n by $f(n)=\sum_{d|n}\frac{\mu(d)}{d(m,\frac{n}{d})}$. Show that each of the following is true:

(i). f is a multiplicative function;

(ii). We have $f(n)=\frac{\varphi(n)}{n(n,m)}\cdot\delta_{n,m}$, where

$$\delta_{n,m}=\begin{cases} 0, & \text{if } \exists \text{ a prime } p|n, \text{ such that } v_p(n)\leq v_p(m);\\ 1, & \text{if for all prime } p|n,\ v_p(n)>v_p(m).\end{cases}$$

(25). Define the arithmetical function $\tilde{\mu}$ as follows:

$$\tilde{\mu}(n)=\begin{cases} \mu(\sqrt{n}), & \text{if } n \text{ is a square;}\\ 0, & \text{otherwise.}\end{cases}$$

Prove that $\tilde{\mu}$ is a multiplicative function.

(26). Define the arithmetical function $\bar{\mu}$ as follows:

$$\bar{\mu}(n)=\begin{cases} \mu(\sqrt[3]{n}), & \text{if } n \text{ is a cube;}\\ 0, & \text{otherwise.}\end{cases}$$

Prove that $\bar{\mu}$ is a multiplicative function.

(27). Let $d \geq 1$ be an integer. Define the arithmetical function $\tilde{\mu}_d$ as follows:

$$\tilde{\mu}_d(n) = \begin{cases} \mu(\sqrt[d]{n}), & \text{if } n \text{ is } d\text{−th power of an integer;} \\ 0, & \text{otherwise.} \end{cases}$$

Prove that $\tilde{\mu}_d$ is a multiplicative function.

(28). Let $d \geq 1$ be an integer and f be an arithmetical function. Define the arithmetical function $\tilde{f}_d$ as follows:

$$\tilde{f}_d(n) = \begin{cases} f(\sqrt[d]{n}), & \text{if } n \text{ is } d\text{−th power of an integer;} \\ 0, & \text{otherwise.} \end{cases}$$

Prove that each of the following is true:

(i). $\tilde{f}_d$ is a multiplicative function.

(ii). $\tilde{f}_d$ is not necessarily completely multiplicative if f is completely multiplicative.

(29). Let f be an arithmetical function and r be a given real number. Define the arithmetical function f^r for any positive integer n by $f^r(n) := f(n)^r$. Show that if f is multiplicative, then so is f^r.

(30). Let x be a real number. Show that each of the following is true:

(i). $[x] + [x + \frac{1}{2}] = [2x]$.

(ii). $[x] + [x + \frac{1}{3}] + [x + \frac{2}{3}] = [3x]$.

(iii). If $n \geq 1$ is an integer, then $\sum_{k=0}^{n-1}[x + \frac{k}{n}] = [nx]$.

(31). Let $f(x) = x - [x] - \frac{1}{2}$. Show that each of the following holds:

(i). $\sum_{k=0}^{n-1} f(x + \frac{k}{n}) = f(nx)$.

(ii). If $m \geq 1$ is an integer and x is a real number, then $|\sum_{n=1}^{m} f(2^n x + \frac{1}{2})| \leq 1$.

(32). Let $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$ be any $2n$ positive integers. Let $n \geq t \geq 3$ be a given integer. Show that if $\gcd(a_{i_1}, ..., a_{i_t}) = \gcd(b_{i_1}, ..., b_{i_t})$ for any $1 \leq i_1 < ... < i_t \leq n$, then we have

$$\frac{a_1 a_2 \cdots a_n}{\text{lcm}(a_1, a_2, \ldots, a_n)} \cdot \prod_{r=2}^{t-1} \prod_{1 \leq i_1 < \ldots < i_r \leq n} (\gcd(a_{i_1}, ..., a_{i_r}))^{(-1)^r}$$

$$= \frac{b_1 b_2 \cdots b_n}{\text{lcm}(b_1, b_2, \ldots, b_n)} \cdot \prod_{r=2}^{t-1} \prod_{1 \leq i_1 < \ldots < i_r \leq n} (\gcd(b_{i_1}, ..., b_{i_r}))^{(-1)^r}.$$

(33). Let $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$ be any $2n$ positive integers. Show that if for any $1 \leq i_1 < i_2 < i_3 \leq n$, we have $\gcd(a_{i_1}, a_{i_2}, a_{i_3}) = \gcd(b_{i_1}, b_{i_2}, b_{i_3})$, then

$$\frac{1}{\prod_{1 \leq i < j \leq n} \gcd(a_i, a_j)} \cdot \frac{a_1 a_2 \cdots a_n}{\text{lcm}(a_1, a_2, \ldots, a_n)} = \frac{1}{\prod_{1 \leq i < j \leq n} \gcd(b_i, b_j)} \cdot \frac{b_1 b_2 \cdots b_n}{\text{lcm}(b_1, b_2, \ldots, b_n)}.$$

(34). Let $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$ be any $2n$ positive integers. Show that if $\gcd(a_i, a_j) = \gcd(b_i, b_j)$ for any $1 \leq i < j \leq n$, then we have

$$\frac{a_1 a_2 \cdots a_n}{\text{lcm}(a_1, a_2, \ldots, a_n)} = \frac{b_1 b_2 \cdots b_n}{\text{lcm}(b_1, b_2, \ldots, b_n)}.$$

(35). Let $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$ be any $2n$ positive integers. Let $n \geq t \geq 3$ be a given integer. Show that if $\text{lcm}(a_{i_1}, ..., a_{i_t}) = \text{lcm}(b_{i_1}, ..., b_{i_t})$ for any

$1 \le i_1 < ... < i_t \le n$, then we have

$$\frac{a_1a_2\cdots a_n}{\gcd(a_1,a_2,\dots,a_n)}\cdot\prod_{r=2}^{t-1}\prod_{1\le i_1<\ldots<i_r\le n}(\mathrm{lcm}(a_{i_1},...,a_{i_r}))^{(-1)^r}$$
$$=\frac{b_1b_2\cdots b_n}{\gcd(b_1,b_2,\dots,b_n)}\cdot\prod_{r=2}^{t-1}\prod_{1\le i_1<\ldots<i_r\le n}(\mathrm{lcm}(b_{i_1},...,b_{i_r}))^{(-1)^r}.$$

(36). Let $a_1, a_2, \dots, a_n$ and $b_1, b_2, \dots, b_n$ be any $2n$ positive integers. Show that if for any $1 \le i_1 < i_2 < i_3 \le n$, we have $\mathrm{lcm}(a_{i_1}, a_{i_2}, a_{i_3}) = \mathrm{lcm}(b_{i_1}, b_{i_2}, b_{i_3})$, then

$$\frac{1}{\prod_{1\le i<j\le n}\mathrm{lcm}(a_i,a_j)}\cdot\frac{a_1a_2\cdots a_n}{\gcd(a_1,a_2,\dots,a_n)}=\frac{1}{\prod_{1\le i<j\le n}\mathrm{lcm}(b_i,b_j)}\cdot\frac{b_1b_2\cdots b_n}{\gcd(b_1,b_2,\dots,b_n)}.$$

(37). Let $a_1, a_2, \dots, a_n$ and $b_1, b_2, \dots, b_n$ be any $2n$ positive integers. Show that if $\mathrm{lcm}(a_i, a_j) = \mathrm{lcm}(b_i, b_j)$ for any $1 \le i < j \le n$, then we have

$$\frac{a_1a_2\cdots a_n}{\gcd(a_1,a_2,\dots,a_n)}=\frac{b_1b_2\cdots b_n}{\gcd(b_1,b_2,\dots,b_n)}.$$

(38). Let f be an arithmetic function defined by $f(n) = [\sqrt{n}] - [\sqrt{n-1}]$. Show that f is multiplicative but not completely multiplicative.

(39). Let f be an arithmetic function defined by $f(n) = [\sqrt[3]{n}] - [\sqrt[3]{n-1}]$. Show that f is multiplicative but not completely multiplicative.

(40). Let $d \ge 2$ be an integer and f be an arithmetic function defined by $f(n) := [\sqrt[d]{n}] - [\sqrt[d]{n-1}]$. Show that f is multiplicative but not completely multiplicative.

(41). Let f be a multiplicative function. Show that:

(i). $f^{-1}(n) = \mu(n)f(n)$ for every squarefree n.

(ii). $f^{-1}(p^2) = f(p)^2 - f(p^2)$ for every prime p.

(42). Show that there is a multiplicative arithmetical function g such that

$$\sum_{k=1}^{n} f(\gcd(k,n)) = \sum_{d|n} f(d)g\Big(\frac{n}{d}\Big)$$

for every arithmetical function f. Using this identity to prove that

$$\sum_{k=1}^{n}\gcd(k,n)\mu(\gcd(k,n)) = \mu(n).$$

Chapter 4. Quadratic Reciprocity

§4.1. Quadratic Residues

Let p be an odd prime and a an integer not divisible by p. Then a is called a *quadratic residue modulo* p if there exists an integer x such that

$$x^2 \equiv a \pmod{p}. \tag{4.1}$$

If this congruence has no solution, then a is called a *quadratic nonresidue modulo* p. Thus, an integer a is a quadratic residue modulo p if and only if $(a,p)=1$ and a has a square root modulo p. We can show that exactly half the congruence classes relatively prime to p have square roots modulo p. We define the *Legendre symbol* for the odd prime p as follows.

Definition. For any integer a,

$$\left(\frac{a}{p}\right) = \begin{cases} 1 & \text{if } (a,p)=1 \text{ and } a \text{ is a quadratic residue modulo } p, \\ -1 & \text{if } (a,p)=1 \text{ and } a \text{ is a quadratic nonresidue modulo } p, \\ 0 & \text{if } p \text{ divides } a. \end{cases}$$

The solvability of congruence (4.1) depends only on the congruence class of a (mod p), that is,

$$\left(\frac{a}{p}\right) = \left(\frac{b}{p}\right) \text{ if } a \equiv b \pmod{p},$$

and so the Legendre symbol is a well-defined function on the congruence classes $\mathbf{Z}/p\mathbf{Z}$. We observe that if p is an odd prime, then the only solutions of the congruence $x^2 \equiv 1 \pmod{p}$ are $x \equiv \pm 1 \pmod{p}$. Moreover, we have

Fact. if $\varepsilon, \varepsilon' \in \{-1,0,1\}$ and $\varepsilon \equiv \varepsilon' \pmod{p}$, then p divides $\varepsilon - \varepsilon'$, and so $\varepsilon = \varepsilon'$. In particular, if $\left(\frac{a}{p}\right) \equiv \varepsilon \pmod{p}$, then $\left(\frac{a}{p}\right) = \varepsilon$.

Theorem 4.1.1. (Euler Criterion) *Let p be an odd prime. For every integer a, we have*

$$\left(\frac{a}{p}\right) \equiv a^{(p-1)/2} \pmod{p}.$$

Proof. If p divides a, then both sides of the congruence are 0. Suppose now that p does not divide a.

Case 1: $\left(\frac{a}{p}\right) = 1$. Then there exists x_0 with $(p, x_0) = 1$ such that $x_0^2 \equiv a \pmod{p}$. So we have $a^{\frac{p-1}{2}} \equiv x_0^{p-1}$. But by Fermat's Little Theorem and $(p,x_0)=1$ we derive that $x_0^{p-1} \equiv 1 \pmod{p}$. Hence the desired result $a^{(p-1)/2} \equiv 1 \pmod{p}$ follows immediately.

Case 2: $\left(\frac{a}{p}\right) = -1$. Take $i_1 \in S := \{1,2,...,p-1\}$. Then there is an unique $j_1 \in S$ such that $i_1 \neq j_1$ and $i_1 j_1 \equiv a \pmod{p}$ because $x^2 \equiv a \pmod{p}$ has no solution modulo p. Take $i_2 \in S \setminus \{i_1, j_1\}$. Then there is an unique $j_2 \in S$ such that $j_2 \neq i_2, i_1, j_1$ and $i_2 j_2 \equiv a \pmod{p}$. Continues in this way. Finally we find $i_1, j_1, ..., i_{\frac{p-1}{2}}, j_{\frac{p-1}{2}} \in S$ such that $\{i_l, j_l | 1 \le l \le \frac{p-1}{2}\} = S$ and for $1 \le l \le \frac{p-1}{2}$, we have

$$i_l j_l \equiv a \pmod{p}.$$

So we get

$$\prod_{l=1}^{\frac{p-1}{2}} i_l j_l \equiv a^{\frac{p-1}{2}} \pmod{p}.$$

Namely

$$(p-1)! \equiv a^{\frac{p-1}{2}} \pmod{p}.$$

On the other hand, Wilson's Theorem tells us that $(p-1)! \equiv -1 \pmod{p}$. So we obtain $a^{\frac{p-1}{2}} \equiv -1 \pmod{p}$ as required.

The proof of Theorem 4.1.1 is complete.

Example. 3 is a quadratic residue modulo the primes 11 and 13 and a quadratic nonresidue modulo the primes 17 and 19, because

$$\left(\frac{3}{11}\right) \equiv 3^5 \equiv 1 \pmod{11},$$

$$\left(\frac{3}{13}\right) \equiv 3^6 \equiv 1 \pmod{13},$$

$$\left(\frac{3}{17}\right) \equiv 3^8 \equiv -1 \pmod{17},$$

$$\left(\frac{3}{19}\right) \equiv 3^9 \equiv -1 \pmod{19}.$$

The next result states that the Legendre symbol is a completely multiplicative arithmetic function.

Theorem 4.1.2. *Let p be an odd prime. Then:*

(i). *If $a \equiv b \pmod{p}$, then $(\frac{a}{p}) = (\frac{b}{p})$;*

(ii). *$(\frac{\cdot}{p})$ is completely multiplicative. That is: For any integers a and b we have*

$$\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right);$$

(iii). $(\frac{a^2}{p}) = 1$.

Proof. We only prove part (ii) here. If p divides a or b, then p divides ab, and

$$\left(\frac{ab}{p}\right) = 0 = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right).$$

If p does not divide ab, then by Theorem 4.1.1, we have

$$\begin{aligned}\left(\frac{ab}{p}\right) &\equiv (ab)^{(p-1)/2} \pmod{p}\\ &\equiv a^{(p-1)/2}b^{(p-1)/2} \pmod{p}\\ &\equiv \left(\frac{a}{p}\right)\left(\frac{b}{p}\right) \pmod{p}.\end{aligned}$$

The result follows immediately from the observation that each side of this congruence is ±1.

Theorem 4.1.2 implies that the Legendre symbol $\left(\frac{\cdot}{p}\right)$ is completely determined by its values at -1,2 and odd primes q. If a is an integer not divisible by p, then we can write

$$a = \pm 2^{r_o} q_1^{r_1} q_2^{r_2} \cdots q_k^{r_k},$$

where $q_1, ..., q_k$ are distinct odd primes not equal to p. Then

$$\left(\frac{a}{p}\right) = \left(\frac{\pm 1}{p}\right)\left(\frac{2}{p}\right)^{r_0}\left(\frac{q_1}{p}\right)^{r_1}\cdots\left(\frac{q_k}{p}\right)^{r_k}.$$

We shall first determine the set of primes p for which -1 is a quadratic residue. By the following result, this depends only on the congruence class of p modulo 4.

Theorem 4.1.3. *Let p be an odd prime number. Then*

$$\left(\frac{-1}{p}\right) = \begin{cases} 1 & \textit{if } p \equiv 1 \pmod 4, \\ -1 & \textit{if } p \equiv 3 \pmod 4). \end{cases}$$

Equivalently,

$$\left(\frac{-1}{p}\right) = (-1)^{(p-1)/2}.$$

Proof. We observe that

$$(-1)^{(p-1)/2} = \begin{cases} 1 & \textit{if } p \equiv 1 \pmod 4, \\ -1 & \textit{if } p \equiv 3 \pmod 4). \end{cases}$$

Applying Theorem 4.1.1 with $a = -1$, we obtain

$$\left(\frac{-1}{p}\right) \equiv (-1)^{(p-1)/2} \pmod p.$$

Again, the theorem follows immediately from the observation that both sides of this congruence are ±1.

Definition. Let p be an odd prime, and let S be a set of $(p-1)/2$ integers. We call S a *Gaussian set* modulo p if $S \cup -S = S \cup \{-s : s \in S\}$ is a reduced system of residues modulo p.

Equivalently, S is a Gaussian set if for every integer a not divisible by p, there exist $s \in S$ and $\varepsilon \in \{1, -1\}$ such that $a \equiv \varepsilon s \pmod p$. Moreover, s and ε are uniquely determined by a. For example, the sets $\{1, 2, ..., (p-1)/2\}$ and $\{2, 4, 6, ..., p-1\}$ are Gaussian sets modulo p for every odd prime p. If S is a Gaussian set, $s, s' \in S$, and $s \equiv \pm s' \pmod p$, then $s = s'$.

Theorem 4.1.4. (Gauss lemma) *Let p be an odd prime, and a an integer not divisible by p. Let S be a Gaussian set modulo p. For every $s \in S$ there exist unique integers $u_a(s) \in S$ and $\varepsilon_a(s) \in \{1, -1\}$ such that*

$$as \equiv \varepsilon_a(s) u_a(s) \pmod p.$$

Moreover we have

$$\left(\frac{a}{p}\right) = \prod_{s \in S} \varepsilon_a(s) = (-1)^m,$$

where m is the number of $s \in S$ such that $\varepsilon_a(s) = -1$.

Proof. Since S is a Gaussian set, for every $s \in S$ there exist unique integers $u_a(s) \in S$ and $\varepsilon_a(s) \in \{1, -1\}$ such that

$$as \equiv \varepsilon_a(s) u_a(s) \pmod p.$$

Let $s, s' \in S$. If $u_a(s) = u_a(s')$, then

$$\begin{aligned} as' &\equiv \varepsilon_a(s') u_a(s') \equiv \varepsilon_a(s') u_a(s) \pmod p \\ &\equiv \varepsilon_a(s') \varepsilon_a(s) \varepsilon_a(s') u_a(s) \pmod p \\ &\equiv \pm as \pmod p. \end{aligned}$$

Dividing by a, we obtain

$$s' \equiv \pm s \pmod p,$$

and so $s' = s$. It follows that the map $u_a : S \to S$ is a permutation of S, and so

$$\prod_{s\in S} s = \prod_{s\in S} u_a(s).$$

Therefore,

$$\begin{aligned} a^{(p-1)/2}\prod_{s\in S} s &\equiv \prod_{s\in S} as \pmod{p} \\ &\equiv \prod_{s\in S} \varepsilon_a(s)u_a(s) \pmod{p} \\ &\equiv \prod_{s\in S} \varepsilon_a(s)\prod_{s\in S} u_a(s) \pmod{p} \\ &\equiv \prod_{s\in S} \varepsilon_a(s)\prod_{s\in S} s \pmod{p} \end{aligned}$$

Dividing by $\prod_{s\in S} s$, we obtain

$$\left(\frac{a}{p}\right) \equiv a^{(p-1)/2} \equiv \prod_{s\in S} \varepsilon_a(s) \pmod{p}.$$

The proof is completed by the observation that the left and right sides of this congruence are ±1. The proof is complete. □

We shall use Gauss's lemma to compute the Legendre symbol $\left(\frac{3}{11}\right)$. Let S be the Gaussian set {2,4,6,8,10}. We have

$$\begin{aligned} 3\cdot 2 &\equiv 6 \pmod{11}, \\ 3\cdot 4 &\equiv (-1)10 \pmod{11}, \\ 3\cdot 6 &\equiv (-1)4 \pmod{11}, \\ 3\cdot 8 &\equiv 2 \pmod{11}, \\ 3\cdot 10 &\equiv 8 \pmod{11}. \end{aligned}$$

The number of $s \in S$ with $\varepsilon_3(s) = -1$ is $m = 2$, and so $\left(\frac{3}{11}\right) = (-1)^2 = 1$, that is, 3 is a quadratic residue modulo 11. Indeed,

$$5^2 \equiv 6^2 \equiv 3 \pmod{11},$$

and so 5 and 6 are the square roots of 3 modulo 11.

Theorem 4.1.5. *Let p be an odd prime. Then*

$$\left(\frac{2}{p}\right) = \begin{cases} 1 & if\ p \equiv \pm 1 \pmod{8}, \\ -1 & if\ p \equiv \pm 3 \pmod{8}. \end{cases}$$

Equivalently,

$$\left(\frac{2}{p}\right) = (-1)^{(p^2-1)/8}.$$

Proof. We apply Gauss'lemma (Theorem 4.1.4) to the Gaussian set $S = \{1, 2, 3, ..., (p-1)/2\}$. Then

$$\{2s : s \in S\} = \{2, 4, 6, ..., p-1\},$$

and

$$\left(\frac{2}{p}\right) = (-1)^m,$$

where m is the number of integers $s \in S$ such that $\varepsilon_2(s) = -1$. If $1 \le 2s \le (p-1)/2$, then $2s \in S$, and so $u_2(s) = 2s$ and $\varepsilon_2(s) = 1$. If $(p+1)/2 \le 2s \le p-1$, then $1 \le p - 2s \le (p-1)/2$, and so $p - 2s \in S$. Since

$$2s \equiv -(p-2s) \pmod{p},$$

it follows that $u_2(s) = p - 2s$ and $\varepsilon_2(s) = -1$, or equivalently,

$$\frac{p+1}{4} \le s \le \frac{p-1}{2}. \tag{4.2}$$

Since every odd prime p is congruence to 1,3,5 or 7 modulo 8, there are four cases to consider.

(i). If $p \equiv 1 \pmod 8$, then $p = 8k+1$, and $s \in S$ satisfies (4.2) if and only if

$$2k + \frac{1}{2} \le s \le 4k,$$

and so $m = 2k$ and $\left(\frac{2}{p}\right) = (-1)^{2k} = 1$.

(ii). If $p \equiv 3 \pmod 8$, then $p = 8k+3$, and $s \in S$ satisfies (4.2) if and only if

$$2k + 1 \le s \le 4k + 1,$$

and so $m = 2k+1$ and $\left(\frac{2}{p}\right) = (-1)^{2k+1} = -1$.

(iii). If $p \equiv 5 \pmod 8$, then $p = 8k+5$ and so $s \in S$ satisfies (4.2) if and only if

$$2k + 1 + \frac{1}{2} \le s \le 4k + 2,$$

and so $m = 2k+1$ and $\left(\frac{2}{p}\right) = (-1)^{2k+1} = -1$.

(iv). If $p \equiv 7 \pmod 8$, then $p = 8k+7$ and so $s \in S$ satisfies (4.2) if and only if

$$2k + 2 \le s \le 4k + 3,$$

and so $m = 2k+2$ and $\left(\frac{2}{p}\right) = (-1)^{2k+2} = 1$.

Finally, we observe that

$$\frac{p^2-1}{8} \equiv 0 \pmod 2 \text{ if } p \equiv 1 \text{ or } 7 \pmod 8$$

and

$$\frac{p^2-1}{8} \equiv 1 \pmod 2 \text{ if } p \equiv 3 \text{ or } 7 \pmod 8.$$

This completes the proof. □

§4.2. Quadratic Reciprocity Law

Let p and q be distinct odd prime. If q is a quadratic residue modulo p, then the congruence

$$x^2 \equiv q \pmod p$$

is solvable. Similarly, if p is a quadratic residue modulo q, then the congruence

$$x^2 \equiv p \pmod q$$

is solvable. There is no obvious connection between these two congruences. One of the greatest discoveries of eighteenth-century mathematics is that there is, in fact, a subtle and powerful relation between them that depends only on the congruence classes of the primes p and q modulo 4. This is expressed in Gauss's celebrated *law of quadratic reciprocity.*

Theorem 4.2.1. (Gauss quadratic reciprocity law) *Let p and q be distinct odd primes. If $p \equiv 1 \pmod 4$ or $q \equiv 1 \pmod 4$, then p is a quadratic residue modulo q if and only if q is a quadratic residue modulo p. If $p \equiv q \equiv 3 \pmod 4$, then p is a quadratic residue modulo q if and only if q is a quadratic nonresidue modulo p. Equivalently,*

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}.$$

Proof. Let

$$S = \{1, 2, ..., (p-1)/2\}$$

and

$$T = \{1, 2, ..., (q-1)/2\}.$$

Then S is a Gaussian set for the prime p, and T is a Gaussian set for the prime q. Let

$$S \times T = \{(s,t) : s \in S, t \in T\}.$$

This is a rectangle of lattice points in R^2 of cardinality

$$|S \times T| = \frac{p-1}{2}\frac{q-1}{2}.$$

We shall count the number m of lattice points (s,t) in this rectangle that lie in the strip defined by the inequality

$$1 \leq pt - qs \leq \frac{p-1}{2}. \tag{4.3}$$

(To understand this proof, it is helpful to choose small primes, for example, $p = 17, q = 13$, and draw pictures of the rectangles $S \times T$ and the regions defined by inequality.)

If $s \in S, t_1, t_2 \in T$, and the lattice points (s, t_1) and (s, t_2) both satisfy (4.3), then

$$p|t_1 - t_2| = |(pt_1 - qs) - (pt_2 - qs)| < \frac{p-1}{2} < p,$$

and so $t_1 = t_2$. It follows that for every $s \in S$ there exists at most one $t \in T$ that satisfies (4.3). If this inequality holds for some $t \in T$, then $s' := pt - qs \in S$ and

$$qs \equiv -s' \pmod p.$$

Using the notation in Gauss's lemma (Theorem 4.1.4), we have $u_q(s) = s'$ and $\varepsilon_q(s) = -1$.

Conversely, if $s \in S$ and $\varepsilon_q(s) = -1$, then there exists $u_q(s) \in S$ such that

$$qs \equiv -u_q(s) \pmod p,$$

and there exists an integer t such that

$$qs = -u_q(s) + pt.$$

Since

$$0 < pt = qs + u_q(s) \leq \frac{q(p-1)}{2} + \frac{p-1}{2} = \frac{(q+1)(p-1)}{2},$$

it follows that

$$1 \leq t \leq \frac{(q+1)(p-1)}{2p} < \frac{q+1}{2}.$$

The prime q is odd, and so

$$1 \leq t \leq \frac{q-1}{2}.$$

Therefore $t \in T$ and the lattice point $(s, t) \in S \times T$ satisfies inequality (4.3). Thus, the number m of lattice points $(s, t) \in S \times T$ that satisfy inequality (4.3) is equal to the number of $s \in S$ such that $\varepsilon_q(s) = -1$. By Gauss's lemma,

$$\left(\frac{q}{p}\right) = (-1)^m.$$

Similarly,

$$\left(\frac{p}{q}\right) = (-1)^n,$$

where n is the number of lattice points $(s, t) \in S \times T$ such that

$$1 \leq qs - pt \leq \frac{q-1}{2},$$

or, equivalently,

$$-\frac{q-1}{2} \leq pt - qs \leq -1. \tag{4.4}$$

Since $pt - qs \neq 0$ for all $s \in S$ and $t \in T$, it follows that

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{m+n},$$

where $m + n$ is the number of lattice points $(s, t) \in S \times T$ such that

$$-\frac{q-1}{2} \leq pt - qs \leq \frac{p-1}{2}. \tag{4.5}$$

Let M denote the number of lattice points $(s, t) \in S \times T$ such that

$$pt - qs > \frac{p-1}{2}$$

and let N denote the number of lattice points $(s, t) \in S \times T$ such that

$$pt - qs < -\frac{q-1}{2}.$$

Then

$$m + n + M + N = |S \times T| = \frac{p-1}{2}\frac{q-1}{2}.$$

We define a map from the set $S \times T$ to itself by reflection:

$$(s, t) \mapsto (s', t'),$$

where

$$s' = \frac{p+1}{2} - s$$

and

$$t' = \frac{q+1}{2} - t.$$

This map is a bijection, since

$$\frac{p+1}{2} - s' = s$$

and

$$\frac{q+1}{2} - t' = t.$$

If $(s,t) \in S \times T$ and

$$pt - qs > \frac{p-1}{2},$$

then $(s',t') \in S \times T$ and

$$\begin{aligned} pt' - qs' &= p\left(\frac{q+1}{2} - t\right) - q\left(\frac{p+1}{2} - s\right) \\ &= \frac{p}{2} - pt - \frac{q}{2} + qs - (pt - qs) + \frac{p-1}{2} - \frac{q-1}{2} \\ &< -\frac{q-1}{2}. \end{aligned}$$

Therefore, $M \leq N$. Similarly, if $(s,t) \in S \times T$ and

$$pt - qs < -\frac{q-1}{2},$$

then $(s',t') \in S \times T$ and

$$pt' - qs' > \frac{p-1}{2},$$

and so $M \geq N$. Therefore $M = N$ and

$$\begin{aligned} \left(\frac{p}{q}\right)\left(\frac{q}{p}\right) &= (-1)^{m+n} = (-1)^{m+n+2M} \\ &= (-1)^{m+n+M+N} = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}. \end{aligned}$$

This completes the proof. □

The quadratic reciprocity law provides an effective method to calculate the value of the Legendre symbol. For example, since $7 \equiv 59 \equiv 3 \pmod 4$ and $59 \equiv 3 \pmod 7$, we have

$$\begin{aligned} \left(\frac{7}{59}\right) &= -\left(\frac{59}{7}\right) = -\left(\frac{3}{7}\right) \\ &= \left(\frac{7}{3}\right) = \left(\frac{1}{3}\right) = 1. \end{aligned}$$

Similarly, since $51 = 3 \cdot 17$ and $97 \equiv 17 \equiv 1 \pmod 4$, we have

$$\begin{aligned} \left(\frac{51}{97}\right) &= \left(\frac{3}{97}\right)\left(\frac{17}{97}\right) = \left(\frac{97}{3}\right)\left(\frac{97}{17}\right) \\ &= \left(\frac{1}{3}\right)\left(\frac{12}{17}\right) = \left(\frac{12}{17}\right) \\ &= \left(\frac{4}{17}\right)\left(\frac{3}{17}\right) = \left(\frac{3}{17}\right) \\ &= \left(\frac{17}{3}\right) = \left(\frac{2}{3}\right) = -1. \end{aligned}$$

Quadratic reciprocity also allows us to determine all primes p for which a given integer a is a quadratic residue. Here are some examples. If $a = 5$, then

$$\left(\frac{5}{p}\right) = \left(\frac{p}{5}\right) = \begin{cases} 1 & \text{if } p \equiv 1,4 \pmod 5, \\ -1 & \text{if } p \equiv 2,3 \pmod 5. \end{cases}$$

Let $a = 7$. If $p \equiv 1 \pmod 4$, then

$$\left(\frac{7}{p}\right) = \left(\frac{p}{7}\right) = \begin{cases} 1 & \text{if } p \equiv 1, 2, 4 \pmod 7, \\ -1 & \text{if} p \equiv 3, 5, 6 \pmod 7. \end{cases}$$

If $p \equiv 3 \pmod 4$,then

$$\left(\frac{7}{p}\right) = -\left(\frac{p}{7}\right) = \begin{cases} 1 & \text{if } p \equiv 3, 5, 6 \pmod 7, \\ -1 & \text{if } p \equiv 1, 2, 4 \pmod 7. \end{cases}$$

Equivalently,

$$\left(\frac{7}{p}\right) = \begin{cases} 1 & \text{if } p \equiv 1, 3, 9, 19, 25, 27 \pmod{28}, \\ -1 & \text{if } p \equiv 5, 11, 13, 15, 17, 23 \pmod{28}. \end{cases}$$

Let $a = 35$. Then

$$\left(\frac{35}{p}\right) = 1$$

if and only if

$$p \equiv 1, 4 \pmod 5 \text{ and } p \equiv 1, 3, 9, 19, 25, 27 \pmod{28}$$

or

$$p \equiv 2, 3 \pmod 5 \text{ and } p \equiv 5, 11, 13, 15, 17, 23 \pmod{28}).$$

This is equivalent to a set of congruence classes modulo 140.

§4.3. Quadratic Residues to Composite Moduli

Let m be an odd positive integer and a an integer relatively prime to m. If the congruence $x^2 \equiv a \pmod m$ is solvable, then a is called a *quadratic residue modulo* m. Otherwise, a is called a *quadratic nonresidue modulo* m. For example, the quadratic residues modulo 7 are 1,2, and 4; the quadratic nonresidues are 3,5 and 6. The only quadratic residue modulo 8 is 1, and the quadratic nonresidues modulo 8 are 3,5 and 7. We shall prove that a is a quadratic residue modulo m if and only if a is a quadratic residue modulo p for every prime p that divides m. The Chinese remainder theorem (see Theorem 2.4.3) implies that it suffices to consider congruence modulo prime powers.

We begin with *Hensel's lemma*, an important result that gives a sufficient condition that a polynomial congruence solvable modulo a prime p will also be solvable modulo p^k for every positive integer k.

Let

$$f(x) = a_n x^n + a_{n-1}^{n-1} + \cdots + a_1 x + a_0$$

be a polynomial with coefficients in a ring R. The *derivation* of $f(x)$ is the polynomial

$$f'(x) = na_n x^n + (n-1)a_{n-1}^{n-1} + \cdots + a_1.$$

If $f(x)$ is a polynomial of degree $n \geq 1$ with coefficients in the ring $\mathbf{Z}$, then the derivation $f'(x)$ has degree $n-1$ and leading coefficient na_n. For example, if $f(x) = x^3 - 5x + 1$, then $f'(x) = 3x^2 - 5$. Moreover,

$$\begin{aligned} f(x+h) &= (x+h)^3 - 5(x+h) + 1 \\ &= (x^3 + 3x^2 h + 3xh^2 + h^3) - (5x + 5h) + 1 \\ &= (x^3 - 5x + 1) + (3x^2 - 5)h + (3x + h)h^2 \\ &= f(x) + f'(x)h + r(x,h)h^2, \end{aligned}$$

where $r(x,h) = 3x + h$.

Theorem 4.3.1. *Let R be a ring and $f(x) = \sum_{i=0}^{n} a_i x^i$ a polynomial with coefficient in R. Then*

$$f(x+h) = f(x) + f'(x)h + r(x,h)h^2.$$

where r(x,h) is a polynomial in the two variables x and h with coefficients in R.

Proof. This is a standard calculation. Expanding $f(x+h)$ by the binomial theorem, we obtain

$$\begin{aligned} f(x+h) &= \sum_{i=0}^{n} a_i (x+h)^i \\ &= \sum_{i=0}^{n} a_i \sum_{j=0}^{i} \left(\begin{array}{c} i \\ j \end{array} \right) x^{i-j} h^j \\ &= \sum_{j=0}^{n} \sum_{i=j}^{n} \left(\begin{array}{c} i \\ j \end{array} \right) a_i x^{i-j} h^j \\ &= \sum_{i=0}^{n} a_i x^i + \sum_{i=1}^{n} i a_i x^{i-1} h + \sum_{j=2}^{n} \sum_{i=j}^{n} \left(\begin{array}{c} i \\ j \end{array} \right) a_i x^{i-j} h^j \\ &= f(x) + f'(x)h + r(x,h)h^2, \end{aligned}$$

where

$$r(x,h) = \sum_{j=2}^{n} \sum_{i=j}^{n} \left(\begin{array}{c} i \\ j \end{array} \right) a_i x^{i-j} h^{j-2}$$

is a polynomial in x and h with coefficients in R. □

Theorem 4.3.2. (Hensel's lemma) *Let p be a prime and $f(x) \in \mathbf{Z}[x]$ be a polynomial of degree n with leading coefficient not divisible by p. If there exists an integer x_1 such that $f(x_1) \equiv 0 \pmod{p}$ and $f'(x_1) \not\equiv 0 \pmod{p}$, then for every integer $k \geq 2$, there exists an integer x_k such that*

$$f(x_k) \equiv 0 \pmod{p^k} \tag{4.6}$$

and

$$x_k \equiv x_{k-1} \pmod{p^{k-1})}. \tag{4.7}$$

Proof. The proof is by induction on k. We begin by constructing x_2. There exist integers u_1 and v_1 such that $f(x_1) = u_1 p$ and $f'(x_1) = v_1 \neq 0 \pmod{p}$. We shall prove that there exists an integer y_1 such that $f(x_1 + y_1 p) \equiv 0 \pmod{p^2}$.

By Theorem 4.3.1, there exists a polynomial $r(x,h)$ with integer coefficients such that

$$\begin{aligned} f(x_1 + y_1 p) &= f(x_1) + f'(x_1) y_1 p + r(x_1, y_1 p) p^2 \\ &= u_1 p + v_1 y_1 p + r(x_1, y_1 p) p^2 \\ &\equiv u_1 p + v_1 y_1 p \pmod{p^2}. \end{aligned}$$

Therefore there exists an integer y_1 such that

$$f(x_1 + y_1 p) \equiv 0 \pmod{p^2}$$

if and only if the linear congruence

$$v_1 y \equiv -u_1 \pmod{p}$$

is solvable. We see that this congruence does not have a solution y_1 because $(v_1, p) = 1$. Let

$$x_2 = x_1 + y_1 p.$$

Then

$$f(x_2) \equiv 0 \pmod{p^2} \text{ and } x_2 \equiv x_1 \pmod{p}.$$

Let $k \geq 3$ and assume that we have constructed integers $x_2, \cdots, x_{k-1}$ such that

$$f(x_i) \equiv 0 \pmod{p^i} \text{ and } x_i \equiv x_{i-1} \pmod{p^{i-1}}.$$

for $i = 2, ..., k-1$. There exists an integer u_{k-1} such that

$$f(x_{k-1}) = u_{k-1} p^{k-1}.$$

Let $f'(x_{k-1}) = v_{k-1}$. Since $x_{k-1} \equiv x_1 \pmod{p}$, it follows that

$$v_{k-1} = f'(x_{k-1}) \equiv f'(x_1) \neq 0 \pmod{p}.$$

Applying Theorem 4.3.1 with $t = x_{k-1}$ and $h = y_{k-1}p^{k-1}$, we obtain

$$\begin{aligned} f(x_{k-1} + y_{k-1}p^{k-1}) &= f(x_{k-1}) + f'(x_{k-1})y_{k-1}p^{k-1} + r(x_{k-1}, y_{k-1}p^{k-1})y_{k-1}^2 p^{2k-2} \\ &\equiv u_{k-1}p^{k-1} + v_{k-1}y_{k-1}p^{k-1} \pmod{p^k}. \end{aligned}$$

It follows that

$$f(x_{k-1} + y_{k-1}p^{k-1}) \equiv 0 \pmod{p^k}$$

if and only if there exists an integer y_{k-1} such that

$$v_{k-1}y_{k-1} \equiv -u_{k-1} \pmod{p}.$$

This last congruence is solvable, since $(v_{k-1}, p) = 1$, and the integer $x_k = x_{k-1} + y_{k-1}p^{k-1}$ satisfies conditions (4.6) and (4.7).

Theorem 4.3.3. *Let p be an odd prime, and let a be an integer not divisible by p. If a is a quadratic residue modulo p, then a is a quadratic residue modulo p^k for every $k \geq 1$.*

Proof. Consider the polynomial $f(x) = x^2 - a$ and its derivative $f'(x) = 2x$. If a is a quadratic residue modulo p, then there exists an integer x_1 such that $x_1 \not\equiv 0 \pmod{p}$ and $x_1^2 \equiv a \pmod{p}$. Then $f(x_1) \equiv 0 \pmod{p}$ and $f'(x_1) \not\equiv 0 \pmod{p}$. By Hensel's lemma, the polynomial congruence $f(x) \equiv 0 \pmod{p^k}$ is solvable for every $k \geq 1$, and so a is a quadratic residue modulo p^k for every $k \geq 1$.

So we get the following result immediately.

Theorem 4.3.4. *Let $(a, m) = 1$. Then a is a quadratic residue modulo m if and only if a is a quadratic residue modulo p for every prime p that divides m.*

§4.4. Pell's equations

In this section, we solve Pell's equation $x^2 - dy^2 = 1$. To do so, we need the following preliminary result.

Proposition 4.4.1. *Let d be a positive square-free integer. Then there are infinitely many pairs (x, y) of integers such that $|x^2 - dy^2| < 1 + 2\sqrt{d}$.*

Proof. First, Theorem 1.6.3 applied to $\xi = \sqrt{d}$ gives us the existence of infinitely many pairs (x, y) of integers with $y > 0$ such that $|x - \sqrt{d}y| < \frac{1}{y}$. So we deduce that

$$|x + \sqrt{d}y| \leq |x - \sqrt{d}y| + 2\sqrt{d}|y| < \frac{1}{y} + 2\sqrt{d}y$$

It then follows that

$$|x^2 - dy^2| = |x + \sqrt{d}y||x - \sqrt{d}y)| < |x + \sqrt{d}y|\frac{1}{y} \leq 2\sqrt{d} + \frac{1}{y^2} \leq 1 + 2\sqrt{d}$$

as desired. Thus Proposition 4.4.1 is proved. □

Theorem 4.4.1. *Let d be a positive square-free integer. Then $x^2 - dy^2 = 1$ has infinitely many integral solutions. Furthermore, there is a solution (x_1, y_1) such that each solution is of the form (x_n, y_n) where $x_n + \sqrt{d}y_n = (x_1 + \sqrt{d}y_1)^n, n \in \mathbf{Z}$.*

Proof. Omitted. □

Exercises for Chapter 4

1. Find all solutions of the congruences $x^2 \equiv 2 \pmod{47}$ and $x^2 \equiv 2 \pmod{53}$.

2. Show that $S = \{3, 4, 5, 9, 10\}$ is a Gaussian set modulo 11. Apply Gauss's lemma to this set to compute the Legendre symbols $\left(\frac{3}{11}\right)$ and $\left(\frac{7}{11}\right)$.

3. Let p be an odd prime. Show that $\{2, 4, 6, ..., p-1\}$ is a Gaussian set modulo p.

4. Use Theorems 4.1.3 and 4.1.5 to find all primes p for which -2 is a quadratic residue.

5. Let p and q be distinct odd prime numbers. Show that

$$\sum_{x_1+...+x_q \equiv q \pmod{p}, 1 \leq x_i \leq p-1} \left(\frac{x_1...x_q}{p}\right) \equiv 1 \pmod{q},$$

where the sum is over all ordered q-tuples of integers $(x_1, ..., x_q)$ such that $x_1 + ... + x_q \equiv q \pmod{p}$ and $1 \leq x_i \leq p-1$ for $i = 1, ..., q$.

6. Use quadratic reciprocity to compute $\left(\frac{19}{101}\right)$. Find an integer x such that $x^2 \equiv 19 \pmod{101}$.

7. Show that the congruence

$$(x^2 - 2)(x^2 - 19)(x^2 - 38) \equiv 0 \pmod{p}$$

has a solution for every prime number p.

8. Let p_1 and p_2 be any two distinct primes. Show that for any prime p, the congruence equation

$$(x^2 - p_1)(x^2 - p_2)(x^2 - p_1p_2) \equiv 0 \pmod{p}$$

always has a solution.

9. Find all primes for which -5 is a quadratic residue.

10. Let p be a prime, $p \neq 3$, and let a be an integer not divisible by p. If the congruence $x^3 \equiv a \pmod{p}$ is solvable, then a is called a *cubic residue modulo p*. Show that if a is a cubic residue modulo p, then a is a cubic residue modulo p^k for every $k \geq 1$.

11. Denote the derivative of the polynomial $f(x)$ by $D(f)(x) = f'(x)$. We define

$$D^{(0)}(f)(x) = f(x),$$

$$D^{(k)}(f)(x) = D(D^{(k-1)}(f))(x) \text{ for } k \geq 1.$$

The polynomial $D^{(k)}(f)$ is called the *kth derivative* of f. Show that if $f(x)$ is a polynomial with integer coefficients, then $D^{(k)}(f)(x) = 0$ if and only if the degree of $f(x)$ is at most $k-1$.

12. Let $f(x)$ and $g(x)$ be polynomials. Show the Leibniz formula:

$$D(f\cdot g)(x)=f(x)\cdot D(g)(x)+D(f)(x)\cdot g(x).$$

13. Let $f(x)$ be a polynomial of degree n. Show Taylor's formula:

$$f(x+h)=\sum_{k=0}^{n}\frac{D^{(k)}(x)}{k!}h^k.$$

14. Let p be an odd prime such that $p\equiv 3 \pmod 4$. Show that each of the following is true:

(i).

$$\sum_{r=1}^{p-1}r^2\left(\frac{r}{p}\right)=p\sum_{r=1}^{p-1}r\left(\frac{r}{p}\right).$$

(ii). Let $q=\frac{p-1}{2}$. Then

$$2\Big(1-2\Big(\frac{2}{p}\Big)\Big)\sum_{r=1}^{q}r\Big(\frac{r}{p}\Big)=p\Big(1-\Big(\frac{2}{p}\Big)\Big)\sum_{r=1}^{q}\Big(\frac{r}{p}\Big).$$

15.

Chapter 5. Primitive Roots

§5.1. Polynomials and Primitive Roots

Let m be a positive integer than 1, and a an integer relatively prime to m. The *order* of a modulo m, denoted by $\nu_m(a)$, is the smallest positive integer d such that $a^d \equiv 1 \pmod{m}$. By Theorem 2.14, $ord_m(a)$ is a divisor of the Euler phi function $\varphi(m)$. The order of a modulo m is also called the *exponent* of a modulo m.

We investigate the least nonnegative residues of the powers of a modulo m. For example, if $m = 7$ and $a = 2$, then

$$2^0 \equiv 1 \pmod{7}, 2^1 \equiv 2 \pmod{7}, 2^2 \equiv 4 \pmod{7}, 2^3 \equiv 1 \pmod{7},$$

and 2 has order 3 modulo 7. If $m = 7$ and $a = 3$, then

$$3^0 \equiv 1 \pmod{7}, 3^1 \equiv 3 \pmod{7}, 3^2 \equiv 2 \pmod{7}, 3^3 \equiv 6 \pmod{7},$$

$$3^4 \equiv 4 \pmod{7}, 3^5 \equiv 5 \pmod{7}, 3^6 \equiv 1 \pmod{7},$$

and 3 has order 6 modulo 7. The powers of 3 from a reduced residue system modulo 7.

The integer a is called a *primitive root of modulo* m if a has order $\varphi(m)$. In this case, the $\varphi(m)$ integers $1, a, a^2, ..., a^{\varphi(m)-1}$ are relatively prime to m and are pairwise incongruent modulo m. Thus, they form a reduced system modulo m. For example, 3 is a primitive root of modulo 7. Similarly, 3 is a primitive root modulo 10, since $\varphi(10) = 4$ and

$$3^0 \equiv 1 \pmod{10}, 3^1 \equiv 3 \pmod{10},$$

$$3^2 \equiv 9 \pmod{10}, 3^3 \equiv 7 \pmod{10}, 3^4 \equiv 1 \pmod{10}.$$

Some modulo do not have primitive roots. There is no primitive root modulo 8, for example, since $\varphi(8) = 4$, but

$$1^2 \equiv 3^2 \equiv 5^2 \equiv 7^2 \equiv 1 \pmod{8}, \tag{5.1}$$

and no integer has order 4 modulo 8.

In this section we prove that every prime p has primitive root. In Section 5.2 we determine all composite modulo m for which there exists primitive roots.

We begin with some remarks about polynomials. Let R be a commutative ring with identity. A *polynomial with coefficient in R* is an expression of the form

$$f(x) = a_m x^m + a_{m-1}x^{m-1} + ... + a_1 x^1 + a_0,$$

where $a_0, a_1, ..., a_m \in R$. The element a_i is called the *coefficient* of the term x^i. The *degree* of the polynomial $f(x)$, denoted by $\deg(f)$, is the greatest integer n such that $a_n \neq 0$, and a_n is called the *leading coefficient.* If $deg(f) = n$, we define $a_i = 0$for $i > n$. Nonzero constant polynomial$f(x) = a_0 \neq 0$ has degree 0. The zero polynomial $f(x) = 0$ has no degree. A *monic polynomial* is a polynomial whose leading coefficient is 1.

We define addition and multiplication of polynomial in the usual way. If $f(x) = \sum_{i=0}^{n} a_i x^i$ and $g(x) = \sum_{j=0}^{m} b_j x^j$, then

$$(f+g)(x) = \sum_{k=0}^{\max(m,n)} (a_k + b_k)x^k$$

and

$$fg(x) = \sum_{k=0}^{mn} c_k x^k,$$

where

$$c_k = \sum_{i+j=k, 0\le i\le n, 0\le j\le m} a_i b_j = \sum_{i=0}^{k} a_i b_{k-i}$$

With this addition and multiplication, the set $R[x]$ of all polynomials with coefficient in R is a commutative ring. Moreover,

$$\deg(f+g) \le max(\deg(f), \deg(g)).$$

If $f, g \in F[x]$ for some field F, then

$$\deg(fg) = \deg(f) + \deg(g),$$

and the leading coefficient of fg is $a_m b_n$.

For every $\alpha \in R$, the *evaluation map* $\Theta_\alpha : R[x] \to R$ defined by

$$\Theta_\alpha(f) = f(\alpha) = a_n\alpha^n + a_{n-1}\alpha^{n-1} + ... + a_1\alpha + a_0$$

is a ring homomorphism, that is, $(f+g)(\alpha) = f(\alpha)+g(\alpha)$ and $(fg)(\alpha) = f(\alpha)g(\alpha)$. The element α is called a *zero* of a *root* of the polynomial $f(x)$ if $\Theta_\alpha(f) = f(\alpha) = 0$.

We say that the polynomial $d(x)$ divides the polynomial $f(x)$ if there exists a polynomial such that $f(x) = d(x)q(x)$.

Theorem 5.1.1. (Division algorithm for polynomials) *Let F be a field. If $f(x)$ and $d(x)$ are polynomials in $F[x]$ and if $d(x) \neq 0$, then there exists unique polynomial $q(x)$ and $r(x)$ such that $f(x) = d(x)q(x) + r(x)$ and either $r(x) = 0$ or the degree of $r(x)$ is strictly smaller than the degree of $d(x)$.*

Proof. Let $d(x) = b_m x^m + ... + b_1 x + b_0$, where $b_m \neq 0$ and $\deg(d) = m$. If $d(x)$ does not divide $f(x)$, then $f - dq \neq 0$ and $\deg(f - dq)$ is a nonnegative integer for every polynomial $q(x) \in F[x]$. Choose $q(x)$ such that $l = \deg(f - dq)$ is minimal, and let

$$r(x) = f(x) - d(x)q(x) = c_l x^l + ... + c_1 x + c_0 \in F[x],$$

where $c_l \neq 0$. We shall prove that $l < m$.

Since F is a field,$b_m^{-1} \in F$. If $l \ge m$, then

$$d(x)b_m^{-1}c_l x^{l-m}$$

is a polynomial of degree l with leading coefficient c_l. Then

$$Q(x) = q(x) + b_m^{-1}c_l x^{l-m} \in F[x],$$

and

$$\begin{aligned} R(x) &= f(x) - d(x)Q(x) \\ &= f(x) - d(x)(q(x) + b_m^{-1}c_l x^{l-m}) \\ &= r(x) - d(x)b_m^{-1}c_l x^{l-m} \end{aligned}$$

is a polynomial of degree at most $l - 1$. This contradicts the minimality of l, and so $l < m$.

Next we prove that the polynomial $q(x)$ and $r(x)$ are unique. Suppose that

$$f(x) = d(x)q_1(x) + r_1(x) = d(x)q_2(x) + r_2(x),$$

where $q_1(x), q_2(x), r_1(x), r_2(x)$ are polynomials in $F[x]$ such that $r_i(x) = 0$ or $\deg(r_i) < \deg(d)$ for $i = 1, 2$. Then

$$d(x)(q_1(x) - q_2(x)) = r_2(x) - r_1(x).$$

If $q_1(x) \neq q_2(x)$, then

$$\deg(d) \leq \deg(d(q_1 - q_2)) = \deg(r_2 - r_1) < \deg(d),$$

which is abused. Therefore, $q_1(x) = q_2(x)$,and so $r_1(x) = r_2(x)$. This completes the proof.

Theorem 5.1.2. *Let $f(x) \in F[x]$, $f(x) \neq 0$, and let $N_0(f)$ denote the number of distinct zeros of $f(x)$ in F. Then $N_0(f)$ does not exceed the degree of $f(x)$, that is,*

$$N_0(f) \leq \deg(f).$$

Proof. We use the division algorithm for polynomials. Let $\alpha \in F$. Dividing$f(x)$ by $x - \alpha$, we obtain

$$f(x) = (x - \alpha)q(x) + r(x),$$

where $r(x) = 0$ or $\deg(r) < \deg(x - \alpha) = 1$, that is, $r(x) = r_0$ is a constant. Letting $x = \alpha$, we see that $r_0 = f(\alpha)$, and so

$$f(x) = (x - \alpha)q(x) + f(\alpha)$$

for every $\alpha \in F$. IN particular, if α is a zero of $f(x)$, then $x - \alpha$ divides $f(x)$.

We prove the theorem by induction on $n = \deg(fx)$. If $n = 0$, then $f(x)$ is a nonzero constant and $N_0(f) = 0$. If $n = 1$, then $f(x) = a_0 + a_1 x$ with $a_1 \neq 0$, and $N_0(f) = 1$ since $f(x)$ has the unique zero $\alpha = -a_1^{-1} a_0$. Suppose that $n \geq 2$ and the theorem is true for all polynomials of degree at most $n - 1$. If $N_0(f) = 0$, we are done. If $N_0(f) \geq 1$, let $\alpha \in F$ be a zero of $f(x)$. Then

$$f(x) = (x - \alpha)q(x),$$

and

$$\deg(q) = n - 1.$$

If β is a zero of $f(x)$ and $\beta \neq \alpha$, then

$$0 = f(\beta) = (\beta - \alpha)q(\beta),$$

and so β is a zero of $q(x)$. Since $\deg(q) = n - 1$, the induction hypothesis implies that

$$N_0(f) \leq 1 + N_0(q) \leq 1 + \deg(q) = n.$$

This completes the proof.

Theorem 5.1.3.*Let G be a finite subgroup of the multiplicative group of a field. Then G is cyclic.*

Proof. Let $|G| = m$. By Theorem 2.5.4 (Lagrange's theorem), if $a \in G$, then the order of a is a divisor of m. For every divisor d of m, Let $\psi(d)$ denote the number of elements of G of order d. If $\psi(d) \neq 0$, then there exists an element a of order d, and every element of the cyclic subgroup $\langle a \rangle$ generated by a satisfies $a^d = 1$. By Theorem 5.1.2, the polynomial $f(x) = x^d - 1 \in F[x]$ has at most d zeros, and so every zero of $f(x)$ belongs to the cyclic subgroup $\langle a \rangle$. In particular, every element of G of order d must belong to $\langle a \rangle$. By Theorem 2.5.7, a cyclic group of d has exactly $\psi(d) = 0$ or $\psi(d) = \varphi(d)$ for every divisor d of m. Since every element of G has order d for some divisor d of m, it follows that

$$\sum_{d|m} \psi(d) = m.$$

By Theorem 2.3.3,

$$\sum_{d|m} \varphi(d) = m,$$

and so $\psi(d) = \varphi(d)$ for every divisor d of m. In particular, $\psi(m) = \varphi(m) \geq 1$, and so G is a cyclic group of order m.

Theorem 5.1.4. *For every prime p, the multiplicative group of the finite field Z/pZ is cyclic. This group has $\varphi(p-1)$ generators. Equivalently, for every prime p, there exists $\varphi(p-1)$ pairwise incongruent primitive roots modulo p.*

Proof. This follows immediately from Theorem 5.1.3, since $|(\mathbf{Z}/p\mathbf{Z})^\times| = p - 1$.

The following table lists the primitive roots for the first six primes.

p	$\varphi(p-1)$	primitive roots
2	1	1
3	1	2
5	2	2,3
7	2	3,5
11	4	2,6,7,8
13	4	2,6,7,11

Let p be a prime, and let g be a primitive root modulo p. If a is an integer not divisible by p, then there exists a unique integer k such that

$$a \equiv g^k \pmod{p}$$

and

$$k \in \{0, 1, ..., p-2\}.$$

This integer k is called the *index* of a with respect to the primitive root g, and is denoted by

$$k = \text{ind}_g(a).$$

If k_1 and k_2 are any integers such that $k_1 \leq k_2$ and

$$a \equiv g^{k_1} \equiv g^{k_2} \pmod{p},$$

then

$$g^{k_2 - k_1} \equiv 1 \pmod{p},$$

and so

$$k_1 \equiv k_2 \pmod{p-1}.$$

If $a \equiv g^k$ $(mod\ p)$ and $b \equiv g^l \pmod{p}$, then $ab \equiv g^k g^l = g^{k+l} \pmod{p}$ and so

$$\text{ind}_g(ab) \equiv k + l \equiv \text{ind}_g(a) + \text{ind}_g(b) \pmod{p-1}.$$

The index map ind_g is also called *discrete logarithm* to the base g modulo p.

For example, 2 is a primitive root modulo 13. Here is a table of $\text{ind}_2(a)$ for $a = 1, ..., 12$:

a	$\text{ind}_2(a)$	a	$\text{ind}_2(a)$
1	0	7	11
2	1	8	13
3	4	9	8
4	2	10	10
5	9	11	7
6	5	12	6

By Theorem 2.5.7, if g is a primitive root modulo p, then g^k is a primitive root if and only if $(k, p-1) = 1$. For example, for $p = 13$ there are $\varphi(12) = 4$ integers k such that $0 \leq k \leq 11$ and $(k, 12) = 1$, namely, $k = 1, 5, 7, 11$, and so the four pairwise incongruent primitive roots modulo 13 are

$$2^1 \equiv 2 \pmod{13}, 2^5 \equiv 6 \pmod{13}, 2^7 \equiv 11 \pmod{13}, 2^{11} \equiv 7 \pmod{13}.$$

§5.2. Primitive Roots to Composite Moduli

In the previous section we proved that primitive roots exist for every prime number. We also observed that primitive roots do not exist for every modulus. For example, congruence (3.1) shows that there is no primitive root modulo 8. The goal of this section is to prove that an integer $m \geq 2$ has a primitive root if and only if $m = 2, 4, p^k$, or $2p^k$, where p is an odd prime and k is a positive integer.

Theorem 5.2.1. *Let m be a positive integer that is not a power of 2. If m has a primitive root, then $m = p^k$ or $2p^k$, where p is an odd prime and k is a positive integer.*

Proof. Let a and m be integers such that $(a, m) = 1$ and $m \geq 3$. Suppose that

$$m = m_1 m_2, \text{ where } (m_1, m_2) = 1, \text{ and } m_1 \geq 3, m_2 \geq 3. \tag{5.2}$$

Then $(a, m_1) = (a, m_2) = 1$. The Euler phi function $\varphi(m)$ is even for $m \geq 3$ (Exercise to U). Let

$$n = \frac{\varphi(m)}{2} = \frac{\varphi(m_1)\varphi(m_2)}{2}.$$

By Euler's theorem,

$$a^{\varphi(m_1)} \equiv 1 \pmod{m_1},$$

and so

$$a^n = (a^{\varphi(m_1)})^{\varphi(m_2)/2} \equiv 1 \pmod{m_1}.$$

Similarly,

$$a^n = (a^{\varphi(m_2)})^{\varphi(m_1)/2} \equiv 1 \pmod{m_2}.$$

Since $(m_1, m_2) = 1$ and $m = m_1 m_2$, we have

$$a^n \equiv 1 \quad (\text{mod } m)),$$

and so the order of a modulo m is strictly smaller than $\varphi(m)$. Consequently, if we can factor m in the form (5.2), then there does not exist a primitive root modulo m. In particular, if m is divisible by two distinct odd primes, then m does not have a primitive root. Similarly, if $m = 2^l p^k$, where $l \geq 2$, then m does not have a primitive root. Therefore, the only moduli $m \neq 2^l$ for which primitive roots can exist are of the form $m = p^k$ or $m = 2p^k$ for some odd prime p.

To prove the converse of Theorem 5.2.1, we use the following result about the exponential increase in the order of an integer modulo prime powers.

Theorem 5.2.2. *Let p be an odd prime, and let $a \neq \pm 1$ be an integer not divisible by p. Let d be the order of a modulo p. Let k_0 be the largest integer such that $a^d \equiv 1 \pmod{p^{k_0}})$. Then the order of a modulo p^k is d for $k = 1, ..., k_0$ and dp^{k-k_0} for $k \geq k_0$.*

Proof. There exists an integer u_0 such that

$$a^d = 1 + p^{k_0} u_0 \text{ and } (u_0, p) = 1. \tag{5.3}$$

Let $1 \leq k \leq k_0$, and let e be the order of a modulo p^k. If $a^e \equiv 1 \pmod{p^k}$, then $a^e \equiv 1 \pmod{p}$, and so d divides e. By (5.3), we have $a^d \equiv 1 \pmod{p^k}$, and so e divides d. It follows that $e = d$.

Let $j \geq 0$. We shall show that there exists an integer u_j such that

$$a^{dp^j} = 1 + p^{j+k_0} u_j \text{ and } (u_j, p) = 1. \tag{5.4}$$

The proof is by induction on j. The assertion is true for $j = 0$ by (5.3). Suppose we have (5.4) for some integer $j \geq 0$. By the binomial theorem, there exists an integer v_j such that

$$\begin{aligned} a^{dp^{j+1}} &= (1 + p^{j+k_0} u_j)^p \\ &= 1 + p^{j+1+k_0} u_j + \sum_{i=2}^{p} \left(\begin{array}{c} p \\ i \end{array} \right) p^{i(j+k_0)} u_j^i \\ &= 1 + p^{j+1+k_0} u_j + p^{j+2+k_0} v_j \\ &= 1 + p^{j+1+k_0} (u_j + p v_j) \\ &= 1 + p^{j+1+k_0} u_{j+1}, \end{aligned}$$

and the integer $u_{j+1} = u_j + p v_j$ is relatively prime to p. Thus (5.4) holds for all $j \geq 0$.

Let $k \geq k_0 + 1$ and $j = k - k_0 \geq 1$. Suppose that the order of a modulo p^{k-1} is dp^{j-1}. Let e_k denote the order of a modulo p^k. The congruence

$$a^{e_k} \equiv 1 \pmod{p^k}$$

implies that

$$a^{e_k} \equiv 1 \pmod{p^{k-1}},$$

and so dp^{j-1} divides e_k. Since

$$a^{dp^{j-1}} = 1 + p^{k-1} u_{j-1} \not\equiv 1 \pmod{p^k},$$

it follows that dp^{j-1} is a proper divisor of e_k.

On the other hand,

$$a^{dp^j} = 1 + p^k u_j \equiv 1 \pmod{p^k},$$

and so e_k divides dp^j. It follows that the order of a modulo p^k is exactly $e_k = dp^j = dp^{k-k_0}$. This completes the proof. □

Theorem 5.2.3. *Let p be an odd prime. If g is a primitive root modulo p, then either g or g+p is a primitive root of p^k for all $k \geq 2$. If g is a primitive root modulo p^k and $g_1 \in \{g, g + p^k\}$ is odd, then g_1 is a primitive root modulo $2p^k$.*

Proof. Let g be a primitive root modulo p. The order of g modulo p is $p - 1$. Let k_0 be the largest integer such that p^{k_0} divides $g^{p-1} - 1$. By Theorem 5.2.2, if $k_0 = 1$, then the order of g modulo p^k is $(p-1)p^{k-1} = \varphi(p^k)$, and g is a primitive root modulo p^k for all $k \geq 1$.

If $k_0 \geq 2$, then

$$g^{p-1} = 1 + p^2 v$$

for some integer v. By the binomial theorem,

$$\begin{aligned}(g+p)^{p-1} &= \sum_{i=0}^{p-1} \left(\begin{array}{c} p-1 \\ i \end{array} \right) g^{p-1-i} p^i \\ &\equiv g^{p-1} + (p-1) g^{p-2} p \pmod{p^2} \\ &\equiv 1 + p^2 v + g^{p-2} p^2 - g^{p-2} p \pmod{p^2} \\ &\not\equiv 1 \pmod{p^2}.\end{aligned}$$

Then $g+p$ is a primitive root modulo p such that

$$(g+p)^{p-1} = 1 + pu_0 \text{ and } (u_0, p) = 1.$$

Therefore, $g+p$ is a primitive root modulo p^k for all $k \geq 1$.

Next we prove that primitive roots exist for all moduli of the form $2p^k$. If g is a primitive root modulo p^k, then $g+p^k$ is also a primitive root modulo p^k. Since p^k is odd, it follows that one of the two integers g and $g+p^k$ is odd, and the other is even. Let g_1 be the odd integer in the set $\{g, g+p^k\}$. Since $(g+p^k, p^k) = (g, p^k) = 1$, it follows that $(g_1, 2p^k) = 1$. The order of g_1 modulo $2p^k$ is not less than $\varphi(p^k)$, which is the order of g_1 modulo p^k, and not greater than $\varphi(2p^k)$. However, since p is an odd prime, we have

$$\varphi(2p^k) = \varphi(p^k),$$

and so g_1 has order $\varphi(2p^k)$ modulo $2p^k$, that is, g_1 is a primitive root modulo $2p^k$. This completes the proof. □

For example, 2 is a primitive root modulo 3. Since 3 is the greatest power of 3 that divides $2^2 - 1$, it follows that 2 is a primitive root modulo 3^k for all $k \geq 1$, and $2 + 3^k$ is a primitive roots modulo powers of 2.

Theorem 5.2.4. *There exists a primitive root modulo $m = 2^k$ if and only if $m = 2$ or 4.*

Proof. we note that 1 is a primitive root modulo 2, and 3 is a primitive root modulo 4. We shall prove that if $k \geq 3$, then there is no primitive root modulo 2^k. Since $\varphi(2^k) = 2^{k-1}$, it suffices to show that

$$a^{2^{k-2}} \equiv 1 \pmod{2^k} \tag{5.5}$$

for a odd and $k \geq 3$. We do this by induction on k. The case $k = 3$ is congruence(5.1). Let $k \geq 3$, and suppose that (5.5) is true. Then

$$a^{2^{k-2}} - 1$$

is divisible by 2^k. Since a is odd, it follows that

$$a^{2^{k-2}} + 1$$

is even. Therefore,

$$a^{2^{k-1}} - 1 = \left(a^{2^{k-2}} - 1\right)\left(a^{2^{k-2}} + 1\right)$$

is divisible by 2^{k+1}, and so

$$a^{2^{k-1}} \equiv 1 \pmod{2^{k+1}}.$$

This completes the induction and the proof of the theorem. □

Let $k \geq 3$. By Theorem 5.2.4, there is no primitive root modulo p^k, that is, there does not exist an odd integer whose order modulo 2^k is 2^{k-1}. However, there do exist odd integers of order 2^{k-2} modulo 2^k.

Theorem 5.2.5. *For every positive integer k, we have*

$$5^{2^k} \equiv 1 + 3 \cdot 2^{k+2} \pmod{2^{k+4}}.$$

Proof. The proof is by induction on k. For $k = 1$ we have

$$5^{2^1} = 25 \equiv 1 + 3 \cdot 2^3 \pmod{2^5}.$$

Similarly, for $k = 2$ we have

$$5^{2^2} = 625 = 148_5 76 \equiv 1 + 3 \cdot 2^4 \pmod{2^6}.$$

If the theorem holds for $k \geq 1$, then there exists an integer u such that

$$5^{2^k} \equiv 1 + 3 \cdot 2^{k+2} + 2^{k+4}u = 1 + 2^{k+2}(3 + 4u).$$

Since $2k + 4 \geq k + 5$, we have

$$\begin{aligned} 5^{2^{k+1}} &= \left(5^{2^k}\right)^2 \\ &= \left(1 + 2^{k+2}(3 + 4u)\right)^2 \\ &\equiv 1 + 2^{k+3}(3 + 4u) \pmod{2^{2k+4}} \\ &\equiv 1 + 3 \cdot 2^{k+3} \pmod{2^{k+5}} \end{aligned}$$

This completes the proof. □

Theorem 5.2.6. *If $k \geq 3$, then 5 has order 2^{k-2} modulo 2^k. If $a \equiv 1 \pmod 4$, then there exists a unique integer $i \in \{0, 1, ..., 2^{k-2} - 1\}$ such that*

$$a \equiv 5^i \pmod{2^k}.$$

If $a \equiv 3 \pmod 4$, then there exists a unique integer $i \in \{0, 1, ..., 2^{k-2} - 1\}$ such that

$$a \equiv -5^i \pmod{2^k}.$$

Proof. In the case $k = 3$, we observe that 5 has order 2 modulo 8, and

$$\begin{aligned} 1 &\equiv 5^0 \pmod 8, \\ 3 &\equiv -5^1 \pmod 8, \\ 5 &\equiv 5^1 \pmod 8, \\ 7 &\equiv -5^0 \pmod 8. \end{aligned}$$

Let $k \geq 4$. By Theorem 5.2.5, we have

$$\begin{aligned} 5^{2^{k-2}} &\equiv 1 + 3 \cdot 2^k \pmod{2^{k+2}} \\ &\equiv 1 \pmod{2^k} \end{aligned}$$

and

$$\begin{aligned} 5^{2^{k-3}} &\equiv 1 + 3 \cdot 2^{k-1} \pmod{2^{k+1}} \\ &\equiv 1 + 3 \cdot 2^{k-1} \pmod{2^k} \\ &\not\equiv 1 \pmod{2^k} \end{aligned}$$

Therefore, 5 has order exactly 2^{k-2} modulo 2^k, and so the integers 5^i are pairwise incongruence modulo 2^k for $i = 0, 1, ..., 2^{k-1} - 1$. Since $5^i \equiv 1 \pmod 4$ for all i, and since exactly half, that is, 2^{k-2}, of the 2^{k-1} odd numbers between 0 and 2^k are congruence to 1 modulo 4, it follows that the congruence

$$5^i \equiv a \pmod{2^k}$$

is solvable for every $a \equiv 1(mod 4)$. If $a \equiv 3 \pmod 4$, then $-a \equiv 1 \pmod 4$ and so the congruence

$$-a \equiv 5^i \pmod{2^k},$$

or equivalently,

$$a \equiv -5^i \pmod{2^k},$$

is solvable. This completes the proof. □

In algebraic language, Theorem 5.2.5 states that for all $k \geq 3$,

$$(\mathbf{Z}/2^k\mathbf{Z})^\times = \langle -1 \rangle \times \langle 5 \rangle \cong \mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/2^{k-2}\mathbf{Z},$$

where $\langle a \rangle$ denotes the cyclic subgroup of $(\mathbf{Z}/2^k\mathbf{Z})^\times$ generated by a for $a = -1$ and $a = 5$.

§5.3. Power Residues

Let m, k and a be integers such that $m \geq 2, k \geq 2$, and $(a, m) = 1$. We say that a is a *kth power residue modulo m* if there exists an integer x such that

$$x^k \equiv a \pmod m.$$

If this congruence has no solution, then a is called a k-th *power nonresidue modulo m*.

Let $k = 2$ and $(a, m) = 1$. If the congruence $x^2 \equiv a \ (mod\ m)$ is solvable, then a is called a *quadratic residue modulo m*. Otherwise, a is called a *quadratic nonresidue modulo m*. For example, the quadratic residues modulo 7 are 1,2, and 4; the quadratic nonresidues are 3,5 and 6. The only quadratic residue modulo 8 is 1, and the quadratic nonresidues modulo 8 are 3,5 and 7.

Let $k = 3$ and $(a, m) = 1$. If the congruence $x^3 \equiv a \pmod m$ is solvable, then a is called a *cubic residue modulo m*. Otherwise, a is called a *cubic nonresidue modulo m*. For example, the cubic residues modulo 7 are 1 and 6; the cubic nonresidues are 2, 3, 4, and 5. The cubic residues modulo 5 are 1, 2, 3 and 4; there are no cubic nonresidues modulo 5.

In this and next two sections we investigate power residues modulo primes. In Section 4.3 we have considered quadratic residues to composite moduli.

Theorem 5.3.1. *Let p be prime, $k \geq 2$, and $d = (k, p-1)$. Let a be an integer not divisible by p. Let g be a primitive root modulo p, then a is a k-th power residue modulo p if and only if*

$$\mathrm{ind}_g(a) \equiv 0 \pmod d$$

if and only if

$$a^{(p-1)/d} \equiv 1 \pmod p.$$

If a is a k-th power residue modulo p, then the congruence

$$x^k \equiv a \pmod p \tag{5.7}$$

has exactly d solutions that are pairwise incongruence modulo p. Moreover, there are exactly $(p-1)/d$ pairwise incongruence k-th power residues modulo p.

Proof. Let $l = \mathrm{ind}_g(a)$, where g is a primitive root modulo p. Congruence (5.7) is solvable if and only if there exists an integer y such that

$$g^y \equiv x \pmod p$$

and

$$g^{ky} \equiv x^k \equiv a \equiv g^l \pmod p.$$

This is equivalent to

$$ky \equiv l \pmod{p-1}. \tag{5.8}$$

This linear congruence in y has a solution if and only if

$$\mathrm{ind}_g(a) = l \equiv 0 \pmod d,$$

where $d = (k, p-1)$. Thus the k-th power residues modulo p are precisely the integers in the $(p-1)/d$ congruence classes $g^{id} + p\mathbf{Z}$ for $i = 0, 1, ..., (p-1)/d - 1$. Moreover,

$$a^{(p-1)/d} \equiv g^{(p-1)l/d} \equiv 1 \pmod p$$

if and only if

$$\frac{(p-1)l}{d} \equiv 0 \pmod{p-1}$$

if and only if

$$\mathrm{ind}_g(a) = l \equiv 0 \pmod d.$$

Finally, if the linear congruence (5.8) is solvable, then by Theorem 2.2 it has exactly d solutions y that are pairwise incongruence modulo $p-1$, and so (5.7) has exactly d solutions $x = g^y$ that are pairwise incongruence modulo p. This completes the proof. □

For example, let $p = 19$ and $k = 13$. Then $d = (k, p-1) = (3, 18) = 3$. We can check that 2 is a primitive root modulo 19, and so a is a cubic residue modulo 19 if and only if 3 divides $\mathrm{ind}_2(a)$. Since $-1 \equiv 2^9 \pmod 3$ and $\mathrm{ind}_2(-1) = 9$, it follows that -1 is a cubic residue modulo 19. The solutions of the congruence $x^3 \equiv -1 \pmod{19}$ are of the form $x \equiv 2^y \pmod{19}$, where $0 \le y \le 17$ and $3y \equiv 9$ (mod 18). Then $y \equiv 3 \pmod 6$, and so $y = 3, 9$ and 15. These give the following three cube roots of -1 modulo 19:

$$8 \equiv 2^3 \pmod{19},$$

$$18 \equiv 2^9 \pmod{19},$$

and

$$12 \equiv 2^{15} \pmod{19}.$$

Corollary 5.3.2. *Let p be an odd prime, and let $k \ge 2$ be an integer such that $(k, p-1) = 1$. If $(a, p) = 1$, then a is a kth power residue modulo p, and the congruence $x^k \equiv a \pmod p$ has a unique solution modulo p.*

Exercises for Chapter 5

1. Compute $\mathrm{ind}_2(27)$ modulo 101.

2. Let p be an odd prime, and let g be a primitive root modulo p. Show that $(p-1)! \equiv g^{\frac{(p-2)(p-1)}{2}} \equiv -1 \pmod p$.

3. Show that if m has one primitive root, then there are exactly $\varphi(\varphi(m))$ pairwise incongruent primitive roots modulo m.

4. Let g and r be primitive roots modulo p. Show that

$$\mathrm{ind}_r(a) \equiv \mathrm{ind}_g(a)\mathrm{ind}_r(g) \pmod{p-1}$$

for every integer a relatively prime to p.

5. Let g be a primitive root modulo the odd prime p. Show that $g^{\frac{p-1}{2}} \equiv -1 \pmod{p}$.

6. Let g be a primitive root modulo the odd prime p. Show that $-g$ is a primitive root modulo p if and only if $p \equiv 1 \pmod{4}$.

7. Show that for every prime $p \geq 5$, we have

$$\sum_{1 \leq i < j \leq p-1} ij \equiv 0 \pmod{p}$$

and

$$\sum_{1 \leq i < j < k \leq p-1} ijk \equiv 0 \pmod{p}.$$

8. Find an integer g that is a primitive root modulo 5^k for all $k \geq 1$. Find a primitive root modulo 10. Find a primitive root modulo 50.

9. Show that if g is a primitive root modulo p^2, then g is primitive root modulo p^k for all $k \geq 2$.

10. Let p be an odd prime. Show that

$$(1+px)^{p^k} \equiv 1 + p^{k+1}x \pmod{p^{k+2}}$$

for every integer x and every nonnegative integer k.

11. Let $\{x\}$ denote the fractional part of x. Compute

$$\left\{ \left(\frac{3}{2} \right)^n \right\}$$

for $n = 1, ..., 10$. Let r_n be the least nonnegative residue of 3^n modulo 2^n. Show that

$$\left\{ \left(\frac{3}{2} \right)^n \right\} = \frac{r_n}{3^n}.$$

12. Find all cubic residues modulo 19.

13. Define the map $f : (\mathbf{Z}/23\mathbf{Z})^\times \to (\mathbf{Z}/23\mathbf{Z})^\times$ by $f(x + 23\mathbf{Z}) = x^3 + 23\mathbf{Z}$. Show that f is an isomorphism of the multiplicative group $(\mathbf{Z}/23\mathbf{Z})^\times$.

(14).

(15). Let $a, b \in \mathbf{Z}^+$ and g be a primitive root modulo p. Show that $a \equiv b \pmod{p-1}$ if $g^a \equiv g^b \pmod{p}$.

(16). Let $a, b, c \in \mathbf{Z}^+$ and g be a primitive root modulo p^c. Show that $a \equiv b \pmod{p^{c-1}(p-1)}$ if $g^a \equiv g^b \pmod{p^c}$.

(17). Let $a, b, m \in \mathbf{Z}^+$ and g be a primitive root modulo m. Show that $a \equiv b \pmod{\varphi(m)}$ if $g^a \equiv g^b \pmod{m}$.

Chapter 6. The Riemann Zeta Function

§6.1 Definition and Convergence

In order to make progress in number theory, it is sometimes necessary to use techniques from other areas of mathematics, such as algebra, analysis or geometry. In this chapter we give some number-theoretic applications of the theory of infinite series. These are based on the properties of the Riemann zeta function ζ(s), which provides a link between number theory and real and complex analysis.

One of the most familiar examples of an infinite series is the harmonic series

$$\sum_{n=1}^{\infty}\frac{1}{n}=1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\cdots .$$

Since number theory is mainly about the positive integers $n = 1, 2, 3\cdots$,it is not surprising that this series is of interest to number theorists. Unfortunately, it diverges, but only just: the sum of the first $n$terms is about $\log n$,and although this tends to $+\infty$ as $n \to \infty$, it does so rather slowly. To make the series converge, without losing its important number-theoretic properties, we replace its general term $1/n$ with the smaller term $1/n^s$, where $s > 1$. This gives rise to the Riemann zeta function, defined by

$$\zeta(s)=\sum_{n=1}^{\infty}\frac{1}{n^s}=1+\frac{1}{2^s}+\frac{1}{3^s}+\cdots . \tag{6.1}$$

Although this function is named after Riemann, who wrote a fundamental paper on its properties in 1859, it was in fact introduced about 120 years earlier by Euler, who showed that it can be expanded as a product

$$\zeta(s)=\prod_{p}\frac{1}{1-p^{-s}}, \tag{6.2}$$

where p ranges over all the primes. This is a very powerful result, since it allows methods of analysis to be applied to the study of prime numbers. Euler regarded $\zeta(s)$ as function of a real variable s to a complex number, so that the even richer theory of complex functions could be used. One of the great unsolved problems in number theory is the Riemann Hypothesis (see Section 9), a conjecture concerning the complex zeros of $\zeta(s)$; a solution of this would resolve many important problems concerning the distribution of prime numbers. Before dealing withe questions of convergence, we will first outline a proof of (6.2), and the show a simple but effective application of this product formula. Each factor on the right-hand side of (6.2) can be expanded as geometric series

$$\frac{1}{1-p^s}=1+p^{-s}+p^{-2s}+\cdots=\sum_{e=0}^{\infty}p^{-es}$$

convergent since $|p^{-s}| = p^{-s} < 1$ for all $s > 0$. If we multiply these series together (and we will justify this later),then the general term in their product has the form

$$p_1^{-e_1 s}\cdots p_k^{-e_k s}=\frac{1}{n^s},$$

where $p_1, \cdots, p_k$ are distinct primes, each $e_i \geq 0$, and $n = p^{e_1}\cdots p^{e_k}$. By the Fundamental Theorem of Arithmetic, every integer $n > 1$ has a unique factorization of this form, so it contributes exactly one summand, equal to $1/n^s$, and hence (6.2) represents $\sum 1/n^s = \zeta(s)$. (We will prove this more rigorously later in the chapter, in Theorem 6.3)

Using (6.2), we can now sketch a quick proof of Theorem 6.6, that there are infinitely many primes: if there were only finitely many primes, then $\zeta(s)$ would approach a finite limit $\prod_p (1-p^{-1})^{-1}$ as $s \to 1$, whereas in fact $\zeta(s) \to +\infty$, as we shall shortly see.

To justify the preceding arguments, we must first consider the convergence of the series (6.1). We will show that it converges for all real $s > 1$, and diverges for all real $s \leq 1$.

Theorem 6.1. *The series (6.1) converges for all real $s > 1$, and diverges for all real $s \leq 1$.*

Proof. Suppose first that $s > 1$. We group the terms together in blocks of length $1, 2, 4, 8, \cdots$, giving

$$\zeta(s) = 1 + (\frac{1}{2^s} + \frac{1}{3^s} + (\frac{1}{4^s} + \cdots + \frac{1}{7^s}) + (\frac{1}{8^s} + \cdots + \frac{1}{15^s}) + \cdots .$$

Now

$$\begin{aligned} \frac{1}{2^s} + \frac{1}{3^s} &\leq \frac{1}{2^s} + \frac{1}{2^s} = \frac{2}{2^s} = 2^{1-s} \\ \frac{1}{4^s} + \cdots + \frac{1}{7^s} &\leq \frac{1}{4^s} + \cdots + \frac{1}{4^s} = \frac{4}{4^s} = (2^{1-s})^2 \\ \frac{1}{8^s} + \cdots + \frac{1}{15^s} &\leq \frac{1}{8^s} + \cdots + \frac{1}{8^s} = \frac{8}{8^s} = (2^{1-s})^3 \end{aligned}$$

and so on,so we can compare (6.1) with the geometric series

$$1 + 2^{1-s} + (2^{1-s})^2 + (2^{1-s})^3 + \cdots$$

This converges since $0 < 2^{1-s} < 1$, and hence so does (6.1) by the Comparison Test. In fact, this argument shows that $1 \leq \zeta(s) \leq f(s)$ for all $s > 1$, where

$$f(s) = \sum_{n=0}^{\infty} (2^{1-s})^n = \frac{1}{1-2^{1-s}}$$

If $s \to +\infty$ then $2^{1-s} \to 0$ and so $f(s) \to 1$, giving

$$\lim_{s \to +\infty} \zeta(s) = 1.$$

We now show that (6.1) diverges for $s < 1$. This is obvious if $s \leq 0$, since then $1/n^s \nrightarrow 0$ as $n \to \infty$, so let us assume that $s > 0$. By grouping the term of (6.1) in blocks of length $1, 1, 2, 4, \cdots$, we have

$$\zeta(s) = 1 + \frac{1}{2^s} + (\frac{1}{3^s} + \frac{1}{4^s}) + (\frac{1}{5^s} + \cdots + \frac{1}{8^s}) + \cdots .$$

If $s \leq 1$ then

$$\begin{aligned} \frac{1}{2^s} &\geq \frac{1}{2} \\ \frac{1}{3^s} + \frac{1}{4^s} &\geq \frac{1}{4} + \frac{1}{4} = \frac{1}{2} \\ \frac{1}{5^s} + \cdots + \frac{1}{8^s} &\geq \frac{1}{8} + \cdots + \frac{1}{8} = \frac{1}{2}, \end{aligned}$$

and so on so (6.1) diverges by comparison with the divergent series $1+\frac{1}{2}+\frac{1}{2}+\frac{1}{2}+\cdots$. In particular, by taking $s = 1$ we see that the harmonic series $\sum 1/n$ diverges. □

There is an alternative proof based on the Integral Test, using the fact that $\int_1^{+\infty} x^{-s}dx$ converges if and only if $s > 1$.

§6.2 Applications to prime numbers

We can now give a more rigorous analytic proof of **Euclid Theorem** stating that *there are infinitely many primes.*

Analytic proof of Euclid's Theorem. Suppose that there are only finitely many primes, say $p_1, \cdots, p_k$. For each prime $p = p_i$, we have $|1/p| < 1$, so there is a convergent geometric series

$$1 + \frac{1}{p} + \frac{1}{p^2} + \frac{1}{p^3} + \cdots = \frac{1}{1 - p^{-1}}.$$

It follows that if we multiply these k different series together, their product

$$\prod_{i=1}^{k}(1 + \frac{1}{p_i} + \frac{1}{p_i^2} + \frac{1}{p_i^3} + \cdots) = \prod_{i=1}^{k} \frac{1}{1 - p_i^{-1}} \tag{6.3}$$

is finite. Now these convergent series all consist of positive terms, so they ate absolutely convergent. It follows that we can multiply out the series in (6.3) and rearrange the terms, without changing the product. If we take a typical term $i/p_1^{e_1}$ form the first series, $i/p_2^{e_2}$ form the second series, and so on, where each $e_i \le 0$, then their product

$$\frac{1}{p_1^{e_1}} \cdot \frac{1}{p_2^{e_2}} \cdot \cdots \cdot \frac{1}{p_k^{e_k}} = \frac{1}{p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}}$$

will represent a typical term in the expansion of (6.3). By the Fundamental Theorem of Arithmetic, every integer $n \ge 1$ has a unique expression

$$n = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k} \quad (e_i \ge 0)$$

as a product of powers of the primes p_i, since we are assuming that these are the only primes; notice that we allow $e_i = 0$, in case n is not divisible by a particular prime p_i. This uniqueness implies that each n contributes exactly one term $1/n$ to (6.3), so the expansion takes the form

$$\prod_{i=1}^{k}(1 + \frac{1}{p_i} + \frac{1}{p_i^2} + \frac{1}{p_i^3} + \cdots) = \sum_{n=1}^{\infty} \frac{1}{n}.$$

The right-hand side is the harmonic series, which is divergent. However, we have seen that the left-hand side is finite, so this contradiction proves that there must be infinitely many primes. □

The next result, also due to Euler, develops this method a little further.

Theorem 6.2. (Euler) *If p_n denotes the n-th prime (in increasing order), then the series*

$$\sum_{n=1}^{\infty} \frac{1}{p_n} = \frac{1}{2} + \frac{1}{3} + \frac{1}{5} + \cdots$$

diverges.

Proof. If $\sum 1/p_n$ converges to a finite sum l, then its partial sums must satisfy

$$\left| \sum_{n=1}^{N} \frac{1}{p_n} - l \right| \le \frac{1}{2}$$

for all sufficiently large N, so that

$$\sum_{n>N} \frac{1}{p_n} \leq \frac{1}{2}$$

for any such N. This implies that the series

$$\sum_{k=1}^{\infty} \Big(\sum_{n>N} \frac{1}{p_n} \Big)^k \tag{6.4}$$

converges by comparison with the geometric series$\sum_{k=1}^{\infty} 1/2^k$. If q denotes the product $p_1 \cdots p_N$ then no integer of the form $qr + 1 (r \geq 1)$ can be divisible by any of $p_1, \cdots, p_N$, so it must be a product of primes p_n for $n > N$ (possibly with repetitions), say

$$qr + 1 = p_{n_1} \cdots p_{n_k}$$

where each $n_i > N$. Then the reciprocal $1/qr + 1$ of each such an integer appears as a summand $1/p_{n_1} \cdots p_{n_k}$ in the expansion of

$$\Big(\sum_{n>N} \frac{1}{p_n} \Big)^k,$$

and hence it appears (just once) as a summand in the expansion of (6.4). since (6.4) converges, it follows that the series

$$\sum_{r=1}^{\infty} \frac{1}{qr+1}$$

which is contained within (6.4), also converges. However this series diverges, since its terms exceed those of the divergent series

$$\sum_{r=1}^{\infty} \frac{1}{qr+1} = \frac{1}{q} \sum_{r=1}^{\infty} \frac{1}{r+1} = \frac{1}{q} \sum_{r=2}^{\infty} \frac{1}{r}.$$

This contradiction shows that $\sum 1/p_n$ must diverge. □

Comments.

(1). It can be shown that

$$\frac{1}{p_1} + \cdots + \frac{1}{p_n} \to +\infty$$

about as fast as $\log \log n$, so this series diverges very slowly indeed.

(2). Theorem 6.2 gives yet another proof that there are infinitely many primes, since a finite series must converge. It also shows that the primes are more densely distributed than the perfect squares: the series $\sum 1/n^2$ converges (by the Integral Test), so $1/n^2 \to 0$ faster than $1/p_n \to 0$ as $n \to \infty$, that is, primes occur more frequently than squares.

We can now use these ideas to give a rigorous proof of the product expansion (6.2).

Theorem 6.3. (Euler) *If* $s > 1$ *then*

$$\zeta(s) = \prod_p \frac{1}{1 - p^{-s}},$$

where the product is over all primes p.

Proof. The method is to consider the product $P_k(s)$ of the factors corresponding to the first k primes, and to show that $P_k(s) \to \zeta(s)$ as $k \to \infty$. Let $p_1, \cdots, p_k$ be the first k primes. Arguing as before with (6.3), we see that if $s > 0$ (so that the geometric series all converge), then

$$P_k(s) = \prod_{i=1}^{k} \frac{1}{1 - p_i^{-s}} = \prod_{i=1}^{k} (1 + \frac{1}{p_i^s} + \frac{1}{p_i^{2s}} + \frac{1}{p_i^{3s}} + \cdots).$$

If we expand this product, the general term is the resulting series is $1/n^s$ where $n = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}$ and each $e_i \le 0$. The fundamental Theorem of Arithmetic implies that each such n contributes just one term to $P_k(s)$, so

$$P_k(s) = \sum_{n \in A_k} \frac{1}{n^s},$$

where $A_k = \{n = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}, e_i \le 0\}$ is the set of integers n whose prime factors are among $p_1, \cdots, p_k$. Each $n \notin A_k$ is divisible by some prime $p > p_k$, and so $n > p_k$. It follows that if $s > 1$ then

$$|P_k(s) - \zeta(s)| = \sum_{n \notin A_k} \frac{1}{n^s} \le \sum_{n > p_k} \frac{1}{n^s} = \zeta(s) - \sum_{n \le p_k} \frac{1}{n^s}.$$

Since $s > 1$, the partial sums of the series $\sum 1/n^s$ converge to $\zeta(s)$, so in particular

$$\sum_{n \le p_k} \frac{1}{n^s} \to \zeta(s)$$

as $k \to \infty$. Thus $|P_k(s) - \zeta(s)| \to 0$ as $k \to \infty$, so $P_k(s) \to \zeta(s)$ are required.

A similar method gives a rigorous proof of the following result. We will also prove this as part of a more general result later in this chapter (see Example 6.4).

Theorem 6.4. *If $s > 1$ then*

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n} = \prod_{p} (1 - p^{-s}) = \frac{1}{\zeta(s)}.$$

§6.3 Evaluating $\zeta(2)$ and $\zeta(2k)$

Having shown that $P = 1/\zeta(2)$, we now prove that

$$\zeta(2) = \frac{\pi^2}{6}. \tag{6.5}$$

Apostol (1983) gives an elementary proof of this, evaluating

$$\int_0^1 \int_0^1 (1 - xy)^{-1} \, dx \, dy$$

in two ways: first by writing $(1 - xy)^{-1} = \sum (xy)^n$ and integrating term by term, and second by using a change of variables to rotate the xy-plane through $\pi/4$ and then using some straightforward trigonometric substitutions. A quicker but less elementary proof is simply to put $x = 1$ in the Fourier series expansion

$$x^2 = \frac{1}{3} + \frac{4}{\pi^2} \sum_{n \ge 1} \frac{(-1)^n}{n^2} \cos(n\pi x)$$

of the function x^2 on the interval $[-1, 1]$; we have $\cos(n\pi x) = (-1^n)$, so (6.5) follows immediately.

Instead, we will give a third proof, which has the advantage of extending to certain other values of $\zeta(s)$.We will use the infinite product expansion

$$\sin z = z\prod_{n\neq 0}(1-\frac{z}{n\pi}) = z\prod_{n\geq 1}(1-\frac{z^2}{n^2\pi^2}), \tag{6.6}$$

proofs of which can be found in books on complex variable theory, e.g. Jones and Singerman (1987,Chapter 3, Section 8). The first product in (6.6) is over all non-zero integers n, and the second product is obtained from the first by pairing the factors corresponding to ±.One can explain (if not rigorously prove) the first product by regarding $\sin z$ as behaving like a polynomial with infinitely many zeros at $z = n\pi (n \in \mathbb{Z})$, so we have a 'factorisation'

$$\sin z = cz\prod_{n\neq 0}(1-\frac{z}{n\pi})$$

with

$$c = \lim_{z\to 0}\frac{\sin z}{z} = 1.$$

By expanding the second product in (6.6) and collecting powers of z, we obtain a power series for $\sin z$ which must coincide with its Taylor series expansion

$$\sin z = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \cdots . \tag{6.7}$$

By comparing coefficients of z^3 in (6.6) and (6.7) we see that

$$-\sum_{n\geq 1}\frac{1}{n^2\pi^2} = -\frac{1}{3!},$$

so multiplying through by $-\pi^2$ we obtain (6.5). □

With the aid of the previous section and a pocket calculator, we immediately deduce

Theorem 6.5. *The probability that two randomly- and independently-chosen integers are coprime is given by* $P = \frac{1}{\zeta(2)} = \frac{6}{\pi^2} = 0.607927101\ldots$.

By Exercise 6.7, this is also the probability that single randomly-chosen integer is square-free.

For many reasons, it would be useful to know the values of $\zeta(s)$ for *all* integers $s \geq 2$(see exercise 6.6, for instance). In 1978 Apéry proved a long-standing conjecture, that $\zeta(3)$ is irrational, but very little else is known about $\zeta(s)$ when s is odd. However, with a little extra work we can use (6.6) to evaluate $\zeta(s)$ for all even integers $s = 2k \geq 2$. Some of the techniques we use require rather careful analytic justification, using such concepts as uniform convergence, but for simplicity we will omit these details.

By taking logarithms in (6.6), we have

$$\log\sin z = \log z + \sum_{n\geq 1}\log\left(1-\frac{z^2}{n2\pi^2}\right),$$

and differentiating term by term we have

$$\cot z = \frac{1}{z} - \sum_{n\geq 1}\frac{2z}{n^2\pi^2}\left(1-\frac{z^2}{n2\pi^2}\right)^{-1}$$

If we use the geometric series to write

$$\frac{2z}{n^2\pi^2}\left(1-\frac{z^2}{n2\pi^2}\right)^{-1}=\frac{2z}{n^2\pi^2}\sum_{k\geq 0}\left(\frac{z^2}{n2\pi^2}\right)^k=2\sum_{k\geq 0}\frac{z^{2k+1}}{n^{2k+2}\pi^{2k+2}}=2\sum_{k\geq 1}\frac{z^{2k-1}}{n^{2k}\pi^{2k}}$$

and then collect powers of z, we get

$$\cot z=\frac{1}{z}-2\sum_{k\geq 1}\sum_{n\geq 1}\frac{z^{2k-1}}{n^{2k}\pi^{2k}}=\frac{1}{z}-2\sum_{k\geq 1}\frac{\zeta(2k)z^{2k-1}}{\pi^{2k}}, \tag{6.8}$$

which is the Laurent series for $\cot z$.

We will now compare (6.8) with a second expansion for $\cot z$. The exponential series

$$e^t=1+t+\frac{t^2}{2!}+\frac{t^3}{3!}+\cdots$$

implies that

$$\frac{e^t-1}{t}=1+\frac{t}{2!}+\frac{t^2}{3!}+\cdots,$$

and the reciprocal of this has a Taylor series expansion which can be written in the form

$$\frac{t}{e^t-1}=\left(1+\frac{t}{2!}+\frac{t^2}{3!}+\cdots\right)^{-1}=\sum_{m\geq 0}\frac{B_m}{m!}t^m \tag{6.9}$$

for certain constants $B_0, B_1, \cdots$, known as the *Bernoulli number*. Now

$$\begin{aligned}\frac{t}{e^t-1} &= \frac{t}{2}(\frac{e^t+1}{e^t-1}-1)\\ &= \frac{t}{2}(\frac{e^{t/2}+e^{-t/2}}{e^{t/2}-e^{-t/2}}-1)\\ &= \frac{t}{2}(\coth\frac{t}{2}-1)\\ &= \frac{t}{2}(i\cot\frac{it}{2}-1)\end{aligned}$$

where $i=\sqrt{-1}$. Putting $z=it/2$ we get

$$\frac{t}{e^t-1}=z\cot z=\frac{z}{i}=z\cot z+iz,$$

so dividing by z and using (6.9) we have

$$\cot z=-i+\frac{1}{z}\sum_{m\geq 0}\frac{B_m}{m!}t^m=-i+\sum_{m\geq 0}\left(\frac{2}{i}\right)^m z^{m-1}.$$

By comparing coefficients with those in (6.8), we see that if $m=2k\geq 2$, then

$$-2\frac{\zeta(2k)}{\pi^{2k}}=\frac{B_{2k}}{(2k)!}\left(\frac{2}{i}\right)^{2k},$$

so that

$$\zeta(2k)=\frac{(-1)^{k-1}2^{2k-1}\pi^{2k}B_{2k}}{(2k)!}. \tag{6.10}$$

Thus

$$\zeta(2)=\pi^2B_2,\quad \zeta(4)=-\frac{\pi^4B_4}{3},\quad \zeta(6)=\frac{2\pi^6B_6}{45},$$

and so on.

To evaluate the Bernoulli numbers, we write (6.9) in the form

$$t = \sum_{m\geq 0} \frac{B_m}{m!}t^m \cdot (e^t - 1) = \sum_{m\geq 0} \frac{B_m}{m!}t^m \cdot \sum_{n\geq 1} \frac{1}{n!}t^n. \tag{6.11}$$

If we put $m + n = r$, we find that the coefficient of t^r in the right-hand side of (6.11) is

$$\sum_{m+n=r} \frac{B_m}{m!n!} = \sum_{m=0}^{r-1} \frac{B_m}{m!(r-m)!} = \frac{1}{r!}\sum_{m=0}^{r-1} \binom{r}{m} B_m.$$

Comparing this with the left-hand side of (6.11), we see that

$$\frac{1}{r!}\sum_{m=0}^{r-1} \binom{r}{m} B_m = \begin{cases} 1 & if\, r = 1, \\ 0 & if\, r > 1. \end{cases} \tag{6.12}$$

For $r = 1, 2, \ldots$, this is an infinite sequence of linear equations

$$\begin{aligned} B_0 &= 1, \\ B_0 + 2B_1 &= 0, \\ B_0 + 3B_1 + 3B_2 &= 0, \end{aligned}$$

and so on, which we can solve in succession to find each B_m. (A more efficient but less elementary method for evaluating Bernoulli numbers is given in Graham *et al.,* 1989,Chapter 6, Section5.) The first few values are

$$B_0 = 1, \quad B_1 = -\frac{1}{2}, \quad B_2 = \frac{1}{6}, \quad B_3 = 0, \quad B_4 = -\frac{1}{30}, \quad B_5 = 0, \quad B_6 = \frac{1}{42},$$

and so on. (In particular, $B_m = 0$ for all odd $m > 1$, reflexting the fact that $\cot z$ is an odd function.) Substituting these values for even m in (6.10), we get

$$\zeta(2) = \frac{\pi^2}{6} = 1.64493406\ldots, \zeta(4) = \frac{\pi^4}{90} = 1.08232323\ldots, \zeta(6) = \frac{\pi^6}{945} = 1.01734306\ldots.$$

The coefficients in the linear equations (6.12) are all rational numbers, so it follows by induction on r that the Bernoulli numbers are all rational. It then follows from (6.10) that $\zeta(2k)$ is rational multiple of π^{2k}. Now a complex number is said to be *algebraic* if it is a root of some non-trivial polynomial with integer coefficients (for instance, $\sqrt{2}$ is a root of $x^2 - 2$); all other complex numbers are called *transcendental.* In 1882 Lindemann proved that π is transcendental; it follows easily that π^{2k}, and hence $\zeta(2k)$ is also transcendental, and are therefore irrational.

§6.4 Dirichlet series

We defined the Riemann zeta function as $\zeta(s) = \sum_{n=1}^{\infty} 1/n^s$, and then saw that $1/\zeta(s) = \sum_{n=1}^{\infty} \mu(n)/n^s$. These are just two examples of an important class of series of this general form.

Definition. If f is an arithmetic function, the its *Dirichlet* series is the series

$$F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s}$$

For convenience, we will often abbreviate this to $F(s) = \sum f(n)/n^s$, with the convention that $\sum$ without limits denotes $\sum_{n=1}^{\infty}$. Just as generating functions $A(x) = \sum a_n x^n$ are useful for studying sequences (a_n) defined by recurrence relations, Dirichlet series $F(s)$ are useful for studying arithmetic functions f, especially those associated with primes and divisors. For instance, in 1837 Dirichlet used series of this type, called L-series, to prove Theorem 6.10, that if a and b are coprime then there are infinitely many primes $p \equiv b \mod (a)$. The arithmetic functions u, N and μ have particularly simple Dirichlet series.

Example 6.1. If $f = u$ then $F(s) = \sum_u (n)/n^s = \sum 1/n^s = \zeta(s)$.

Example 6.2. If $f = N$ then $F(s) = \sum N(n)/n^s = \sum n/n^s = \sum 1/n^{s-1} = \zeta(s-1)$.

Example 6.3. If $f = \mu$ then $F(s) = \sum \mu(n)/n^s = 1/\zeta(s)$ by Theorem 6.4.

The next result helps to explain the importance of Dirichlet series: multiplication of Dirichlet series corresponds to Dirichlet product of arithmetic functions.

Theorem 6.6. *Suppose that*

$$F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s}, \qquad G(s) = \sum_{n=1}^{\infty} \frac{g(n)}{n^s} \quad \text{and} \quad H(s) = \sum_{n=1}^{\infty} \frac{h(n)}{n^s},$$

where $h = f * g$, Then

$$H(s) = F(s)G(s)$$

for all s such that $F(s)$ and $G(s)$ both converge absolutely.

Proof. If $F(s)$ and $G(s)$ both converge absolutely, then we can multiply these series and rearrange their terms to give

$$F(s)G(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s} \cdot \sum_{n=1}^{\infty} \frac{g(n)}{n^s} = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \frac{f(m)g(n)}{(mn)^s} = \sum_{k=1}^{\infty} \sum_{mn=k} \frac{f(m)g(n)}{k^s}$$
$$= \sum_{k=1}^{\infty} \frac{(f*g)(k)}{k^s} = \sum_{k=1}^{\infty} \frac{h(k)}{k^s} = H(s).$$

Example 6.4 If we take $f = \mu$ and $g = u$, then $h = f * g = \mu * u = I$ by our definition of μ (Chapter 8,Sections 3 and 6). Now $I(n) = 0$ for all $n > 1$, so $H(s) = \sum I(n)/n^s = 1$ for all s. We have $F(s) = \sum \mu(n)/n^s$, and $G(s) = \sum u(n)/n^s = \sum 1/n^s = \zeta(s)$, both absolutely convergent for $s > 1$; hence Theorem 6.6 gives

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} \cdot \zeta(s) = 1,$$

so that

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = \frac{1}{\zeta(s)}$$

for all $s > 1$, proving part of Theorem 6.4.

Example 6.5. Let $f = \varphi$ and $g = u$. As before, $G(s) = \zeta(s)$ is absolutely convergent for $s > 1$. Now $1 \leq \phi(n) \leq n$ for all n, so $F(s) = \sum \phi(n)/n^s$ is absolutely convergent by comparison with $\sum n/n^s = \zeta(s-1)$ for $s - 1 > 1$, that is, for $s > 2$. Thus Theorem 6.6 is valid for $s > 2$. Now Theorem 5.8 gives $\phi * u = N$,so

$$\sum_{n=1}^{\infty} \frac{\phi(n)}{n^s} \cdot \zeta(s) = \sum_{n=1}^{\infty} \frac{N(n)}{n^s} = \sum_{n=1}^{\infty} \frac{n}{n^s} = \zeta(s-1)$$

and hence

$$\sum_{n=1}^{\infty} \frac{\phi(n)}{n^s} = \frac{\zeta(s-1)}{\zeta(s)}$$

for all $s > 2$.

§6.5 Euler product

Many Dirichlet series have product expansions analogous to that in Theorem 6.3, in which the factors are indexed by the primes. these are called *Euler product.* First we need to consider a stronger form multiplicativity. Let's recall that an arithmetic function f is *completely multiplicative* if $f(mn) = f(m)f(n)$ for all positive integers m and n.

Example 6.6. The function N, u and I are completely multiplicative, whereas the multiplicative function μ and φ are not.

Theorem 6.7. (i). *If f is multiplicative, and $\sum_{n=1}^{\infty} f(n)$ is absolutely convergent, then*

$$\sum_{n=1}^{\infty} f(n) = \prod_p (1 + f(p) + f(p^2) + \cdots).$$

(ii). *If f is completely multiplicative, and $\sum_{n=1}^{\infty} f(n)$ is absolutely convergent, then*

$$\sum_{n=1}^{\infty} f(n) = \prod_p \frac{1}{1 - f(p)}.$$

In each case, p ranges over all the primes.

Proof. (i). The proof follows that used for Theorem 6.3. Let $p_1, \cdots p_k$ be the first k primes, and let

$$P_k = \prod_{i=1}^{k} (1 + f(p_i) + f(p_i^2) + \cdots).$$

The general term in the expansion of P_k is $f(p_1^{e_1}) \cdots f(p_l^{e_k}) = f(p_1^{e_1} \cdots p_k^{e_k})$, because f is multiplicative. Thus

$$P_k = \sum_{n \in A_k} f(n)$$

where $A_k = \{n | n = p_1^{e_1} \cdots p_k^{e_k}, \quad e_i \geq 0\}$. We have

$$|P_k - \sum_{n=1}^{\infty} f(n)| = | \sum_{n \notin A_k} f(n)| \leq \sum_{n \notin A_k} |f(n)| \leq \sum_{n \geq p_k} |(f(n)|,$$

since $n \geq p_k$ for each $n \notin A_k$. Now $\sum_{n=1}^{\infty} |f(n)|$ converges, so as $k \to \infty$ we have $\sum_{n \geq p_k} |f(n)| \to 0$ and hence $|P_k - \sum_{n=1}^{\infty} f(n)| \to 0$; thus $P_k \to \sum_{n=1}^{\infty} f(n)$ as $k \to \infty$, as required.

(ii). If f is completely multiplicative, then $f(p^e) = f(p)^e$ for each prime-power p^e, so part (a) gives

$$\begin{aligned} \sum_{n=1}^{\infty} f(n) &= \prod_p (1 + f(p) + f(p^2) + \cdots) \\ &= \prod_p (1 + f(p) + f(p)^2 + \cdots) \\ &= \prod_p \frac{1}{1 - f(p)} \end{aligned}$$

We can apply this result to Dirichlet series:

Corollary 6.8. *Suppose that $\sum_{n=1}^{\infty} f(n)/n^s$ converges absolutely. If f is multiplicative, the*

$$\sum_{n=1}^{\infty}\frac{f(n)}{n^s}=\prod_p\left(1+\frac{f(p)}{p^s}+\frac{f(p^2)}{p^{2s}}\cdots\right),$$

and if f is completely multiplicative, then

$$\sum_{n=1}^{\infty}\frac{f(n)}{n^s}=\prod_p\frac{1}{1-f(p)p^{-s}}.$$

Proof. In Theorem 6.7, We simply replace $f(n)$ with $f(n)n^{-s}$, which is multiplicative (or completely multiplicative) if and only if $f(n)$ is. □

Example 6.7. The function u is completely multiplicative, so as a special case of Theorem 6.3, we get that

$$\zeta(s)=\sum_{n=1}^{\infty}\frac{1}{n^s}=\prod_p\frac{1}{1-p^{-s}}$$

for all $s>1$.

Example 6.8. The Mӧbius function $\mu(n)$ is multiplicative, with $\mu(p)=-1$ and $\mu(p^e)=0$ for all $e\geq 2$, so

$$\sum_{n=1}^{\infty}\frac{\mu(n)}{n^s}=\prod_p\left(1+\frac{\mu(p)}{p^s}+\frac{\mu(p^2)}{p^{2s}}+\cdots\right)=\prod_p(1-p^{-s})$$

for all $s>1$. Inverting the factors in this product we obtain $1/\zeta(s)$ by the previous example, so this completes the proof of Theorem 6.4 which we promised earlier.

§6.6 Complex variables

In considering Dirichlet series $F(s)=\sum f(n)/n^s$, such as the Riemann zeta function $\sum 1/n^s$, we have assumed (often implicitly) that the variable s is real. For many purposes, this is adequate, but for some more advanced applications it is necessary to allow s to be complex. The advantage of this is that functions of a complex variable are often easier to deal with than those of a real variable: in particular, their domains of definition can often be extended by analytic continuation, and they can be integrated by the calculus of residues, techniques which are not available if we restrict to real variables.

Our earlier results on Dirichlet series and Euler products all extend to the case where s is a complex variable, provided we have absolute convergence. We therefore need to consider the subset of the complex plane $\mathbb{C}$ on which a Dirichlet series converges absolutely on a disc (which may be the whole plane or single point), a Dirichlet series has a half-plane of absolute convergence, which may be the whole plane or empty set.

Following the traditional (if slightly bizarre) notation we put

$$s=\sigma+it\in\mathbb{C}\qquad\text{where}\qquad\sigma,t\in\mathbb{R}$$

Then $n^s+n^{\sigma+it}=n^{\sigma}\cdot n^{it}=n^{\sigma}\cdot e^{it\log(n)}$ with $n^{\sigma}>0$. and $|e^{it\log(n)}|=1$, so $|n^s|=n^{\sigma}$. Now suppose that $F(s)$ converges absolutely (that is $\sum|f(n)/n^s|$ converges) at some point $s=a+ib\in\mathbb{C}$; if $\sigma\geq a$ then

$$\left|\frac{f(n)}{n^{\sigma+it}}\right|=\left|\frac{f(n)}{n^{\sigma}}\right|\leq\left|\frac{f(n)}{n^{a}}\right|=\left|\frac{f(n)}{n^{a+ib}}\right|,$$

so$\sum f(n)/n^{\sigma+it}$converges absolutely by the Comparison Test. This implies

Theorem 6.9. *Suppose that $\sum_{n=1}^{\infty}|f(n)/n^s|$ neither converges for all $s \in \mathbb{C}$, nor diverges for all $s \in \mathbb{C}$. Then there exists $\sigma_a \in \mathbb{R}$ such that $\sum_{n=1}^{\infty}|f(n)/n^s|$ converges for all $s = \sigma + it$ with $\sigma > \sigma_a$, and diverges for all $s = \sigma + it$ with $\sigma < \sigma_a$.*

Proof. We take σ_a to be the least upper bound of all $a \in \mathbb{R}$ such that $\sum |f(n)/n^s|$ diverges at $s = a + ib$; by the preceding argument this coincides with the greatest lower bound of all $a \in \mathbb{R}$ such that$\sum |f(n)/n^s|$ converges at $s = a + ib$. □

Definition. We call σ_a the *abscissa of absolute convergence* of $F(s)$, and $\{s = \sigma + it \in \mathbb{C} | \sigma > \sigma_a\}$ its half-plane of absolute convergence.

Note that the theorem says nothing about the behavior of $F(s)$ when $\sigma = \sigma_a$. Note also that there are two other extreme possibilities, not covered by the theorem:$F(s)$ may converge absolutely for all $s \in \mathbb{C}$, or for no $s \in \mathbb{C}$; we then write$\sigma_a = -\infty$ or $+\infty$ respectively. A similar but more complicated argument shows that there exists $\sigma_c \leq \sigma_a$, called the abscissa of convergence, such that $F(s)$ converges for $\sigma > \sigma_c$ and diverges for $\sigma < \sigma_c$; if $\sigma_c < \sigma_a$, then convergence is conditional for $\sigma_c < \sigma < \sigma_a$.

Example 6.9. Theorem 6.1 states that $\sum 1/n^s$ converges (absolutely) for all real $s > 1$ and diverges for $s \leq 1$. This series therefore has $\sigma_a = \sigma_c = 1$, so it converges absolutely for all $s = \sigma + it \in \mathbb{C}$ with $\sigma > 1$, and diverges for $\sigma < 1$. Similarly, $\sum (-1)^n/n^s$ has $\sigma_a = 1$, but in this case $\sigma_c = 0$ since the series converges for all $s > 0$ by the Alternating Test (Appendix C), but diverges for $s \leq 0$.

Example 6.10. If f is bounded, say $|f(n)| \leq M$ for all n, then $|f(n)/n^s| \leq M/n^{\sigma}$ where $s = \sigma + it$, so $\sum f(n)/n^s$ converges absolutely whenever $\sigma > 1$, by comparison with $\sum M/n^{\sigma}$. (It may converge absolutely for smaller σ, depending on the particular function f.) This applies to $f = \mu$ for example, with $M = 1$. More generally, if there are constants M and k such that $|f(n)| \leq Mn^k$ for all n, then $\sum f(n)/n^s$ converges absolutely for $\sigma > 1+k$ by comparison with $\sum Mn^k/n^{\sigma}$. Now $|\phi(n)| \leq n$ for all n, so taking $k = 1$ we see that $\sum \phi(n)/n^s$ converges absolutely for $\sigma > 2$.

A complex function $F(s)$ is said to be analytic if it is differentiable with respect to s.

Theorem 6.10. *A Dirichlet series $\sum_{n=1}^{\infty} f(n)/n^s$ represents an analytic function $F(s)$ for $\sigma > \sigma_c$, with derivative $F'(s) = \sum_{n=1}^{\infty} f(n)\log(n)/n^s$.*

Proof. (Outline proof) For each $n \geq 1$, the function $f(n)/n^s = f(n)e^{-s\log(n)}$ is analytic for all s (since the exponential function e^s is analytic), with derivative $\varphi(n)\log(n)/n^s$. One now shows that $\sum f(n)/n^s$ converges uniformly on all compact (closed, bounded) subset of the half-plane $\sigma.\sigma_c$, and then quotes the theorem that a uniformly convergent series of analytic functions has an analytic sum, which may be differentiated term by term. For full details, see Apostol (1976, Chapter 11, Section 7). □

For example, the series $\sum 1/n^s$ defines an analytic function $\zeta(s)$ on the half-plane $\sigma > 1$. Riemann used analytic continuation to extend the domain of a simple pole at $s = 1$ (this means that $(s-1)\zeta(s)$ is analytic at $s = 1$, so that $\zeta(s)$ diverges there like$1/(s-1)$). Note that we do not claim that the series $\sum 1/n^s$ converges outside the half-plane $\sigma > 1$: what Riemann showed is that there is a function $\zeta(s)$ which is analytic for all $s \neq 1$, and which agrees for instance the geometric series $1 + z + z^2 + \cdots$ converges (absolutely) for all z with $|z| < 1$, and with this disc of convergence its sum is given by $(1-z)^{-1}$; however, this function $(1-z^{-1}$ is analytic for all $a \neq 1$, even though the series diverges for $|z| \geq 1$.

Riemann showed that the extended function $\zeta(s)$ has zeros at $s = -2, -4, -6, \cdots$; these are called the *trivial zeros.* and he showed that the remaining non-trivial zeros all lie in the critical strip $0 \leq \sigma \leq 1$. the celebrated Riemann Hypothesis is the conjecture that these non-trivial zeros all lie on the line $\sigma = 1/2$. A great deal is now known about the location of the zeros of $\zeta(s)$: For instance, Hardy show in 1914 that there are infinitely many on the line $\sigma = 1/2$. Despite strong evidence in its favor, the Riemann Hypothesis is still unproved; since many conjectures about the distribution of prime numbers depend on this result, the resolution of this problem remains one of the greatest challenges of numbers theory.

Exercises for Chapter 6

(1). Use a similar argument to outline a proof that

$$\prod_p (1-p^{-s}) = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s},$$

where μ is the Möbius function, and hence show that

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = \frac{1}{\zeta(s)}.$$

(2). Show that if $s > 1$ then $\zeta(s) \geq (1 + f(s))/2$, where $f(s) := \frac{1}{1-2^{1-s}}$, and deduce that $\zeta(s) \to +\infty$ as $s \to 1$.

(3). Show that $P_k(1) \to +\infty$ as $k \to \infty$, and deduce that for each $\varepsilon > 0$ there exists n such that $\varphi(n)/n < \varepsilon$.(See Exercise 5.11 in Chapter 5 for a similar result, and for a probabilistic interpretation of this.)

(4). Prove Theorem 6.4.

(5). Show that an integer point $(x, y) \neq (0, 0)$ is visible from O if and only if x and y are coprime.

(6). For each integer $s \geq 2$, let $P(s)$ denote the probability that s randomly- and independently-chosen integers have greatest common divisor 1 (so $P = P(2)$). Give three arguments to show that $P(s)$ is given by the formulae

$$\frac{1}{\zeta(s)}, \qquad \prod_p (1-p^{-s}), \qquad \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s}.$$

(7). Prove (in three different ways) that a single randomly-chosen integer x is square-free with probability $P = 1/\zeta(2)$. (Hint: consider $Sq(x)$, the largest square factor of x.)

(8). For each integer $s \geq 2$, calculate (in three different ways) the probability $Q(s)$ that a randomly-chosen integer x should be s-th power-free, that is, divisible by no s-th power greater than 1.

(9). Assuming Lindemann's result, prove the remarks in the last sentence.

(10). Show that

$$\sum_{n=1}^{\infty} \frac{\tau(n)}{n^s} = \zeta(s)^2$$

for all $s > 1$.

(11). Express $\sum_{n=1}^{\infty} \sigma_k(n)/n^s$ in terms of the Riemann zeta function, where $\sigma_k(n) = \sum_{d|n} d^k$.

(12). Liouville's function λ is defined by

$$\lambda(p_1^{e_1}\cdots p_k^{e_k}) = (-1)^{e_1+\cdots+e_k}$$

where $p_1, \cdots, p_k$ are distinct primes. Show that

$$\sum_{d|n} \lambda(d) = \begin{cases} 1 & \text{if } n \text{ is a perfect square,} \\ 0 & \text{otherwise,} \end{cases}$$

and hence show that

$$\sum_{n=1}^{\infty} \frac{\lambda(n)}{n^s} = \frac{\zeta(2s)}{\zeta(s)}$$

for all $s > 1$.

(13). Let $\nu(n)$ be the number of distinct primes dividing n (so that $\nu(60) = 3$, for instance). Show that

$$\sum_{n=1}^{\infty} \frac{\nu(n)}{n^s} = \zeta(s) \sum_{p} \frac{1}{p^s},$$

where p ranges over the set of primes. For which real s is this valid?

(14). Find the Euler product expansion for the Dirichlet series $\sum_{n=1}^{\infty} |\mu(n)|/n^s$, and hence show that $\sum_{n=1}^{\infty} |\mu(n)|/n^s = \zeta(s)/\zeta(2s)$ for $s > 1$. Deduce that if $n > 1$ then $\sum_d \lambda(d) = 0$, where d ranges over the divisors of n such that n/d is square-free. (Here λ is Liouvill's function, defined in Exercise 6.12.)

(15). Show that $\prod_{p \le x}(1 - p^{-1}) \to 0$ as $x \to +\infty$.

(16). Find examples of Dirichlet series for which $\sigma_a = -\infty$ and $\sigma_a = +\infty$.

(17). Show that $\zeta'(s) = -\sum \log(n)/n^s$ and $-\zeta'(s)/\zeta(s) = \sum \Lambda(n)/n^s$ for all s with $\sigma > 1$, where Λ is Mangoldt's function (see Exercise 8.24).

(18). Let $\tau_k(n)$ be the number of k-tuples $(d_1, \cdots, d_k)$ of positive integers d_i such that $d_1 \cdots d_k = n$(so that $\tau_2 = \tau$, for instance). Show that $\sum_{n=1}^{\infty} \tau_k(n)/n^s = \zeta(s)^k$ for all $s = \sigma + it$ with $\sigma > 1$.

(19). Show that $\sum_{n=1}^{\infty} \tau(n)^2/n^s = \zeta(s)^4/\zeta(2s)$ for $\sigma > 1$.

(20). Recall that $\pi(x)$ is the number of primes $p \le x$. Show that if $q(x)$ denotes the number of square-free integers $m \le x$, then

$$2^{\pi(x)} \ge q(x) \ge x\left(2 - \frac{\pi^2}{6}\right),$$

and hence

$$\begin{aligned} \pi(x) &\ge \log_2 x + \log_2\left(2 - \frac{\pi^2}{6}\right) \\ &= \frac{\log x}{\log 2} + \log_2\left(2 - \frac{\pi^2}{6}\right) \end{aligned}$$

(This estimate for $\pi(x)$ is very weak: for instance it gives $\pi(10^9) \ge 28$, whereas in fact $\pi(10^9) \approx 5 \times 10^7$.)

(21). Let $f_k(n)$ denote the number of subgroups of finite index n in the group $\mathbb{Z}^k$. Express the Dirichlet series $F_k(s) = \sum_{n=1}^{\infty} f_k(n)/n^s$ of f_k in term of the Riemann zeta function. For which $s \in \mathbb{C}$ is your expression valid.

Chapter 7. Integer Matrices with Applications

§7.1. Hermite and Smith normal forms of the 2×2 integer matrices

We first discuss 2×2 matrices in order to describe the content of this chapter. Part of this discussion already appeared in Chapter 13 but for the sake of completeness and ease of understanding we shall repeat slightly.

Throughout this section by a matrix we mean the square matrix

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tag{7.1}$$

where a, b, c, d are integers, and we call them the elements of the matrix M. If all the elements are zero, then we call the matrix a *null* matrix and we denote it by 0. The quantity

$$ad - bc$$

is called the *determinant* of the matrix M. If determinant is ± 1, then we call M a *modular matrix*, and if the determinant is 1, then we call M a *positive modular matrix*. If the determinant is zero, then we say that M is *singular*, otherwise we say that M is *non-singular*.

The *product* of two matrices

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad B = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}$$

is defined to be the matrix

$$\begin{pmatrix} aa_1 + bc_1 & ab_1 + bd_1 \\ ca_1 + dc_1 & cb_1 + dd_1 \end{pmatrix}, \tag{7.2}$$

which is denoted by AB. It is clear that the determinant of AB is the product of the determinants of A and B, and hence the product of two (positive) modular matrices is a (positive) modular matrices.

Let k be an integer. We define

$$k \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} ka & kb \\ kc & kd \end{pmatrix}$$

The matrix

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

is called the *unit matrix*. For any matrix M we always have $MI = IM = M$. If $AB = I$, then we call B the inverse of A and we denote it by A^{-1}. It is easy to see that inverse of a modular matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ always exists, and that

$$A^{-1} = \pm \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

where we take the positive sign if and only if A is a positive modular matrix. it is also clear that $AA^{-1} = A^{-1}A = I$. Again, from taking the determinants of both sides of the equation $AB = I$, we see that if A has an inverse, then A must be a modular matrix. Therefore a necessary and sufficient condition for the existence of A^{-1} is that A be a modular matrix.

There are two very important positive modular matrices, namely

$$S = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \tag{7.3}$$

and

$$T = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{7.4}$$

it is easy to verify that, for any integer m,

$$S^m = \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} \tag{7.5}$$

and

$$T^2 = -I. \tag{7.6}$$

Theorem 7.1.1 *Any positive modular matrix can be expressed as a product of the matrices S and T. In other words the group of positive modular matrices can be generated by S and T.*

Proof. Let

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{7.7}$$

be a positive modular matrix. If $a = 0$, then $b \neq 0$, and so from

$$\begin{pmatrix} 0 & b \\ c & d \end{pmatrix} T = \begin{pmatrix} -b & 0 \\ -d & c \end{pmatrix}$$

we see that we may assume that $a \neq 0$. Again from $MT^2 = -M$ we may further assume that $a > 0$. We can also suppose that

$$0 \leq b < a, \tag{7.8}$$

since we can choose an integer q such that $0 \leq aq + b < a$, and the matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & q \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a & aq+b \\ c & cq+d \end{pmatrix} \tag{7.9}$$

satisfies the condition (7.8).

We now proceed by induction on a. If $a = 1$, then, by (7.8), $b = 0$ and hence $d = 1$ giving

$$\begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -c \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = TS^{-c}T^{-1}$$

so that the matrix (7.7) is a product of S and T.

Suppose now that, for $0 < a < k$, all matrices (7.7) satisfying the condition (7.8) and products of S and T. Then the positive modular matrix

$$\begin{pmatrix} k & l \\ s & t \end{pmatrix}, \qquad 0 \leq l < k$$

(since $k > 1$, we see that l must be positive) satisfies

$$\begin{pmatrix} k & l \\ s & t \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} l & -k \\ t & -s \end{pmatrix}$$

and we see from the method of (7.9) that the right hand side of this equation is product of S and T. The inductive argument is complete. □

Note : Positive modular matrices can also be expressed as a product of

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}. \tag{7.10}$$

This is because

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Theorem 7.1.2. *Any modular matrix can be expressed as a product of the matrices*

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \textit{and} \quad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}. \tag{7.11}$$

That is the group of modular matrices can be generated by these two matrices.

Proof. If a modular matrix M is not positive, then $M \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is positive. It follows from the note above that any modular matrix is expressible as a product of the three matrices

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

But

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

so that the theorem follows. □

Definition 1. Let M and N be two matrices. Suppose that there is a modular matrix U such that

$$M = UN.$$

then we say that N is *left associated* to M, and we denote this by $M \overset{\mathrm{L}}{=} N$.

Clearly left association has the following three properties: (i) $M \overset{\mathrm{L}}{=} N$(reflexive); (ii) if $M \overset{\mathrm{L}}{=} N$, then $N \overset{\mathrm{L}}{=} M$(symmetric); (iii) if $M \overset{\mathrm{L}}{=} N, N \overset{\mathrm{L}}{=} P$, then $M \overset{\mathrm{L}}{=} P$(transitive).

A similar definition can be given for *right associated*.

Theorem 7.1.3. *Any matrix is left associated to a matrix of the form*

$$\begin{pmatrix} a & 0 \\ c & d \end{pmatrix}, \qquad a \geq 0, \qquad d \geq 0; \tag{7.12}$$

if $a > 0$, then $0 \leq c < a$.

Proof. Corresponding to the matrix

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

there are integers r, s such that

$$rb + sd = 0, (r, s) = 1.$$

Now there are integer u, v such that $rv - su = 1$ so that

$$U = \begin{pmatrix} r & s \\ u & v \end{pmatrix}$$

is a positive modular matrix, and

$$UM = \begin{pmatrix} a_1 & 0 \\ c_1 & d_1 \end{pmatrix}.$$

If $a_1 \leq 0$, then we multiply this matrix by $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ which will give a matrix with $a_1 \geq 0$, and similarly we can make $d_1 \geq 0$. Therefore every matrix is left association to a matrix of the form

$$\begin{pmatrix} a & 0 \\ c & d \end{pmatrix}, \qquad a \geq 0, \qquad d \geq 0.$$

If $a > 0$, then we can choose q so that $0 \leq qa + c < a$, and from

$$\begin{pmatrix} 1 & 0 \\ q & 1 \end{pmatrix}\begin{pmatrix} a & 0 \\ c & d \end{pmatrix} = \begin{pmatrix} a & 0 \\ qa+c & d \end{pmatrix}$$

we see that theorem is proved. □

Definition 2. We call the matrix in (7.12) *the normal form of Hermite.*

Theorem 7.1.4. *The normal form of Hermite for a non-singular matrix is unique.*

Proof. We first note that the normal form of Hermite $\begin{pmatrix} a & 0 \\ c & d \end{pmatrix}$ for a non-singular matrix cannot have a or d equal to zero. Now if

$$\begin{pmatrix} s & t \\ u & v \end{pmatrix}\begin{pmatrix} a & 0 \\ c & d \end{pmatrix} = \begin{pmatrix} a_1 & 0 \\ c_1 & d_1 \end{pmatrix}, \qquad sv - tu = \pm 1,$$

then , from $td = 0$, we have $t = 0$. Also, from $sa = a_1 > 0, vd = d_1 > 0$ as $sv = \pm 1$ we see that $s = v = 1$. Finally, from $ua + c = c_1, 0 \leq c < 1, 0 \leq c_1 < a_1 = a$ we see that $u = 0$. the theorem is proved. □

Exercise. Investigate the situation for a singular matrix.

Definition 3. Let there be two modular matrices U and V such that

$$UMV = N.$$

Then we say that M and N are *equivalent*, and we write $M \sim N$. Clearly, being equivalent has the three properties of being reflexive, symmetric, and transitive.

Theorem 7.1.5. *Any matrix is equivalent to a matrix of the form*

$$\begin{pmatrix} a_1 & 0 \\ 0 & a_1 a_2 \end{pmatrix}, \qquad a_1 \geq 0, \qquad a_2 \geq 0 \tag{7.13}$$

.

Proof. Consider the matrix

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Since the theorem becomes trivial if M is the null matrix we can assume that $a \neq 0$ and indeed we can even assume that $a > 0$. We first prove that M must be equivalent to a matrix of the form

$$\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}, \qquad a_1 | (b_1, c_1, d_1).$$

We use induction on a, the case $a = 1$ being trivial. When $a > 1$ and $a \nmid b$, we can choose q so that $0 < aq + b < a$ and we consider

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} q & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} aq+b & * \\ * & * \end{pmatrix},$$

where the leading element is a positive integer less than a. If $a \mid b$ and $a \nmid c$, then we choose q' such that $0 < aq' + c < a$, and we consider

$$\begin{pmatrix} q' & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} aq'+c & * \\ * & * \end{pmatrix},$$

where the leading element is once again a positive integer less than a. Finally, if $a \mid (b, c)$, but $a \nmid d$, we let $c = c'a$ so that

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -c' & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & (1-c')b+d \\ * & * \end{pmatrix},$$

and $a \nmid \{(1-c')b + d\}$ which reduces back to the case when $a \nmid b$. The inductive argument is now complete.

Now $a_1 \mid (b_1, c_1, d_1)$. We let $b_1 = a_1b_2, c_1 = a_1c_2$,and $d_1 = a_1d_2$, and we consider

$$\begin{pmatrix} 1 & 0 \\ -c_2 & 1 \end{pmatrix} \begin{pmatrix} a_1 & a_1b_2 \\ a_1c_2 & a_1d_2 \end{pmatrix} \begin{pmatrix} 1 & -b_2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a_1 & 0 \\ 0 & a_1(d_2 - b_2c_2) \end{pmatrix}$$

where we can assume that $a_1 > 0$, since otherwise we can multiply by$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$. Similarly we can assume that $a_2 = d_2 - b_2c_2 \geq 0$. The theorem is proved. □

Definition 4. We call the matrix in (7.13) *the normal form of Smith.*

We summarize our result as follows: By Theorem 1.2 any modular matrix is a product of the matrices

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

From

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} c & b \\ a & d \end{pmatrix}$$

and

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} b & a \\ d & c \end{pmatrix}$$

we see that the effect of multiplying by$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$or its inverse is merely the interchanging of the two rows or the two columns of the matrix. Again, from

$$\begin{pmatrix} 1 & \pm 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a \pm c & b \pm d \\ c & d \end{pmatrix}$$

and

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & \pm 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a & b \pm a \\ c & d \pm c \end{pmatrix},$$

we see that effect of multiplying by$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$or by its inverse$\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$is the addition or subtraction of the second row to the first row, or the first column to the second column of the matrix. We call these operations here the elementary transformations of the matrix. We can there restate Theorem 1.5 as follows: We can use elementary transformation to reduce a given matrix to the normal form of Smith.

Now the greatest common factor of the elements of a matrix is invariant under an elementary transformation, and so from Theorem 1.5 we have $(a, b, c, d) = a_1$. Also

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc = \pm a_1^2 a_2.$$

Therefore we have

Theorem 7.1.6 *The normal form of Smith for a given matrix is unique.*

§7.2. The product of matrices

Let $a_{11}, a_{12}, \cdots, a_{mn}$ be integers. We call the array

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

an m by n matrix and we sometimes denote it by $A^{(m,n)}$. If $m = n$, then we denote it by $A^{(n)}$and we call it a square matrix of order n. Let B be an n by l matrix

$$B = \begin{pmatrix} b_{11} & \cdots & b_{1l} \\ b_{21} & \cdots & b_{2l} \\ \cdots & \cdots & \cdots \\ b_{n1} & \cdots & b_{nl} \end{pmatrix}.$$

We define the *product matrix* of A and B by

$$AB = C = \begin{pmatrix} c_{11} & \cdots & c_{1l} \\ c_{21} & \cdots & c_{2l} \\ \cdots & \cdots & \cdots \\ c_{m1} & \cdots & c_{ml} \end{pmatrix}, \qquad C_{rs} = \sum_{n} a_{rt} b_{ts} (r = 1, \cdots, m; s = 1, \cdots, l). \tag{7.14}$$

We see from the the definition that the product matrix of A and B exists only when the number of columns in A is the same as the number of rows in B. Note also that, when AB and BA both exist, they may be different. If $AB = BA$, then we say that A and B *commute*. However we always have $(AB)D = A(BD)$ whenever either side of this equation exists.

If A and B are square matrices, then the determinant of AB is the product of the determinants of A and B. A square matrix whose determinant is zero is called a *singular* matrix, otherwise we call it a *non-singular* matrix. *Modular matrices* are those square matrices whose determinant equal ±1 and *positive modular matrices* are those whose determinants equal 1. Clearly the product of two (positive) modular matrices is a (positive) modular matrix.

The square matrix

$$A = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

where each element not on the main diagonal is zero is called a *diagonal matrix*, and we denote it simply by $\Lambda = [\lambda_1, \lambda_2, \cdots, \lambda_n]$. If $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 1$then,

$$\Lambda = I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

and we call I the *unit matrix.* Clearly we have $AI = IA = A$ for any square matrix A of order n.

If the square matrices A and B satisfy AB=I, then we call B the *inverse* of A and we denote it by A^{-1}.

Consider a square matrix $A(= A^{(n)})$. By the $cofactor$ of the element a_{ij} we mean the determinant of the square matrix of order $(n-1)$ obtained by removing the *i-th* row and the *j-th* column of A. If we attach the sign $(-1)^{i+j}$ to the cofactor of a_{ij} then we call it the *algebraic cofactor* of a_{ij} and we denote it by A_{ij}. Let

$$A_0 = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix},$$

that is the matrix obtained from A by replacing each element a_{rs} with the algebraic cofactor A_{rs} of a_{rs}, is called the *adjoint matrix* of A. It is not difficult to prove that

$$AA_0 = A_0A = aI,$$

where a is the determinant of A. It follows that if A is modular matrix, then its inverse exists, and that $A^{-1} = \pm A_0$. Conversely, if A has an inverse, then it must be a modular matrix.

If $AB = I$, then from $B = (\pm A_0A)B = \pm A_0(AB) = \pm A_0$ we see that the inverse is unique and that $AA^{-1} = A^{-1}A = I$. Also, if A and B both have inverses, then $(AB)^{-1} = B^{-1}A^{-1}$.

A 1 by n matrix $(x_1, \cdots, x_n)$, where we no longer restrict the elements to be integers is called a *vector*, and we write $x = (x_1, \cdots, x_n)$. We should take care that this notation here is not to be confused with the greatest common factor symbol $(x_1, \cdots, x_n) = d$. We shall use the convention that $(x_1, \cdots, x_n)$ by itself always represents a vector, while $(x_1, \cdots, x_n) = d$ means the greatest common factor of $x_1, \cdots, x_n$. Also we shall always use the letters x and y to denote a vector with n terms.

The equation

$$y = xB \qquad (B = B^{(n,l)}) \tag{7.15}$$

represents the system of linear equations

$$y_i = \sum_{j=1}^{n} x_j b_{ji} \qquad 1 \le i \le l.$$

If $n = l$ and B is non-singular, then (7.15) is called a *transformation*. Corresponding to integers $x_1, \cdots, x_n$ the transformation gives integers $y_1, \cdots, y_n$, but not conversely. However, if B is a modular matrix, then when $y_1, \cdots, y_n$ are integers, the numbers $x_1, \cdots, x_n$ must also be integers. In this case we call (7.15) a *modular transformation*.

Example 1. Let $r \neq 1$, and $y_1 = -x_r, y_r = x_1, y_i = x_i (i \neq 1, \quad i \neq r)$. This is a modular transformation whose corresponding matrix is obtained from I by multiplying the first row by -1 and then interchanging it with the *r-th* row (or multiplying the *r-th* column by -1 and then interchanging it with the first column). We denote this matrix by E_r so that

$$E_r = \begin{array}{@{}c@{}c} \left(\begin{array}{cccccc} 0 & 0 & \cdots & 1 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ -1 & 0 & \cdots & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 \end{array}\right) & \begin{array}{c} \\ \\ \\ \quad r \\ \\ \\ \end{array} \\ \begin{array}{cccccc} & & & r & & \end{array} & \end{array} \qquad (7.16)$$

Example 2. Let $r \neq 1$, and $y_i = x_i (i \neq r), y_r = x_r + x_1$. This too is a modular transformation and its corresponding matrix is

$$V_r = \begin{array}{@{}c} \left(\begin{array}{cccccc} 1 & 0 & \cdots & 1 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 \end{array}\right), \\ \begin{array}{cccccc} & & & r & & \end{array} \end{array} \qquad (7.17)$$

that is the matrix obtained from I by adding the *r-th* row to the first row (or adding the first column to the *r-th* column).

It is easy to prove that V_r is representable as a product of V_2 and E_i. In fact, if $r > 2$, then

$$V_r = E_2 E_r E_2 V_2 E_2 E_r E_2. \qquad (7.18)$$

The proof is as follows: Let

$$t = \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{pmatrix},$$

so that

$$E_2 t = \begin{pmatrix} t_2 \\ -t_1 \\ t_3 \\ \vdots \\ t_n \end{pmatrix}, \quad E_r E_2 t = \begin{array}{@{}c@{}c} \left(\begin{array}{c} t_r \\ -t_1 \\ t_3 \\ \vdots \\ -t_2 \\ \vdots \\ t_n \end{array}\right) & \begin{array}{c} \\ \\ \\ \\ \quad r \\ \\ \\ \end{array} \end{array}, \cdots,$$

$$E_2E_rE_2V_2E_2E_rE_2t=\left(\begin{array}{c}t_1+t_r\\ t_2\\ \vdots\\ t_n\end{array}\right).$$

But

$$V_rt=\left(\begin{array}{c}t_1+t_r\\ t_2\\ \vdots\\ t_n\end{array}\right),$$

so that (7.18) follows.

Example 3. For fixed distinct r and s we let $y_i=x_i(i\neq s)$ and $y_s=x_s+x_r$. Then this is also a modular transformation whose matrix is obtained from I by adding the *s-th* row to the *r-th* row (or adding the *r-th* column to the *s-th* column). We denote this matrix by V_{rs} so that

$$V_{rs}=\begin{array}{cc}\left(\begin{array}{cccccccc}1&0&\cdots&0&\cdots&0&\cdots&0\\ \cdots&\cdots&\cdots&\cdots&\cdots&\cdots&\cdots&\cdots\\ 0&0&\cdots&1&\cdots&1&\cdots&0\\ \cdots&\cdots&\cdots&\cdots&\cdots&\cdots&\cdots&\cdots\\ 0&0&\cdots&0&\cdots&1&\cdots&0\\ \cdots&\cdots&\cdots&\cdots&\cdots&\cdots&\cdots&\cdots\\ 0&0&\cdots&0&\cdots&0&\cdots&1\end{array}\right)&\begin{array}{c}\\ \\ r\\ \\ s\\ \\ \\ \end{array}\\ \begin{array}{cccccccc}\ &\ &\ &r&\ &s&\ &\ \end{array}&\end{array}. \qquad (7.19)$$

When $s>1, V_{rs}=E_r^{-1}V_sE_r$, and $V_{r1}=E_r^{-1}V_r^{-1}E_r$. Therefore V_{rs} can also be represented as a product of V_2 and $E_2,\cdots,E_n$.

The matrices $V_{rs}(1\leq r\leq n, 1\leq s\leq n, r\neq s)$together with all the products formed by them forms a group which we denote by $\mathfrak{M}_n$. We saw , from the note following Theorem 1.1, that the group $\mathfrak{M}_2$, generated by the matrices $V_{21}=\left(\begin{array}{cc}1&0\\ 1&1\end{array}\right)$ and $V_{12}=\left(\begin{array}{cc}1&1\\ 0&1\end{array}\right)$, is identical with the group of all 2 by 2 positive modular matrices. We now prove the corresponding result for n by n positive modular matrices.

Theorem 7.2.1. *The group $\mathfrak{M}_n$ is the group of all n by n positive modular matrices.*

It is clear that each matrix in $\mathfrak{M}_n$ is a positive modular matrix so that we only have to prove that every positive modular matrix is in $\mathfrak{M}_n$, that is every positive modular matrix can be expressed as a product of the matrices V_{rs}. For this purpose we shall first establish the following two theorems.

Theorem 7.2.2 *If $(x_1,\cdots,x_n)=d$, then there exists $U\in\mathfrak{M}_n$ such that*

$$(x_1,\cdots,x_n)U=(d,0,\cdots,0).$$

Proof. Consider first the case $n=2$. If $(x_1,x_2)=d$, then there are integers r and s such that $rx_1+sx_2=d,(r,s)=1$. We take $u=-x_2/d, v=x_1/d$ so that $vx_2+ux_1=0,\ vr-us=1$. Thus

$$(x_1,x_2)\left(\begin{array}{cc}r&u\\ s&v\end{array}\right)=(d,0)$$

and $P = \begin{pmatrix} r & u \\ s & v \end{pmatrix}$ is a positive modular matrix. Since we already know that $P \in \mathfrak{M}_2$ by the note following Theorem 1.1, the case $n = 2$ is proved.

We now proceed by induction on n. Let $(x_{n-1}, x_n) = d_1$, so that there exists $P \in \mathfrak{M}_2$ such that

$$(x_{n-1}, x_n)P = (d_1, 0).$$

Let

$$V^{(n)} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & r & u \\ 0 & 0 & \cdots & s & v \end{pmatrix} = \begin{pmatrix} I^{(n-2)} & 0 \\ 0 & P \end{pmatrix}.$$

It is easy to see that $V^{(n)} \in \mathfrak{M}_n$ and that

$$(x_1, \cdots, x_n)V^{(n)} = (x_1, \cdots, x_{n-2}, d_1, 0).$$

From the induction hypothesis we have $V^{(n-1)} \in \mathfrak{M}_{n-1}$ and that

$$(x_1, \cdots, x_{n-2}, d_1)V^{(n-1)} = (d, 0, \cdots, 0).$$

We now let

$$V_1^{(n)} = \begin{pmatrix} V^{(n-1)} & 0 \\ 0 & 1 \end{pmatrix}$$

so that

$$(x_1, \cdots, x_n)V^{(n)}V_1^{(n)} = (d, 0, \cdots, 0).$$

It is easy to see that

$$U = V^{(n)}V_1^{(n)} \in \mathfrak{M}_n$$

so that the theorem is proved.

Theorem 7.2.3. *Let $(a_{11}, a_{12}, \cdots, a_{1n}) = d$. Then there is a matrix in $\mathfrak{M}_n$ whose first row is*

$$(\frac{a_{11}}{d}, \frac{a_{12}}{d}, \cdots, \frac{a_{1n}}{d}).$$

Proof. By Theorem 2.2 there is a matrix U in $\mathfrak{M}_n$ such that

$$(a_{11}, a_{12}, \cdots, a_{1n})U = (d, 0, \cdots, 0)$$

and so the matrix U^{-1} is a suitable candidate. □

Proof of Theorem 7.2.1. The case $n = 2$ is already established. We now use induction on n. Let

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

be any positive modular matrix. Clearly $(a_{11}, a_{12}, \cdots, a_{1n}) = 1$. On multiplying by the matrix U in Theorem 2.3 we have

$$AU = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ a'_{21} & a'_{22} & \cdots & a'_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a'_{n1} & a'_{n2} & \cdots & a'_{nn} \end{pmatrix}$$

The matrix

$$V = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -a'_{21} & 1 & 0 & \cdots & 0 \\ -a'_{31} & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -a'_{n1} & 0 & 0 & \cdots & 1 \end{pmatrix}$$

is in $\mathfrak{M}_n$, and

$$VAU = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & a'_{22} & a'_{23} & \cdots & a'_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & a'_{n2} & a'_{n3} & \cdots & a'_{nn} \end{pmatrix}. \tag{7.20}$$

From the induction hypothesis, the matrix

$$\begin{pmatrix} a'_{22} & a'_{23} & \cdots & a'_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a'_{n2} & a'_{n3} & \cdots & a'_{nn} \end{pmatrix}$$

is in $\mathfrak{M}_{n-1}$, and so the matrix

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & a'_{22} & a'_{23} & \cdots & a'_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & a'_{n2} & a'_{n3} & \cdots & a'_{nn} \end{pmatrix}$$

is in $\mathfrak{M}_n$. From (7.20) we see that the theorem follows.

§7.3. The number of generators for modular matrices

We proved in §1 that any 2 by 2 positive modular matrix can be expressed as product of the matrices $V_{21} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$and $V_{12} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. We now discuss the general case, and ask for the matrices whose products give all possible n by n positive modular matrices-that is we want to know the generators of the group $\mathfrak{M}_n$.

From the definition for $\mathfrak{M}_n$ any matrix in it is a product of V_{rs}, and from the previous section we know that each V_{rs} is expressible as a product of the following n matrices:

$$E_2 = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ -1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}, \cdots, \qquad E_n = \begin{pmatrix} 0 & 0 & 0 & \cdots & 1 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -1 & 0 & 0 & \cdots & 0 \end{pmatrix},$$

$$V_2 = \begin{pmatrix} 1 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Thus $\mathfrak{M}_n$ can be generates by the n matrices $E_2, E_3, \cdots, E_n, V_2$.

Let

$$U_1 = \begin{pmatrix} 0 & 0 & \cdots & 0 & (-1)^{(n-1)} \\ 1 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}$$

It is not difficult to prove that each of $E_2, E_3, \cdots, E_n$ is expressible as a product of U_1 and E_2. In fact, we have

$$\begin{aligned} E_r &= (E_2U_1)^{r-2}E_2(E_2U_1)^{n-r+1}, && \text{if } n \text{ is even,} \\ E_r &= (E_2^{-1}U_1)^{r-2}E_2(E_2^{-1}U_1)^{n-r+1}, && \text{if } n \text{ is odd, } r \text{ is even,} \\ E_r &= (E_2^{-1}U_1)^{r-2}E_2^{-1}(E_2^{-1}U_1)^{n-r+1}, && \text{if } n \text{ and } r \text{ is odd.} \end{aligned} \tag{7.21}$$

The proof of (7.21) is similar to that of (2.5).

Thus $\mathfrak{M}_n$ can be generated by the three matrices U_1, V_2, E_2. If we write

$$U^* = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix},$$

then it is easy to verify that $E_2 = U^{*-1}V_2U^{*-1}$, so that $\mathfrak{M}_n$ can also be generated by the three matrices

$$U_1 = \begin{pmatrix} 0 & 0 & \cdots & 0 & (-1)^{(n-1)} \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix},$$

$$U_2 = V_2 = \begin{pmatrix} 1 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}, U^* = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \tag{7.22}$$

When $n = 2$ we saw that $\mathfrak{M}_n$ can actually be generated by the *two* matrices $U_1 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and $U_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. We now ask whether $\mathfrak{M}_n (n \geq 3)$ can also be generated by U_1 and U_2 ; that is whether U^* is expressible as a product of U_1 and U_2. We first examine the cases $n = 3$ and 4.

(1). For $n = 3$, we have

$$U_1 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad U_2 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad U^* = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

In the following we call the positive for the *i-th* row and the *j-th* column the "position (i, j)". Consider the operation of multiplying U_2 by U_1 on the left and U_1^{-1} on the right. We see from

$$S = U_1U_2U_1^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, T = U_1^2U_2(U_1^{-1})^2 = U_1^{-1}U_2U_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix},$$

$$U_1^3U_2(U_1^{-1})^3 = U_2 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

that successive applications of the above operation will leave the elements in the main diagonal invariant, whereas the element 1 not on the main diagonal will take up the successive positions (1, 2), (2, 3), (3, 1). similarly the elements in the three positions(3, 2), (1, 3), (2, 1) will be permuted along a rail as shown in the diagram.

Consequently in order to obtain the element 1 in the position (2, 1)we have first to produce this element in one of the position (1, 3) or (3, 2). Now if we multiply T by U_2^{-1} on the left and U_2 on the right, it will give rise to the element 1 in the position (3, 2); that is

$$U_2^{-1}TU_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

The operation of multiplying by U_1^{-1} on the left and U_1 on the right will make the element 1 in the position (3,2) in the matrix $U_2^{-1}TU_2$ move to the position (2,1), that is

$$W = U_1^{-1}U_2^{-1}TU_2U_1 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Therefore we need only to annihilate the element 1 in the position (2, 3)to give the required matrix U^*, and this can be accomplished by S^{-1} on the left;that is

$$S^{-1}W = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = U^*.$$

Therefore, for $n = 3$, we have

$$U^* = U_1U_2^{-1}U_1U_2^{-1}U_1^{-1}U_2U_1U_2U_1. \qquad (.723)$$

(2). For $n = 4$ we have

$$U_1 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \qquad U_2 = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad U* = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Similarly to the case $n = 3$, we start with

$$T = U_1^{-1}U_2U_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}.$$

We can produce the element -1 in the position (4,2) by multiplying T by U_2^{-1} on the left and U_2 on the right;that is

$$U_2^{-1}TU_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & -1 & 0 & 1 \end{pmatrix}.$$

Again, the operation of multiplying by U_1^{-1} on the left and U_1 on the right will move the element -1 from the position (4,2) to the position (3,1); that is

$$U_1^{-1}(U_2^{-1}TU_2)U_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{7.24}$$

Performing the first operation of multiplying by U_2^{-1} on the left and U_2 on the right will now produce the element -1 in the position (3,2); that is

$$U_2^{-1}(U_1^{-1}U_2^{-1}TU_2U_1)U_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & -1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Performing the second operation of multiplying by U_1^{-1} on the left and U_1 on the right will now move the element -1 in the position (3, 2) to the position (2, 1); that is

$$W = U_1^{-1}(U_2^{-1}U_1^{-1}U_2^{-1}TU_2U_1U_2)U_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

At this point we observe that the elements of the matrix below the main diagonal matches those of U^{*-1}, and the problem now is the annihilation of the elements 1 above the main diagonal.

From (7.24) we have

$$S = U_1^{-1}(U_1^{-1}U_2^{-1}TU_2U_1)U_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

and hence

$$S^{-1}W = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = U^{*-1}.$$

Therefore, for $n = 4$, we have

$$\begin{aligned} U^{*-1} &= U_1^{-1}U_1^{-1}U_2^{-1}U_1^{-1}U_2^{-1}U_1U_2U_1U_1U_1^{-1}U_2^{-1}U_1^{-1}U_2^{-1}U_1^{-1} \\ &\times \quad U_2U_1U_2U_1U_2U_1 \end{aligned} \tag{7.25}$$

If we write U=U_2U_1, then (7.23) and (7.25) become

$$U^* = U_1^{-1}U^{-1}U_1U_1U^{-1}U_1^{-1}U^2 \qquad (n = 3),$$

$$U^{*-1} = U_1^{-1}(U^{-1})^2U_1UU_1(U^{-1})^2U_1^{-1}U^3 \qquad (n = 4), \tag{7.26}$$

and in general we have

$$U^{*(-1)^{(n-1)}} = U_1^{-1}(U^{-1})^{n-2}U_1U^{n-3}U_1(U^{-1})^{n-2}U_1^{-1}U^{n-1} \tag{7.27}$$

The reader can follow the proof of (2.5) to prove (.727). Therefore we have

Theorem 7.3.1. *The group $\mathfrak{M}_n$ of positive modular matrices can be generated by the two matrices*

$$U_1 = \begin{pmatrix} 0 & 0 & \cdots & 0 & (-1)^{(n-1)} \\ 1 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}, \qquad U_2 = \begin{pmatrix} 1 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

In other words, any positive modular matrix is expressible as a product of U_1 and U_2.

Any modular matrix which is not positive will become so on multiplying by

$$U_3 = \begin{pmatrix} -1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Therefore we have

Theorem 3.2 *the group of all modular matrices can be generated by the three matrices U_1, U_2 and U_3. In other words any modular matrix is expressible as a product of the matrices U_1, U_2 and U_3.*

§7.4. Left association

Definition 1. Let A and B be two square matrices. Suppose that there is a modular matrix U such that

$$A = UB.$$

then we say that B is *left associated* to A , and we denote this by $A \overset{\mathrm{L}}{=} B$.

Clearly left association is reflexive, symmetric and transitive.

Theorem 7.4.1. *Any square matrix is left associated to a matrix of the form*

$$\begin{pmatrix} b_{11} & 0 & 0 & \cdots & 0 & 0 \\ b_{21} & b_{22} & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ b_{n-1,1} & b_{n-1,2} & b_{n-1,3} & \cdots & b_{n-1,n-1} & 0 \\ b_{n1} & b_{n2} & b_{n3} & \cdots & b_{n,n-1} & b_{nn} \end{pmatrix}, \qquad (7.28)$$

where $b_{vv} \geq 0$, Also if $b_{vv} > 0$, then $0 \leq b_{iv} < b_{vv} (i > v)$.

Proof. The case $n = 2$ has already been proved (Theorem 1.3). We now proceed by induction on n. Let

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

be any square matrix. If there is a non-zero element in the last column of the matrix A, then we let $(a_{1n}, a_{2n}, \cdots, a_{nn}) = b_{nn}$. There are integers $b_1, b_2, \cdots, b_n$ such that

$$b_1 a_{1n} + b_2 a_{2n} + \cdots + b_n a_{nn} = b_{nn}, (b_1, b_2, \cdots, b_n) = 1.$$

By Theorem 7.2.3 there is a modular matrix V whose first row is $(b_1, b_2, \cdots, b_n)$. We interchange the first row of V with its *n-th* row to give a modular matrix U

whose *n-th* row is $(b_1, b_2, \cdots, b_n)$. We then have

$$A \overset{\mathrm{L}}{=} UA = \begin{pmatrix} a'_{11} & a'_{12} & \cdots & a'_{1n} \\ a'_{21} & a'_{22} & \cdots & a'_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a'_{n1} & a'_{n2} & \cdots & b_{nn} \end{pmatrix}.$$

It is easy to see that $a'_{1n}, \cdots, a'_{n-1,n}$ are linear combination of $a_{1n}, a_{2m}, \cdots, a_{nn}$ and are therefore divisible by b_{nn}. Therefore

$$A \overset{\mathrm{L}}{=} \begin{pmatrix} 1 & 0 & \cdots & 0 & -\frac{a'_{1n}}{b_{nn}} \\ 0 & 1 & \cdots & 0 & -\frac{a'_{2n}}{b_{nn}} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix} \begin{pmatrix} a'_{11} & a'_{12} & \cdots & a'_{1n} \\ a'_{21} & a'_{22} & \cdots & a'_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a'_{n1} & a'_{n2} & \cdots & b_{nn} \end{pmatrix}$$

$$= \begin{pmatrix} a''_{11} & \cdots & a''_{1,n-1} & 0 \\ a''_{21} & \cdots & a''_{2,n-1} & 0 \\ \cdots & \cdots & \cdots & \cdots \\ a''_{n-1,1} & \cdots & a''_{n-1,n-1} & 0 \\ a''_{n1} & \cdots & a''_{n,n-1} & b_{nn} \end{pmatrix}. \tag{7.29}$$

The above still holds even when all the elements in the last column of A are zero, except that we have $b_{nn} = 0$. It follows from the induction hypothesis that

$$A \overset{\mathrm{L}}{=} \begin{pmatrix} b_{11} & 0 & \cdots & 0 & 0 \\ b_{21} & b_{22} & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ b_{n-1,1} & b_{n-1,2} & \cdots & b_{n-1,n-1} & 0 \\ b'_{n1} & b'_{n2} & \cdots & b'_{n,n-1} & b_{nn} \end{pmatrix},$$

where $b_{vv} \geq 0, b_{iv} = 0 (i < v)$, and if $b_{vv} > 0$, then $0 \leq b_{iv} < b_{vv} (1 \leq v < i \leq n-1)$.

If $b_{n-1,n-1} > 0$, then there exists an integer q_{n-1} such that

$$0 \leq q_{n-1} b_{n-1,n-1} + b'_{n,n-1} < b_{n-1,n-1}.$$

Therefore

$$A \overset{\mathrm{L}}{=} \begin{pmatrix} b_{11} & 0 & \cdots & 0 & 0 \\ b_{21} & b_{22} & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ b_{n-1,1} & b_{n-1,2} & \cdots & b_{n-1,n-1} & 0 \\ b''_{n1} & b''_{n2} & \cdots & b''_{n,n-1} & b_{nn} \end{pmatrix},$$

where $b''_{ni} = q_{n-1} b_{n-1,i} + b'_{ni} (1 \leq i \leq n-1), 0 \leq b''_{n,n-1} < b_{n-1,n-1}$. The theorem follows from repeated applications of this. □

Definition 2. We call a square matrix of the form (7.28) the *normal form of Hermite.*

Exercise. Prove that the normal form of Hermite for a non-singular square matrix is unique.

§7.5. Invariant factors and elementary divisors

Definition 1. Let $A (= A^{(m,n)})$ and $B (= B^{(m,n)})$be two matrices. Suppose that there are two modular matrices $U (= U^{(m)})$, $V (= V^{(n)})$such that

$$A = UBV.$$

Then we say that A and B are *equivalent* and we write $A \sim B$.

Clearly equivalence has the three properties of being reflexive, symmetric and transitive.

Theorem 7.5.1. *Any matrix $A(= A^{(m,n)})$ must be equivalent to a matrix of the form*

$$\begin{pmatrix} d_1 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & d_1d_2 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & d_1d_2d_3 & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & d_1d_2\cdots d_m & 0 & \cdots & 0 \end{pmatrix} \qquad (m \le n) \quad (7.30)$$

or

$$\begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_1d_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & d_1d_2\cdots d_n \\ 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \qquad (m \ge n) \qquad (7.31)$$

where $d_i \ge 0$.

Proof. Let $A = (a_{11}, a_{12}, \cdots, a_{1k})$ be a 1 by k matrix where k is any positive integer $(k > 1)$. By Theorem 2.2 there is a positive modular U such that

$$AU = (d, 0, 0, \cdots, 0)$$

and so the required result is proved. Also, from

$$U' \begin{pmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{1k} \end{pmatrix} = \begin{pmatrix} d \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

where U' is the transposed matrix of U, we see that the theorem also holds for k by matrices.

We now proceed by induction on the number of rows of the matrix A. Let A is any given matrix. If $A = 0$, then the result is trivial. If $A \neq 0$, then we may assume that $a_{11} \neq 0$ and indeed we can even assume that $a_{11} > 0$. We first prove that must be equivalent to a matrix of the form:

$$A \sim A_1 = \begin{pmatrix} a'_{11} & a'_{12} & \cdots & a'_{1n} \\ a'_{21} & a'_{22} & \cdots & a'_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a'_{m1} & a'_{m2} & \cdots & a'_{mn} \end{pmatrix}, \quad a'_{11} \mid a'_{ii} \quad (1 \le i \le m, 1 \le j \le n).$$

This is clearly so if $a_{11} = 1$. When $a_{11} > 1$, if $a_{11} \nmid a_{i_0j_0}$ then we can move $a_{i_0j_0}$ to other of the positions occupied by a_{12}, a_{21}, a_{22}, by means of row or column interchanging. Therefore, using the method of proof for Theorem 1.5, we can change the leading element to a positive integer which is less than a_{11}, and an inductive argument completes the first part of proof.

Now from

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ -\frac{a'_{21}}{a'_{11}} & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ -\frac{a'_{m1}}{a'_{11}} & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} a'_{11} & a'_{12} & \cdots & a'_{1n} \\ a'_{21} & a'_{22} & \cdots & a'_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a'_{m1} & a'_{m2} & \cdots & a'_{mn} \end{pmatrix}$$

$$\times \quad \begin{pmatrix} 1 & -\frac{a'_{12}}{a'_{11}} & \cdots & -\frac{a'_{1n}}{a'_{11}} \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = \begin{pmatrix} a'_{11} & 0 & \cdots & 0 \\ 0 & a''_{22} & \cdots & a''_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & a''_{m2} & \cdots & a''_{mn} \end{pmatrix},$$

we have

$$A \sim \begin{pmatrix} a'_{11} & 0 & \cdots & 0 \\ 0 & a''_{22} & \cdots & a''_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & a''_{m2} & \cdots & a''_{mn} \end{pmatrix}.$$

Therefore, from the induction hypothesis, we have

$$A \sim \begin{pmatrix} a'_{11} & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & d'_2 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & d'_2 \cdots d'_m & 0 & \cdots & 0 \end{pmatrix} \qquad (m \leq n) \qquad (7.32)$$

or

$$A \sim \begin{pmatrix} a'_{11} & 0 & \cdots & 0 \\ 0 & d'_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & d'_2 \cdots d'_n \\ 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \qquad (m \geq n). \qquad (7.33)$$

Since $a'_{11} \mid a'_{ii}$, and d'_2 is a linear combination of the elements of A_1, it follows that $a'_{11} \mid d'_2$. If we let $a'_{11} = d_1, d'_2 = d_1 d_2, d'_3 = d_3, d'_4 = d_4, \cdots$, then the theorem follows from (7.32) and (7.33). □

Definition 2. We call matrices of the form (7.30) or (7.31) the *normal forms of Smith*.

In the proof of Theorem 5.1 the operations that we use are: the interchange of rows (or columns), the addition of an integer multiple of a row (or column) to another row (or column); the multiplication by -1 to a row (or column). We call these operations the *elementary operation* of matrices. We can therefore restate Theorem 5.1 as follows: any matrix can be reduced to the normal form of Smith by elementary operations.

After the interchange of two rows (or columns) or the multiplication by -1 to a row (or column), the i by i sub-determinants of the resulting matrix are either the same as the i by i sub-determinants of the original matrix, or differ by their signs only. Again if we add an integer multiple of a row (or column)to another row (or column) the i by i sub-determinants of the resulting matrix are either the same as the i by i sub-determinants of the original matrix, or the i by i sub-determinants with the addition of an integer multiple of i by i sub-determinants. It follows that the greatest common factor of all the i by i sub-determinants of a matrix is invariant under any elementary transformation. Therefore we have

Theorem 7.5.2 *Let $A \sim B$. Then the greatest common factor of the i by i sub-determinants of the two matrices A and B are the same.*

Meanwhile we see from (1) and (2) that

$$h_i = d_1 \cdot d_1 d_2 \cdots d_1 \cdots d_i,$$

are the greatest common factor of the i by i sub-determinants of A. Therefore we have

Theorem 7.5.3. *The normal form of Smith for a matrix is unique.*

Definition 3. Let the non-zero elements of the normal form of Smith in (7.30) and (7.31) for a matrix A be

$$d_1, d_1 d_2, \cdots, d_1, \cdots, d_k \qquad (k \leq min(m,n)).$$

We call these numbers the *invariant factors* of A of orders $1, 2, \cdots, k$ respectively. The number k is called the *rank* of the matrix A. Let

$$d_1 \cdots d_i = p_1^{e_{i1}} \cdots p_{l_i}^{e_{il_i}} (e_{ij} > 0, 1 \leq i \leq k, l_{i-1} \leq l_i)$$

be the standard prime factorization of an invariant factor. We call the prime power $p_j^{e_{ij}}$ an *elementary divisor* of the matrix A.

It is easy to see that the indices of the elementary divisors satisfy $e_{kj} \geq e_{k-1,j} \geq e_{k-2,j} \geq \cdots \qquad (1 \leq j \leq l)$. It also follows from the definition that is two matrices have the same invariant factors, then they have the same rank and the same elementary divisors. Conversely if the ranks are the same and the elementary divisors are same, then the invariant factors are the same. Therefore we have

Theorem 7.5.4. *A necessary and sufficient condition for two m by n matrices to be equivalent is that they should have the same ranks and the same elementary divisors.*

§7.6. Applications

Let us consider the solutions to the system of linear equations

$$y_i = \sum_{j=1}^{n} x_j a_{ji} \qquad (1 \leq i \leq m, n \geq m), \tag{7.34}$$

with integer coefficients, and given integers y_i-that is we consider the integer solutions to

$$y = xA, \qquad y = (y_1, \cdots, y_m), \qquad x = (x_1, \cdots, x_n),$$

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}. \tag{7.35}$$

We saw in the previous section that there two modular matrices $U(=U^{(n)})$ and $V(=V^{(m)})$ such that

$$UAV=\begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_1d_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & d_1\cdots d_m \\ 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}=D. \tag{7.36}$$

We now let $yV=y^*=(y'_1,\cdots,y'_m)$, $\qquad xU^{-1}=x^*=(x'_1,\cdots,x'_n)$, so that, from (7.35),

$$y^*=x^*D, \tag{7.37}$$

or

$$y'_i=d_1d_2\cdots d_ix'_i \qquad (1\le i\le m). \tag{7.38}$$

A necessary and sufficient condition for (7.34) to have a solution is that (7.38) has a solution. If $d_1\cdots d_k\neq 0, d_{k+1}=0$, then a necessary and sufficient condition for (7.38) to have a solution is that

$$d_1\cdots d_i\mid y'_i \quad (1\le i\le k),\quad y'_{k+1}=\cdots=y'_m=0. \tag{7.39}$$

From (7.39) we have

$$\begin{pmatrix} U & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} A \\ y \end{pmatrix}V=\begin{pmatrix} D \\ y^* \end{pmatrix}. \tag{7.40}$$

Now, if (7.39) holds, then we have, by (7.40), that

$$\begin{pmatrix} A \\ y \end{pmatrix}\sim\begin{pmatrix} D \\ 0 \end{pmatrix}; \tag{7.41}$$

conversely, if (8) holds, then $\begin{pmatrix} D \\ y^* \end{pmatrix}\sim\begin{pmatrix} D \\ 0 \end{pmatrix}$, and from Theorem 5.2 we have

$$d_1\mid y'_1, d_1d_2\mid y'_2,\cdots,d_1\cdots d_k\mid y'_k, \qquad y'_{k+1}=\cdots=y'_m=0,$$

which is formula (7.39). Therefore a necessary and sufficient condition for (7.34) to have a solution is that (7.41) holds; that is, we have

Theorem 7.6.1. *A necessary and sufficient condition for the system (7.34) to have a solution is that there are two matrices A and* $\begin{pmatrix} A \\ y \end{pmatrix}$ *with the same invariant factors.*

If (7.35) holds, then we have

$$x'_1=\frac{y'_1}{d_1},\quad x'_2=\frac{y'_2}{d_1d_2},\quad\cdots,\quad x'_k=\frac{y'_k}{d_1\cdots d_k}. \tag{7.42}$$

This means that $x'_1,x'_2,\cdots,x'_k$ are uniquely determined, and $x'_{k+1},\cdots,x'_n$ can be any integers. Thus, if $t_1,\cdots t_{n-k}$ are $n-k$ arbitrary integers, then

$$x_i=\sum_{j=1}^{k}x'_ju_{ji}+\sum_{l=1}^{n-k}t_lu_{k+l,i}=x_i^{(0)}+\sum_{l=1}^{n-k}t'_lu_{k+l,i} \qquad (1\le i\le n), \tag{7.43}$$

where $x_1^{(0)},\cdots,x_n^{(0)}$ is set of solution to (7.34) when $t_1=t_2=\cdots=t_{n-k}=0$.

Exercises for Chapter 7