Basic Number Theory

基 础 数 论

胡 勇
huy@sustech.edu.cn

2021 年 8 月 24 日

Contents

Preface　前言

本讲义根据作者于 2018 年和 2019 年秋季学期在南方科技大学讲授《初等数论》课程时的讲稿整理而成. 2018 级的李昀升同学对讲义进行了仔细的阅读和校对, 多次指出错漏之处, 在此向他表示特别的感谢.

Notation and Conventions
关于记号和术语的约定

Throughout these notes, the following notation and conventions will be in force. 本讲义将一直采用如下约定的记号和术语.

(0.1) Set theory 关于集合

1. Given two subsets A and B, their *difference set* (差集) is denoted by $A \setminus B := \{x \in A \,|\, x \notin B\}$.
2. The set of *natural numbers* is denoted by $\mathbb{N} = \{0, 1, \dots\}$. 自然数集记为$\mathbb{N}$, 约定其中包含 0.
   非零自然数构成的集合记为$\mathbb{N}^* = \mathbb{N} \setminus \{0\} = \{1, 2, \dots\}$.
3. The cardinality of a set S is denoted by $\#S$, $|S|$ or $\mathrm{Card}(S)$.
4. $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$; and similarly, $\mathbb{Q}^* = \mathbb{Q} \setminus \{0\}$; $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$.
5. $\mathbb{R}_+^* = \mathbb{R}_{>0} := \{x \in \mathbb{R} \,|\, x > 0\}$.
   Similarly, $\mathbb{Q}_+^* = \mathbb{Q}_{>0} := \{x \in \mathbb{Q} \,|\, x > 0\}$.
6. If $a \leq b$ are integers, $[\![a, b]\!] := [a, b] \cap \mathbb{Z} = \{k \in \mathbb{Z} \,|\, a \leq k \leq b\}$.

(0.2) Matrices 关于矩阵

$\mathbf{M}_n(R)$: set of $n \times n$ matrices with entries in a ring R.
$\mathbf{GL}_n$: general linear group.
$\mathbf{SL}_n$: special linear group.

(0.3) Others 其他

$A := B$ means A is defined to be B.
$A =: B$ means A will be denoted by B.

更多内容以后适时补充.

Chapter 1

Divisibility of Integers

1.1 The set of integers and mathematical induction

1.1.1 The well ordering property and the integral part

(1.1) Notation.

- $\mathbb{N}$ = the set of natural numbers = $\{0, 1, 2, \dots\}$.
- $\mathbb{N}^* = \mathbb{N} \setminus \{0\} = \{1, 2, 3, \dots\}$.
- $\mathbb{Z}$ = the set of integers = $\{0, \pm 1, \pm 2, \dots\}$.
- $\mathbb{Q}$ = the set of rational numbers = $\{p/q \,|\, p, q \in \mathbb{Z}, q \neq 0\}$.
- $\mathbb{R}$ = the set of real numbers.
- $\mathbb{C}$ = the set of complex numbers.
- $[\![a, b]\!] = \{x \in \mathbb{Z} \,|\, a \leq x \leq b\}$. ■

The following statement may look obvious :

(1.2) The well ordering principle. 良序原理 *Every nonempty subset of* $\mathbb{N}^*$ *has a smallest element. That is, if* $A \subseteq \mathbb{N}^*$ *and* $A \neq \emptyset$*, then there exists* $a \in A$ *such that* $a \leq x$*, for all* $x \in A$*.*

For a nonempty subset S of $\mathbb{R}$, we say that S has the ***well ordering property*** 良序性质 if every nonempty subset of S has a smallest element.

The well ordering property of $\mathbb{N}^*$ should be thought of as one of the axioms defining the set of positive integers, or it may be derived from a set of axioms in which it is not induced.

For a nice account of axiomatic constructions of the number system (from integers to real numbers), the interested readers may consult Terence Tao's book [Tao]. ■

Definition 1.3. A ***rationed number*** (有理数) is a real number that can be written in the form p/q, where $p, q \in \mathbb{Z}$ and $q \neq 0$. An ***irrational number*** (无理数) is a real number that is not a rational number. ■

Example 1.4. $\sqrt{2}$ is irrational.

Proof. Assume the contrary , i.e., $\sqrt{2} = p/q$ for some integers p, q which we may assume positive. Then

$$S := \{k \cdot \sqrt{2} \mid k \text{ and } k \cdot \sqrt{2} \text{ are both positive integers}\}$$

is a nonempty subset of $\mathbb{N}^*$ (S is nonempty because our assumption implies $q\sqrt{2} = p \in S$). By the well ordering property, S has a least element, say $s = t\sqrt{2}$, with t, $s \in \mathbb{N}^*$.

Then

$$(s-t)\sqrt{2} = s\sqrt{2} - t\sqrt{2} = s\sqrt{2} - s = t\sqrt{2} \cdot \sqrt{2} - s = 2t - s \in \mathbb{Z}, \text{ because } s\,,\, t \in \mathbb{Z}\,.$$

Further, since $s - t = t\sqrt{2} - t = t\left(\sqrt{2} - 1\right)$ and $\sqrt{2} - 1 > 0$, $t > 0$, we have $s - t > 0$ and hence $(s-t)\sqrt{2} > 0$. It follows that $s - t \in \mathbb{N}^*$ and $(s-t)\sqrt{2} \in \mathbb{N}^*$, showing that $(s-t)\sqrt{2} \in S$.

But $(s-t)\sqrt{2} - s = (2t-s) - s = 2(t-s) < 0$. Hence $(s-t)\sqrt{2}$ is an element in S which is smaller than s. This contradicts the minimality of s in S. The result is thus proved by contradiction. □

(1.5) The principle of Archimedes (阿基米德原理) *For any fixed positive number ε and any real number x, there exists a unique integer $M\varepsilon \le x < (M+1)\varepsilon$.*

Just as the well ordering property for the integers, the principle of Archimedes for the reals is actually part of a collection of axioms that characterize the set $\mathbb{R}$ of all real numbers.

Applying the principle with $\varepsilon = 1$, we see that there is a unique integer M such that $M \le x < M+1$. This is the greatest integer less than or equal to x. It will be denoted by $[x]$ or $\lfloor x \rfloor$, and called the ***integral part*** (or ***integer part***) (整数部分) of x. The function $x \mapsto [x] = \lfloor x \rfloor$ is called the ***floor function*** (地板函数) or the ***greatest integer function*** (取整函数) . ■

1.1.2 Mathematical induction

The principle of Mathematical induction is a basic tool for proving results about integers.

(1.6) The principle of mathematical induction. (数学归纳法原理) A set S of positive integers must be equal to the whole set $\mathbb{N}^*$, if it has the following properties:

(a) $1 \in S$;

(b) $\forall n \in \mathbb{N}^*$, $n \in S$ implies $n + 1 \in S$.

Proof. Suppose $S \neq \mathbb{N}^*$, which means that the complement $A := \mathbb{N}^* \setminus S$ is a nonempty subset of $\mathbb{N}^*$. By the well ordering property, A has a least element m. By assumption (a), $m > 1$. Thus $n := m - 1 \in \mathbb{N}^*$. Since $n \in \mathbb{N}^*$ is strictly smaller than m and m is the least element of A, we have $n \notin A$. In others words, $n \in S$. By property(b), this implies that $m = n + 1 \in S$, a contradiction (since $m \in A = \mathbb{N}^* \setminus S$). □

Exercise 1.7. Prove the following: Let S be a subset of $\mathbb{Z}_{\ge a} := \{x \in \mathbb{Z} | x \ge a\}$, where a is a fixed integer. Then $S = \mathbb{Z}_{\ge a}$ if the following two conditions hold:

1. $a \in S$;
2. $\forall n \in \mathbb{Z}$, $n \in S$ implies $n + 1 \in S$. ■

The next result is a variant of the principle of induction stated in (1.6).

(1.8) The second principle of mathematical induction. (第二数学归纳法原理) A subset S of $\mathbb{N}^*$ must be equal to $\mathbb{N}^*$ itself if the following conditions hold:

(a) $1 \in S$;

(b) For every $n \in \mathbb{N}^*$, $[\![1, n]\!] \subseteq S$ implies $n+1 \in S$.

Proof. Let $A := \mathbb{N}^* \setminus S$. If $S \neq \mathbb{N}^*$, then A is nonempty, so by the well ordering property, A has a least element m. This implies that, $[\![1, m-1]\!] \subseteq S = \mathbb{N}^* \setminus A$, noticing that $n := m-1 \in \mathbb{N}^*$ by (a). Now property (b) implies that $m = n+1 \in S$, contradicting the fact $m \in A$. □

1.2 Divisibility and prime numbers

1.2.1 Divisibility of integers

Definition 1.9. Let a and b be integers. We say that b is ***divisible*** by a (b 被 a 整除) or a ***divides*** (整除) b, or a is a ***divisor*** (因子) of b, or b is a ***multiple*** (倍数) of a, if there exists $c \in \mathbb{Z}$ such that $b = ac$. We write $a \,|\, b$ to mean "a divides b".

Warning: Do not confuse the notations $a \,|\, b$ and a/b. ■

Proposition 1.10. *Let a, b, $c \in \mathbb{Z}$.*

(1) If $a \,|\, b$ and $b \,|\, c$, then $a \,|\, c$.

(2) If $c \,|\, a$ and $c \,|\, b$, then $c \,|\, (ma + nb)$ for all m, $n \in \mathbb{Z}$.

(3) If $a \,|\, b$ and $b \neq 0$, then $|a| \leq |b|$.

Proof. Exercise. □

The following simple fact can be very useful.

Theorem 1.11 (The division algorithm (带余除法))**.** *Suppose $a \in \mathbb{Z}$ and $b \in \mathbb{N}^*$. There are unique integers q and r such that*

$$a = b \cdot q + r \quad \text{with} \quad 0 \leq r < b\,.$$

In the above equation, we call q the ***quotient*** (商), r the ***remainder*** (余数), a the ***dividend*** (被除数) and b the ***divisor*** (除数) of the division.

In the above theorem we assumed $b > 0$. This is in fact not a serious restriction.

Corollary 1.12. *Let $a, c \in \mathbb{Z}$ with $c \neq 0$. Then there are unique integers q and r such that*

$$a = c \cdot q + r \text{ with } 0 \leq r < |c|$$

Proof. Apply Theorem 1.11 with $b = |c|$. □

By the uniqueness statement in 1.12, we have:

Corollary 1.13. *Let a, $c \in \mathbb{Z}$ with $c \neq 0$.*

Then $c \,|\, a$ if and only if the remainder r is 0 in the division algorithm $a = cq + r$.

Proof of Theorem 1.11. Set $q = \left\lfloor \frac{a}{b} \right\rfloor$ and $r := a - bq$. Then we have

$$q \leq \frac{a}{b} < q + 1$$

which is clearly equivalent to

$$0 \leq r = a - bq < b\,.$$

This proves the existence of the pair (q, r) with required properties.

To show that the values for the quotient q and the remainder r are unique, suppose that we have another pair $(q', r') \in \mathbb{Z} \times \mathbb{Z}$ such that

$$a = b \cdot q' + r' \quad \text{and} \quad 0 \le r' < b\,.$$

Then

$$\begin{aligned} 0 = a - a = (bq + r) - (bq' + r') &= (q - q') \cdot b + (r - r') \\ \textit{i.e.,} \qquad (r - r') &= (q' - q)b. \end{aligned}$$

This implies $b \,|\, (r - r')$. If $r - r' \neq 0$, then by Prop. 1.10 (3), we must have $b < |r - r'|$. But this is impossible, since the assumptions

$$0 \le r < b \quad \text{and} \quad 0 \le r' < b$$

imply $|r - r'| < b$. (The distance of any two points in the interval $[0\,,\, b)$ is strictly smaller than b.)

The above discussions show that $r - r' = 0$, i.e., $r = r'$. That $q = q'$ follows from the equation $(r - r') = (q' - q)b$ since $b \neq 0$. □

1.2.2 Representations of integers in different bases

In daily life, we use decimal notation to represent integers. For instance, when we write out the integer 37465, we mean

$$3 \cdot 10^4 + 7 \cdot 10^3 + 4 \cdot 10^2 + 6 \cdot 10^1 + 5 \cdot 10^0\,.$$

We say that 10 is the ***base*** (基) of this numbering system.

With the development of computer sciences, bases other than 10 (especially base 2, base 8 and 16) have been more and more extensively used.

Let us give a proof that every integer $b > 1$ can be used as a base.

Theorem 1.14. *Let b be an integer with $b > 1$. Then every positive integer n can be written uniquely in the form*

$$n = a_k b^k + a_{k-1} b^{k-1} + \cdots + a_1 b + a_0$$

where $k \in \mathbb{N}$, each $a_j \in [\![0, b-1]\!]$ for every $j = 1, 2, 3, \cdots, k$ and $a_k \neq 0$.

The above expression is referred to the **base b expansion** *(以 b 为基的展开式) of the integer n, or the* **representation in base b** *(以 b 为基的表示) of n. The numbers a_j are called the* **digits** *(数字) of the expansion.*

To distinguish representations of integers with different bases, we will often write $n = (a_k a_{k-1} \cdots a_0)_b$ to mean the integer n has a base b expansion as above.

Proof. Uniqueness: Suppose that we have two such expressions. That is,

$$\begin{aligned} n &= a_k b^k + a_{k-1} b^{k-1} + \cdots + a_1 b + a_0 \\ &= c_k b^k + c_{k-1} b^{k-1} + \cdots + c_1 b + c_0 \end{aligned} \tag{1.14.1}$$

with $a_j, c_j \in [\![0\,,\, b-1]\!]$. Here we may add initial terms with zero coefficients to one of the two expansions to have the number of terms become the same. (So only one of a_k and c_k is required to be nonzero, the

other can be 0 or not.)

If we consider the division algorithm of n by b, then from the two expansions in (1.14.1) we see that both a_0 and c_0 are the remainder of division. So, by the uniqueness of the remainder, we have $a_0 = c_0$. Then we get

$$\begin{aligned} n_1 := \frac{n - a_0}{b} &= a_k b^{k-1} + a_{k-1} b^{k-2} + \cdots + a_2 b + a_1 \\ &= \frac{n - c_0}{b} = c_k b^{k-1} + c_{k-1} b^{k-2} + \cdots + c_2 b + c_1 \,. \end{aligned}$$

This shows that a_1 and c_1 are both the remainder of n_1 divided by b. Hence $a_1 = c_1$.

Continuing in the way, we get $a_j = c_j$ for every $j = 0, 1, \cdots, k$. This proves the uniqueness of the expansion.

To prove the existence, we may use the second principle of mathematical induction. If $n < b$ (e.g. $n = 1$), then $n = a_0$ with $k = 0$ and $a_0 \in [\![1\,,\, b-1]\!]$ is a base b representation of n.

Now suppose that the integers $1, 2, \cdots, n$ all have a base b expansion. We want to prove that $n+1$ also has a base b expansion. We divide $n + 1$ by b to obtain the equation

$$n + 1 = qb + r \quad \text{with} \quad r \in [\![0\,,\, b-1]\!] \,.$$

If $q = 0$, then taking $k = 0$ and $a_0 = n + 1$ we are done. If $q \neq 0$, then the above equation implies that $q \in [\![1\,,\, n]\!]$. So, by the induction hypothesis, q has a representation in base b:

$$q = c_s b^s + c_{s-1} b^{s-1} +, \cdots + c_1 b + c_0 \,.$$

Thus,

$$n + 1 = qb + r = c_s b^{s+1} + c_{s-1} b^s +, \cdots + c_1 b^2 + c_0 b + r$$

Taking $k = s + 1$, $a_0 = r$ and $a_j = c_{j-1}$ for all $j \in [\![1, k]\!]$, we get a base b representation of $n + 1$. This finishes the proof. □

We call base 10 notation, our conventional way of writing integers, ***decimal*** (十进制) notation. Base 2 expansions are called ***binary*** (二进制) expansions, base 8 expansions are called ***octal*** (八进制) expansions, and base 16 expansions are called ***hexadecimal*** (十六进制) or ***hex*** for short.

(1.15) The proof of Theorem 1.14 provides a method of finding the base b expansion $(a_k a_{k-1} \cdots a_1 a_0)_b$ of a positive integer n:

First divide n by b. The remainder is the digit a_0. The quotient is $q_0 := \left[\frac{n}{b}\right]$.

In the second step, divide q_0 by b, obtaining the remainder a_1 and the quotient $q_1 := \left[\frac{q_0}{b}\right]$.

We continue this process, successively dividing the quotient by the base b, to obtain the digits $a_0, a_1, \cdots$, etc. The algorithm terminates as soon as a quotient of 0 is obtained.

To illustrate the procedure, let us consider the base 2 expansion of the integer 1864.

We do the division by 2 repeatedly, replacing the dividend each time with the quotient, and we

stop when we come to a quotient which equals 0:

$$\begin{aligned}
1864 &= 932 \cdot 2 + 0 \\
932 &= 466 \cdot 2 + 0 \\
466 &= 233 \cdot 2 + 0 \\
233 &= 116 \cdot 2 + 1 \\
116 &= 58 \cdot 2 + 0 \\
\\
58 &= 29 \cdot 2 + 0 \\
29 &= 14 \cdot 2 + 1 \\
14 &= 7 \cdot 2 + 0 \\
7 &= 3 \cdot 2 + 1 \\
3 &= 1 \cdot 2 + 1 \\
1 &= 0 \cdot 2 + 1
\end{aligned}$$

To obtain the base 2 expansion of 1864, with the above computational results we can simply take the remainders of there divisions in reverse order. We obtain $(1864)_{10} = (11101001000)_2$. ■

The reader should have been familiar with basic arithmetic operations (addition, subtraction, multiplication and division) with integers for long, but perhaps only in base 10. Our goal now is to show that performing these operations in other bases can actually work with essentially the same method.

(1.16) Here is an example of how we do addition.

In base 10:

	1	2	3	4
+	5	6	7	8
		(+1)	*(+1)*	
	6	9	1	2

i.e., $(1234)_{10} + (5678)_{10} = (6912)_{10}$

When carries happen, denoted by (+1) in italics, note that 1 means 1×10 because 10 is the base

In base 4:

	1	2	1	3
+	3	2	1	2
(+1)	*(+1)*		*(+1)*	
1	1	0	3	1

i.e., $(1213)_4 + (3212)_4 = (11031)_4$. ■

(1.17) Subtraction.

In base 10:

	(−1)	*(−1)*	*(−1)*	
	5	1	3	4
−	2	6	7	8
	2	4	5	6

i.e., $(5134)_{10} - (2678)_{10} = (2456)_{10}$

When a borrow happen, indicated by (-1) in italics, 1 means 1×10 here!

In base 2:

		(−1)			
	1	1	0	1	1
−	1	0	1	1	0
			1	0	1

i.e., $(11011)_2 - (10110)_2 = (101)_2$

When a borrow happen, indicated by (-1) in italics, 1 means 1×2 here! ■

(1.18) Multiplication.

First look at a multiplication in base 10:

			2	3	4
		×	5	6	7
			(+2)	*(+2)*	
		1	6	3	8
		(+2)	*(+2)*		
	1	4	0	4	
	(+2)	*(+2)*			
1	1	7	0		
1	3	2	6	7	8

This is calculating 234×7

This is calculating 234×6 but with a ***shifting*** to the left by 1 place

From the above example,we see that there are only 2 principle to memorize:

- Use the "shifting" principle, computing each time a multiplication by a 1-digit number.
- Express the product obtained each time in the given base b and pay attention to "carries".

Now in base 8:

			2	3	4
		×	5	6	7
			(+3)	*(+3)*	
		2	1	0	4
		(+2)	*(+3)*		
	1	6	5	0	
	(+2)	*(+2)*			
1	4	1	4		
	(+1)	*(+1)*			
1	6	2	2	0	4

because $28 = (34)_8$, $3 \times 7 + 3 = (30)_8$, etc.

i.e., $(234)_8 \times (567)_8 = (162204)_8$. ■

(1.19) Division.

In base 10:

$$
\begin{array}{rcccccccccc}
 & & & & 8 & 0 & 7 & 2 & \leftarrow & \leftarrow & \leftarrow \\
\cline{5-8}
231 &) & 1 & 8 & 6 & 4 & 7 & 5 & 5 & & \\
 & - & 1 & 8 & 4 & 8 & & & & & \\
\cline{3-6}
 & & & & 1 & 6 & 7 & 5 & & & \\
 & & & - & 1 & 6 & 1 & 7 & & & \\
 & & & & & & 5 & 8 & 5 & & \\
 & & & & & - & 4 & 6 & 2 & & \\
\cline{6-8}
 & & & & & & & 1 & 2 & 3 &
\end{array}
$$

Here the first digit 8 is the largest $q \in [\![1, 9]\!]$ such that $(231)_{10} \times (q)_{10} < (1864)_{10}$

i.e., $(1864755)_{10} = (231)_{10} \times (8072)_{10} + (123)_{10}$, the quotient being $(8072)_{10}$ and the remainder being $(123)_{10}$.

In base 9:

$$
\begin{array}{rcccccccccc}
 & & & & 7 & 5 & 0 & 1 & \leftarrow & \leftarrow & \leftarrow \\
\cline{5-8}
231 &) & 1 & 8 & 6 & 4 & 7 & 5 & 5 & & \\
 & - & 1 & 7 & 3 & 7 & & & & & \\
\cline{3-6}
 & & & 1 & 2 & 6 & 7 & & & & \\
 & & - & 1 & 2 & 6 & 5 & & & & \\
\cline{4-7}
 & & & & & & 2 & 5 & 5 & & \\
 & & & & & - & 2 & 3 & 1 & & \\
\cline{6-8}
 & & & & & & & 2 & 4 & &
\end{array}
$$

Here the first digit 7 is the largest $q \in [\![1, 8]\!]$ such that $(231)_9 \times (q)_9 < (1864)_9$

i.e., $(1864755)_9 = (231)_9 \times (7501)_9 + (24)_9$, the quotient being $(7501)_9$ and the remainder being $(24)_9$.

To summarize, to perform a division the key is to know how to determine the quotient when it is a 1-digit number: If n is the dividend and d is the divisor, then the quotient (when it has only 1 digit in the given base b) is the largest $q \in [\![1, b-1]\!]$ such that $d \cdot q < n$. In practice, when the numbers n and d are expressed in base b, one has to be very careful with his calculation in order not to be confused with the base 10 computation! ■

1.2.3 Greatest common divisor and Euclidean algorithm

(1.20) Let $a \in \mathbb{Z}$. Recall (cf (1.9)) that a *divisor* of a is an integer $d \in \mathbb{Z}$ such that there is another integer c such that $a = cd$. If $a \neq 0$, then the absolute value $|d|$ of any divisor of a is bounded by $|a|$.

If a and b are two integers, an integer d is called a ***common divisor*** (公因子) of a and b iff it is both a divisor of a and a divisor of b. When a and b are not both 0, they have a largest common divisor (why?), which we call the ***greatest common divisor*** (最大公因子) of a and b, and we denote it by $\gcd(a, b)$. ■

Definition 1.21. Suppose that a , $b \in \mathbb{Z}$ are not both 0. We say that a and b are ***relatively prime*** or ***coprime*** (互素) if $\gcd(a, b) = 1$. ■

Proposition 1.22. *Let a and b be integers that are not both 0.*

(1) $\gcd(a, b) = \gcd(\pm a, \pm b) > 0$.

(2) If $d = \gcd(a, b)$, then $\frac{a}{d}$ and $\frac{b}{d}$ are relatively prime.

Proof. This is an easy exercise. □

The following is a simple but very useful lemma.

Lemma 1.23. *Let $a\,, b \in \mathbb{Z}$ with $b \neq 0$. Suppose that $a = k \cdot b + r$ for some $k\,, r \in \mathbb{Z}$.*

Then the set of common divisors of a and b is the same as the set of common divisors of b and r. In particular, $\gcd(a\,, b) = \gcd(b\,, r)$.

Proof. An easy exercise left to the reader. □

Using the above lemma, we can prove a famous algorithm which computes a greatest common divisor in a highly efficient manner.

(1.24) (**Euclidean algorithm** (欧几里得算法), or "辗转相除法" in Chinese litterature.) Let $a\,, b \in \mathbb{Z}$ with $b \neq 0$. By iterating the division algorithm, we obtain a sequence of quotients $q_1, q_2, \cdots$ and remainders $r_1, r_2, \cdots$ such that

$$\text{(1.24.1)} \qquad \begin{cases} a = q_1 \cdot b + r_1 \\ b = q_2 \cdot r_1 + r_2, \\ r_1 = q_3 \cdot r_2 + r_3, \\ \cdots\cdots\cdots \\ \cdots\cdots\cdots \\ r_{n-2} = q_n \cdot r_{n-1} + \boxed{r_n} \\ r_{n-1} = q_{n+1} \cdot r_n + 0 \end{cases}$$

where r_n is the last nonzero remainder. Such an r_n exists since

$$|b| > |r_1| > |r_2| > \cdots$$

is a strictly decreasing sequence of natural numbers if the remainders are nonzero, and such a sequence cannot continue indefinitely. Exhibited in another way, the algorithm reads as follows:

quotients	a
	b
q_1	r_1
q_2	r_2
$\vdots$	$\vdots$
$\vdots$	r_{n-2}
q_{n-1}	r_{n-1}
q_n	$\boxed{r_n}$
q_{n+1}	0

Here r_n is the last nonzero remainder, and the sample blocks of the above table look like

	dividend
	divisor
quotient	remainder

Apply Lemma 1.23 repeatedly, we find easily that

$$\gcd(a,b) = \gcd(b,r_1) = \gcd(r_1,r_2) = \cdots = \gcd(r_n,0) = r_n\,.$$

In other words, $\gcd(a,b) = r_n$ = the last nonzero remainder in the above algorithm. ■

Example 1.25. The following table shows how we do the computation of $\gcd(47121188013, 47121192136)$.

	$a = 47121192136$
	$b = 47121188013$
1	4123
11428859	2356
1	1767
1	589
3	0

A most important application of the Euclidean algorithm is:

Theorem 1.26 (Bezout's identity)**.** *Let $a\,,\ b \in \mathbb{Z}$, not both 0. Let $d = \gcd(a\,,\ b)$.*
Then there exists $(u\,,\ v) \in \mathbb{Z} \times \mathbb{Z}$ such that

$$d = au + bv$$

*That is, the gcd of a and b is a $\mathbb{Z}$-linear combination of a and b. (By a $\mathbb{Z}$-***linear combination** *of a and b we mean any element of the set $\{ax + by \,|\, x\,,\ y \in \mathbb{Z}\}$.)*

Corollary 1.27. *Let $a\,,\ b \in \mathbb{Z}$, not both 0, and let $d = \gcd(a,b)$.*
Then

$$\{\mathbb{Z}\text{-linear combinations of } a \text{ and } b\} = \{\text{integral multiples of } d\} =: d\mathbb{Z}\,.$$

Therefore, d in the smallest positive integer that is a $\mathbb{Z}$-linear combination of a and b.

Proof. If z is a $\mathbb{Z}$-linear combination of a and b, then there exists $x\,,\ y \in \mathbb{Z}$ such that $z = ax + by$. Since $d = \gcd(a,b)$ is a common divisor of a and b, it is a divisor of $ax + by$. That is, z is a multiple of d.

Conversely, since d is a $\mathbb{Z}$-linear combination of a and b by Thm. 1.26, it follows easily that every multiple of d is a $\mathbb{Z}$-linear combination of a and b. □

Corollary 1.28. *Let $a\,,\ b \in \mathbb{Z}$, not both 0. Let $d \in \mathbb{N}^*$. Then $d = \gcd(a,b)$ if and only if the following two conditions hold:*

(1) d is a common divisor of a and b;

(2) every common divisor of a and b is a divisor of d.

Proof. If $d = \gcd(a\,,\,b)$, then clearly (1) holds. If c is a divisor of a and b, then by Thm. 1.26, c is a divisor of $d = \gcd(a, b)$.

Conversely, suppose that d is a positive integer satisfying (1) and (2). We compare it with $D := \gcd(a, b)$. By (2), we have $D \,|\, d$ since D is a common divisor of a and b. In particular,$D \leq d$. On the other hand, (1) tells us that d is a common divisor of a and b, so d is less than or equal to the largest common divisor of a and b. That is, $d \leq D$. So we get $d = D := \gcd(a\,,\,b)$, as desired. □

Proof of Thm. 1.26. With out loss of generality, we may assume $b \neq 0$. Using the Euclidean algorithm we obtain (cf. (1.24.1))

$$\begin{cases} a = q_1 \cdot b + r_1 \\ b = q_2 \cdot r_1 + r_2, \\ r_1 = q_3 \cdot r_2 + r_3, \\ \cdots\cdots\cdots \\ \cdots\cdots\cdots \\ r_{n-2} = q_n \cdot r_{n-1} + r_n \\ r_{n-1} = q_{n+1} \cdot r_n + 0 \end{cases}$$

with $r_n = d = \gcd(a, b)$. By an easy induction, we can deduce that for every $k \in [\![1, n]\!]$, r_k is a $\mathbb{Z}$-linear combination of a and b. In particular, $r_n = d = \gcd(a\,,\,b)$ is one. □

(1.29) Our proof of Theorem 1.26 also suggests a method of computing integers $u, v \in \mathbb{Z}$ satisfying the Bezout identify. Indeed, if we express a and b as $\mathbb{Z}$-linear combination of themselves, we have

$$\begin{cases} a = 1 \cdot a + 0 \cdot b \\ b = 0 \cdot a + 1 \cdot b \end{cases}$$

Putting

$$\begin{aligned} u_{-1} &= 1 \qquad & v_{-1} &= 0 \\ u_0 &= 0 \qquad & v_1 &= 1 \end{aligned}$$

We get

$$\begin{cases} a = u_{-1}a + v_{-1}b \\ b = u_0 a + v_0 b \end{cases}$$

This, substituted into the system (1.24.1), yields

$$\begin{aligned} r_1 &= (u_{-1}a + v_{-1}b) - q_1(u_0 a + v_0 b) \\ &= (u_{-1} - q_1 u_0)a + (v_{-1} - q_2 v_0)b \\ &= u_1 a + v_1 b \end{aligned}$$

Continuing this procedure, we obtain

$$\begin{aligned} r_2 &= b - q_2 r_1 \\ &= (u_0 a + v_0 b) - q_2(u_1 a + v_1 b) \\ &= (u_0 - q_2 v_1)a + (v_0 - q_2 v_1)b \\ &=: u_2 a + v_2 b \end{aligned}$$

and so on.
This amounts to saying that if we define recursively two sequences

$$\begin{cases} u_k = u_{k-2} - q_k u_{k-1} \\ v_k = v_{k-2} - q_k v_{k-1} \end{cases} \qquad \forall k \geq 1 \tag{1.29.1}$$

then the integers $u = u_n$ and $v = v_n$ satisfy the relation

$$d = r_n = u_n a + v_n b = au + bv$$

This method is called the ***extended Euclidean algorithm***, and can be illustrated by the following table: ■

	a	$u_{-1} = 1$	$v_{-1} = 0$
	b	$u_0 = 0$	$v_0 = 1$
q_1	r_1	$u_1 = u_{-1} - q_1 \cdot u_0$	$v_1 = v_{-1} - q_1 \cdot v_0$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
	r_{n-2}	u_{n-2}	v_{n-2}
q_{n-1}	r_{n-1}	u_{n-1}	v_{n-1}
q_n	r_n	$u_n = u_{n-2} - q_n \cdot u_{n-1}$	$v_n = v_{n-2} - q_n \cdot v_{n-1}$
q_{n+1}	0		

Example 1.30.

	$a = 47121192136$	1	0
	$b = 47121188013$	0	1
1	4123	1	-1
11428859	2356	-11428859	11428860
1	1767	11428860	-11428861
1	589	-22857719	22857721
3	0		

Thus

$$\begin{aligned} 589 &= \gcd(47121192136, 47121188013) \\ &= 47121192136 \times \underline{(-2857719)} + 47121188013 \times \underline{22857721} \\ &= au + bv \end{aligned}$$

■

1.2.4 Prime numbers and the fundamental theorem of arithmetic

Definition 1.31. A ***prime number*** (or simply a ***prime***) (素数) is an integer $p > 1$ that has no positive divisors other than 1 and p itself. If an integer $n > 1$ is not prime, it is called a ***composite number*** (合数). ■

Lemma 1.32. *Every positive integer greater than* 1 *has a divisor which is a prime, called a* **prime divisor** *(素因子).*

Proof. We prove the lemma by contradiction. Suppose the set

$$A := \{n \in \mathbb{N} \,|\, n > 1 \text{ and } n \text{ has no prime divisor}\}$$

is nonempty. Then, by the well ordering property, A has a least element, say $n \in A$. Since n is a divisor of itself and by assumption n has no prime divisor, n is not a prime. Therefore we can write $n = ab$, with $1 < a < n$, $1 < b < n$, and $a\,, b \in \mathbb{Z}$. Since $a < n$ and n is minimal in A, we have $a \notin A$. This means that a has a prime divisor. But that prime divisor is also a divisor of n, thus contradicting the hypothesis $n \in A$. The lemma is thus proved. □

Theorem 1.33 (Euclid)**.** *There are infinitely many primes.*

Proof. Suppose that there are only finitely many primes, $p_1, p_2, \cdots, p_n$. Consider the integer

$$Q = p_1 p_2 \cdots p_n + 1\,.$$

By Lemma 1.32, Q has a prime divisor p. This prime must be one of the p_i's. Hence $p \,|\, p_1 p_2 \cdots p_n$. Then it follows that p divides $Q - p_1 p_2 \cdots p_n = 1$, which is absurd because $p > 1$. □

Now we prove two lemmas that will enable us to prove the fundamental theorem of arithmetic, and that are of interest on their own.

Lemma 1.34. *Let $a\,, b \in \mathbb{Z}$, not both 0. Assume that a and b are relatively prime, i.e., $\gcd(a, b) = 1$. Then for any $c \in \mathbb{Z}$, one has*

$$a \,|\, c \iff a \,|\, bc\,.$$

Proof. Let us prove $a \,|\, bc \Rightarrow a \,|\, c$, the other implication being trivial.

By Bezout's identity, there exist integers $u, v \in \mathbb{Z}$ such that

$$1 = \gcd(a, b) = au + bv\,.$$

Multiplying both sides by c, we obtain

$$c = a(cu) + (bc)v\,.$$

Then the assumption $a|bc$ clearly shows that the right hand side is divisible by a. Hence $a \,|\, c$. □

Lemma 1.35. *Let p be a prime and $a_1, \cdots a_n \in \mathbb{Z}\,(n \geq 1)$.*

Then

$$p \,|\, a_1 a_2 \cdots a_n \iff p \,|\, a_i \text{ for some } i \in [\![1, n]\!]\,.$$

Proof. Let us assume p divides the product $a_1 a_2 \cdots a_n$ and show that $p \,|\, a_i$ for some i. By induction, we may reduce to the case $n = 2$. Now by assumption $p \,|\, a_1 a_2$.

Notation: For any $a\,, b \in \mathbb{Z}$, we write $a \nmid b$ to mean a does not divide b.

If $p \nmid a_1$, then p is not a common divisor of a_1 and p. Hence $d := \gcd(p, a_1)$ is not p. But $d \,|\, p$. Since p is prime, the only possibility is that $d = 1$. This means that $\gcd(p\,, a_1) = 1$. Thus, the assumption $p \,|\, a_1 a_2$ implies $p \,|\, a_2$, by Lemma 1.34. □

Theorem 1.36 (Fundamental theorem of arithmetic 算术基本定理)**.** *Every positive integer greater than 1 can be written uniquely as a product of primes, with the prime factors in the product written in nondecreasing order.*

That is, for every $n > 1$, $n \in \mathbb{N}^*$, one can find prime numbers $p_1, p_2, \cdots, p_m$ such that

$$n = p_1 \cdot p_2 \cdots p_m \quad \text{and} \quad p_1 \leq p_2 \leq \cdots \leq p_m \tag{1.36.1}$$

and such an expression is unique.

By combining all the factors of same value into a power, we may rewrite (1.36.1) in the form

$$n = p_1^{e_1} p_2^{e_2} \cdots p_s^{e_s}, \quad \text{where} \quad e_i \in \mathbb{N}^*, \ p_i \text{ are primes and } p_1 < p_2 < \cdots < p_s\,. \tag{1.36.2}$$

We say that (1.36.1) or (1.36.2) is the ***factorization into prime factors*** (or simply, ***prime factorization***) of the integer $n > 1$.

Proof. First prove the uniqueness. Suppose that we have two factorizations

$$n = p_1 p_2 \cdots p_m \quad \text{and} \qquad n = q_1 q_2 \cdots q_r$$

where p_i and q_i are all primes and

$$p_1 \leq p_2 \leq \cdots \leq p_m\,, \quad q_1 \leq q_2 \leq \cdots \leq q_r\,.$$

Then it is clear that $p_1 \,|\, q_1 q_2 \cdots q_r$. Hence, by Lemma 1.35, $p_1 \,|\, q_j$ for some $j \in [\![1, r]\!]$. Since q_j is a prime and $p_1 > 1$, we must have $p_1 = q_j$. Removing this common factor in the above two factorizations of n, we obtain

$$p_2 p_3 \cdots p_m = q_1 q_2 \cdots q_{j-1} q_{j+1} \cdots q_r\,.$$

Arguing in the same way as above, we can show that the numbers m and r must be the same, and that the primes p_i and q_j must agree in pairs. This proves the uniqueness of the expression (1.36.1).

To prove the existence, we use the second principle of mathematical induction. In the case $n = 2$, we may choose $m = 1$ and $p_i = 2$ to obtain (1.36.1)

Now assume $n > 2$ and all the integers $2, 3, \ldots, n-1$ have a factorization into prime factors. If n is a prime ,then taking $m = 1$ and $p_1 = n$ finishes the proof.

If n is composite, then $n = ab$ for some $a\,,\ b \in [\![2, n-1]\!]$. By the induction hypothesis, both a and b have a factorization into prime factors. Therefore, the product $n = ab$ also has such a factorization. □

(1.37) Theoretically, prime factorizations can be used to determine the gcd of any pair of positive integers. Indeed, given $a\,,\ b \in \mathbb{N}^*$, let their prime factorizations be

$$\begin{cases} a = p_1^{e_1} p_2^{e_2} \cdots p_m^{e_m} \\ b = p_1^{f_1} p_2^{f_2} \cdots p_m^{f_m} \end{cases} \tag{1.37.1}$$

where the exponents e_i and f_i are integers ≥ 0, and $\{p_1\,, p_2\,, \cdots, p_m\}$ is the set of all prime divisors of ab. (In this way we may arrange that the primes occurring in the factorizations of a and b are the same, at the price that some exponents must be allowed to equal 0.)

Then we have

$$\gcd(a, b) = p_1^{\min(e_1, f_1)} \cdot p_2^{\min(e_2, f_2)} \cdots p_m^{\min(e_m, f_m)}. \tag{1.37.2}$$

The proof of this is based on the following fact:

For any d, $n \in \mathbb{N}^*$, d divides n if and only if for every prime number p, the number of times that p appears in the prime factorization of d in the form (1.36.1) is less than or equal to the number of times that p appears in the prime factorization of n. (You should give a proof of this fact using the fundamental theorem of arithmetic!)

More generally, for any finite number of integers $a_1, \cdots a_n \in \mathbb{Z}$ with $n \geq 1$, that are not all 0, we can define their ***greatest common divisor*** by

$$\gcd(a_1, \cdots a_n) := \text{largest common divisors of all the } a_i.$$

When the a_i's are all positive, one can find an analog of the formula (1.37.2) which determines the above gcd by means of the prime factorizations of the a_i's. We leave it to the reader to find out and prove this formula. ■

Exercise 1.38. Given a finite number of nonzero integers $a_1, \cdots, a_n \in \mathbb{Z}$, prove that the set

$$\{M \in \mathbb{N}^* \mid M \text{ is a multiple of } a_i \text{ for every } i \in [\![1, n]\!]\}$$

has a smallest element. We call it the ***least common multiple*** (lcm) (最小公倍数) of $a_1, \cdots, a_n$, and denote it by $\operatorname{lcm}(a_1, \cdots a_n)$ (or $[a_1, \cdots, a_n]$).

Then find a formula that determines $\operatorname{lcm}(a_1, \cdots, a_n)$ in terms of the prime factorizations of the a_i's, when each $a_i > 0$. ■

Let us give a few more applications of the fundamental theorem arithmetic.

Proposition 1.39. *For every prime p, the real number $\sqrt{p}$ is irrational.*

Proof. Assume the contrary. Then $\sqrt{p} = a/b$ for some positive integers a and b. Replacing a by $\frac{a}{\gcd(a,b)}$ and b by $\frac{b}{\gcd(a,b)}$ if necessary, we may assume that $\gcd(a, b) = 1$ (Prop. 1.22 (2)). Then the relation $\sqrt{p} = a/b$ yields

$$a^2 = p \cdot b^2$$

By the fundamental theorem of arithmetic, the number of times that p can occur in the prime factorization of a square must be even. Thus, for a^2, this number is even, while for $p \cdot b^2$ that number has to be odd. Therefore $a^2 = pb^2$ is impossible. □

One can also use the following more general result to prove the result of (1.39).

Theorem 1.40. *Let $x \in \mathbb{R}$ be a zero of a polynomial*

$$X^n + c_{n-1}X^{n-1} + \cdots + c_1X + c_0\,, \quad \text{where each } c_i \in \mathbb{Z}\,.$$

Then x is either an integer or irrational.

Proof. Suppose that x is rational. Then we may write $x = \frac{a}{b}$, with $a\,, b \in \mathbb{Z}$ relatively prime. By assumption,

$$0 = x^n + c_{n-1}x^{n-1} + \cdots + c_1x + c_0 = \left(\frac{a}{b}\right)^n + c_{n-1}\left(\frac{a}{b}\right)^{n-1} + \cdots + c_1\frac{a}{b} + c_0$$

which yields, after multiplication by b^n,

$$0 = a^n + c_{n-1}a^{n-1}b + \cdots + c_1ab^{n-1} + c_0b^n\,.$$

All the terms other than a^n on the right is a multiple of b, so we have $b \,|\, a^n$. Since we assumed $\gcd(a\,,\,b) = 1$, using Lemma 1.34 and induction on n we find that $b \,|\, a$. Hence $x = \frac{a}{b}$ is an integer. □

Example 1.41. Let $a \in \mathbb{N}^*$ be a positive integer that is not the m-th power of any integer, where m is an integer ≥ 2. Then $\sqrt[m]{a}$ is irrational, by Thm. 1.40. Consequently, such numbers as $\sqrt{2}$, $\sqrt[5]{3}$, $\sqrt[6]{10}$, etc are irrational. ■

We end up this section with the following lemma, which will be used later.

Lemma 1.42. *Let m and n be relatively prime positive integers, and let d be a positive divisors of mn. Then, there is a unique pair $(d_1, d_2) \in \mathbb{N} \times \mathbb{N}$ such that*

$$d_1 \,|\, m\,,\ d_2 \,|\, n \quad \textit{and} \quad d = d_1 \cdot d_2$$

In other words, the map

$$\begin{aligned} \{\textit{positive divisors of } m\} \times \{\textit{positive divisors of } n\} &\longrightarrow \{\textit{positive divisors of } mn\} \\ (d_1\,,\,d_2) &\longmapsto d_1 \cdot d_2 \end{aligned}$$

is bijective, provided that $\gcd(m, n) = 1$.

Proof. Let $m = {p_1}^{m_1} {p_2}^{m_2} \cdots {p_s}^{m_s}$ and $n = {q_1}^{n_1} {q_2}^{n_2} \cdots {q_t}^{n_t}$ be the prime factorizations of m and n, where m_i, $n_j \in \mathbb{N}^*$, p_i and q_j are primes.

Since m and n are relatively prime, the sets $\{p_1\,,\,\cdots\,,\,p_s\}$ and $\{q_1\,,\,\cdots\,,\,q_t\}$ are disjoint. Thus the prime factorization of mn is

$$mn = p_1^{m_1} p_2^{m_2} \cdots p_s^{m_s} q_1^{n_1} q_2^{n_2} \cdots q_t^{n_t}$$

Thus, if d is a divisor of mn, then

$$d = p_1^{e_1} p_2^{e_2} \cdots p_s^{e_s} q_1^{f_1} q_2^{f_2} \cdots q_t^{f_t}$$

with $e_i \in [\![0, m_i]\!]$ and $f_j \in [\![0, n_j]\!]$.

Thus, taking

$$d_1 = p_1^{e_1} p_2^{e_2} \cdots p_s^{e_s} \quad \text{and} \quad d_2 = q_1^{f_1} q_2^{f_2} \cdots q_t^{f_t}$$

we get the factorization $d = d_1 d_2$ in the desired form.

If there is another factorization $d = d_1' \cdot d_2'$, then on the one hand, we have $\gcd(d_1, d_2') = 1$ since $\gcd(d_1, d_2') \,|\, \gcd(m, n) = 1$ and by assumption $d_1 \,|\, m$ and $d_2' \,|\, n$. Thus, from the obvious relation $d_1 \,|\, d_1 d_2 = d_1' \cdot d_2'$ we conclude that $d_1 \,|\, d_1'$, by Lemma 1.34. Similarly, $d_2 \,|\, d_2'$. If $d_1 < d_1'$, then we would have

$$d = d_1 d_2 < d_1' d_2 \leq d_1' d_2' = d$$

a contradiction. So we must have $d_1 = d_1'$ and $d_2 = d_2'$. This completes the proof. □

1.3 p-adic valuation of integers

1.3.1 The p-adic valuation

(1.43) Let p be a prime number. For each $n \in \mathbb{Z}$, the ***p-adic valuation*** $v_p(n)$ of n in defined as follows:

$$\begin{aligned} &\text{If } n = 0\,, \quad \text{we put } v_p(0) = +\infty \\ &\text{If } n \neq 0\,, \quad \text{we define } v_p(n) := \text{the unique integer } r \in \mathbb{N} \text{ such that} \\ &\qquad\qquad p^r \text{ exactly divides } n\,, \text{ i.e., } p^r \mid n \text{ but } p^{r+1} \nmid n \\ &\qquad\qquad = \text{the largest integer } r \in \mathbb{N} \text{ such that } p^r \mid n\,. \end{aligned} \tag{1.43.1}$$

Thus, for every $r \in \mathbb{N}$, we have

$$v_p(n) \geq r \iff p^r \mid n\,. \tag{1.43.2}$$

Suppose $n \neq 0$. Then, by the fundamental theorem of arithmetic, $v_p(n)$ is nothing but the exponent of p in the prime factorization of $|n|$. This yields the following useful description of $v_p(n)$:

$$v_p(n) = r \iff n = p^r \cdot a \text{ for some } a \in \mathbb{Z} \text{ relatively prime to } p\,, \tag{1.43.3}$$

when $n \neq 0$ and $r \in \mathbb{N}$. ■

Proposition 1.44. *Let p be a prime number and $m\,, n \in \mathbb{Z}$.*

Then we have:

(0) $v_p(n) = +\infty \iff n = 0$.

(1) $v_p(n) = 0 \iff p \nmid n \iff \gcd(p, n) = 1$.

(2) $v_p(mn) = v_p(m) + v_p(n)$.

(Here we use the convention that $(+\infty) + a = a + (+\infty) = +\infty$, $\forall a \in \mathbb{Z} \cup \{+\infty\}$)

(3) $v_p(m + n) \geq \min(v_p(m), v_p(n))$.

(4) If $v_p(m) \neq v_p(n)$, then $v_p(m + n) = \min\{v_p(m), v_p(n)\}$.

Proof. (0) and (1): This is clear from the definition, by the fundamental theorem of arithmetic.

(2) Suppose $s = v_p(m)$ and $r = v_p(n)$. (Here we may assume $mn \neq 0$, for otherwise the relation holds trivially.) By (1.43.3), $m = p^s \cdot b$ and $n = p^r \cdot a$ for some $a, b \in \mathbb{Z}$ with $\gcd(a\,, p) = gcd(b\,, p) = 1$. Then by the fundamental theorem of arithmetic we have $\gcd(ab\,, p) = 1$. Since $mn = (p^s \cdot b)(p^r \cdot a) = p^{s+r}(ab)$, it follows from (1.43.3) that

$$v_p(mn) = s + r = v_p(m) + v_p(n)\,.$$

(3) and (4): We may assume $mn \neq 0$. Without loss of generality, we may assume $r = v_p(n) \leq s = v_p(m)$.

By (1.43.3),

$$m = p^s \cdot b \text{ and } n = p^r \cdot a\,, \quad \text{with } a\,, b \in \mathbb{Z} \text{ satisfying } p \nmid ab\,.$$

Thus,

$$m+n = p^r(a+p^{s-r}\cdot b)$$

In view of (1.43.2), this implies $v_p(m+n) \geq r = \min(v_p(m), v_p(n))$. If moreover $r \neq s$, i.e., $r < s$, then

$$\gcd(a+p^{s-r}b\,,\,p) = \gcd(a\,,\,p) = 1$$

since $p^{s-r}b$ is a multiple of p in this case. Therefore, (1.43.3) shows that

$$v_p(m+n) = r = \min(v_p(m), v_p(n))$$

when $r = v_p(n) < s = v_p(m)$. □

According to the easy lemma below, p-adic valuations are a very useful tool to prove divisibility relations between integers.

Lemma 1.45. *Let $m, n \in \mathbb{Z}$. Then*

$$m \,|\, n \iff \textit{for every prime } p\,,\ v_p(m) \leq v_p(n)\,.$$

Proof. This is an easy consequence of the fundamental theorem of arithmetic and the definition of p-adic valuations. In fact, we have

$$x = \prod_p p^{v_p(x)}\,,\ \forall\, x \in \mathbb{N}^* \tag{1.45.1}$$

where p runs over the set of all prime numbers and, by Prop. 1.44 (1), $v_p(x) = 0$ except for a finite number of primes p (which depend on the integer x). □

1.3.2 Factorization of $n!$

Now we prove a very nice formula that determines the p-adic valuation of the factorial $n!$ of a positive integer n.

Theorem 1.46. *Let $n \in \mathbb{N}^*$ and let p be a prime number. Then*

$$v_p(n!) = \sum_{j=1}^{+\infty} \left[\frac{n}{p^j}\right] \tag{1.46.1}$$

In particular,

$$\frac{n-p}{p-1} - \log_p(n) < v_p(n!) < \frac{n}{p-1} \tag{1.46.2}$$

Proof. Let $k = [\log_p(n)]$, so that $p^k \leq n < p^{k+1}$. For every $a \in [\![1, n]\!]$, we have $v_p(a) \leq k$. Thus, if we define

$$I_r := \{a \in [\![1, n]\!] \,|v_p(a) = r\}\,,\ \forall\, r \in [\![1, k]\!]$$

then by Prop. 1.44 (2),

$$v_p(n!) = \sum_{a=1}^{n} v_p(a) = \sum_{r=1}^{k} r \cdot |I_r|$$

To compute $|I_r|$, notice that

$$\begin{aligned} I_r =& \{a \in [\![1,n]\!] \,|v_p(a) \geq r\} \setminus \{a \in [\![1,n]\!] \,|v_p(a) \geq r+1\} \\ \overset{(1.43.2)}{=}& \{a \in [\![1,n]\!] \ : \ p^r \,|\, a\} \setminus \{a \in [\![1,n]\!] \ : \ p^{r+1} \,|\, a\}\,. \end{aligned}$$

Hence,

$$\begin{aligned} |I_r| =& \text{ number of multiples of } p^r \text{ in } [\![1,n]\!] \\ & - \text{ number of multiples of } p^{r+1} \text{ in } [\![1,n]\!] \\ =& \left[\frac{n}{p^r}\right] - \left[\frac{n}{p^{r+1}}\right] \end{aligned}$$

Thus,

$$\begin{aligned} v_p(n!) &= \sum_{r=1}^{k} r \cdot |I_r| = \sum_{r=1}^{k} r \cdot \left(\left[\frac{n}{p^r}\right] - \left[\frac{n}{p^{r+1}}\right] \right) \\ &= \sum_{r=1}^{k} \left(r \cdot \left[\frac{n}{p^r}\right] - (r+1)\left[\frac{n}{p^{r+1}}\right] + \left[\frac{n}{p^{r+1}}\right] \right) \\ &= \sum_{r=1}^{k} r \cdot \left[\frac{n}{p^r}\right] - \sum_{t=2}^{k+1} t \cdot \left[\frac{n}{p^t}\right] + \sum_{t=2}^{k+1} \left[\frac{n}{p^t}\right] \\ &= \left[\frac{n}{p}\right] - (k+1)\left[\frac{n}{p^{k+1}}\right] + \sum_{t=2}^{k} \left[\frac{n}{p^t}\right] + \left[\frac{n}{p^{k+1}}\right] \\ &= \sum_{t=1}^{k} \left[\frac{n}{p^t}\right] - k\left[\frac{n}{p^{k+1}}\right] \end{aligned}$$

For all $j \geq k+1 = [\log_p(n)] + 1$, since $0 < n < p^{k+1} \leq p^j$ we have $[\frac{n}{p^j}] = 0$. So we obtain

$$v_p(n!) = \sum_{j=1}^{k} \left[\frac{n}{p^j}\right] = \sum_{j=1}^{+\infty} \left[\frac{n}{p^j}\right]$$

proving (1.46.1). The formula (1.46.2) follows from (1.46.1), in view of the obvious inequalities

$$\sum_{j=1}^{+\infty} \left[\frac{n}{p^j}\right] = \sum_{j=1}^{k} \left[\frac{n}{p^j}\right] \leq \sum_{j=1}^{k} \frac{n}{p^j} < \sum_{j=1}^{+\infty} \frac{n}{p^j} = n \cdot \frac{\frac{1}{p}}{1-\frac{1}{p}} = \frac{n}{p-1}$$

and

$$\sum_{j=1}^{+\infty} \left[\frac{n}{p^j}\right] = \sum_{j=1}^{k} \left[\frac{n}{p^j}\right] \geq \sum_{j=1}^{k} \left(\frac{n}{p^j} - 1\right) = \frac{n}{p} \cdot \frac{1-(\frac{1}{p})^k}{1-\frac{1}{p}} - k = \frac{n-\frac{n}{p^k}}{p-1} - k > \frac{n-p}{p-1} - \log_p(n)$$

This finishes the proof. □

As an application of Thm. 1.46, we show by an example how to find the prime factorization of $n!$.

Example 1.47. Let us factorize 20! in to prime factors.

It suffices to compute $v_p(20!)$ for every prime not exceeding 20, i.e., for all $p \in \{2,3,5,7,11,13,17,19\}$.

It is clear that $v_p(20!) = 1$, for $p = 11, 13, 17, 19$. Using (1.46.1) we get

$$\begin{aligned} v_2(20!) &= \left[\frac{20}{2}\right] + \left[\frac{20}{4}\right] + \left[\frac{20}{8}\right] + \left[\frac{20}{16}\right] \\ &= 10 + 5 + 2 + 1 = 18 \\ v_3(20!) &= \left[\frac{20}{3}\right] + \left[\frac{20}{9}\right] + \left[\frac{20}{27}\right] = 6 + 2 = 8 \\ v_5(20!) &= \left[\frac{20}{5}\right] = 4 \\ v_7(20!) &= \left[\frac{20}{7}\right] = 2 \end{aligned}$$

So we have $20! = 2^{18} \cdot 3^8 \cdot 5^4 \cdot 7^2 \cdot 11 \cdot 13 \cdot 17 \cdot 19$. ■

Proposition 1.48. *Let $n_1, \cdots, n_s \in \mathbb{N}$ and $n = n_1 + n_2 + \cdots + n_s$.*
Then

$$\frac{n!}{n_1! \cdot n_2! \cdot \cdots \cdot n_s!} \in \mathbb{N}^* .$$

Proof. We use the following elementary property of the floor function:

$$[x + y] \geq [x] + [y] , \quad \forall\, x , y \in \mathbb{R} .$$

This implies that

$$\left[\frac{n}{p^j}\right] = \left[\frac{n_1 + \cdots + n_s}{p^j}\right] \geq \left[\frac{n_1}{p^j}\right] + \left[\frac{n_2}{p^j}\right] + \cdots + \left[\frac{n_s}{p^j}\right]$$

for every prime p and every integer j. Taking the sum over all $j \geq 1$, we obtain

$$v_p(n!) \geq v_p(n_1!) + v_p(n_2!) + \cdots + v_p(n_s!)$$

by (1.46.1). □

(1.49) In combinatorial mathematics, the integer $\frac{n!}{n_1! n_2! \cdots n_s!}$ is known as a multinomial coefficient. It is equal to each of the following numbers:

- the coefficient of the monomial $X_1^{n_1} \cdots X_s^{n_s}$ in the polynomial $(X_1 + X_2 + \cdots + X_s)^n$;
- the number of ways one can distribute n distinguishable balls into s distinguishable boxes, such that the i-th box receives n_i balls.

In the special case $s = 2$,

$$\frac{n!}{n_1! n_2!} = \frac{n!}{n_1!(n - n_1)!} = \begin{pmatrix} n \\ n_1 \end{pmatrix} = \begin{pmatrix} n \\ n_2 \end{pmatrix}$$

is our familiar notion of binomial coefficient. ■

(1.50) Here are some examples of corollaries of Prop. 1.48:

(1) $(10!)^{10}$ divides $100!$;

(2) for every $x \in \mathbb{Z}$ and $m \in \mathbb{N}^*$, $x(x+1)\cdots(x+m-1)$ is a multiple of $m!$.

In fact, if $x > 0$, then $\frac{x(x+1)\cdots(x+m-1)}{m!}$ is an integer by Prop. 1.48. If $x < 0$, then

$$\begin{aligned}\frac{x(x+1)\cdots(x+m-1)}{m!} &= \frac{(-1)^m(-x)(-x-1)\cdots(-x-m+1)}{m!} \\ &= (-1)^m \frac{(-x)!}{m!(-x-m)!} \in \mathbb{Z}\,.\end{aligned}$$

So we see that the product of any set of m consecutive integers is a multiple of $m!$. ■

1.4 Pythagorean triples

1.4.1 Finding solutions with elementary methods

Thanks to the famous Pythagorean theorem, the equation

$$x^2 + y^2 = z^2$$

in the unknowns x, y and z is one of the most interesting equations that have been studied since the earliest history of number theory.

(1.51) By a ***Pythagorean triple*** (毕达哥拉斯三元组) we mean a triple of positive integers satisfying the equation

$$x^2 + y^2 = z^2\,.$$

A Pythagorean triple (x, y, z) in called ***primitive*** (本原的) if $\gcd(x, y, z) = 1$.

Clearly, for any Pythagorean triple $(x\,,\,y\,,\,z)$, by setting $x' = \frac{x}{d}\,,\, y' = \frac{y}{d}\,,\, z' = \frac{z}{d}$ with $d = \gcd(x, y, z)$, we can obtain a primitive Pythagorean triple $(x'\,,\,y'\,,\,z')$. Consequently, all Pythagorean triples can be found by taking integral multiples of primitive pythagorean triples. ■

To determine all the primitive Pythagorean triples, we shall need several lemmas.

Lemma 1.52. *Let $(x\,,\,y\,,\,z)$ be a primitive Pythagorean triple. Then*

(1) $\gcd(x\,,\,y) = \gcd(y\,,\,z) = \gcd(x\,,\,z) = 1$.

(2) One and only one of the two integers x and y is even.

Proof. (1) It is sufficient to show that no prime number p divides two of the three integers x, y, z. If it were not the case, then p divides the square of the third, by the equation $x^2 + y^2 = z^2$. Since p is a prime, p divides the square n^2 of an integer n if and only if p divides n. Therefore, if p divides two of $x\,,\,y$ and z, then $p \mid \gcd(x, y, z)$. But this contradicts the assumption that the triple $(x\,,\,y\,,\,z)$ is primitive.

(2) x and y cannot be both even, because $\gcd(x, y) = 1$ by (1). If x and y are both odd, then the relation $x^2 + y^2 = z^2$ shows that z^2 is even, and hence z is even. So we can write

$$x = 2m + 1,\ y = 2n + 1 \ \text{ and } \ z = 2r$$

foe some $m, n, r \in \mathbb{N}$. Then the relation

$$x^2 + y^2 = z^2$$

yields

$$4m^2 + 4m + 4n^2 + 4n + 2 = 4r^2 \,.$$

This would imply $4 \,|\, 2$, which is absurd. □

Lemma 1.53. *Let $r\,,\ s \in \mathbb{N}^*$ be relatively prime. If rs is a perfect square (i.e., $rs = t^2$ for some $t \in \mathbb{Z}$), then both r and s are perfect squares.*

Proof. Note that a positive integer $x \in \mathbb{N}^*$ is a perfect square if and only if in its prime factorization

$$x = \prod_p p^{v_p(x)}$$

the exponents $v_p(x)$ are all even. The assumption $\gcd(r\,,\ s) = 1$ means that

$$\min(v_p(r), v_p(s)) = 0\,, \quad \text{for all primes } p \tag{1.53.1}$$

and it follows that

$$v_p(rs) = v_p(r) + v_p(s) = \max(v_p(r)\,,\ v_p(s)) \tag{1.53.2}$$

Since rs is a perfect square, (1.53.2) tells us that

$$\max(v_p(r), v_p(s)) \text{ is even for all primes } p\,.$$

Together with (1.53.1), this shows that both $v_p(r)$ and $v_p(s)$ are even, for every prime p. Hence, both r and s are perfect squares, as desired. □

Theorem 1.54. *Let $(x\,,\ y\,,\ z) \in \mathbb{N}^* \times \mathbb{N}^* \times \mathbb{N}^*$ be a triple of positive integers with y even.*

Then the triple $(x\,,\ y\,,\ z)$ is a primitive Pythagorean triple if and only if there are relatively prime positive integers $m\,,\ n$ such that the following conditions hold

(1) $m > n$, and exactly one of m and n is even;

(2) $x = m^2 - n^2$, $y = 2mn$ and $z = m^2 + n^2$.

Proof. First assume that $x\,,\ y$ and z satisfy conditions (1) and (2) for some relatively prime integers $m\,,\ n \in \mathbb{N}^*$. We need prove that $(x\,,\ y\,,\ z)$ is a primitive Pythagorean triple. That $(x\,,\ y\,,\ z)$ is a Pythagorean triple follows easily from (2). We must show that $\gcd(x\,,\ y\,,\ z) = 1$.

Suppose that $d = \gcd(x\,,\ y\,,\ z) > 1$. Then d has a prime divisor p. Now $p \,|\, x = m^2 - n^2$ implies that $p \neq 2$, since $m^2 - n^2$ is odd by (1). From $p \,|\, 2mn = y$ it follows that $p \,|\, m$ or $p \,|\, n$. Together with $p \,|\, x = m^2 - n^2$ this would force p to divide both m and n. This leads to a contradiction to the assumption $\gcd(m\,,\ n) = 1$. We have thus proved that $(x\,,\ y\,,\ z)$ is a primitive Pythagorean triple.

Now we assume that $(x\,,\ y\,,\ z)$ is primitive Pythagorean triple. We want to find relatively prime integers $m\,,\ n \in \mathbb{N}^*$ satisfying (1) and (2).

Since we assumed y even, by Lemma 1.52 (1), the integers x and z must be odd. Thus, the integers

$$r := \frac{z + x}{2} \quad \text{and} \quad s = \frac{z - x}{2}$$

have the property that

$$r \cdot s = \frac{z^2 - x^2}{4} = \frac{y^2}{4} = \left(\frac{y}{2}\right)^2$$

Hence rs is a perfect square. Moreover, we have $\gcd(r,\, s) = 1$ because $d := \gcd(r,\, s)$ divides both $r + s = z$ and $r - s = x$, and hence $d \mid \gcd(x\,,\, z) = 1$. Consequently, we can find $m\,,\, n \in \mathbb{N}^*$ such that $\frac{z+x}{2} = r = m^2$ and $\frac{z-x}{2} = s = n^2$. This clearly implies $z = m^2 + n^2$, $x = m^2 - n^2$ and $y = 2\sqrt{rs} = 2mn$, proving (2).

Since $r = \frac{z+x}{2} > s = \frac{z-x}{2}$, we have $m > n$. Any common divisor of m and n must divide $z = m^2 + n^2$ and $x = m^2 - n^2$, so we have $\gcd(m\,,\, n) = 1$, by the fact $\gcd(x\,,\, z) = 1$. Finally, if m and n are both odd or both even, then $z = m^2 + n^2$ and $x = m^2 - n^2$ are both even, which contradicts $\gcd(x\,,\, z) = 1$. This show that m and n satisfy condition (1). The theorem is thus proved. □

1.4.2 Rational points on the unit circle

Our aim in this subsection is to give a geometric proof of Thm. 1.54. We start with some elementary observations.

(1.55) Let $a\,,\, b\,,\, c \in \mathbb{Z}$ be such that $a^2 + b^2 = c^2$. If $c = 0$, then trivially $a = b = 0$. We say that $(0\,,\, 0\,,\, 0)$ is the trivial solution to the equation $a^2 + b^2 = c^2$. If a triple $(a\,,\, b\,,\, c) \in \mathbb{Z}^3$ is a nontrivial solution, then clearly

$$(x\,,\, y) := \left(\frac{a}{c}\,,\, \frac{b}{c}\right) \in \mathbb{Q}^2$$

is a point on the unit circle

$$C := \left\{(x\,,\, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\right\}\,.$$

We call the set

$$C(\mathbb{Q}) := \{(x, y) \in \mathbb{Q}^2 \mid x^2 + y^2 = 1\}$$

the set of ***rational points*** (有理点) on the unit circle C. We have seen that every nontrivial solution to $a^2 + b^2 = c^2$ in $\mathbb{Z}^3$ determines a rational point of C. Conversely, suppose that $(x\,,\, y) \in C(\mathbb{Q})$. By choosing the least common denominator of the rational numbers x and y, we may write $x = \frac{a}{c}$ and $y = \frac{b}{c}$ for some integers $a\,,\, b\,,\, c \in \mathbb{Z}$. The equation $x^2 + y^2 = 1$ show that $a^2 + b^2 = c^2$.

These discussions tell us that finding all Pythagorean triples is almost the same as finding all points of $C(\mathbb{Q})$. ■

To describe the set $C(\mathbb{Q})$, there is a simple but interesting geometric method which we now explain.

(1.56) The point $A = (-1\,,\, 0)$ is obviously a rational point of unit circle. For any other point $B = (x\,,\, y)$ in $C(\mathbb{Q})$, the line AB has a rational slope. That is,

$$t := \frac{y - 0}{x - (-1)} = \frac{y}{x + 1} \in \mathbb{Q}\,.$$

So, we have a map

$$\begin{aligned} \rho\,:\, C(\mathbb{Q})\backslash\{A\} &\longrightarrow \mathbb{Q} \\ B = (x\,,\, y) &\longmapsto \frac{y}{x+1} = \text{ slope of the line } AB\,. \end{aligned}$$

By drawing a picture (or by some tedious computation), one finds easily that the map ρ is injective. That is, for different points $B_1, B_2 \in C(\mathbb{Q})\backslash\{A\}$, the slopes of AB_1 and AB_2 are distinct.

Now we show that ρ is also surjective. Namely, for every $t \in \mathbb{Q}$, the line through the point A with slope t intersects the unit circle C at exactly one more point B, and $B \in C(\mathbb{Q})$. In fact the line in

question has equation $y = t(x+1)$. Solving the system

$$\begin{cases} x^2 + y^2 = 1 \\ y = t(x+1) \end{cases}$$

shows immediately that the only solutions are

$$(-1\,,\,0) \quad \text{and} \quad \left(\frac{1-t^2}{1+t^2}\,,\,\frac{2t}{1+t^2}\right)\,.$$

The point we are looking for is given by

$$B = \left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right)\,.$$

■

We have thus obtained the following:

Theorem 1.57. *The set of rational points on the unit circle is given by*

$$\begin{aligned} C(\mathbb{Q}) &= \{(-1,0)\} \bigcup \left\{ \left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right) \,\middle|\, t \in \mathbb{Q} \right\} \\ &= \left\{ \left(\frac{1-t^2}{1+t^2}\,,\,\frac{2t}{1+t^2}\right) \,\middle|\, t \in \mathbb{Q} \cup \{\infty\} \right\} \end{aligned}$$

where we use the convention that

$$\left(\frac{1-t^2}{1+t^2}\,,\,\frac{2t}{1+t^2}\right) = (-1\,,\,0)\,,\ \textit{when } t = \infty\,.$$

Writing $t \in \mathbb{Q}$ in the form $t = \frac{m}{n}$ with $m \in \mathbb{Z}$ and $n \in \mathbb{N}^*$, we find

Corollary 1.58.

$$\begin{aligned} &\{\textit{solutions to } x^2 + y^2 = z^2 \textit{ in } \mathbb{Z}^3 \textit{ with } z \geq 0 \textit{ and } y \textit{ even}\} \\ =&\left\{(m^2 - n^2\,,\,2mn\,,\,m^2 + n^2) \mid m \in \mathbb{Z}\,,\, n \in \mathbb{N}\right\}. \end{aligned}$$

This shows that all Pythagorean triples can be found in a geometric way.

1.4.3 Fermat's method of infinite descent

Previously we showed that the ***diophantine equation*** (丢番图方程)*

$$x^2 + y^2 = z^2$$

has infinitely many solutions in nonzero integers x, y and z.

What happens when we replace the exponent 2 in this equation with a larger integer? The answer to this question is the famous ***Fermat's last theorem***. We refer the reader to the textbook, material on the internet, or any other sources for all the legend about this "big" theorem and all mathematicians who have contributed to this problem.

*by a diophantine equation we mean a polynomial equation with integer coefficients to which we are only interested in the integral solutions

In this course, we are only able to provide a proof of very special case of Fermat's last theorem. It makes use of the method of infinite descent devised by Fermat.

Theorem 1.59. *The diophantine equation*

$$x^4 + y^4 = z^2$$

has no solution in nonzero integers x, y and z. Therefore, the equation $x^4 + y^4 = z^4$ has no nonzero integer solutions either.

Proof. Suppose that the equation $x^4 + y^4 = z^2$ has a solution in nonzero integers. Since changing the sign of each variable does not affect the validity the equation, we may assume that we have a solution (x_0, y_0, z_0) with x_0, y_0, z_0 all positive.

Now consider the set

$$A := \{n \in \mathbb{N}^* \,|\, \text{for some positive integers } x,\, y \text{ one has } x^4 + y^4 = n^2\}\,.$$

By our assumption, A is nonempty. So, the well ordering property tells us that A has a least element. That is, there is a triple $(x_0, y_0, z_0) \in \mathbb{N}^{*3}$ such that

$$x_0^4 + y_0^4 = z_0^2$$

and that for any other triple $(x, y, z) \in \mathbb{N}^{*3}$, if $z < z_0$, then (x, y, z) does not satisfy the equation $x^4 + y^4 = z^2$.

To prove the theorem by contradiction, the strategy of Fermat's infinite descent is to show the following statement:

Given any solution $(x_0, y_0, z_0) \in \mathbb{N}^{*3}$ to the equation, one can find another solution $(x_1, y_1, z_1) \in \mathbb{N}^{*3}$ with $z_1 < z_0$.

First, we may assume that our solution (x_0, y_0, z_0) satisfies $\gcd(x_0, y_0) = 1$. For, if $d = \gcd(x_0,\, y_0) > 1$, then from $d^4 \,|\, x_0^4$ and $d^4 \,|\, y_0^4$ we can deduce that $d^4 \,|\, z_0^2 = x_0^4 + y_0^4$. This implies that $d^2 | z_0$ (Prove this as an exercise !). Thus, putting

$$x_1 = \frac{x_0}{d}\,,\; y_1 = \frac{y_0}{d} \;\text{ and }\; z_1 = \frac{z_0}{d^2}$$

we get a solution (x_1, y_1, z_1) with $z_1 < z_0$. This gives the desired result needed in the method of infinite descent.

So we may assume $\gcd(x_0, y_0) = 1$. Note that (x_0^2, y_0^2, z_0) is a Pythagorean triple. Furthermore, we have $\gcd(x_0^2, y_0^2) = 1$ (by the assumption $\gcd(x_0, y_0) = 1$). Hence, (x_0^2, y_0^2, z_0) is a primitive Pythagorean triple. By Thm. 1.54, we can find relatively prime integers m, $n \in \mathbb{N}^*$, of different parity, such that

$$\begin{cases} x_0^2 = m^2 - n^2 \\ y_0^2 = 2mn \\ z_0 = m^2 + n^2 \end{cases}$$

where we have interchanged x_0^2 and y_0^2, if necessary, to make y_0^2 even. The first equation shows that

$$x_0^2 + n^2 = m^2\,.$$

Since $\gcd(m,n)=1$, (x_0,n,m) is again a primitive Pythagorean triple. In particular, m is odd. Using Thm. 1.54 once again, we can find r, $s\in\mathbb{N}^*$, with $\gcd(r,s)=1$, of different parity, such that

$$\begin{cases} x_0 = r^2 - s^2 \\ n = 2rs \\ m = r^2 + s^2 \end{cases}$$

Because m is odd and $\gcd(m,\,n)=1$, we know that $\gcd(n,\,2m)=1$. Since $2mn=y_0^2$ is a square, Lemma 1.53 implies that $m=z_1^2$ and $2n=w^2$ for some z_1, $w\in\mathbb{N}^*$. Hence w must be even, so that $w=2v$ for some $v\in\mathbb{N}^*$. Now

$$n = 2rs = \frac{w^2}{2} = \frac{(2v)^2}{2} = 2v^2$$

yields $v^2=rs$. Since $\gcd(r,s)=1$, we many apply Lemma 1.53 once again to conclude $r=x_1^2$ and $s=y_1^2$ for some relatively prime $x_1,y_1\in\mathbb{N}^*$. Now

$$x_1^4 + y_1^4 = r^2 + s^2 = m = z_1^2.$$

Obviously $z_1^2=m<2mn=y_0^2<z_0^2$. This completes the proof. □

Chapter 2

Congruences and Applications

2.1 Introduction to congruences

2.1.1 Congruences and systems of residues

Definition 2.1. Let $m \in \mathbb{N}^*$. Two integers a and b are said to be ***congruent modulo*** m (模 m 同余) if $m \mid a-b$. When it is the case, we write $a \equiv b \pmod{m}$, otherwise we write $a \not\equiv b \pmod{m}$. The integer m is called the ***modulus*** (模) of the congruence. ■

We now show that congruence satisfies a number of important properties.

Theorem 2.2. *Let $m \in \mathbb{N}^*$. Congruence modulo m is an equivalence relation on the set $\mathbb{Z}$. Namely, for all integers a, b, c one has:*

(1) (Reflexivity) $a \equiv a \pmod{m}$;

(2) (Symmetry) $a \equiv b \pmod{m}$ if and only if $b \equiv a \pmod{m}$;

(3) (Transitivity) If $a \equiv b \pmod{m}$ and $b \equiv c \pmod{m}$, then $a \equiv c \pmod{m}$.

Proof. An easy exercise. □

Theorem 2.3. *Let $m \in \mathbb{N}^*$ and $a, b, c, d \in \mathbb{Z}$. If $a \equiv b \pmod{m}$ and $c \equiv d \pmod{m}$, then*

$$\begin{aligned} a + c &\equiv b + d \pmod{m} \\ a - c &\equiv b - d \pmod{m} \\ ac &\equiv bd \pmod{m} \end{aligned}$$

Proof. Only the last congruence is less obvious. Suppose $k = \frac{a-b}{m} \in \mathbb{Z}$ and $r = \frac{c-b}{m} \in \mathbb{Z}$. Then

$$\frac{ac-bd}{m} = \frac{ac-bc}{m} + \frac{bc-bd}{m} = c \cdot k + b \cdot r \in \mathbb{Z}$$

So we have $ac \equiv bd \pmod{m}$. □

Proposition 2.4. *Let $m \in \mathbb{N}^*$ and $A, B, c \in \mathbb{Z}$. Suppose that $a = \frac{A}{c} \in \mathbb{Z}$ and $b = \frac{B}{c} \in \mathbb{Z}$.*

If $A \equiv B \pmod{m}$, then

$$a \equiv b \left(\mathrm{mod}\ \frac{m}{\gcd(c\,,\, m)} \right) .$$

Proof. From $m \,|\, (A-B) = c(a-b)$ we see that

$$\frac{m}{d} \,\Big|\, \frac{c}{d}(a-b)\,, \quad \text{where } d = \gcd(c\,,\, m)\,.$$

Since $\gcd\left(\frac{m}{d}\,,\,\frac{c}{d}\right) = 1$, it follows from Lemma 1.34 that $\frac{m}{d} \,|\, a-b$. This means that $a \equiv b \pmod{\frac{m}{d}}$. □

Proposition 2.5. *Let $m_1, \cdots, m_r \in \mathbb{N}^*$ and $a, b \in \mathbb{Z}$. Then the following conditions are equivalent:*

1. *$a \equiv b \pmod{m_i}$ for every $i \in [\![1, r]\!]$.*
2. *$a \equiv b \pmod{M}$ where $M = \operatorname{lcm}(m_1, \cdots, m_r)$.*

Proof. If $a \equiv b \pmod{M}$, we have $M \,|\, a-b$ and hence $m_i \,|\, a-b$ for every i, since each $m_i \,|\, M$.

Conversely suppose that $a \equiv b \pmod{m_i}$ for every i. This means that $c := a-b$ is a common multiple of the m_i. Use the division algorithm to write

$$c = q \cdot M + r \quad \text{with } q\,,\, r \in \mathbb{Z} \text{ and } 0 \le r < M\,.$$

Since M is a common multiple of the m_i's, it follows that the remainder r is also a common multiple. Then we must have $r = 0$, because $r < M$ and by definition, M is the smallest positive common multiple of the m_i's. Hence $c = q \cdot M$ is a multiple of M. This proves that $a \equiv b \pmod{M}$, as required. □

Corollary 2.6. *Let $m_1, \cdots, m_r \in \mathbb{N}^*$ be pairwise relatively prime. Then for all $a, b \in \mathbb{Z}$, one has*

$$a \equiv b \pmod{m_i}\,,\ \forall\, i \in [\![1, r]\!] \iff a \equiv b \pmod{M}$$

where $M = m_1 m_2 \cdots m_r$.

Proof. As an easy exercise, the reader should prove that $\operatorname{lcm}(m_1, \cdots, m_r) = m_1 m_2 \cdots m_r$ when the m_i's are pairwise relatively prime. Then apply Prop. 2.5. □

(2.7) Fix $m \in \mathbb{N}^*$. For any $a \in \mathbb{Z}$, the division algorithm gives rise to a unique expression

$$a = bm + r \ \text{ with } b\,,\, r \in \mathbb{Z} \text{ and } 0 \le r < m\,.$$

We say that r is the result of ***reducing*** a ***modulo*** m (a 模 m 的约化) or the ***least nonnegative residue*** (最小非负剩余) of a modulo m. When $r > 0$, we also say that it is the ***least positive residue*** (最小正剩余) of a modulo m.

We shall write $r = a \text{ MOD } m$ to mean that r is the least nonnegative residue of a modulo m. (In our textbook, "MOD" is denoted by "mod" in bold.)

We may consider "MOD" as a map

$$\mathbb{Z} \longrightarrow [\![0, m-1]\!]\ ; \quad a \longmapsto a \text{ MOD } m \tag{2.7.1}$$

which is sometimes called the ***reduction mod*** m (模 m 的约化). ■

The easy proof of the next proposition is left to the reader.

Proposition 2.8. *Let $m \in \mathbb{N}^*$ and $a, b \in \mathbb{Z}$. Then*

$$a \equiv b \pmod{m} \iff a \text{ MOD } m = b \text{ MOD } m\,.$$

Definition 2.9. Let $m \in \mathbb{N}^*$. A ***complete system*** (or ***set***) ***of residues modulo*** m (模 m 的完全剩余系) is a set of integers such that every integer is congruent modulo m to exactly one integer of the set.

The set $\{0, 1, 2, \cdots, m-1\} = [\![0,\, m-1]\!]$ is clearly a complete system of residues modulo m. It is called the ***set of least nonnegative residues*** (最小非负剩余系) mod m. ■

Lemma 2.10. *Let $m \in \mathbb{N}^*$ and $S \subseteq \mathbb{Z}$. Then the following conditions are equivalent:*

1. *S is a complete system of residues modulo m.*
2. *Every integer is congruent to at least one element of S, and distinct elements of S are incongruent modulo m.*
3. *The map*

$$\rho \,:\, S \longrightarrow [\![0, m-1]\!] \;; \qquad x \longmapsto x \text{ MOD } m \,. \tag{2.10.1}$$

   is bijective.

4. *$|S| = m$ and and distinct elements of S are incongruent modulo m.*

Proof. (1) ⇒ (2). The first assertion in (2) is clearly part of the definition of complete system of residues. If a, $b \in S$ are distinct, then they cannot be congruent modulo m because otherwise the integer $a \in S$ is congruent to at least two members of S (i.e., $b \in S$ and $a \in S$ itself).

(2) ⇒ (1). It remains to check that for any $x \in \mathbb{Z}$, there are at most one element a in S such that $x \equiv a \pmod m$. If there were another, say $b \in S$, such that $x \equiv b \pmod m$, then we have $a \equiv x \equiv b \pmod m$, by Thm. 2.2. Since we have assumed that distinct members of S are incongruent, it follows that $a = b$.

(2) ⇔ (3). According to Prop. 2.8, to say that distinct members of S are incongruent mod m is to say that ρ is an injective map.

For each $j \in [\![0,\, m-1]\!]$, the first hypothesis in (2) says nothing but the following: there exists $a \in S$ such that $a \equiv j \pmod m$. Since $0 \le j \le m-1$, this j must be the least nonnegative residue of a mod m, i.e., $j = \rho(a)$. Therefore, the first statement in (2) is equivalent to the surjectivity of ρ.

(3) ⇔ (4). As we said above, the second statement in (4) is equivalent to the injectivity of ρ, by Prop. 2.8. This together with the assumption $|S| = m = |\{0, 1, \ldots, m-1\}|$ is equivalent to the bijectivity of ρ. □

Theorem 2.11. *Let $m \in \mathbb{N}^*$ and let $\{r_1, \cdots, r_m\}$ be a complete system of residues mod m.*

Then for any $a, b \in \mathbb{Z}$ with $\gcd(a, m) = 1$, the integers

$$ar_1 + b\,,\; ar_2 + b\,,\; \cdots\,,\; ar_m + b$$

form a complete system of residues mod m.

Proof. According to Lemma 2.10, it is sufficient to show that

$$ar_i + b \not\equiv ar_j + b \pmod m$$

for all distinct $i, j \in [\![1, m]\!]$. Without loss of generality, we may assume $i = 1$, $j = 2$. Notice that

$$\begin{aligned} ar_1 + b \equiv ar_2 + b \pmod{m} &\iff \frac{(ar_1 + b) - (ar_2 + b)}{m} \in \mathbb{Z} \\ &\iff \frac{a(r_1 - r_2)}{m} \in \mathbb{Z} \iff m \mid a(r_1 - r_2) \\ \text{since } \gcd(a, m) = 1 \quad &\iff m \mid (r_1 - r_2) \iff r_1 \equiv r_2 \pmod{m}. \end{aligned}$$

But r_1 and r_2 are incongruent modulo m, since they are elements in a complete system of residues mod m. We have thus proved $ar_1 + b \not\equiv ar_2 + b \pmod{m}$, as desired. □

2.1.2 Linear congruences

We fix a positive integer m.

(2.12) By a ***linear congruence*** (***equation***) (线性同余方程) we mean a congruence of the form

$$ax \equiv b \pmod{m} \tag{2.12.1}$$

where $a, b \in \mathbb{Z}$ and x is an unknown integer. We first note that if $x_0 \in \mathbb{Z}$ is a solution to (2.12.1) and $x_1 \in \mathbb{Z}$ is congruent to x_0 modulo m, then we have

$$ax_1 \equiv ax_0 \equiv b \pmod{m}.$$

Therefore, x_1 is also a solution to (2.12.1). The problem of finding all solutions to a linear congruence mod m is thus reduced to finding all solutions in any given complete system of residue modulo m.

We shall say that the linear congruence (2.12.1) has N ***incongruent solutions modulo*** m (N 个模 m 不同余的解) if in a complete system of residues modulo m there are N elements satisfying the congruence. ■

Our goal now is to determine when the congruence has solutions, and in case it does, to tell exactly the number of incongruent solutions mod m.

Theorem 2.13. *Consider the linear congruence*

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

where $m \in \mathbb{N}^*$ and $a, b \in \mathbb{Z}$ are given, and x is unknown. Let $d = \gcd(a, m)$.

Then the above congruence has solutions if and only if $d \mid b$.

When $d \mid b$, the congruence has exactly d incongruent solutions modulo m, and these solutions are all congruent modulo $\frac{m}{d}$.

Proof. If the congruence has a solution $x \in \mathbb{Z}$, then $b - ax = my$ for some $y \in \mathbb{Z}$. Thus $b = ax + my$ is a $\mathbb{Z}$-linear combination of a and m. By Corollary 1.27, b is a multiple of $d = \gcd(a, m)$. Conversely, if $d \mid b$, then by Corollary 1.27 again, b is a $\mathbb{Z}$-linear combination of a and m. This is, $b = ax + my$ for some $x, y \in \mathbb{Z}$. This yields $ax \equiv b \pmod{m}$, whence a solution to the congruence in question.

Now suppose $d \mid b$. We know that there exists $x_0 \in \mathbb{Z}$ satisfying $ax_0 \equiv b \pmod{m}$. For any $x \in \mathbb{Z}$,

$$\begin{aligned} ax \equiv b \pmod{m} &\iff ax_1 \equiv ax_0 \pmod{m} \iff m \mid a(x - x_0) \\ &\iff \frac{m}{d} \,\Big|\, \frac{a}{d}(x - x_0) \\ \text{since } \gcd\left(\frac{m}{d}, \frac{a}{d}\right) = 1 \quad &\iff \frac{m}{d} \,\Big|\, x - x_0 \,. \end{aligned}$$

On the one hand, this shows that all integer solutions to $ax \equiv b \pmod{m}$ are congruent to x_0 modulo $\frac{m}{d}$.

On the other hand, we see that in the integer interval $[\![x_0, x_0 + m - 1]\!]$, which is a complete system of residues mod m by Thm. 2.11, the solutions to the congruence are precisely elements of the set

$$\begin{aligned} S :&= \left\{ x \in [\![x_0, x_0 + m - 1]\!] \;\middle|\; x \equiv x_0 \left(\mathrm{mod}\, \frac{m}{d}\right) \right\} \\ &= \left\{ x_0 + \frac{m}{d} \cdot k \;\middle|\; 0 \le \frac{m}{d} \cdot k \le m - 1 \right\} \end{aligned}$$

Therefore, the number of incongruent solutions m is equal to $|S| = d$. This completes the proof. □

Corollary 2.14. *The congruence $ax \equiv b \pmod{m}$ has a unique solution modulo m if and only if $\gcd(a, m) = 1$.*

Corollary 2.15. *Let $m \in \mathbb{N}^*$ and $a \in \mathbb{Z}$. Then the following are equivalent:*

(a) The congruence $ax \equiv 1 \pmod{m}$ has a solution.

(b) The congruence $ax \equiv 1 \pmod{m}$ has a unique solution modulo m.

(c) $\gcd(a, m) = 1$.

(2.16) Given $a \in \mathbb{Z}$, an integer solution x of the congruence $ax \equiv 1 (mod\ m)$ is called an ***inverse of a mod*** m (a 模 m 的逆). By Corollary 2.15, such an inverse exists if and only if $\gcd(a, m) = 1$. When it is the case, any two inverses of a mod m must be congruent mod m, by Coro. 2.14. To find such an inverse explicitly (when it exists), one can use the extended Euclidean algorithm to find $u, v \in \mathbb{Z}$ such that

$$1 = \gcd(a, m) = au + mv \,.$$

Then clearly u is an inverse of a mod m. ■

Proposition 2.17. *Let p be a prime. An integer $a \in \mathbb{Z}$ is its own inverse modulo p if and only if*

$$a \equiv 1 \pmod{p} \quad \text{or} \quad a \equiv -1 \pmod{p} \,.$$

Proof. An exercise left to the reader. □

Theorem 2.18. *Let $a, b, c \in \mathbb{Z}$ be nonzero integers.*

Then the equation $ax + by = c$ has integer solutions if and only if $\gcd(a, b) \mid c$.

If $(x_0, y_0) \in \mathbb{Z} \times \mathbb{Z}$ is a particular solution, then a general solution has the following form:

$$\text{(2.18.1)} \qquad \begin{cases} x = x_0 + \frac{b}{\gcd(a,b)} t \\ \\ y = y_0 - \frac{a}{\gcd(a,b)} t \end{cases} \qquad \text{where } t \in \mathbb{Z}$$

Proof. To prove the first assertion, note that

$$\begin{aligned} & ax+by=c \text{ has integer solutions} \\ \iff & \text{the congruence } ax\equiv c \pmod{b} \text{ has solutions.} \\ & \text{(Here obviously, we may assume } b>0) \\ \text{by Theorem (2.13)} \iff & \gcd(a,\,b)\,|\,c\,. \end{aligned}$$

Alternatively, we may use corollary(1.17) to get

$$\begin{aligned} & ax+by=c \text{ has integer solutions} \\ \iff & c \text{ is a } \mathbb{Z}\text{-linear combination of } a \text{ and } b \\ \iff & \gcd(a,\,b)\,|\,c\,. \end{aligned}$$

Now put $d=\gcd(a,b)$ and suppose that (x_0,y_0) is an integer solution to $ax+by=c$. Then x_0 is a solution of the congruence $ax\equiv c \pmod{b}$. Similarly, if (x,y) is another solution to $ax+by=c$, then x is a solution to $ax\equiv c \pmod{b}$. By the last assertion of Thm. 2.13, we have

$$x=x_0 \pmod{\frac{b}{d}}$$

which means that $x=x_0+\frac{b}{d}\cdot t$ for some $t\in\mathbb{Z}$. Then, from the equation $ax+by=c$ we can solve out

$$y=y_0-\frac{a}{d}t\,.$$

So we find that all solution to $ax+by=c$ satisfy (2.18.1).

Conversely, one verifies directly that all integers x,y satisfying (2.18.1) must be a solution to $ax+by=c$. This completes the proof of the theorem. □

2.2 The Chinese Remainder Theorem and Euler's Phi-function

2.2.1 Statement and proof of the theorem

Perhaps motivated by a number of famous mathematical puzzles, mathematicians in ancient China contributed a lot to the solution of systems of congruences in only one unknown, but relative to different moduli. The outcome of their work is the following remarkable theorem:

Theorem 2.19 (Chinese Remainder Theorem 中国剩余定理)**.** *Let $m_1,\cdots,m_r\in\mathbb{N}^*$ be* **pairwise relatively prime**. *Then, for all integers $a_1,\cdots,a_r\in\mathbb{Z}$, the system of congruences*

$$\begin{cases} x\equiv a_1 \pmod{m_1} \\ \quad\vdots \\ \quad\vdots \\ x\equiv a_r \pmod{m_r} \end{cases} \tag{2.19.1}$$

has a unique solution modulo $M:=m_1m_2\cdots m_r$.

Proof. The uniqueness follows from the following exercise (cf. [Ros, §3.5, Exercise 37]): for any $c \in \mathbb{Z}$,

$$m_i \,|\, c \quad \text{for all } i \iff \operatorname{lcm}(m_1, \cdots, m_r) \,|\, c\,.$$

Here, $\operatorname{lcm}(m_1 m_2 \cdots m_r) = M = m_1 m_2 \cdots m_r$ since the m_i's are pairwise coprime by assumption.

To prove the existence, we put

$$N_i = \frac{M}{m_i} = \prod_{\substack{1 \le j \le r \\ j \neq i}} m_j\,, \qquad \text{for each } i = 1\,,\, 2\,,\, \cdots\,,\, r\,.$$

Then $\gcd(N_i, m_i) = 1$. So we can find an inverse y_i of N_i mod m_i, i.e., $N_i \cdot y_i \equiv 1 \pmod{m_i}$. Since $m_j \,|\, N_i$ for $j \neq i$, the integer $e_i := N_i \cdot y_i$ has the property:

$$\begin{cases} e_i \equiv 1 \pmod{m_i} \\ e_i \equiv 0 \pmod{m_j}\,,\ \forall\, j \neq i \end{cases}$$

One can then check easily that

$$x := a_1 e_1 + a_2 e_2 + \cdots + a_r e_r$$

is an integer solution to (2.19.1). □

(2.20) Now we illustrate the use of the Chinese Remainder theorem by solving the following system that arises from an ancient Chinese puzzle:

$$\begin{cases} x \equiv 2 \pmod 3 \\ x \equiv 3 \pmod 5 \\ x \equiv 2 \pmod 7 \end{cases}$$

According to the method of our proof of Theorem 2.19, we shall first find integers $e_1, e_2, e_3 \in \mathbb{Z}$ such that

		mod 3	mod 5	mod 7
e_1	$\equiv$	1	0	0
e_2	$\equiv$	0	1	0
e_3	$\equiv$	0	0	1

To do so, consider $N_1 = 5 \times 7 = 35$. By mental calculation (or by using the extended Euclidean algorithm), we can find an inverse y_1 of N_1 mod 3. For instance, $y_1 = 2$ is a possible choice. So we may take $e_1 = N_1 y_1 = 70$.

Using the same method, we see that e_2 and e_3 can be chosen to be $e_2 = 21$ and $e_3 = 15$. Thus,

$$x_0 = 2e_1 + 3e_2 + 2e_3 = 233 \ \text{ is a solution}\,.$$

Any $x \in \mathbb{Z}$ satisfying $x \equiv 233 \pmod{3 \times 5 \times 7 = 105}$ is also a solution. The smallest positive solution is $x = 23$. ■

2.2.2 The ring of congruence classes

We continue to work with a fixed $m \in \mathbb{N}^*$.

(2.21) Given an integer $a \in \mathbb{Z}$, we define

$$[a]_m := \{x \in \mathbb{Z} \mid x \equiv a \mod m\}\,. \tag{2.21.1}$$

That is, $[a]_m$ is the ***set*** of all integers that are congruent to a modulo m. If is called the ***congruence class*** 同余类 or ***residue class*** 剩余类 of a mod m.

(A very popular notation for this congruence class is $\overline{a}$. However, in our textbook [Ros] the author often uses the notation $\overline{a}$ to denote an inverse of a modulo m. To avoid confusion and to make the modulus appear in the notation, we write $[a]_m$ for this congruence class.)

If b is another integer, we have

$$[a]_m = [b]_m \iff a \equiv b \pmod m\,. \tag{2.21.2}$$

(You should prove this by yourself, as an easy exercise!)

We now define

$$\mathbb{Z}/m\mathbb{Z} := \text{the set of all congruence classes modulo } m = \{[a]_m \mid a \in \mathbb{Z}\}\,. \tag{2.21.3}$$

Thus, $\mathbb{Z}/m\mathbb{Z}$ is a set whose elements are themselves sets. If $\alpha \in \mathbb{Z}/m\mathbb{Z}$, then by definition, there is an integer $a \in \mathbb{Z}$ such that $\alpha = [a]_m$ =the congruence class of a mod m. We say that a is a ***representative*** 代表元 of the congruence class α.

Notice that α has infinitely many representatives, because for every $b \in \mathbb{Z}$ that is congruent to the given representative a modulo m, we have

$$[b]_m = [a]_m = \alpha, \qquad \text{by (2.21.2)}$$

showing that b is another representative of α. Also, it follows from (2.21.2) that for any complete system S of residues modulo m, we have a bijection

$$S \xrightarrow{\sim} \mathbb{Z}/m\mathbb{Z}\,; \quad x \longmapsto [x]_m\,.$$

In other words, for any complete system s of residues mod m, we have

$$\mathbb{Z}/m\mathbb{Z} = \{[x]_m \mid x \in S\}\,. \tag{2.21.4}$$

In particular

$$\mathbb{Z}/m\mathbb{Z} = \{[0]_m, [1]_m, \cdots, [m-1]_m\} \text{ and } \quad |\mathbb{Z}/m\mathbb{Z}| = m\,. \tag{2.21.5}$$

■

(2.22) The importance of the set $\mathbb{Z}/m\mathbb{Z}$ roots mostly in the fact that, just as numbers, elements of $\mathbb{Z}/m\mathbb{Z}$ can be used to do basic arithmetic operations: addition, multiplication, and so on.

Let us show how the addition and the multiplication are defined for elements of $\mathbb{Z}/m\mathbb{Z}$.

Let $\alpha \in \mathbb{Z}/m\mathbb{Z}$ and $\beta \in \mathbb{Z}/m\mathbb{Z}$ be given. We define their ***sum*** $\alpha + \beta$ and their ***product*** $\alpha \cdot \beta$ in the following way: First, we choose a representation $a \in \mathbb{Z}$ for α and a representative $b \in \mathbb{Z}$ for β. Namely, $\alpha = [a]_m$ is the congruence class of a mod m, and similarly $\beta = [b]_m$. Then we do the addition $a + b$ in $\mathbb{Z}$,

and define $\alpha+\beta$ to be the congruence class $[a+b]_m$, which is a new element of $\mathbb{Z}/m\mathbb{Z}$. Similarly, the product $\alpha\cdot\beta$ is defined as the congruence class of the integer $a\cdot b$, i.e., $\alpha\cdot\beta=[ab]_m$.

Here is a question about the above definitions of $\alpha+\beta$ and $\alpha\cdot\beta$. In the way we described above, the congruence classes $[a+b]_m$ and $[ab]_m$ are obtained by means of a particular choice of representatives of α and β. The question is when we choose some other representatives a' and b' for α and β respectively, do the resulting congruence classes $[a'+b']_m$ and $[a'\cdot b']_m$ agree with $[a+b]_m$ and $[a'\cdot b]_m$ respectively? If the answer is always positive, we say that the sum $\alpha+\beta$ and product $\alpha\cdot\beta$ are ***well defined*** 良好定义的. Indeed, this is true by the criterion (2.21.2) and Thm. 2.3. ■

(2.23) Now we see that the set $\mathbb{Z}/m\mathbb{Z}$ are equipped with two operations: an addition and a multiplication, which we denote by $+$ and $\bullet$ respectively.

The triple $(\mathbb{Z}/m\mathbb{Z}\,,+\,,\bullet)$ has the following properties:

(1) The addition $+$

(a) is commutative: $\forall\,\alpha\,,\beta\in\mathbb{Z}/m\mathbb{Z}$, $\alpha+\beta=\beta+\alpha$;

(b) is associative: $\forall\,\alpha\,,\beta\,,\gamma\in\mathbb{Z}/m\mathbb{Z}$, $(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$;

(c) has an identity: the element $[0]_m$ satisfies $[0]_m+\alpha=\alpha$ for all $\alpha\in\mathbb{Z}/m\mathbb{Z}$;

(d) every element has an additive inverse:

$$\forall\,\alpha\in\mathbb{Z}/m\mathbb{Z}\,,\ \exists\,\beta\in\mathbb{Z}/m\mathbb{Z}\quad\text{ such that }\alpha+\beta=[0]_m\ .$$

(2) The multiplication $\bullet$

(a) is commutative: $\forall\alpha\,,\beta\in\mathbb{Z}/m\mathbb{Z}$, $\alpha\cdot\beta=\beta\cdot\alpha$;

(b) is associative:$\forall\alpha\,,\beta\,,\gamma\in\mathbb{Z}/m\mathbb{Z}$, $(\alpha\cdot\beta)\cdot\gamma=\alpha\cdot(\beta\cdot\gamma)$;

(c) is distributive with respect to the addition :

$$\forall\alpha\,,\beta\,,\gamma\in\mathbb{Z}/m\mathbb{Z}\,,\quad\alpha\cdot(\beta+\gamma)=\alpha\cdot\beta+\alpha\cdot\gamma\text{ and }(\beta+\gamma)\cdot\alpha=\beta\cdot\alpha+\gamma\cdot\alpha$$

(d) has an identity: the element $[1]_m$ satisfies $[1]_m\cdot\alpha=\alpha=\alpha\cdot[1]_m$, $\forall\alpha\in\mathbb{Z}/m\mathbb{Z}$.

We summarize all the above properties by saying that the triple $(\mathbb{Z}/m\mathbb{Z},+,\bullet)$ is a ***commutative ring*** 交换环. We often omit the mention of the two operators $+$ and $\bullet$, and simply say that $\mathbb{Z}/m\mathbb{Z}$ is a commutative ring.

It is easy to show that the additive identity and the multiplicative identity are unique. That is, if $\beta\in\mathbb{Z}/m\mathbb{Z}$ is such that $\alpha+\beta=\alpha=\beta+\alpha$, $\forall\alpha\in\mathbb{Z}/m\mathbb{Z}$, then we must have $\beta=[0]_m$; if β satisfies $\alpha\cdot\beta=\alpha=\beta\cdot\alpha$, $\forall\alpha\in\mathbb{Z}/m\mathbb{Z}$, then we must have $\beta=[1]_m$.

Also, for every $\alpha\in\mathbb{Z}/m\mathbb{Z}$, its additive is unique. We denote it by $-\alpha$. We can define a ***subtraction*** 减法 on $\mathbb{Z}/m\mathbb{Z}$ by setting

$$\alpha-\beta:=\alpha+(-\beta)\,.$$

Of course, this is the same as defining

$$[a]_m-[b]_m:=[a-b]_m$$

because obviously $-[b]_m=[-b]_m$. ■

Now we give the definition of a general commutative ring.

Definition 2.24. A ***commutative ring*** 交换环 is a triple $(A, +, \bullet)$ consisting of a nonempty A, and two maps

$$A \times A \to A; \qquad (x, y) \longmapsto x + y$$

and

$$A \times A \to A; \qquad (x, y) \longmapsto x \cdot y\,,$$

which we call ***addition*** 加法 and ***multiplication*** 乘法 respectively, such that the following conditions hold:

(1) The addition is commutative, has an identity element, and every element has an additive inverse. That is,

   (a) $\forall\, \alpha\,, \beta \in A$, $\alpha + \beta = \beta + \alpha$;

   (b) $\forall\, \alpha\,, \beta\,, \gamma \in A$, $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$;

   (c) $\exists$ an element $0_A \in A$ such that $\alpha + 0_A = \alpha = 0_A + \alpha, \forall \alpha \in A$;

   (d) $\forall\, \alpha \in A$, $\exists\, \beta \in A$ such that $\alpha + \beta = \beta + \alpha = 0_A$.

(2) The multiplication is commutative, associative, distributive with respect to addition, and has an identity element. That is,

   (a) $\forall\, \alpha\,, \beta \in A\,$, $\alpha \cdot \beta = \beta \cdot \alpha$;

   (b) $\forall\, \alpha\,, \beta\,,\ \gamma \in A$, $(\alpha \cdot \beta) \cdot \gamma = \alpha \cdot (\beta \cdot \gamma)$;

   (c) $\forall\, \alpha\,,\ \beta\,, \gamma \in A$, $\alpha \cdot (\beta + \gamma) = \alpha \cdot \beta + \alpha \cdot \gamma$, $(\beta + \gamma) \cdot \alpha = \beta \cdot \alpha + \gamma \cdot \alpha$;

   (d) $\exists$ an element $1_A \in A$ such that $\alpha \cdot 1_A = \alpha = 1_A \cdot \alpha, \forall\, \alpha\ \in A$.

As an exercise, one can prove

- the additive identity and the multiplicative identity are unique; we often denote them by 0_A and 1_A (or simply 0 and 1) respectively;
- the additive inverse of every $\alpha \in A$ is unique; it is denoted by $-\alpha$.

We may thus define a ***subtraction*** 减法 on A by

$$\alpha - \beta := \alpha + (-\beta)\,.$$

When the operations $\times$ and $\bullet$ are clear from the context, we often say simply that A is a commutative ring. Since in this course we will not discuss non-commutative rings at all, we will simply say "ring" to mean "commutative ring". ■

(2.25) There are even more natural examples than the ring $(\mathbb{Z}/m\mathbb{Z}, +, \bullet)$: For each of the following sets, $(A,\ +\,,\ \bullet)$ is a ring, where

$$A = \mathbb{Z}\,,\ \mathbb{Q}\,,\ \mathbb{R} \text{ or } \mathbb{C}$$

and the operations $+$ and $\bullet$ are the usual addition and multiplication of numbers. ■

(2.26) In the definition of a ring A, not all elements of A are required to have a multiplication inverse. If some element $\alpha \in A$ does have one, i.e., $\alpha \cdot \beta = 1_A$ for some $\beta \in A$, then we say that α is a ***unit*** 単位 in A, or that α is ***invertible*** 可逆 in A. The element β such that $\alpha \cdot \beta = 1_A$, when it exists, is unique (you should prove this as an exercise !) and will be denote by α^{-1} and called the ***inverse*** 逆元 (or more precisely, ***multiplication inverse*** 乗法逆) of α. We will write

$$A^* = \{\text{units in } A\} = \{\alpha \in A \mid \alpha \text{ has an inverse in } A\}\,.$$

For example, if $A = \mathbb{Z}$, then $A^* = \{\pm 1\}$; if $A = \mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$, then $A^* = A \setminus \{0\}$. ■

Proposition 2.27. *Let A be the ring $\mathbb{Z}/m\mathbb{Z}$ and let $\alpha \in A$. Let $a \in \mathbb{Z}$ be a representative of the congruence class α, i.e., $[a]_m = \alpha$.*

1. *The following conditions are equivalent for an element $\beta \in A = \mathbb{Z}/m\mathbb{Z}$:*
   (a) *$\alpha \cdot \beta = 1_A$, i.e., β is an inverse of α in A;*
   (b) *every representative $b \in \mathbb{Z}$ of β satisfies $ab \equiv 1 \pmod m$;*
   (c) *some representative $b \in \mathbb{Z}$ of the congruence class β satisfies $ab \equiv 1 \pmod m$.*
2. *α is a unit in A if and only if the congruence $ax \equiv 1 \pmod m$ has a solution, if and only if $\gcd(a, m) = 1$.*

Proof. (1) First we assume $\beta = \alpha^{-1}$ and prove that every representative $b \in \mathbb{Z}$ of β satisfies $ab \equiv 1 \pmod m$. In fact, the product $\alpha \cdot \beta$ is by definition the congruence class $[ab]_m$. The element $1_A \in \mathbb{Z}/m\mathbb{Z}$ is the congruence class $[1]_m$. Therefore,

$$\alpha\beta = 1_A \iff [ab]_m = [1]_m\,.$$

The latter condition is equivalent to $ab \equiv 1 \pmod m$, by (2.21.2).
Since the definition of the product $\alpha \cdot \beta$ is independent of the choice of representatives, if for some representative $b \in \mathbb{Z}$ of β the condition $ab \equiv 1 \pmod m$ holds, then we have $\alpha \cdot \beta = [1]_m = 1_A$, i.e., β is an inverse of α.

We have thus proved $(c) \Rightarrow (a) \Rightarrow (b)$. The implication $(b) \Rightarrow (c)$ is trivial.

(2) If α is a unit, then the representatives of α^{-1} are solutions to the congruence $ax \equiv 1 \pmod m$, by (1).

Conversely, if $b \in \mathbb{Z}$ is a solution to $ax \equiv 1 \pmod m$. then the element $\beta := [b]_m \in \mathbb{Z}/m\mathbb{Z}$ is an inverse of $\alpha = [a]_m$, by (1) again. We already know that $\gcd(a, m) = 1$ is a necessary and sufficient condition for $ax \equiv 1 \pmod m$ to have a solution. So, this finishes the proof. □

Corollary 2.28. *We have*

$$(\mathbb{Z}/m\mathbb{Z})^* = \{[a]_m \in \mathbb{Z}/m\mathbb{Z} \mid a \in \mathbb{Z}\,,\ \gcd(a, m) = 1\}\,. \tag{2.28.1}$$

If S is a complete system of residues modulo m, then we also have

$$(\mathbb{Z}/m\mathbb{Z})^* = \{[a]_m \in \mathbb{Z}/m\mathbb{Z} \mid a \in S\,,\ \gcd(a, m) = 1\}\,. \tag{2.28.2}$$

In particular,

$$(\mathbb{Z}/m\mathbb{Z})^* = \{[a]_m \in \mathbb{Z}/m\mathbb{Z} \mid a \in [\![0, m-1]\!]\,,\ \gcd(a, m) = 1\}\,. \tag{2.28.3}$$

Proof. (2.28.1) follows from Prop. 2.27 (2). It implies (2.28.2) because

$$\mathbb{Z}/m\mathbb{Z} = \{[a]_m \mid a \in S\}\,, \text{ by (2.21.4)}\,.$$

Finally, (2.28.3) is a special case of (2.28.2), since $[\![0, m-1]\!]$ is a complete system of residues modulo m. □

Example 2.29. Let us compute two concrete examples.

(1) Let $m = 9$. Then

$$(\mathbb{Z}/9\mathbb{Z})^* = \{[1]_9, [2]_9, [4]_9, [5]_9, [7]_9, [8]_9\} = \{\pm[1]_9, \pm[2]_9, \pm[4]_9\}\,.$$

The inverses of these units are

$$\begin{array}{lll} {[1]_9}^{-1} = [1]_9, & {[2]_9}^{-1} = [5]_9 = -[4]_9, & {[4]_9}^{-1} = [7]_9 = -[2]_9 \\ {[5]_9}^{-1} = [2]_9, & {[7]_9}^{-1} = [4]_9, & {[8]_9}^{-1} = [8]_9\,. \end{array}$$

(2) Let $m = 12$. we have

$$(\mathbb{Z}/12\mathbb{Z})^* = \{\pm[1]_{12}, \pm[5]_{12}\}\,.$$

Their inverses are

$$(\pm[1]_{12})^{-1} = \pm[1]_{12}\,, \quad (\pm[5]_{12})^{-1} = \pm[5]_{12}$$

Note that for every $\alpha \in (\mathbb{Z}/12\mathbb{Z})^*$, we have $\alpha^{-1} = \alpha$. ■

Note that if p is a prime number, then every integer in $[\![1, p-1]\!]$ is relatively prime to p. So we get the following corollary of (2.28.3):

Corollary 2.30. *Let p be a prime. Then*

$$(\mathbb{Z}/p\mathbb{Z})^* = \{[1]_p, [2]_p\,, \cdots, [p-1]_p\} = \{\alpha \in \mathbb{Z}/p\mathbb{Z} \mid \alpha \neq [0]_p\}\,.$$

(2.31) Let us return to the situation of a general ring A. The set A^* of units in A has the following properties with respect to the multiplication:

- If $\alpha \in A^*$, then its inverse α^{-1} is also a unit, i.e., $\alpha^{-1} \in A^*$.
- $1_A \in A^*$, because 1_A is its own inverse.
- If $\alpha\,,\ \beta \in A^*$, then $\alpha \cdot \beta \in A^*$, because $\beta^{-1} \cdot \alpha^{-1}$ is an inverse of $\alpha\beta$.

This means that we have a multiplication map

$$A^* \times A^* \to A^*; \qquad (x, y) \longmapsto x \cdot y\,,$$

which satisfies the following properties:

(1) the multiplication is associative and commutative;

(2) there is a unique *identity element* e such that

$$\alpha \cdot e = \alpha = e \cdot \alpha \text{ for all } \alpha\,;$$

(3) every element α has an inverse.

We express this collection of facts by saying that the set A^* together with its multiplication is a ***multiplicative group*** 乘法群. To emphasize its relationship with the ring$(A, +, \cot)$, we say that A^* is the ***group of units*** 单位群 or ***group of invertible elements*** 可逆元乘法群 of the ring A. ■

(2.32) Let us specialize to the case $A = \mathbb{Z}/m\mathbb{Z}$. We denote by

$$\begin{aligned}
\phi(m) :&= \left|(\mathbb{Z}/m\mathbb{Z})^*\right| = \left|\{\alpha \in \mathbb{Z}/m\mathbb{Z} \mid \alpha \text{ is invertible}\}\right| \\
\text{by (2.28.3)} &= \left|\{[a]_m \in \mathbb{Z}/m\mathbb{Z} \mid a \in [\![0, m-1]\!], \gcd(a,m) = 1\}\right| \\
\text{by (2.28.2)} &= \left|\{[a]_m \in \mathbb{Z}/m\mathbb{Z} \mid a \in [\![1, m-1]\!], \gcd(a,m) = 1\}\right| \\
&= \text{number of natural numbers in the interval } [1,\, m] \text{ that are relatively prime to } m\,.
\end{aligned}$$

The function $n \in \mathbb{N}^* \mapsto \phi(n) \in \mathbb{N}^*$ is called the ***Euler phi-function*** (or ***Euler ϕ-function***) 欧拉 phi 函数 (欧拉 ϕ-函数).

Note that $\phi(1) = 1$. ■

Example 2.33. Here are some values of the Euler ϕ-function.

(1) According to Example 2.29, we have $\phi(9) = 6$ and $\phi(12) = 4$.

(2) By Corollary 2.30, for every prime p we have $\phi(p) = p - 1$. ■

2.2.3 Reduced residue systems

(2.34) Let R be a set of integers. We say that R is a ***reduced residue system***or ***reduced system of residues***) modulo m 模 m 的既约剩余系 if the map

$$\rho : R \to \mathbb{Z}/m\mathbb{Z}; \qquad x \mapsto [x]_m$$

is injective and its image is equal to $(\mathbb{Z}/m\mathbb{Z})^*$. In other words, R is a reduced residue system mod m if and only if the following conditions hold:

- For all $x, y \in R, x \neq y$ implies $[x]_m \neq [y]_m$ (or equivalently, $x \not\equiv y (\mathrm{mod}\ m)$).
- For every $x \in R$, the element $[x]_m \in \mathbb{Z}/m\mathbb{Z}$ is invertible, i.e., $[x]_m \in (\mathbb{Z}/m\mathbb{Z})^*$, or equivalently, $\gcd(x, m) = 1$.
- For every $\alpha \in (\mathbb{Z}/m\mathbb{Z})^*$, there exists a (unique) $x \in R$ such that $[x]_m = \alpha$. That is, α has a unique representative in the set R.

According to the above definition, ρ induces a bijection

$$R \xrightarrow{\sim} (\mathbb{Z}/m\mathbb{Z})^*; \qquad x \longmapsto [x]_m\,. \tag{2.34.1}$$

Therefore, any reduced system of residues mod m has exactly $\phi(m)$ elements. ■

Proposition 2.35. *Let $m \in \mathbb{N}^*$ and $R \subseteq \mathbb{Z}$. The following statements are equivalent:*

1. *R is a reduced residue system modulo m.*

2. $|R| = \phi(m)$, *every* $x \in R$ *is relatively prime to* m *and the elements of* R *are pairwise incongruent mod* m.

3. $|R| = \phi(m)$, *and every* $\alpha \in (\mathbb{Z}/m\mathbb{Z})^*$ *has a representative in* R.

Proof. To see (1) implies (2) and (3), use the definition (which is equivalent to the bijection (2.34.1) essentially).

Conversely, assuming (2) or (3), think of establishing the bijection (2.34.1). We leave it to the reader to phrase the details of the proof in a proper manner. □

Example 2.36. It is clear that $\{1\,,\,3\,,\,5\,,\,7\}$ is a reduced residue system mod 8. The set $\{-3\,,\,-1\,,\,1\,,\,3\}$ is again a reduced residue system mod 8, since it has the same number of elements, all numbers in it are relatively prime to 8 and pairwise incongruent mod 8. ■

Theorem 2.37. *Suppose* $\{x_1, \cdots, x_r\}$ *is a reduced residue system modulo* m.

Then for any integer a *such that* $\gcd(a\,,\,m) = 1$, *the set* $\{ax_1,\, ax_2,\, \cdots,\, ax_r\}$ *is a reduced residue system modulo* m.

Proof. First of all, we have $r = \phi(m)$. To prove the theorem we use the criterion given in statement (2) of Prop. 1.35. So we only need to check

- that ax_i is relatively prime to m for each i,

  and

- that $ax_i \not\equiv ax_j \pmod{m}$ for distinct indices i, j.

The first assertion follows from the assumption

$$\gcd(a, m) = 1 = \gcd(x_i, m)\,.$$

(For x_i, we use the fact that every number in a reduced system of residues in relatively prime to the modulus.)

To prove the second assertion, notice that we can find $b \in \mathbb{Z}$ such that $ba = 1 \pmod{m}$ since $\gcd(a, m) = 1$. Thus

$$ax_i \equiv ax_j \pmod{m} \Longrightarrow (ba)x_i \equiv (b \cdot a)x_j \pmod{m} \Longrightarrow 1 \cdot x_i \equiv 1 \cdot x_j\,.$$

But as distinct elements of a reduced residue system, x_i and y_j are not congruent mod m, for $i \neq j$. So we are done. □

Remark 2.38. We can rephrase our proof of Thm. 2.37 in a more fancy language:

Saying $\{x_1, \cdots, x_r\}$ is a reduced residue system mod m means that $r = \phi(m)$ and

$$\{[x_1]_m, \cdots, [x_r]_m\} = (\mathbb{Z}/m\mathbb{Z})^*\,.$$

The assumption $\gcd(a, m) = 1$ tells us that $[a]_m \in (\mathbb{Z}/m\mathbb{Z})^*$. We know that $(\mathbb{Z}/m\mathbb{Z})^*$ is a multiplication group, so the product of any two elements in $(\mathbb{Z}/m\mathbb{Z})^*$ is again an element of $(\mathbb{Z}/m\mathbb{Z})^*$. Hence $[a]_m \cdot [x_i]_m = [ax_i]_m$ belongs to $(\mathbb{Z}/m\mathbb{Z})^*$ for each $i = 1, 2, \cdots, r = \phi(m)$. Moreover, in the multiplicative group $(\mathbb{Z}/m\mathbb{Z})^*$ the element $\alpha = [a]_m$ has an inverse, i.e., some $\beta \in (\mathbb{Z}/m\mathbb{Z})^*$ satisfying $\beta \cdot \alpha = [1]_m$. Thus,

$$\begin{aligned}
&\alpha \cdot [x_i]_m = \alpha \cdot [x_j]_m && \text{in } (\mathbb{Z}/m\mathbb{Z})^* \\
\Longrightarrow\ &\beta \cdot (\alpha \cdot [x_i]_m) = \beta \cdot (\alpha \cdot [x_j]_m) && \text{in } (\mathbb{Z}/m\mathbb{Z})^* \\
\Longrightarrow\ &(\beta \cdot \alpha) \cdot [x_i]_m = (\beta \cdot \alpha) \cdot [x_j]_m && \text{in } (\mathbb{Z}/m\mathbb{Z})^* \\
\Longrightarrow\ &[1]_m \cdot [x_i]_m = [1]_m \cdot [x_j]_m && \text{in } (\mathbb{Z}/m\mathbb{Z})^* \\
\Longrightarrow\ &[x_i]_m = [x_j]_m && \text{in } (\mathbb{Z}/m\mathbb{Z})^*
\end{aligned}$$

Since $[x_i]_m \neq [x_j]_m$ for $i \neq j$, we have

$$[ax_i]_m = \alpha \cdot [x_i]_m \neq \alpha \cdot [x_j]_m = [ax_j]_m \,.$$

We have thus shown that

$$(\mathbb{Z}/m\mathbb{Z})^* = \{[ax_1]_m, \cdots, [ax_r]_m\}$$

with $r = \phi(m) = \left|(\mathbb{Z}/m\mathbb{Z})^*\right|$. This means precisely that $\{ax_1, \cdots, ax_r\}$ is a reduced residue system modulo m. ■

2.2.4 Some special congruences

Theorem 2.39 (Euler)**.** *Let $m \in \mathbb{N}^*$. For every $a \in \mathbb{Z}$ with $\gcd(a, m) = 1$, one has $a^{\phi(m)} \equiv 1 \pmod{m}$.*

Proof. We provide a proof that uses the language of congruence classes. We denote by α the congruence class $[a]_m$. Then $\alpha^{\phi(m)} = \left[a^{\phi(m)}\right]_m$. Hence,

$$a^{\phi(m)} = 1 \pmod{m} \iff \alpha^{\phi(m)} = [1]_m \quad \text{in } \mathbb{Z}/m\mathbb{Z} \,.$$

To prove the latter equality, notice that the assumption $\gcd(a, m) = 1$ guarantees that $\alpha = [a]_m$ is invertible in the ring $(\mathbb{Z}/m\mathbb{Z})^*$, that is, α is a element of the multiplication group $(\mathbb{Z}/m\mathbb{Z})^*$.

Let $[x_1]_m, \cdots, [x_r]_m$ denote the elements of $(\mathbb{Z}/m\mathbb{Z})^*$, where $r = \phi(m) = \left|(\mathbb{Z}/m\mathbb{Z})^*\right|$. By Thm. 2.37, the collection

$$[ax_1]_m = \alpha \cdot [x_1]_m \,, \cdots, \, [ax_r] = \alpha \cdot [x_r]_m$$

is another representation of all the elements of $(\mathbb{Z}/m\mathbb{Z})^*$. So, if we put

$$\beta := \prod_{i=1}^{r} [x_i]_m \in \left(\mathbb{Z}/m\mathbb{Z}\right)^* , \tag{2.39.1}$$

which is the product of all the elements of the group$(\mathbb{Z}/m\mathbb{Z})^*$, then we have

$$\beta := \prod_{i=1}^{r} (\alpha \cdot [x_i]_m) \tag{2.39.2}$$

because the right hand side is also the product of all the elements of $(\mathbb{Z}/m\mathbb{Z})^*$. From (2.39.2) it follows that

$$\prod_{i=1}^{r} (\alpha \cdot [x_i]_m) = \alpha^r \cdot \prod_{i=1}^{r} [x_i]_m = \alpha^r \cdot \beta = \alpha^{\phi(m)} \cdot \beta \,.$$

This combined with (2.39.2) yields

$$[1]_m \cdot \beta = \beta = \alpha^{\phi(m)} \cdot \beta \quad \text{in } (\mathbb{Z}/m\mathbb{Z})^* \,. \tag{2.39.3}$$

Every element in a multiplication group has an inverse. So β has an inverse $\beta^{-1} \in (\mathbb{Z}/m\mathbb{Z})^*$. Multiplying both sides of (2.39.3) by β^{-1} yields immediately

$$[1]_m = \alpha^{\phi(m)} = \left[a^{\phi(m)}\right]_m$$

which is exactly what should be proved. □

Corollary 2.40 (Fermat's little theorem)**.** *Let p be a prime number and let $a \in \mathbb{Z}$ be such that* $\gcd(a, p) = 1$.

Then $a^{p-1} \equiv 1 \pmod{p}$. For every $x \in \mathbb{Z}$, we have $x^p \equiv x \pmod{p}$.

Proof. The first assertion follows from Thm. 2.39, since $\phi(p) = p - 1$ (cf. Example 2.33 (2)).

For the second assertion, if x is relatively prime to p, then $x^{p-1} \equiv 1 \pmod{p}$, Therefore

$$x^p = x^{p-1} \cdot x \equiv 1 \cdot x = x \pmod{p} \,.$$

If x is not relatively prime to p, then $p \,|\, x$ (since $\gcd(x, p) = p$ in this case). Thus $x \equiv 0 \pmod{p}$. It follows that

$$x^p \equiv 0^p = 0 \equiv x \pmod{p} \,.$$

This completes the proof. □

Remark 2.41. Another way to state Fermat's little theorem is as follows: For any prime p,

$$\begin{aligned} \alpha^{p-1} &= [1]_m \,, \qquad \forall\, \alpha \in (\mathbb{Z}/p\mathbb{Z})^* \\ \beta^p &= \beta \,, \qquad \forall\, \beta \in \mathbb{Z}/p\mathbb{Z} \,. \end{aligned}$$

■

(2.42) Application. Suppose $\gcd(a, m) = 1$. Then Euler's theorem implies that $a^{\phi(m)-1}$ is an inverse of a mod m, or in other words, $\left[a^{\phi(m)-1}\right]_m = {[a]_m}^{-1}$. This provides us a method for solving a linear congruence $ax \equiv b \pmod{m}$, since

$$[x]_m = {[a]_m}^{-1} \cdot [ax]_m = {[a]_m}^{-1} \cdot [b]_m = \left[a^{\phi(m)-1} \cdot b\right]_m$$

namely

$$x \equiv a^{\phi(m)-1} \cdot b \pmod{m}$$

For example, the congruence $3x \equiv 7 \pmod{10}$ has solution

$$x \equiv 3^{\phi(10)-1} \cdot 7 \pmod{10} \,.$$

It is easy to verify that

$$\{i \in [\![1, 10]\!] \mid \gcd(i, 10) = 1\} = \{1, 3, 7, 9\} \,,$$

so $\phi(10) = 4$. Hence the solution to $3x \equiv 7 \pmod{10}$ is

$$x \equiv 3^{4-1} \cdot 7 = 3^3 \cdot 7 = 27 \cdot 7 \equiv 7 \cdot 7 \equiv 9 \pmod{10}.$$

■

Remark 2.43. You may wonder whether the above method remains valid if $\gcd(a, m) = d > 1$. The thing is that we can first check whether $d \mid b$ or not. If not, we know that the congruence $ax \equiv b \pmod{m}$ has no solution. If yes, one can show easily that

$$ax \equiv b \pmod{m} \iff \frac{a}{d}x \equiv \frac{b}{d} \left(\mathrm{mod}\, \frac{m}{d}\right).$$

Now the congruence

$$\frac{a}{d}x \equiv \frac{b}{d} \left(\mathrm{mod}\, \frac{m}{d}\right).$$

satisfies the hypothesis $\gcd\left(\frac{a}{d}, \frac{m}{d}\right) = 1$, so we can then apply the method of (2.42). For example, the congruence

$$9x \equiv 21 \pmod{30}$$

is equivalent to

$$3x \equiv 7 \pmod{10}.$$

So its solution is $x \equiv 9 \pmod{10}$ by (2.42). Notice however that if one wants to express the solutions to $9x \equiv 21 \pmod{30}$ in terms of their residues modulo 30 (instead of modulo 10), then there are three possibilities in total: $x \equiv 9$, 19 or 29 $\pmod{30}$. ■

(2.44) Here is another example of how Euler's theorem can be used.

We want to find the last digit of the integer 3^{1000} in the decimal notation. This is the same as finding the integer $a \in [\![0, 9]\!]$ such that $\left[3^{1000}\right]_{10} = [a]_{10}$ in $\mathbb{Z}/10\mathbb{Z}$. So we want to complete α^{1000} in $\mathbb{Z}/10\mathbb{Z}$ for $\alpha = [3]_{10}$. By Euler's theorem, we have

$$\alpha^4 = \alpha^{\phi(10)} = [1]_{10}.$$

Thus

$$\alpha^{1000} = (\alpha^4)^{250} = ([1]_{10})^{250} = [1]_{10}.$$

This means that $\left[3^{1000}\right]_{10} = [1]_{10}$, i,e., the last digit of 3^{1000} is 1. ■

(2.45) Let us consider the same question for 2^{1000}. The difference with the previous example is that 2 is not relatively prime to 10. So we cannot conclude $2^{\phi(10)} \equiv 1 \pmod{10}$, which is obviously wrong.

However, the key idea is similar: we can use phenomena of periodicity ! We first compute small powers of 2 and we find that

$$2^5 \equiv 2 \pmod{10}.$$

Thus

$$\begin{aligned}\left[2^{1000}\right]_{10} &= \left[2^5\right]_{10}{}^{200} = [2]_{10}{}^{200} = \left[2^5\right]_{10}{}^{40} \\ &= [2]_{10}{}^{40} = \left[2^5\right]_{10}{}^{8} = [2]_{10}{}^{8} \\ &= \left[2^5\right]_{10}\left[2^3\right]_{10} = [2]_{10} \cdot \left[2^3\right]_{10} = [16]_{10} = [6]_{10}.\end{aligned}$$

Hence the last digit of 2^{1000} is 6.

There is an alternative, perhaps simpler approach which relies on the Chinese remainder theorem. In fact,

$$2^{1000} \equiv a \pmod{10} \iff \begin{cases} 2^{100} \equiv a \pmod{2} \\ 2^{100} \equiv a \pmod{5} \end{cases}$$

Obviously, $2^{1000} = 0^{1000} = 0 \pmod 2$ and since $2^{\phi(5)} = 2^4 \equiv 1 \pmod 5$, we have

$$\left[2^{1000}\right]_5 = \left[2^4\right]_5^{250} = [1]_5^{250} = [1]_5$$

i.e., $2^{1000} \equiv 1 \pmod 5$. So we need only to find the integer $a \in [\![0, 9]\!]$ such that

$$\begin{cases} a \equiv 2^{1000} \equiv 0 \pmod{2} \\ a \equiv 2^{1000} \equiv 1 \pmod{5} \end{cases}$$

It is immediately seen that $a = 6$. ■

The following result, known as Wilson's theorem, is another very useful congruence relation.

Theorem 2.46 (Wilson)**.** *If p is a prime number, then $(p-1)! \equiv -1 (\mathrm{mod}\ p)$.*

Proof. The result is evident if $p = 2$. So we may assume p is an odd prime. Since (by Corollary 2.30) $(\mathbb{Z}/p\mathbb{Z})^* = \{[i]_p \mid i = 1, \ldots, p-1\}$, what we want to show is equivalent to the following equality

$$\prod_{\alpha \in (\mathbb{Z}/p\mathbb{Z})^*} \alpha = [-1]_p$$

in the group $(\mathbb{Z}/p\mathbb{Z})^*$ (or equivalently, in the ring $\mathbb{Z}/p\mathbb{Z}$).

Now observe that for any $\alpha \in (\mathbb{Z}/p\mathbb{Z})^*$, its inverse is not equal to itself unless $\alpha = [1]_p$ or $\alpha = [p-1]_p$ (cf. Prop. 2.17). The set

$$S := (\mathbb{Z}/p\mathbb{Z})^* \setminus \{[1]_p, [p-1]_p\} = \{[2]_p, \ldots, [p-2]_p\}\,.$$

has an even number of elements. In fact, $|S| = p - 3$. By what we have just said, for any $\alpha \in S$, we have $\alpha^{-1} \in S$ and α, α^{-1} are two different elements of S. Moreover, if α, $\beta \in S$ and $\beta \notin \{\alpha, \alpha^{-1}\}$, then $\{\alpha, \alpha^{-1}\}$ and $\{\beta, \beta^{-1}\}$ are disjoint sets. Therefore, the set S can be partitioned into $\frac{p-3}{2}$ subsets, each of which consists precisely of two mutually inverse elements. So the product of all the elements in S is $[1]_p$. It follows that

$$\prod_{\alpha \in (\mathbb{Z}/p\mathbb{Z})^*} \alpha = [1]_p \cdot [p-1]_p \cdot \prod_{\alpha \in S} \alpha = [1]_p \cdot [p-1]_p \cdot [1]_p = [-1]_p$$

proving the desired result. □

2.2.5 Ring theoretic interpretation of the Chinese remainder theorem

Our aim now is to give a ring theoretic interpretation of the Chinese Remainder theorem.

(2.47) First recall that (cf. (2.19)) the Chinese remainder theorem (CRT) asserts the following:

If $m, n \in \mathbb{N}^*$ are relatively prime, then for any $a, b \in \mathbb{Z}$, the system

$$\begin{cases} x \equiv a \pmod{m} \\ x \equiv b \pmod{n} \end{cases} \tag{2.47.1}$$

has integer solutions and its solutions are all congruent mod mn.

Using the language of congruence classes, the system may be rewritten as

$$\begin{cases} [x]_m = [a]_m \\ [x]_n = [b]_n \end{cases} \tag{2.47.2}$$

So, the CRT may be stated in the following way:

Suppose $m, n \in \mathbb{N}^*$ are relatively prime. Then, for any $\alpha \in \mathbb{Z}/m\mathbb{Z}$ and any $\beta \in \mathbb{Z}/n\mathbb{Z}$, there exist integers $x \in \mathbb{Z}$ such that

$$[x]_m = \alpha \qquad \text{and} \qquad [x]_n = \beta\,.$$

Moreover, all these integers lie in the same congruence class modulo mn. This means that the element $[x]_{mn} \in \mathbb{Z}/m\mathbb{Z}$ is uniquely determined by the system (2.47.2).

The above discussions suggest that the CRT can be thought of as some kind of bijective correspondence between the sets $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ and $\mathbb{Z}/mn\mathbb{Z}$. ■

Now we want to describe explicitly a bijective map between $\mathbb{Z}/mn\mathbb{Z}$ and $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$.

(2.48) Let $M, m \in \mathbb{N}^*$ be such that $m \,|\, M$. We define a map

$$\theta \;:\; \mathbb{Z}/M\mathbb{Z} \longrightarrow \mathbb{Z}/m\mathbb{Z}\,, \qquad [x]_M \longmapsto [x]_m\,. \tag{2.48.1}$$

Indeed, if $[x]_M = [y]_M$, then we have $x \equiv y \pmod{M}$, which means that $M \,|\, (x-y)$. Since $m \,|\, M$, we have $m|(x-y)$, i.e., $x \equiv y \pmod{m}$, or equivalently, $[x]_m = [y]_m$. This shows that the map θ in (2.48.1) is well defined. That is, starting from an element $\alpha \in \mathbb{Z}/m\mathbb{Z}$, we can choose any representative $x \in \mathbb{Z}$ of α. We let $\theta(\alpha)$ be the congruence class of the integer x, but this time modulo m, so that the result $[x]_m$ is an element of $\mathbb{Z}/m\mathbb{Z}$. What we have done above is the proof of the fact that $\theta(\alpha)$ depends only on the congruence class $\alpha \in \mathbb{Z}/M\mathbb{Z}$, but not on the choice of the representative $x \in \mathbb{Z}$.

Let us look at a concrete example: Let $M = 6$ and $m = 3$. Then the map

$$\theta \;:\; \mathbb{Z}/6\mathbb{Z} \longrightarrow \mathbb{Z}/3\mathbb{Z}\,; \quad [x]_6 \longmapsto [x]_3 \tag{2.48.2}$$

satisfies

$$\begin{aligned} \theta([0]_6) = \theta([3]_6) = [0]_3 = [3]_3 \in \mathbb{Z}/3\mathbb{Z}\,, \\ \theta([1]_6) = \theta([4]_6) = [1]_3 = [4]_3 \in \mathbb{Z}/3\mathbb{Z}\,, \\ \theta([2]_6) = \theta([5]_6) = [2]_3 = [5]_3 \in \mathbb{Z}/3\mathbb{Z}\,. \end{aligned}$$

Notice that the map θ in (2.48.2) is not injective, because two integers in the same congruence class mod 3 (e.g. 0 and 3) need not be in the same congruence class when the modulus is 6.

However, when two integers belong to the same congruence class modulo 6, they must be in the same congruence class modulo 3. This is why the map θ in (2.48.2) is well defined. ■

(2.49) Now let us consider two positive integers $m, n \in \mathbb{N}^*$. Put $M = mn$. Then by (2.48), there are

well defined maps

$$\theta_1 : \qquad \mathbb{Z}/mn\mathbb{Z} = \mathbb{Z}/M\mathbb{Z} \longrightarrow \mathbb{Z}/m\mathbb{Z}\,; \quad [x]_{mn} \longmapsto [x]_m$$

and

$$\theta_2 : \qquad \mathbb{Z}/mn\mathbb{Z} = \mathbb{Z}/M\mathbb{Z} \longrightarrow \mathbb{Z}/n\mathbb{Z}\,; \quad [x]_{mn} \longmapsto [x]_n\,.$$

Gathering them together we obtain a map

$$\theta = (\theta_1, \theta_2) : \qquad \begin{array}{c} \mathbb{Z}/mn\mathbb{Z} \longrightarrow \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \\ [x]_{mn} \longmapsto ([x]_m\,, [x]_n) \end{array} \tag{2.49.1}$$

■

The following is a restatement of the CRT:

Theorem 2.50 (Chinese Remainder Theorem)**.** *Suppose $m, n \in \mathbb{N}^*$ are relatively prime. Then the map θ in (2.49.1) is bijective. In other words, given any $\alpha \in \mathbb{Z}/m\mathbb{Z}$ and $\beta \in \mathbb{Z}/n\mathbb{Z}$, there exists a unique $\gamma \in \mathbb{Z}/mn\mathbb{Z}$ such that*

$$\theta_1(\gamma) = \alpha \quad \text{and} \quad \theta_2(\gamma) = \beta$$

(If $x \in \mathbb{Z}$ is a representative of γ, i.e., $[x]_{mn} = \gamma$, then the above condition means that $[x]_m = \alpha$ and $[x]_m = \beta$.)

Example 2.51. Let $\alpha = [1]_3 \in \mathbb{Z}/3\mathbb{Z}$ and $\beta = [2]_4 \in \mathbb{Z}/4\mathbb{Z}$. Then for the map

$$\theta : \mathbb{Z}/12\mathbb{Z} \longrightarrow \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$$

we have

$$\theta([10]_{12}) = (\alpha, \beta) \in \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$$

since $[10]_3 = [1]_3$ and $[10]_4 = [2]_4$. ■

(2.52) If m and n are not relatively prime, then the map α in (2.49.1) is neither injective nor surjective. For example, if $m = 4$ and $n = 6$, we get the map

$$\theta : \qquad \begin{array}{c} \mathbb{Z}/24\mathbb{Z} \longrightarrow \mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/6\mathbb{Z} \\ [x]_{24} \longmapsto ([x]_4\,, [x]_6) \end{array}$$

It is not hard to see that

$$\theta([0]_{24}) = ([0]_4, [0]_6) = ([12]_4, [12]_6) = \theta([12]_{24})$$

but $[0]_{24} \neq [12]_{24}$. So this θ is not injective.

If we take $\alpha = [1]_4$ and $\beta = [2]_6$, then the element $(\alpha, \beta) \in \mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/6\mathbb{Z}$ is not in the image of the map θ. (Can you explain why?) ■

(2.53) Let us assume $\gcd(m, n) = 1$, so that the map

$$\theta : \mathbb{Z}/mn\mathbb{Z} \longrightarrow \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$$

is bijective. If we take an element $[x]_{mn}$ in $(\mathbb{Z}/mn\mathbb{Z})^*$, which means that $\gcd(x, mn) = 1$, then we have

$$\theta([x]_{mn}) = ([x]_m, [x]_n) \in (\mathbb{Z}/m\mathbb{Z})^* \times (\mathbb{Z}/n\mathbb{Z})^*$$

because

$$\gcd(x, mn) = 1 \Longrightarrow \begin{cases} \gcd(x, m) = 1 \\ \gcd(x, n) = 1 \end{cases}$$

Conversely, suppose $\alpha \in (\mathbb{Z}/m\mathbb{Z})^*$ and $\beta \in (\mathbb{Z}/n\mathbb{Z})^*$. Let $\gamma \in \mathbb{Z}/mn\mathbb{Z}$ be the unique element such that $\theta(\gamma) = (\alpha, \beta)$. Let $x \in \mathbb{Z}$ be a representative of θ, i.e., $[x]_{mn} = \theta$. Then $\theta([x]_{mn}) = \theta(\gamma) = (\alpha, \beta)$. This tells us that $\alpha = [x]_m$, $\beta = [x]_n$. Now the hypothesis $\alpha \in (\mathbb{Z}/m\mathbb{Z})^*$ means that $\gcd(x, m) = 1$. Similarly, $\gcd(x, n) = 1$. Now

$$\left.\begin{matrix} \gcd(x, m) = 1 \\ \gcd(x, n) = 1 \end{matrix}\right\} \Longrightarrow \gcd(x, mn) = 1$$

(Can you prove this ?) Therefore $\gamma = [x]_{mn}$ belongs to $(\mathbb{Z}/mn\mathbb{Z})^*$. All the above shows that for any $\gamma \in \mathbb{Z}/mn\mathbb{Z}$,

$$\gamma \in (\mathbb{Z}/mn\mathbb{Z})^* \iff \theta(\gamma) \in (\mathbb{Z}/m\mathbb{Z})^* \times (\mathbb{Z}/n\mathbb{Z})^* \,.$$

Therefore, we can deduce a bijection

$$\theta' \,:\quad (\mathbb{Z}/mn\mathbb{Z})^* \xrightarrow{\sim} (\mathbb{Z}/m\mathbb{Z})^* \times (\mathbb{Z}/n\mathbb{Z})^* \,;\quad [x]_{mn} \longmapsto ([x]_m, [x]_n) \tag{2.53.1}$$

from the bijection $\theta \,:\, \mathbb{Z}/mn\mathbb{Z} \xrightarrow{\sim} \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$. ■

Corollary 2.54. *Euler's ϕ-function is multiplicative. That is, if $m, n \in \mathbb{N}^*$ are relatively prime, then*

$$\phi(mn) = \phi(m) \cdot \phi(n) \,.$$

Proof. Use the bijection (2.53.1). □

Lemma 2.55. *Let p be a prime number and $r \in \mathbb{N}^*$. Then*

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

Proof.

$$\begin{aligned} \phi(p^r) &= \left|\{a \in [\![1,\, p^r]\!] \mid \gcd(a, p^r) = 1\}\right| \\ &= p^r - \left|\{a \in [\![1, p^r]\!] \mid \gcd(a, p^r) \neq 1\}\right| \\ &= p^r - \left|\{a \in [\![1, p^r]\!] \mid \ a \text{ and } p^r \text{ have a common prime divisor}\}\right| \\ &= p^r - \left|\{a \in [\![1, p^r]\!] \ \mid \ p \text{ divides } a\}\right| \\ &= p^r - \text{the number of multiples of } p \text{ in } \ [\![1, p^r]\!] \\ &= p^r - \left[\frac{p^r}{p}\right] = p^r - p^{r-1} \,. \end{aligned}$$

□

Corollary 2.56. *Let $n = {p_1}^{e_1} \cdots {p_r}^{e_r}$ be the prime factorization of a natural number $n > 1$. (Here $e_i \in \mathbb{N}^*$ and p_i are distinct primes.)*

Then

$$\phi(n) = \prod_{i=1}^{r} p_i^{e_i - 1}(p_i - 1)$$

and

$$\frac{\phi(n)}{n} = \prod_{i=1}^{r} \left(1 - \frac{1}{p_i}\right) .$$

2.3 Some applications and complements

2.3.1 Divisibility tests

Theorem 2.57. *Let b be a positive integer ≥ 2 and let $n = (a_k \cdot \cdots \cdot a_1 a_0)_b$ be the base b expansion of a positive integer n. Let $d \in \mathbb{N}^*$.*

1. *Suppose $d \,|\, b$. Then for every $j \in [\![0, k]\!]$,*

$$d^j \mid n \iff d^j \,\Big|\, (a_{j-1} \cdots a_1 a_0)_b .$$

2. *Suppose $d \,|\, (b-1)$. Then*

$$d \mid n \iff d \,\Big|\, \sum_{i=0}^{k} a_i .$$

3. *Suppose $d \,|\, (b+1)$. Then*

$$d \mid n \iff d \,\Big|\, \sum_{i=0}^{k} (-1)^i a_i .$$

Proof. (1) Since $d \,|\, b$, we have $d^j \,|\, b^j$. Hence $b^j \equiv b^{j+1} \cdots \equiv b^k \equiv 0 \pmod{d^j}$. Thus

$$\begin{aligned} n = \sum_{i=0}^{k} a_i b^i &= \sum_{j \leq i \leq k} a_i b^i + (a_{j-1} a_{j-2} \cdots a_1 a_0)_b \\ &\equiv \sum_{j \leq i \leq k} a_i \cdot 0 + (a_{j-1} a_{j-2} \cdots a_1 a_0)_b \\ &\equiv (a_{j-1} a_{j-2} \cdots a_1 a_0)_b \pmod{d^j} \end{aligned}$$

This shows that

$$n \equiv 0 \pmod{d^j} \iff (a_{j-1} \cdots a_1 a_0)_b \equiv 0 \pmod{d^j} .$$

The proofs of (2) and (3) use similar ideas. We leave it to the reader to finish the proof. □

Example 2.58. Let $b = 10$ and $n = (a_k \cdots a_1 a_0)_{10}$.

(1) For $d = 2$ or 5,

$$d^j \mid n \iff d^j \mid (a_{j-1} \cdots a_1 a_0)_{10} .$$

For example,

$$8 \mid n \iff 8 \mid (a_2 a_1 a_0)_{10} \quad \text{and} \quad 25 \mid n \iff 25 \mid (a_1 a_0)_{10} .$$

(2) For $d = 3$ or 9, we have

$$d \,|\, n \iff d \mid (a_k + \cdots + a_1 + a_0) .$$

(3) For $d = 11$, we have

$$d \,|\, n \iff d \,|\, a_0 - a_1 + a_2 - a_3 + \cdots + (-1)^k a_k \ .$$

The above criteria provide quick methods of testing divisibility by numbers like 9, 11 etc. in the decimal notation. ■

(2.59) Using the fact $1001 = 7 \times 11 \times 13$ we can deduce the following divisibility criterion for any divisor of 1001.

Let $n = (a_k \cdots a_1 a_0)_{10}$. Then the base 1000 expansion of n will be

$$n = b_0 + b_1 \cdot 1000 + b_2 \cdot 1000 + \cdots$$

where

$$b_0 = a_0 + a_1 \cdot 10 + a_2 \cdot 10^2 = (a_2 a_1 a_0)_{10}$$

$$b_1 = a_3 + a_4 \cdot 10 + a_5 \cdot 10^2 = (a_5 a_4 a_3)_{10}$$

$$\cdots\cdots$$

Therefore

$$\begin{aligned} n &\equiv b_0 - b_1 + b_2 \cdots \qquad (\text{mod } 1001) \\ &\equiv (a_2 a_1 a_0)_{10} - (a_5 a_4 a_3)_{10} + \cdots \qquad (\text{mod } 1001) \end{aligned}$$

So, for any divisor d of 1001 (e.g. $d = 7, 13, 77, 91$, etc.), to check the divisibility of n by d we can proceed as follows:

staring from the rightmost digit in the decimal expansion of n, we form 3-digit integers

$$b_0 = (a_2 a_1 a_0)_{10}, \ \ b_1 = (a_5 a_4 a_3)\,, \ \cdots$$

by using successive strings of 3 digits in the expansion of n. Then n is divisible by d if and only if the alternating sum

$$b_0 - b_1 + b_2 \cdots$$

is divisible by d. For example, let $n = 59358208$. The integers formed from successive blocks of 3 digits of n are

$$b_0 = 208, \quad b_1 = 358 \quad \text{and} \quad b_2 = 59\,.$$

The alternating sum $b_0 - b_1 + b_2$ is divisible by 7 and 13, but not by 11. Hence

$$7 \,|\, n, \ \ 13 \,|\, n, \ \ \text{but} \ \ 11 \nmid n\,.$$

■

2.3.2 Round-Robin tournaments

We skip this section in this year's course.

2.3.3 Pseudo primes

We skip this section in this year's course.

Chapter 3

Primitive Roots and Applications

3.1 Order of integers in modular arithmetic

In this section we fix a positive integer $n > 1$.

3.1.1 The order of an integer residue class

(3.1) Let $a \in \mathbb{Z}$ be relatively prime to n. Then by Euler's theorem, $a^{\phi(n)} \equiv 1 \pmod{n}$. Therefore the set

$$A := \{r \in \mathbb{N}^* \mid a^r = 1 \pmod{n}\}$$

is a nonempty subset of $\mathbb{N}$. By the well ordering property, A has a smallest element. It is called the ***order of*** a ***mudulo*** n, and denoted by $\operatorname{ord}_n a$ or $\operatorname{ord}[a]_n$.

Notice that the integer $\operatorname{ord}_n a \in \mathbb{N}^*$ depends only on the congruence class, i.e., if $b \in \mathbb{Z}$ is another integer relatively prime to n such that $a \equiv b \pmod{n}$ then $\operatorname{ord}_n a = \operatorname{ord}_n b$.

We may thus define an order function

$$\operatorname{ord} \ : \ (\mathbb{Z}/n\mathbb{Z})^* \longrightarrow \mathbb{N} \, ; \quad \alpha = [a]_n \longmapsto \operatorname{ord}_n(\alpha) := \operatorname{ord}_n [a]_n \ .$$

So $\operatorname{ord}(\alpha)$ will denote the order of any integer in the congruence class α, whenever $\alpha \in (\mathbb{Z}/n\mathbb{Z})^*$.

Using that $(\mathbb{Z}/n\mathbb{Z})^*$ is a multiplication group with identify element $1 := [1]_n$, we may rephrase the definition of the order function as follows:

(3.1.1) $$\operatorname{ord}(\alpha) := \text{ the smallest positive integer } r \text{ such that } \alpha^r = 1 \text{ in } (\mathbb{Z}/n\mathbb{Z})^* \, .$$

Clearly, if $k \in \mathbb{N}^*$ is a multiple of $\operatorname{ord}(\alpha)$, then

$$\alpha^k = \left(\alpha^{\operatorname{ord}(\alpha)}\right)^{\frac{k}{\operatorname{ord}(\alpha)}} = 1^{\frac{k}{\operatorname{ord}(\alpha)}} = 1 \ \in \ (\mathbb{Z}/n\mathbb{Z})^* \, .$$

The next proposition shows that the converse is also true. ■

Proposition 3.2. *Let $\alpha \in (\mathbb{Z}/n\mathbb{Z})^*$. Then for every $k \in \mathbb{N}^*$, one has*

$$\alpha^k = 1 \iff \operatorname{ord}(\alpha) \mid k \, .$$

In particular, $\operatorname{ord}(\alpha) \mid \phi(n)$ *for all* $\alpha \in (\mathbb{Z}/n\mathbb{Z})^*$.

Proof. Use the division algorithm to write

$$k = q \cdot \operatorname{ord}(\alpha) + r \quad \text{with } r\,,\ q \in \mathbb{N} \text{ and } 0 < r \leq \operatorname{ord}(\alpha)\,.$$

Then

$$\alpha^k = \alpha^{q\cdot\operatorname{ord}(\alpha)+r} = \left(\alpha^{\operatorname{ord}(\alpha)}\right)^q \cdot \alpha^r = \alpha^r$$

Hence $\alpha^k = 1$ if and only if $\alpha^r = 1$. Since $r < \operatorname{ord}(\alpha)$, if follows from (3.1.1) that

$$\alpha^r = 1 \iff r = 0 \iff \operatorname{ord}(\alpha) \mid k\,.$$

The first assertion is thus proved. The second assertion is a consequence of the first one, because $\alpha^k = 1$ holds for $k = \phi(n)$, by Euler's theorem. □

If one prefers to the language of integers, Prop. 3.2 may be restated as the following:

Proposition 3.3. *Let* $\alpha \in \mathbb{Z}$ *be relatively prime to* n*. Then for every* $k \in \mathbb{N}^*$,

$$\alpha^k \equiv 1 \pmod{n} \iff \operatorname{ord}_n(\alpha) \mid k\,.$$

In particular, $\operatorname{ord}_n a \mid \phi(n)$.

Example 3.4. To compute $\operatorname{ord}_{17} 5$, we note that $\phi(17) = 16$. As a divisor of $\phi(17)$, the only positive values of $\operatorname{ord}_{17} 5$ are $1, 2, 4, 8$ and 16. We compute the powers $5^2, 5^4, 5^8, 5^{16}$ to find

$$\begin{aligned} 5^2 = 25 \equiv 8 \pmod{17},\quad 5^4 \equiv 8^2 = 64 \equiv 13 \equiv -4 \pmod{17} \\ 5^8 \equiv (-4)^2 = 16 \equiv -1 \pmod{17} \\ 5^{16} \equiv (-1)^2 = 1 \pmod{17} \end{aligned}$$

So we have $\operatorname{ord}_{17}\ 5 = 16 = \phi(17)$. ■

Definition 3.5. If $a \in \mathbb{Z}$ satisfies $\gcd(a, n) = 1$ and $\operatorname{ord}_n a = \phi(n)$, we say that a is a ***primitive root*** 原根 ***modulo*** n, or ***of*** n. When such an integer exists, we say that n *has a primitive root.* ■

Theorem 3.6. *Let* $a \in \mathbb{Z}$ *be a primitive root of* n*. Then the set*

$$\left\{a, a^2, a^3, \cdots, a^{\phi(n)}\right\}$$

is a reduced residue system modulo n*. Equivalently,*

$$(\mathbb{Z}/n\mathbb{Z})^* = \{[a]_n^i \mid i = 1, 2, \cdots, \phi(n)\}$$

In the language of abstract algebra, the above equality may expressed as saying that the multiplication group $(\mathbb{Z}/n\mathbb{Z})^*$ *is a* **cyclic group** 循环群*, with a* **generator** 生成元 *given by* $\alpha = [a]_n$.

The proof of the above theorem will use the following lemma, which is of interest on its own.

Lemma 3.7. *Let* $a \in \mathbb{Z}$ *be relatively prime to* n*. for any* $i, j \in \mathbb{N}$*, one has*

$$a^i \equiv a^j \pmod{n} \iff i \equiv j \pmod{\operatorname{ord}_n a}\,.$$

Proof. Without loss of generality, we may assume $i \geq j$. Let $\alpha = [a]_n \in (\mathbb{Z}/n\mathbb{Z})^*$. Then

$$\begin{aligned} a^i \equiv a^j \pmod{n} &\iff \alpha^i = \alpha^j \text{ in } (\mathbb{Z}/n\mathbb{Z})^* \\ &\iff \alpha^i \cdot (\alpha^{-1})^j = \alpha^j \cdot (\alpha^{-1})^j \iff \alpha^{i-j} = 1 \text{ in } (\mathbb{Z}/n\mathbb{Z})^* \\ \text{by (3.2)} &\iff \operatorname{ord}_n a = \operatorname{ord}(\alpha) \mid (i-j) \iff i \equiv j \pmod{\operatorname{ord}_n a} \end{aligned}$$

where $\alpha^{-1} \in (\mathbb{Z}/n\mathbb{Z})^*$ is the inverse of $\alpha \in (\mathbb{Z}/n\mathbb{Z})^*$. This prove the lemma. □

Proof of Thm. 3.6. We want to show that if a is a primitive root of n, then the set

$$R := \{a, a^2, a^3, \cdots a^{\phi(n)}\}$$

is a reduced residue system modulo n. Since $\gcd(a^i, n) = 1$ for every $i \in [\![1, n]\!]$ and the number of elements in R is correct (i.e., equal to $\phi(n)$), it suffices to show that (cf. Prop. 2.35)

$$a^i \not\equiv a^j \pmod{n} \text{ for distinct } i\,,\ j \in [\![1\,,\ \phi(n)]\!]\ .$$

But this is immediate from Lemma 3.7, because $\operatorname{ord}_n a = \phi(n)$ (by the assumption that a is a primitive root) and distinct numbers in $[\![1, \phi(n)]\!]$ are incongruent modulo $\phi(n)$. □

When n possesses a primitive root, it usually has more than one primitive roots in a reduced system of residues. This will be clear from the following result:

Proposition 3.8. *Let $m = \operatorname{ord}_n a$, where $a \in \mathbb{Z}$ is relatively prime to n. Then for every $r \in \mathbb{N}$ one has*

$$\operatorname{ord}_n(a^r) = \frac{m}{\gcd(r, m)} = \frac{\operatorname{ord}_n(a)}{\gcd(r, \operatorname{ord}_n a)}\ .$$

Proof. Write $d = \gcd(r, m) = \gcd(r, \operatorname{ord}_n a)$. For any $k \in \mathbb{N}$, we have

$$\begin{aligned} \iff (a^r)^k \equiv 1 \pmod{n} &\iff a^{rk} \equiv 1 \pmod{n} \\ &\iff \operatorname{ord}_n a = m \mid rk \iff \frac{m}{d} \,\Big|\, \frac{r}{d} \cdot k \\ \text{since } \gcd\left(\frac{m}{d}\,,\ \frac{r}{d}\right) = 1 &\iff \frac{m}{d} \,\Big|\, k\,. \end{aligned}$$

It thus follows that

$$\operatorname{ord}_n(a^r) = \min\{k \in \mathbb{N}^* \mid (a^r)^k \equiv 1 \pmod{n}\} = \min\left\{k \in \mathbb{N}^* \,:\, \frac{m}{d} \mid k\right\}$$

whence the equality $\operatorname{ord}_n(a^r) = \frac{m}{d}$. □

Corollary 3.9. *Suppose $a \in \mathbb{Z}$ is a primitive root of n. Then for every $r \in \mathbb{N}$, the following statements are equivalent:*

1. *a^r is a primitive root of n;*
2. *$\gcd(r, \phi(n)) = 1$.*

Proof. By assumption, $\operatorname{ord}_n a = \phi(n)$. Thus, from Prop. 3.8 we see that

$$\operatorname{ord}_n(a^r) = \frac{\phi(n)}{\gcd(r, \phi(n))}\ .$$

This number is equal to $\phi(n)$ precisely when $\gcd(r, \phi(n)) = 1$. □

Corollary 3.10. *If n has a primitive root, then it has exactly $\phi(\phi(n))$ incongruent primitive roots, that is, any complete system of residues modulo n contains precisely $\phi(\phi(n))$ primitive roots of n.*

Proof. Since primitive roots are required to be relatively prime to n, we can restrict to a reduced residue system. The orders of integers mod n only depend on their congruence classes mod n. So in any reduced residue system we will find the same number of primitive roots.

By Thm. 3.6, we can choose any primitive root a and consider the reduced residue system

$$R = \{a^i \mid 1 \leq i \leq \phi(n)\}.$$

Then, Corollary 3.9 shows that the number of primitive roots n in R equals the number of integers $i \in [\![1, \phi(n)]\!]$ such that $\gcd(i, \phi(n)) = 1$. The last number is equal to $\phi(\phi(n))$, according to the definition of Euler's ϕ-function. □

Remark 3.11. In group theoretic terminology, Corollary 3.10 says that when the multiplication group $(\mathbb{Z}/n\mathbb{Z})^*$ is cyclic, it has a total number of $\phi(\phi(n))$ generators. The reader will see in his *Abstract algebra* course that, for any cyclic group with N elements, the number of generators is equal to $\phi(N)$. The essential idea of the proof is actually the same as in the situation we discussed here. ■

3.1.2 Primitive roots for primes

Our objective now is to show that every prime has a primitive root. To do this, we first need to study polynomial congruence.

(3.12) Let$f(x)$ be a polynomial with integer coefficient of degree $m \geq 0$, i.e.,

$$f(x) = a_m x^m + a_{m-1} x^{m-1} + \cdots + a_1 x + a_0$$

with each $a_i \in \mathbb{Z}$ and $a_m \neq 0$. An integer c is called a ***root of*** f ***modulo*** n (where $n > 1$ is a fixed positive integer) if $f(c) \equiv 0 \pmod{n}$. Clearly, if c is a root of f modulo n, then every integer congruent to c modulo n is also a root of f modulo n.

We will say that f ***has*** N ***incongruent roots modulo*** n if any complete system of residues modulo n contains N roots of f modulo n. For example,

$$f(x) = x^2 + x + 1 \text{ has exactly 2 incongruent roots modulo 7, namely, } x \equiv 2 \text{ or } 4 \pmod{7}$$
$$f(x) = x^2 + 2 \text{ has no roots modulo 5}.$$

If p is a prime number, $f(x) = x^{p-1} - 1$ has exactly $p - 1$ incongruent roots modulo p, by Fermat's little theorem. ■

Theorem 3.13 (Lagrange)**.** *Let*

$$f(x) = a_m x^m + a_{m-1} x^{m-1} + \cdots + a_1 x + a_0$$

be a polynomial with integer coefficients of degree $m \geq 1$. Let p be a prime such that $p \nmid a_m$.

Then $f(x)$ has at most m incongruent roots modulo p.

Proof. We induct on the degree m of the polynomial f. If $m = 1$, we have $f(x) = a_1x + a_0$ with $p \nmid a_1$ by assumption. Thus

$$f(x) \equiv 0 \pmod{p} \iff a_1x \equiv -a_0 \pmod{p}.$$

The number of incongruent solutions to the last linear congruence is $1 = \gcd(a_1\,,\, p)$.

Now suppose $m > 1$ and assume that the theorem is true for polynomials of degree $m - 1$ with leading coefficient relatively prime to p.

Let $f(x)$ be a polynomial of degree m as in the statement of the theorem. Assume that f has $m + 1$ incongruent roots mod p, say, c_0, c_1, $\cdots$, c_m. Then we have

$$f(x) - f(c_0) = a_m x^m + \cdots + a_1x + a_0 - (a_m {c_0}^m + \cdots + a_1c_0 + a_0) = \sum_{i=1}^{m} a_i(x^i - {c_0}^i)\,.$$

Note that for each $i \geq 1$,

$$\begin{aligned} x^i - {c_0}^i &= (x - c_0)(x^{i-1} + x^{i-2}c_0 + x^{i-3}{c_0}^2 + \cdots + x{c_0}^{i-2} + {c_0}^{i-1}) \\ &= (x - c_0)h_i(x), \end{aligned}$$

where $h_i(x)$ is a polynomial with integer coefficients of degree $i - 1$ and leading coefficient 1. Thus

$$f(x) - f(c_0) = \sum_{i=1}^{m} a_i(x^i - {c_0}^i) = (x - c_0)g(x)$$

where

$$g(x) = \sum_{i=1}^{m} a_i h_i(x)$$

has degree $m - 1$ and leading coefficient a_m, which is not divisible by p.

We now show that $c_1, c_2, \cdots, c_m$ are all roots of g mod p, so that this heads to a contradiction to the theorem in the degree $m - 1$ case. Indeed, since $f(c_k) \equiv f(c_0) \equiv 0 \pmod{p}$, for all $k \in [\![1, m]\!]$, we have

$$(c_k - c_0)g(c_k) \equiv 0 \pmod{p}\,, \quad \forall\, k \in [\![1, m]\!]\,.$$

By assumption, $c_k \not\equiv c_0 \pmod{p}$, for $k \geq 1$. So $c_k - c_0$ has an inverse mod p. Therefore

$$(c_k - c_0)g(c_k) \equiv 0 \pmod{p}$$

implies

$$g(c_k) \equiv 0 \pmod{p}\,.$$

This proves our claim, and the theorem thus follows by induction. □

Proposition 3.14. *let p be a prime and let d be a positive divisor of $p - 1$.*

Then the polynomial $x^d - 1$ has exactly d incongruent roots mod p.

Proof. Let $p - 1 = de$. Then

$$\begin{aligned} x^{p-1} - 1 = (x^d)^e - 1 &= (x^d - 1)(x^{d(e-1)} + x^{d(e-2)} + \cdots + x^d + 1) \\ &=: (x^d - 1)g(x)\,. \end{aligned}$$

Let N_1, N_2 and N_3 be respectively the numbers of incongruent roots mod p of the polynomials $x^{p-1} - 1$,

$x^d - 1$ and g. From Fermat's little theorem we see that $N_1 = p - 1$. Lagrange's theorem tells us that $N_2 \leq d$ and $N_3 \leq \deg(g) = d(e-1) = p - 1 - d$. Thus, $N_2 + N_3 \leq d + (p - 1 - d) = p - 1 = N_1$.

On the other hand, since p is a prime, the factorization $x^{p-1} - 1 = (x^d - 1)g(x)$ shows that any root of $x^{p-1} - 1 \bmod p$ is either a root of $x^d - 1 \bmod p$ or a root of $g \bmod p$. Therefore, $N_1 = p - 1 \leq N_2 + N_3$. It follows that $N_2 + N_3 = p - 1$ and hence the two inequalities

$$N_2 \leq d\,,\ N_3 \leq p - 1 - d$$

must be both equalities. This finishes the proof. □

Lemma 3.15. *Let p be a prime and let d be a positive divisor of $p - 1$.*

Then

$$\#\{c \in [\![1, p-1]\!] \mid \operatorname{ord}_p c = d\} \leq \phi(d)\,.$$

Proof. Of course, we may assume there does exist $a \in [\![1, p-1]\!]$ such that $\operatorname{ord}_p a = d$. Then the integers

$$a, a^2, \cdots, a^d$$

are pairwise incongruent mod p, because by Lemma 3.7,

$$a^i \equiv a^j \pmod p \iff d = \operatorname{ord}_p a \mid i - j\,, \quad \forall\, i, j \in \mathbb{N}\,.$$

Since $a^d \equiv 1 \pmod p$, the integers $a, a^2, \cdots, a^d$ are all roots of $x^d - 1 \bmod p$. From Prop. 3.14 we can conclude that every root of $x^d - 1 \bmod p$ is congruent to precisely one of the integers $a, a^2, \cdots, a^d$.

Thus, if $c \in [\![1, p-1]\!]$ satisfies $c^d \equiv 1 \pmod p$ (or equivalently, $\operatorname{ord}_p c \mid d$), then $c \equiv a^k$ for some unique $k \in [\![1, d]\!]$. Now we have

$$\operatorname{ord}_p(a^k) = \operatorname{ord}_p c = d \iff \gcd(k, d) = 1$$

by Prop. 3.8. Therefore,

$$\#\{c \in [\![1, p-1]\!] \mid \operatorname{ord}_p c = d\} \leq \#\{k \in [\![1, d]\!] \mid \gcd(k, d) = 1\} = \phi(d)\,.$$

This completes the proof. □

Theorem 3.16. *Let p be a prime. Then for each positive divisor d of $p - 1$,*

$$\#\{c \in [\![1, p-1]\!] \mid \operatorname{ord}_p c = d\} = \phi(d)\,.$$

In particular, there exists $c \in [\![1, p-1]\!]$ with $\operatorname{ord}_p c = p - 1$, i.e., p has a primitive root.

Proof. Write

$$A_d = \{c \in [\![1, p-1]\!] \,|\, \operatorname{ord}_p c = d\}$$

for each positive divisor d of $\phi(p) = p - 1$. By Prop. 3.3 we have

$$[\![1, p-1]\!] = \bigcup_{d \mid p-1} A_d$$

where d runs over all positive divisors of $p-1$. So we have

$$p-1 = \sum_{d \,|\, p-1} |A_d| \,.$$

Lemma 3.15 shows that $|A_d| \leq \phi(d)$ for each $d \,|\, p-1$. It is now sufficient to prove

$$\sum_{d \,|\, p-1} \phi(d) = p-1 \,.$$

This will be treated in Lemma 3.17 below. □

Lemma 3.17. *Let $n \geq 1$ be a positive integer. Then*

$$n = \sum_{d \,|\, n} \phi(d) = \sum_{d \,|\, n} \phi\left(\frac{n}{d}\right)$$

where d runs over all positive divisors of n.

Proof. As d runs through the positive divisors of n, the integer $\frac{n}{d}$ also runs through all the positive divisors of n. Hence

$$\sum_{d \,|\, n} \phi(d) = \text{ sum of the values of } \phi \text{ at all the positive divisors of } n = \sum_{d \,|\, n} \phi\left(\frac{n}{d}\right) \,.$$

Let us prove the equality

$$n = \sum_{d \,|\, n} \phi\left(\frac{n}{d}\right) \,.$$

We put

$$A_d := \{a \in [\![1, n]\!] \,|\, \gcd(a, n) = d\} \text{ for each } d \,|\, n \,.$$

Then

$$[\![1, n]\!] = \bigcup_{d \,|\, n} A_d \,.$$

So it suffices to show $|A_d| = \phi\left(\frac{n}{d}\right)$. Indeed, for $a \in [\![1,\, n]\!]$,

$$\gcd(a, n) = d \iff a = a' \cdot d \text{ for some } a' \in \left[\!\left[1\,,\, \frac{n}{d}\right]\!\right] \text{ and } \gcd\left(a'\,,\, \frac{n}{d}\right) = 1 \,.$$

Hence

$$\#A_d = \#\left\{a' \in \left[\!\left[1\,,\, \frac{n}{d}\right]\!\right] \;\middle|\; \gcd\left(a'\,,\, \frac{n}{d}\right) = 1\right\} = \phi\left(\frac{n}{d}\right) \,.$$

This completes the proof. □

3.2 Numbers having primitive roots

3.2.1 Prime powers

In the previous section, we showed that every prime has a primitive root. Now we study the case of prime powers and show that every power of an odd prime has a primitive root.

Proposition 3.18. *Let p be a prime > 2 and let r be a primitive root of p.*

Then either r or $r+p$ is a primitive root of p^2.

Proof. Let $n = \operatorname{ord}_{p^2} r$. On the one hand, we have $n \mid \phi(p^2) = p(p-1)$. On the other hand, we have $r^n \equiv 1 \pmod{p^2}$ and hence $r^n \equiv 1 \pmod{p}$. This implies $p-1 = \operatorname{ord}_p(r) \mid n$.

It follows that $n = \operatorname{ord}_{p^2} r$ is either $p-1$ or $p(p-1)$. In the latter case, we obtain $\operatorname{ord}_{p^2} r = \phi(p^2)$, so r is a primitive root of p^2.

Let us assume $\operatorname{ord}_{p^2} r = p-1$. We want to show that in this case $s := r+p$ is a primitive root of p^2. Since $s \equiv r \pmod{p}$, s is also a primitive root of p. So as above we have

$$\operatorname{ord}_{p^2}(s) = p-1 \quad \text{or} \quad \operatorname{ord}_{p^2}(s) = p(p-1) = \phi(p^2)\,.$$

It is sufficient to prove that $\operatorname{ord}_{p^2}(s) \neq p-1$.

We may use the binomial formula to compute

$$\begin{aligned} s^{p-1} &= (r+p)^{p-1} \\ &= r^{p-1} + (p-1)\cdot r^{(p-2)}p + \binom{p-1}{2} r^{p-3}\cdot p^2 + \cdots + \binom{p-1}{p-2} r p^{p-2} + p^{p-1} \\ &\equiv r^{p-1} + (p-1)r^{p-2}p \qquad \pmod{p^2} \end{aligned}$$

Recall that we have assumed $\operatorname{ord}_{p^2}(r) = p-1$. So $r^{p-1} \equiv 1 \pmod{p^2}$. Therefore,

$$s^{p-1} \equiv 1 + p(p-1)r^{p-2} = 1 - pr^{p-2} + p^2\cdot r^{p-2} \equiv 1 - pr^{p-2} \pmod{p^2}\,.$$

Then we find immediately that

$$s^{p-1} \equiv 1 \pmod{p^2} \iff p\cdot r^{p-2} \equiv 0 \pmod{p^2} \iff r^{p-2} \equiv 0 \pmod{p}\,.$$

But r is relatively prime to p. So $r^{p-2} \not\equiv 0 \pmod{p}$. Thus, we see that $s^{p-1} \not\equiv 1 \pmod{p^2}$ and hence $\operatorname{ord}_{p^2}(s) \neq p-1$. This finishes the proof. □

Theorem 3.19. *Let p be a prime > 2. If r is a primitive root of p^2, then r is a primitive root of p^k for all $k \in \mathbb{N}^*$. In particular, for every $k \in \mathbb{N}^*$, the integer p^k has a primitive root.*

Proof. We first prove the result for $k = 1$. Recall that by Fermat's little theorem. $\operatorname{ord}_p r \mid \phi(p) = p-1$. We need to show that for all $s \in [\![1, p-2]\!]$, $r^s \not\equiv 1 \pmod{p}$. If $r^s = 1 + tp$ for some $t \in \mathbb{Z}$, then

$$\begin{aligned} r^{sp} = (1+tp)^p &= 1 + \binom{p}{1} tp + \binom{p}{2} (tp)^2 + \cdots \\ &\equiv 1 + p\cdot tp + 0 + \cdots + 0 \pmod{p^2} \\ &\equiv 1 \pmod{p^2} \end{aligned}$$

This yields $\operatorname{ord}_{p^2}(r) = \phi(p^2) = p(p-1) \leq sp < (p-1)p$, a contradiction. So we must have $\operatorname{ord}_p(r) = p-1$. we shall assume now $k \geq 2$ and put $n = \operatorname{ord}_{p^k}(r)$. As in the proof of (3.18), we have

$$p-1 = \phi(p) = \operatorname{ord}_p(r) \;\big|\; n \;\big|\; \phi(p^k) = p^{k-1}(p-1)\,.$$

Hence $n = (p-1)p^j$ for some $j \in [\![0, k-1]\!]$. We now claim that

$$r^{p^{k-2}(p-1)} \not\equiv 1 \pmod{p^k}\,, \quad \text{when } k \geq 2\,, \tag{3.19.1}$$

so that the case $j \le k-2$ is excluded and we will get $n = \mathrm{ord}_{p^k}(r) = p^{k-1}(p-1)$ as desired.

We prove our clam by induction on k. The case $k=2$ is true because by assumption $\mathrm{ord}_{p^2}(r) = p(p-1) = \phi(p^2)$ (so that $r^{p-1} \not\equiv 1 \pmod{p^2}$). Now suppose (3.19.1) is already true for some $k \ge 2$. We want to show

$$r^{p^{k-1}(p-1)} \not\equiv 1 \pmod{p^{k+1}}\,. \tag{3.19.2}$$

To this end we consider the integer $s := r^{p^{k-2}(p-1)}$. We have $s = r^{p^{k-2}(p-1)} = r^{\phi(p^{k-1})} \equiv 1 \pmod{p^{k-1}}$ by Euler's theorem. Hence $s = 1 + dp^{k-1}$ for some $d \in \mathbb{Z}$. Now (3.19.1) shows that $p \nmid d$. We have

$$\begin{aligned} r^{p^{k-1}(p-1)} &= s^p = (1+dp^{k-1})^p \\ &= 1 + p\cdot dp^{k-1} + \binom{p}{2}(dp^{k-1})^2 + \cdots \\ &\equiv 1 + dp^k \pmod{p^{k+1}} \end{aligned}$$

since $p(k-1) > \cdots > 3(k-1) \ge k+1$ and

$$p^{k+1} = p\cdot p^k \,\Big|\, \binom{p}{2}\cdot p^{2k-2} \qquad \left(\text{because } p \,\Big|\, \binom{p}{2} \text{ and } k \le 2k-2\right).$$

Thus,

$$r^{p^{k-1}(p-1)} \equiv 0 \pmod{p^{k+1}} \iff d\cdot p^k \equiv 0 \pmod{p^{k+1}} \iff p \,|\, d\,.$$

We know that $p \nmid d$. Hence $r^{p^{k-1}(p-1)} \not\equiv 1 \pmod{p^{k+1}}$, thus proving (3.19.2) as required. This proves our claim, and as was discussed previously, the theorem follows from our claim. □

We now turn to the case of powers of 2. Obviously, 3 is a primitive root for both 2 and 4. So it remains to consider the higher powers of 2.

Theorem 3.20. *Let k be an integer ≥ 3.*

1. *For every odd integer a, one has*

$$\mathrm{ord}_{2^k}(a) \,\Big|\, 2^{k-2} = \phi(2^k)/2\,.$$

   Consequently, 2^k has no primitive roots.

2. $\mathrm{ord}_{2^k}(5) = 2^{k-2} = \phi(2^k)/2$.

Proof. (1) The desired conclusion means that

$$a^{2^{k-2}} \equiv 1 \pmod{2^k}\,, \qquad \text{for all } k \ge 3\,. \tag{3.20.1}$$

We prove this by induction, the case $k=3$ being easily checked by direct computation.

Now suppose (3.20.1) holds for some $k \ge 3$. We want to prove

$$a^{2^{k-1}} \equiv 1 \pmod{2^{k+1}}\,. \tag{3.20.2}$$

We may write $a^{2^{k-2}} = 1 + d\cdot 2^k$ for some $d \in \mathbb{Z}$ by (3.20.1). Then

$$a^{2^{k-1}} = (a^{2^{k-2}})^2 = (1+d\cdot 2^k)^2 = 1 + d\cdot 2^{k+1} + d^2\cdot 2^{2k}\,.$$

Now $2k \geq k+1$. So the above immediately shows that (3.20.2) is true.
(2) We only need to show $5^{2^{k-3}} \not\equiv 1 \pmod{2^k}$. The strategy is to prove the stronger relation

$$5^{2^{k-3}} \equiv 1 + 2^{k-1} \pmod{2^k}\,.$$

The proof is similar to precious proofs by induction. So we leave it to the reader to finish the proof. □

3.2.2 The general case

In this section we determine precisely which positive integers have primitive roots. In addition to the case of powers of odd primes, we now prove the following:

Theorem 3.21. *Let p be an odd prime and $t \in \mathbb{N}^*$. Let r be a primitive root of p^t.*

1. *If r is odd, then r is also a primitive root of $2p^t$.*
2. *If r is even, then $r + p^t$ is a primitive root of $2p^t$.*
3. *The integer $2p^t$ always has a primitive root.*

Proof. Clearly (3) follows from (1) and (2), in view of Thm.3.19. When r is an even primitive root of p^t, the integer $r + p^t$ is an odd primitive root of p^t. So (2) is a consequence of (1).

It remains to prove (1). Assume r is odd. Then the congruences

$$r^k \equiv 1 \pmod{p^t} \quad \text{and} \quad r^k \equiv 1 \pmod{2p^t}$$

are equivalent for all $k \in \mathbb{N}^*$, because $r^k \equiv 1 \pmod 2$ holds automatically. It follows that

$$\mathrm{ord}_{2p^t}(r) = \mathrm{ord}_{p^t}(r) = \phi(p^t)\,.$$

Since $\phi(2p^t) = \phi(2)\phi(p^t) = \phi(p^t)$, the result is proved. □

Proposition 3.22. *Let n be an integer > 1. If n has a primitive root, then*

$$n = p^t \quad \text{or} \quad n = 2p^t \quad \text{for some prime } p \text{ and some } t \in \mathbb{N}^*\,.$$

Proof. Let $n = {p_1}^{t_1} \cdots {p_m}^{t_m}$ be the factorization into prime power factors of n, i.e., each $t_i \in \mathbb{N}^*$ and the p_i's are distinct primes. Then for any integer a relatively prime to n, we have

$$a^{\phi({p_i}^{t_i})} \equiv 1 \pmod{p^{t_i}}\,, \text{ for each } i = 1, \cdots, m.$$

Putting

$$u := \mathrm{lcm}(\phi({p_1}^{t_1})\,, \cdots\,, \phi({p_m}^{t_m}))$$

we have

$$a^u \equiv 1 \pmod{p^{t_i}}\,, \text{ for all } i$$

whence $a^u \equiv 1 \pmod n$.

Therefore $\mathrm{ord}_n\, a \leq u$. If n has a primitive root, then we must have $\phi(n) \leq u$, namely,

$$\phi({p_1}^{t_1}) \cdots \phi({p_m}^{t_m}) \leq \mathrm{lcm}(\phi({p_2}^{t_1}), \cdots \phi({p_m}^{t_m}))\,.$$

This inequality can hold only if the integers $\phi({p_1}^{t_1}), \cdots, \phi({p_m}^{t_m})$ are pairwise relatively prime. If $m \geq 3$, or $m = 2$ and both p_1 and p_2 are odd, then at least 2 of the numbers $\phi({p_i}^{t_i})$ are even. So this is not allowed and we must have either $m = 1$, which means n is a prime power, or $m = 2$ and one of p_1 and p_2 is 2. In the latter case $n = 2^s p^t$ for some odd prime p and $t, s \in \mathbb{N}^*$. In order that

$$1 = \gcd(\phi(2^s), \phi(p^t)) = \gcd(2^{s-1}, (p-1)p^{t-1})$$

we must have $s = 1$, i.e., $n = 2p^t$.

This completes the proof of the proposition. □

Summarizing (3.19)–(3.22), we obtain:

Theorem 3.23. *let n be a positive integer > 1. Then, n has a primitive root if and only if*

$$n = 2\,,\ 4\,,\ p^t \ \textit{or}\ 2p^t\,,$$

where p is an odd prime and $t \in \mathbb{N}^*$.

3.3 Applications and complements

3.3.1 Primality tests using orders of integers

Let n be an integer ≥ 2. If n is a prime, then Fermat's little theorem asserts that for every integer x that is relatively prime to n, we have

$$x^{n-1} \equiv 1 \pmod{n}\,.$$

Note however that, when n is composite number, it is still possible that for some x coprime to n, the above congruence still holds. For example, if $n = 341 = 11 \times 31$, the integer $x = 2$ satisfies $x^{n-1} = 2^{340} \equiv 1 \pmod{n}$.

In general, if b is a positive integer greater than 1, we say a positive integer n is a ***pseudo prime to the base*** b 以 b 为基的伪素数, if n is a composite number and if $b^{n-1} \equiv 1 \pmod{n}$ (so in particular $\gcd(b, n) = 1$).

Despite the existence of pseudo primes, we can prove the following result:

Theorem 3.24 (Lucas' converse of Fermat's little theorem)**.** *Let n be an integer ≥ 2. Suppose that there exists an integer x satisfying the following conditions:*

1. $x^{n-1} \equiv 1 \pmod{n}$ *(in particular, we have* $\gcd(x, n) = 1$*);*
2. *For every prime divisor q of $n - 1$,*

$$x^{(n-1)/q} \not\equiv 1 \pmod{n}\,.$$

Then n is a prime.

Proof. We consider $r = \operatorname{ord}_n x$. By (1), we have $r \mid (n-1)$. If $r < n-1$, then $n-1 = rk$ for some $k \in \mathbb{N}$, $k > 1$. Let q be a prime divisor of k. Then

$$x^{(n-1)/q} = x^{rk/q} = (x^r)^{\frac{k}{q}} \equiv 1 \pmod{n}$$

since $x^r = x^{\operatorname{ord}_n(x)} \equiv 1 \pmod{n}$. This leads to a contradiction to condition (2). Therefore, we have $\operatorname{ord}_n x = n-1$. But we also have $\operatorname{ord}_n x \mid \phi(n)$. So we get $n-1 = \operatorname{ord}_n x \leq \phi(n)$. The inequality $\phi(n) \leq n-1$ is clear. Hence $\phi(n) = n-1$. This implies that n is a prime number. □

Corollary 3.25. *Let n be an odd integer > 2. Suppose there exists an integer x such that*

1. $x^{(n-1)/2} \equiv -1 \pmod{n}$;
2. *For every odd prime divisor q of $n-1$,*

$$x^{(n-1)/q} \not\equiv 1 \pmod{n}.$$

Then n is a prime.

Proof. The assumptions here imply that conditions (1) and (2) in Thm. 3.24 hold. □

Another partial converse of Fermat's little theorem is the following:

Theorem 3.26 (Pocklington's primality test)**.** *Let n be a positive integer and suppose that $n-1 = F \cdot R$ for some $F, R \in \mathbb{N}^*$ with $F > R$ and $\gcd(F, R) = 1$. Suppose that there exists $a \in \mathbb{Z}$ such that*

1. $a^{n-1} \equiv 1 \pmod{n}$;
2. *For all prime divisor q of F,*

$$\gcd(a^{(n-1)/q} - 1, n) = 1.$$

(*So, we have in particular $a^{(n-1)/q} \not\equiv 1 \pmod{n}$.*)

Then n is a prime.

Proof. Assume the contrary. Then n has a prime divisor $p \leq \sqrt{n}$. Condition (1) implies that $\operatorname{ord}_p a \mid (n-1)$. Hence $n-1 = t \cdot \operatorname{ord}_p a$ for some $t \in \mathbb{N}^*$.
We claim that $F \mid \operatorname{ord}_p a$. Indeed, since F is a divisor of $n-1 = t \cdot \operatorname{ord}_p a$, it suffices to prove that $\gcd(F, t) = 1$. We need only to show that for every prime divisor q of F, $q \nmid t$.

If $q \mid t$, then

$$a^{(n-1)/q} = a^{t \cdot \operatorname{ord}_p(a)/q} = (a^{\operatorname{ord}_p a})^{\frac{t}{q}} \equiv 1 \pmod{p}.$$

It follows that p is a common divisor of $a^{(n-1)/q} - 1$ and n, contradicting the condition (2). Our claim is thus proved, i.e., we have $F \mid \operatorname{ord}_p a$. In particular,

$$F \leq \operatorname{ord}_p a \leq p - 1 < p.$$

Now

$$n - 1 = F \cdot R < F^2 \quad (\text{since } F > R \text{ by assumption})$$

This implies $n \leq F^2 < p^2$, whence $p > \sqrt{n}$. But this contradicts the initial hypothesis $p \leq \sqrt{n}$. The theorem is thus proved by contradiction. □

Theorem 3.27 (Proth's primality test)**.** *Let $n = k \cdot 2^m + 1$ where $k \in \mathbb{N}$ is an odd natural number, $m \in \mathbb{N}$ and $k < 2^m$.*

If there exists $a \in \mathbb{Z}$ such that $a^{(n-1)/2} \equiv -1 \pmod{n}$, then n is a prime.

Proof. We apply Pocklington's test with $F = 2^m$ and $R = k$. It is sufficient to check that

$$\gcd(a^{(n-1)/2} - 1, n) = 1 \,.$$

Indeed, if $d = \gcd(a^{(n-1)/2} - 1, n)$, then d is a divisor of $a^{(n-1)/2} + 1$ because by assumption $n \mid a^{(n-1)/2} + 1$. Therefore

$$d \,\Big|\, \big((a^{(n-1)/2} + 1) - (a^{(n-1)/2} - 1)\big)\,, \quad \textit{i.e.,}\ d \mid 2.$$

But as a divisor of $n = k \cdot 2^m + 1$, d must be odd. So we get $d = 1$ as desired. □

3.3.2 Universal exponents and power residues

(3.28) Let $n \in \mathbb{N}^*$. A ***universal exponent*** 通用指数 of n is a positive integer $u \in \mathbb{N}^*$ such that $a^u \equiv 1 \pmod{n}$ for all $a \in \mathbb{Z}$ relatively prime to n. By Prop. 3.3,

$$\operatorname{ord}_n a \mid u \tag{3.28.1}$$

for every $a \in \mathbb{Z}$ relatively prime to n and every universal exponent u of n.

By Euler's theorem, $\phi(n)$ is a universal exponent. So the set of all universal exponents of n is a nonempty subset of $\mathbb{N}^*$. By the well ordering principle, there is a ***smallest universal exponent*** of n. We denote it by $\lambda(n)$. As a special case of (3.28.1), $\operatorname{ord}_n a \mid \lambda(n)$. ■

Example 3.29. Suppose $n \geq 2$ has prime factorization

$$n = p_1^{t_1} p_2^{t_2} \cdots p_m^{t_m}$$

where p_i are distinct primes and $t_i \in \mathbb{N}^*$. Then $u = \operatorname{lcm}(\phi(p_1^{t_1}), \cdots, \phi(p_m^{t_m}))$ is a universal exponent of n. In fact, when $a \in \mathbb{Z}$ is relatively prime to n, we have

$$a^{\phi({p_i}^{t_i})} \equiv 1 \pmod{p_i^{t_i}}$$

and since $\phi({p_i}^{t_i}) \mid u$,

$$a^u \equiv 1 \pmod{p_i^{t_i}}\,.$$

It follows that $n = p_1^{t_1} \cdots p_m^{t_m}$ divides $a^u - 1$, i.e., $a^u \equiv 1 \pmod{n}$. ■

Proposition 3.30. *Let $n \in \mathbb{N}^*$.*

1. *$\lambda(1) = 1 = \phi(1)$, $\lambda(2) = 1 = \phi(2)$, $\lambda(4) = 2 = \phi(2)$.*
2. *If $n = 2^t$ with $t \geq 3$, then $\lambda(n) = 2^{t-2}$.*
3. *If $n = p^t$ with p an odd prime and $t \in \mathbb{N}^*$, then $\lambda(n) = \phi(n) = p^{t-1}(p-1)$.*

Proof. That $\lambda(1) = 1$ follows directly from the definition. If $n \geq 2$ and n has a primitive root r, then any universal exponent should be a multiple of $\phi(n)$. Hence $\lambda(n) = \phi(n)$. In view of Thm. 3.23, we deduce the formulas in (1) and (3). If $n = 2^t$ with $t \geq 3$, then by Thm. 3.20, $2^{t-2} = \phi(n)/2$ is a universal exponent and since $\operatorname{ord}_n 5 = 2^{t-2}$, $\lambda(n) = 2^{t-2}$. □

Proposition 3.31. *Let $n \geq 2$ and let $n = 2^{t_0} p_1^{t_1} \cdots p_m^{t_m}$ be its prime factorization (where $t_i \in \mathbb{N}$ and $p_1, \cdots, p_m$ are distinct odd primes). Then*

$$\begin{aligned} \lambda(n) &= \operatorname{lcm}(\lambda(2^{t_0}), \lambda({p_1}^{t_1}), \cdots, \lambda({p_m}^{t_m})) \\ &= \operatorname{lcm}(\lambda(2^{t_0}), \phi({p_1}^{t_1}), \cdots, \phi({p_m}^{t_m})) \end{aligned}$$

Moreover, there exists $a \in \mathbb{Z}$ such that $\operatorname{ord}_n a = \lambda(n)$. Therefore, $\lambda(n)$ is the largest possible value of the orders of integers modulo n.

Proof. Let us write

$$M = \operatorname{lcm}(\lambda(2^{t_0}),\ \phi(p_1^{t_1}),\ \cdots,\ \phi(p_m^{t_m}))\,.$$

Let us first prove that M is a universal exponent of n, so that $\lambda(n) \leq M$.

Take any $b \in \mathbb{Z}$ relatively prime to n. We want to show $b^M \equiv 1 \pmod{n}$. Indeed, we have

$$b^{\lambda(2^{t_0})} \equiv 1 \pmod{2^{t_0}},\ b^{\phi(p_i^{t_i})} = 1 \pmod{p_i^{t_i}},\ \forall\, i \in [\![1, m]\!]\ .$$

Hence

$$b^M \equiv 1 \pmod{2^{t_0}} \quad \text{and} \quad b^M = 1 \pmod{p_i^{t_i}},\ \forall\, i \in [\![1, m]\!]\ .$$

This gives $b^M \equiv 1 \pmod{n}$ as desired.

Now we prove that there exists an integer a relatively prime to n such that $\operatorname{ord}_n a = M$. Since $\operatorname{ord}_n a$ divides any universal exponent (by (3.28.1)), this will show that every universal exponent is a multiple of M and $\lambda(n) \geq M$.

To see this, notice that each $p_i^{t_i}$ has a primitive root r_i and if we put

$$r_0 = \begin{cases} 3 & \text{if } t_0 \leq 2 \\ 5 & \text{if } t_0 \geq 3 \end{cases}$$

then $\operatorname{ord}_{2^{t_0}}(r_0) = \lambda(2^{t_0})$. By the Chinese remainder theorem, the system

$$\begin{cases} x \equiv r_0 \pmod{2^{t_0}} \\ x \equiv r_1 \pmod{p_1^{t_1}} \\ \qquad \vdots \\ x \equiv r_m \pmod{p_m^{t_m}} \end{cases}$$

has a solution $a \in \mathbb{Z}$. We claim that $\operatorname{ord}_n a = M$ holds for this integer a.

Indeed, if $N \in \mathbb{N}^*$ has the property $a^N \equiv 1 \pmod{n}$ then for every prime divisor p of n and $t = v_p(n)$, we have $a^N \equiv 1 \pmod{p^t}$, whence $\operatorname{ord}_{p^t}(a) \mid N$. From the congruence system to which a is a solution we see that

$$\begin{cases} \operatorname{ord}_{2^{t_0}}(a) = \operatorname{ord}_{2^{t_0}}(r_0) = \lambda(2^{t_0}) \\ \operatorname{ord}_{p_i^{t_i}}(a) = \operatorname{ord}_{p_i^{t_i}}(r_i) = \phi(p_i^{t_i}),\ \forall\, i \in [\![1, m]\!] \end{cases}$$

Hence $\lambda(p^t) \mid N$ and it follows that

$$M = \operatorname{lcm}\big(\lambda(2^{t_0}),\ \phi(p_1^{t_1}),\ \cdots,\ \phi(p_m^{t_m})\big) \mid N\ .$$

This being true for all $N \in \mathbb{N}^*$ such that $a^N \equiv 1 \pmod{n}$, it follows that $\operatorname{ord}_n(a) = M$. □

Example 3.32. Let $n = 180 = 2^2 \cdot 3^2 \cdot 5$. Then

$$\lambda(180) = \operatorname{lcm}(\lambda(4), \phi(3^2), \phi(5)) = \operatorname{lcm}(2, 6, 4) = 12\,.$$

To find an integer $a \in \mathbb{Z}$ with $\operatorname{ord}_{180}(a) = 12$, we can first check that 3 is a primitive root of 5 and 2 is a primitive root of 9. According to our proof of Prop. 3.31, it is now sufficient to solve the system

$$\begin{cases} x \equiv 3 \pmod 4 \\ x \equiv 2 \pmod 9 \\ x \equiv 3 \pmod 5 \end{cases}$$

One solution is $x = 83$. Therefore, for $a = 83$, we have $\operatorname{ord}_{180}(a) = 12 = \lambda(180)$. □

Definition 3.33. Let $k, m \in \mathbb{N}^*$ with $m \geq 2$. Let $b \in \mathbb{Z}$ be relatively prime to m. We say that b is a ***k-th power residue*** (k 次幂剩余) of m if the congruence $x^k \equiv b \pmod m$ has a solution, or equivalently, if there exists an element $\alpha \in \mathbb{Z}/m\mathbb{Z}$ such that $\alpha^k = [b]_m \in \mathbb{Z}/m\mathbb{Z}$. □

Note that according to our convention, we only consider integers that are relatively prime to m when we talk about k-th power residues of m.

For simplicity, here we only study k-th power residues in the simplest case: the case where m has a primitive root.

Theorem 3.34. *Suppose that m has a primitive root r. Let $b \in \mathbb{Z}$ be relatively prime to m, $k \in \mathbb{N}^*$ and $d = \gcd(k, \phi(m))$.*

1. *The following are equivalent:*
   - *(a) The congruence $x^k \equiv b \pmod m$ has a solution. i.e., b is a k-th power residue of m.*
   - *(b) $\operatorname{ord}_m(b)$ divides $\frac{\phi(m)}{d} = \frac{\phi(m)}{\gcd(k,\phi(m))}$.*
   - *(c) $b^{\frac{\phi(m)}{d}} \equiv 1 \pmod m$.*
2. *When the congruence $x^k \equiv b \pmod m$ has a solution, it has precisely d incongruent solutions modulo m.*

Proof. We have $[b]_m \in (\mathbb{Z}/m\mathbb{Z})^*$, and by Thm. 3.6,

$$(\mathbb{Z}/m\mathbb{Z})^* = \left\{ [r]^j \;\middle|\; j = 1\,, 2\,, \cdots, \phi(m) \right\}\,.$$

So there is a unique integer $i \in [\![1, \phi(m)]\!]$ such that $[b]_m = [r]_m^i$.

We have

$$\operatorname{ord}_m(b) = \operatorname{ord}_m(r^i) = \frac{\operatorname{ord}_m(r)}{\gcd(i,\, \operatorname{ord}_m(r))} = \frac{\phi(m)}{\gcd(i, \phi(m))}\,.$$

(1) The equivalence (b)⇔(c) follows from the fundamental property of $\operatorname{ord}_m(b)$ (Prop. 3.3).

It remains to prove (a)⇔(c).

If $x^k \equiv b \pmod m$ for some $x \in \mathbb{Z}$, then

$$b^{\frac{\phi(m)}{d}} \equiv (x^k)^{\frac{\phi(m)}{d}} = (x^{\phi(m)})^{\frac{k}{d}} \equiv 1 \pmod m\,.$$

Here the last congruence follows from the facts that $x^{\phi(m)} \equiv 1 \pmod{m}$ (Euler's theorem) and that $d = \gcd(k, \phi(m))$ divides k. This shows (a)⇒(c).

Conversely, assume $b^{\frac{\phi(m)}{d}} \equiv 1 \pmod{m}$. This means

$$\operatorname{ord}_m b = \frac{\phi(m)}{\gcd(i, \phi(m))} \text{ divides } \frac{\phi(m)}{d}$$

or equivalently, d divides $\gcd(i, \phi(m))$. Thus $d = \gcd(k, \phi(m))$ divides i.

It follows that the congruence equation $ky \equiv i \pmod{\phi(m)}$ has integer solutions. We may thus choose $j \in [\![1, \phi(m)]\!]$ such that $kj \equiv i \pmod{\phi(m)}$. Now since $[r^{\phi(m)}]_m = [1]_m$, the integer $x := r^j$ satisfies

$$[x]_m^k = [r^{jk}]_m = [r^i]_m = [b]_m \,,$$

showing that b is a k-th power residue of m.

(2) The question is to find the number of elements $\alpha \in (\mathbb{Z}/m\mathbb{Z})^*$ such that $\alpha^k = [b]_m$. Since

$$(\mathbb{Z}/m\mathbb{Z})^* = \left\{ [r]^j \;\middle|\; j = 1, 2, \cdots, \phi(m) \right\} ,$$

this is equivalent to finding the number of integers $j \in [\![1, \phi(m)]\!]$ such that $[r^j]_m^k = [b]_m = [r^i]_m$. Since

$$[r^j]_m^k = [r^i]_m \iff kj \equiv i \pmod{\phi(m)} \,,$$

we are reduced to counting the number

$$\begin{aligned} &\#\{j \in [\![1, \phi(m)]\!] \mid kj \equiv i \pmod{\phi(m)}\} \\ &= \text{the number of incongruent solutions of the congruence equation } ky \equiv i \pmod{\phi(m)} \,. \end{aligned}$$

But we know that this last number is equal to $\gcd(k, \phi(m)) = d$ (Theorem 2.13). This completes the proof. □

Corollary 3.35. *Let p be a prime, $k \in \mathbb{N}^*$ and $a \in \mathbb{Z}$ relatively prime to p.*

Then a is a k-th power residue of p (i.e., $x^k \equiv a \pmod{p}$ has a solution) if and only if

$$a^{\frac{p-1}{\gcd(k, p-1)}} \equiv 1 \pmod{p} \,.$$

Example 3.36. Let us show that 5 is not a 6-th power residue of 17. Indeed,

$$5^{\frac{17-1}{\gcd(6, 17-1)}} = 5^{\frac{16}{2}} = 5^8 \equiv -1 \not\equiv 1 \pmod{17} \,.$$

So the result follows from Corollary 3.35. ■

Example 3.37. We give a concrete example of computation of primitive roots, roots of k-th power residues, etc. We will do the following:

1. Find a primitive root of 17.
2. Solve the congruence $7^x \equiv 6 \pmod{17}$.
3. Solve the congruence $6x^{12} \equiv 11 \pmod{17}$.

Solution. (1). This is to find an integer $a \in [\![1, 16]\!]$ such that $\mathrm{ord}_{17}(a) = 16$. Clearly $\mathrm{ord}_{17}(1) = 1$.

Next we compute

$$2^2 \equiv 4,\ 2^4 \equiv 16 = -1,\ 2^8 \equiv (-1)^2 = 1\,.$$

Hence, $\mathrm{ord}_{17}(2) = 8$. (Note that $\mathrm{ord}_{17}(a)$ is always a divisor of 16, so we don't need to compute the other powers of 2.)

Once we know the order of 2 mod 17 is 8, we can determine all the elements of order 8 mod 17. In fact, by Thm. 3.16, there are $\phi(8) = 4$ elements of order 8 in $(\mathbb{Z}/17\mathbb{Z})^*$. These elements are given by $[2]^i$, with $i \in [\![1, 8]\!]$ relatively prime to 8, i.e.,

$$[2]^1 = [2]\,,\ [2]^3 = [8]\,,\ \ [2]^5 = [15]\,,\ [2]^7 = [9]\,.$$

This suggests that primitive roots should be chosen from $[\![1, 16]\!] \setminus \{1, 2, 8, 9, 15\}$.

Let us now try the integer 3. Note that $3^2 = 9$ has order 8, as we have just found above. Therefore, $\mathrm{ord}_{17}(3) = 2 \times \mathrm{ord}_{17}(3^2) = 16$. So we see that 3 is a primitive root of 17. (If one wishes to find all the primitive roots of 17 in $[\![1, 16]\!]$, he just needs to consider the least positive residues of the powers 3^i, for $i \in [\![1, 16]\!]$ relatively prime to 16.)

(2) The idea is to first express $[7]_{17}$ and $[6]_{17}$ as powers of the residue class of the primitive root 3, i.e., to find $i,\ j \in [\![1,\ 16]\!]$ such that $3^i \equiv 7 \pmod{17}$ and $3^j \equiv 6 \pmod{17}$.

To this end, we compute the residues of powers of 3:

$$\begin{aligned}
&3^2 = 9,\ 3^3 \equiv 10,\ 3^4 \equiv 30 = 13,\ 3^5 \equiv 13 \times 3 = (-4) \times 3 \equiv 5\\
&3^6 \equiv 5 \times 3 = 15 \equiv -2,\ 3^7 \equiv (-2) \times 3 \equiv 11,\ 3^8 \equiv (-6) \times 3 = -1 \equiv 16\\
&3^9 \equiv (-1) \times 3 \equiv 14,\ 3^{10} \equiv (-3) \times 3 \equiv 8,\ 3^{11} \equiv 8 \times 3 \equiv 7\\
&3^{12} \equiv 7 \times 3 \equiv 4,\ 3^{13} \equiv 4 \times 3 = 12,\ 3^{14} \equiv 12 \times 3 \equiv (-5) \times 3 \equiv 2\\
&3^{15} \equiv 2 \times 3 = 6,\ 3^{16} \equiv 1
\end{aligned}$$

Hence, $i = 11$, $j = 15$. (For j, perhaps a slightly simpler way is to notice that $3^6 \equiv -2, 3^9 \equiv -3$, hence $3^{15} \equiv 6$.)

Now the congruence $7^x \equiv 6 \pmod{17}$

$$\begin{aligned}
7^x \equiv 6 \pmod{17} &\iff 3^{11x} \equiv 3^{15} \pmod{17}\\
&\iff \mathrm{ord}_{17}(3) \,|\, 11x - 15 \iff 11x \equiv 15 \pmod{16}\\
&\iff -5x \equiv 15 \pmod{16} \iff x \equiv -3 \equiv 13 \pmod{16}\,.
\end{aligned}$$

So we conclude that the solutions to $7^x \equiv 6 (\mathrm{mod} 17)$ are the integers x satisfying $x \equiv 13 (\mathrm{mod}\ 16)$.

(3) Again the idea is to express everything in terms of powers of the primitive root 3. Since the residue class $[x]$ must be of the form $[3]^y$ for some $y \in [\![1,\ 16]\!]$, using the computational results obtained above, we find

$$\begin{aligned}
6x^{12} \equiv 11 \pmod{17} &\iff 3^{15} \cdot 3^{12y} \equiv 3^7 \pmod{17}\\
&\iff 15 + 12y \equiv 7 \pmod{16} \iff \cdots \iff y \equiv 2 \pmod{4}\,.
\end{aligned}$$

Therefore, the solutions to the initial congruence $6x^{12} \equiv 11 \pmod{17}$ are $x \equiv 3^2,\ 3^6\,,\ 3^{10}\,,\ 3^{14} \pmod{17}$. Namely, $x \equiv 9,\ 5,\ 18$ or $2 \pmod{17}$. ■

Chapter 4

Quadratic Residues

4.1 Quadratic residues and nonresidues

(4.1) Let m be a integer ≥ 2 and $k \in \mathbb{N}^*$. Let $a \in \mathbb{Z}$ be relatively prime to m. Recall that $a \in \mathbb{Z}$ is called a ***k-th power residue*** of m if the congruence $x^2 \equiv a \pmod{m}$ has an integer solution.

In this chapter we study the case where $k = 2$ and $m = p$ is a prime. For any $a \in \mathbb{Z}$ with $\gcd(a, p) = 1$, we say that a is a ***quadratic residue*** 二次剩余 of p if $x^2 \equiv a \pmod{p}$ has integer solutions. Otherwise we say a is a ***quadratic nonresidue*** 二次非剩余 of p. Note that when we talk about quadratic residues or nonresidues, only integers that are coprime to p are considered.

The case $p = 2$ is not an interesting case, because obviously every odd integer is a quadratic residue of 2. ■

The following lemma is a special case of Thm. 3.34 (2), but one can also prove it directly in a simple way.

Lemma 4.2. *Let p be an odd prime and $a \in \mathbb{Z}$ relatively prime to p.*

Then the congruence $x^2 \equiv a \pmod{p}$ has either no solutions or precisely 2 incongruent solutions modulo p.

Proof. If x_0 is a solution, then

$$\begin{aligned} x^2 \equiv a \pmod{p} &\iff x^2 = x_0^2 \mod p \iff (x + x_0)(x - x_0) \equiv 0 \pmod{p} \\ &\iff x \equiv x_0 \text{ or } x \equiv -x_0 \pmod{p}. \end{aligned}$$

Notice that $x_0^2 \equiv a \not\equiv 0$ implies that $x_0 \not\equiv 0 \pmod{p}$. Since p is odd, we have $2x_0 \not\equiv 0$. This shows that $-x_0 \not\equiv x_0 \pmod{p}$. Therefore, x_0 and $-x_0$ are 2 incongruent solutions to the congruence $x^2 \equiv a \pmod{p}$. This completes the proof. □

Theorem 4.3. *Let p be an odd prime.*

In any reduced residue system of p (e.g. the set $[\![1, \cdots, p-1]\!]$), there are exactly $\frac{p-1}{2}$ quadratic residues and $\frac{p-1}{2}$ quadratic nonresidues of p.

Moreover, the congruence classes of the quadratic residues are given by

$$[1^2]_p, [2^2]_p, [3^2]_p, \cdots, \left[\left(\frac{p-1}{2}\right)^2\right]_p \in (\mathbb{Z}/p\mathbb{Z})^*$$

Proof. The set of congruence classes $[a]_p$, when a runs through the quadratic residues in a reduced residue system, is nothing but the image $\mathrm{Im}(f)$ of the map

$$f \,:\, (\mathbb{Z}/p\mathbb{Z})^* \longrightarrow (\mathbb{Z}/p\mathbb{Z})^* \,; \quad \alpha = [x]_p \longmapsto \alpha^2 = \left[x^2\right]_p \,.$$

By Lemma 4.2, for each $\beta \in \mathrm{Im}(f)$, the inverse image $f^{-1}(\beta) := \{\alpha \in (\mathbb{Z}/p\mathbb{Z})^* \,|\, f(\alpha) = \beta\}$ has exactly 2 elements , Therefore

$$\left|(\mathbb{Z}/p\mathbb{Z})^*\right| = 2 \cdot |\mathrm{Im}(f)|$$

Hence, the number of quadratic residues (in a reduced residues system) is equal to

$$|\mathrm{Im}(f)| = \frac{1}{2}\left|(\mathbb{Z}/p\mathbb{Z})^*\right| = \frac{p-1}{2}$$

and the number of quadratic nonresidues is

$$\left|(\mathbb{Z}/p\mathbb{Z})^*\right| - \text{the number of quadratic residues} = (p-1) - \frac{p-1}{2} = \frac{p-1}{2} \,.$$

To determine the congruence classes of the quadratic residues is to find the image of f. Clearly, $\left[x^2\right]_p \in \mathrm{Im}(f)$ for all $x \in \left[\!\left[1, \frac{p-1}{2}\right]\!\right]$. We need only to show that for distinct numbers $x, y \in \left[\!\left[1, \frac{p-1}{2}\right]\!\right]$, $\left[x^2\right]_p \neq \left[y^2\right]_p$.

Indeed, $\left[x^2\right]_p = \left[y^2\right]_p$ means $x \equiv y$ or $x \equiv -y \pmod p$. When $x \neq y$ and $x, y \in \left[\!\left[1, \frac{p-1}{2}\right]\!\right]$, we have $x \not\equiv y \pmod p$. Moreover, $x + y \in [\![2, p-1]\!]$, so $x + y \not\equiv 0 \pmod p$, i.e., $x \not\equiv -y \pmod p$. This proves the desired result. □

We now introduce a very useful notation associated with quadratic residues.

Definition 4.4. Let p be an odd prime and $a \in \mathbb{Z}$. The ***Legendre symbol*** 勒让德符号 of a with respect to p, denoted by $\left(\frac{a}{p}\right)$, is defined by

$$\left(\frac{a}{p}\right) = \begin{cases} 0 & \text{if } p \,|\, a \\ 1 & \text{if } p \nmid a \text{ and } a \text{ is a quadratic residue of } p \\ -1 & \text{otherwise} \end{cases}$$

Clearly, $\left(\frac{a}{p}\right)$ depends only on the congruence class $[a]_p$. ■

Example 4.5. Let $p = 11$. Then

$$1^2 \equiv 1 \,,\; 2^2 \equiv 4 \,,\; 3^2 \equiv 9 \,,\; 4^2 \equiv 5 \,,\; 5^2 \equiv 3 \pmod{11} \,.$$

This shows that the quadratic residue classes are

$$[1]_{11} \,,\; [3]_{11} \,,\; [4]_{11} \,,\; [5]_{11} \text{ and } [9]_{11} \,.$$

So we have

$$\left(\frac{1}{11}\right) = \left(\frac{3}{11}\right) = \left(\frac{4}{11}\right) = \left(\frac{5}{11}\right) = \left(\frac{9}{11}\right) = 1$$

and

$$\left(\frac{2}{11}\right) = \left(\frac{6}{11}\right) = \left(\frac{7}{11}\right) = \left(\frac{8}{11}\right) = \left(\frac{10}{11}\right) = -1 \,.$$

(Recall that $\left(\frac{11}{11}\right) = 0$.) ■

The next result is a special case of Corollary 3.35.

Theorem 4.6 (Euler's criterion)**.** *Let p be an odd prime and $a \in \mathbb{Z}$. Then*

$$a^{\frac{p-1}{2}} \equiv \left(\frac{a}{p}\right) \pmod{p}\,.$$

Proof. If $p \mid a$, then $a \equiv 0 \pmod{p}$ and $a^{\frac{p-1}{2}} \equiv 0 \pmod{p}$. Since $\left(\frac{a}{p}\right) = 0$, the conclusion is true in this case.

Now suppose $p \nmid a$. If $a \equiv x^2 \pmod{p}$ for some $x \in \mathbb{Z}$, then $\left(\frac{a}{p}\right) = 1$ and

$$a^{\frac{p-1}{2}} \equiv (x^2)^{\frac{p-1}{2}} \equiv x^{p-1} \equiv 1 \pmod{p}$$

by Fermat's little theorem.

If $\left(\frac{a}{p}\right) = -1$, then $x^2 \equiv a \pmod{p}$ has no solution. Thus, if r is a primitive root of p and $a \equiv r^i \pmod{p}$, then i must be odd. Therefore,

$$\operatorname{ord}_p(a) = \operatorname{ord}_p(r^i) = \frac{\phi(p)}{\gcd(i, \phi(p))} = \frac{p-1}{\gcd(i, p-1)}$$

cannot divide $\frac{p-1}{2}$. Hence $a^{\frac{p-1}{2}} \not\equiv 1 \pmod{p}$. Since $(a^{\frac{p-1}{2}})^2 \equiv 1 \pmod{p}$, it follows that

$$a^{\frac{p-1}{2}} \equiv -1 = \left(\frac{a}{p}\right) \pmod{p}\,.$$

The theorem is thus proved. □

Corollary 4.7. *Let p be an odd prime.*

(1) For all $a, b \in \mathbb{Z}$, we have $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$.

(2)

$$\left(\frac{-1}{p}\right) = \begin{cases} 1 & \text{if } p \equiv 1 \pmod{4} \\ -1 & \text{if } p \equiv -1 \pmod{4} \end{cases}$$

Proof. Since $p > 2$, for integers $M, N \in \{0, \pm 1\}$ we have $M = N$ if and only if $M \equiv N \pmod{p}$. So the results are immediate from Theorem 4.6. □

Proposition 4.8. *Let p be an odd prime. Then*

$$\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}} = \begin{cases} 1 & \text{if } p \equiv \pm 1 \pmod{8} \\ -1 & \text{if } p \equiv \pm 3 \pmod{8} \end{cases}$$

To prove Prop. 4.8 we need a lemma due to Gauss.

Lemma 4.9 (Gauss)**.** *Let p be an odd prime and $a \in \mathbb{Z}$ relatively prime to p. For each $k \in \left[\!\left[1, \frac{p-1}{2}\right]\!\right]$, let $r_k = ak$ MOD p be the smallest positive residue of ak mod p. (That is, r_k is the unique integer in*

$[\![1, p-1]\!]$ *such that* $ak \equiv r_k \pmod p$*). Let*

$$s := \#\left\{k \in \left[\!\left[1, \frac{p-1}{2}\right]\!\right] \;\middle|\; r_k > \frac{p}{2}\right\}.$$

Then

$$\left(\frac{a}{p}\right) = (-1)^s.$$

Proof. By Euler's criterion

$$\begin{aligned}
\left(\frac{a}{p}\right) = (-1)^s &\iff a^{\frac{p-1}{2}} \equiv (-1)^s \pmod p \\
&\iff a^{\frac{p-1}{2}} \cdot \left(\frac{p-1}{2}\right)! \equiv (-1)^s \left(\frac{p-1}{2}\right)! \pmod p \\
&\iff \prod_{k=1}^{\frac{p-1}{2}} (ak) \equiv (-1)^s \left(\frac{p-1}{2}\right)! \pmod p \\
&\iff \prod_{k=1}^{\frac{p-1}{2}} r_k \equiv (-1)^s \left(\frac{p-1}{2}\right)! \pmod p
\end{aligned}$$

Now let

$$\{u_1, \cdots, u_s\} = \left\{r_k \;\middle|\; k = 1, \cdots, \frac{p-1}{2} \text{ and } r_k > \frac{p}{2}\right\}$$

and

$$\begin{aligned}
\{v_1, \cdots, v_t\} &= \left\{r_k \mid k = 1, \cdots, \frac{p-1}{2}\right\} \setminus \{u_1, \cdots, u_s\} \\
&= \left\{r_k \mid k = 1, \cdots, \frac{p-1}{2} \text{ and } r_k < \frac{p}{2}\right\}.
\end{aligned}$$

Thus $s + t = (p-1)/2$ and

$$\{p - u_1, \cdots, p - u_s\} \cup \{v_1, \cdots, v_t\} \subseteq \left[\!\left[1, \frac{p-1}{2}\right]\!\right].$$

Since

$$\prod_{k=1}^{\frac{p-1}{2}} r_k = u_1 \cdots u_s \cdot v_1 \cdots v_t \equiv (-1)^s (p-u_1) \cdots (p-u_s) \cdot v_1 \cdots v_t$$

to conduce the proof we need only to show

$$(p-u_1) \cdots (p-u_s) \cdot v_1 \cdots v_t = \left(\frac{p-1}{2}\right)!. \tag{4.9.1}$$

Since every term in the product of the left hand side is an integer in $[\![1, \frac{p-1}{2}]\!]$, it is sufficient to show that they are pairwise incongruent mod p.

Clearly, $(p-u_i) \not\equiv (p-u_j)$ and $v_i \not\equiv v_j \pmod p$ for distinct induces i and j, because the integers $a, 2a, \cdots, \frac{p-1}{2}a$ have distinct residues mod p. If $(p-u_i) \equiv v_j \pmod p$ for some $i \in [\![1, s]\!]$ and $j \in [\![1, t]\!]$,

then there exist distinct $k_1, k_2 \in [\![1, \frac{p-1}{2}]\!]$ such that

$$-k_1 a \equiv p - k_1 a \equiv p - u_i a \equiv v_j \equiv k_2 a \pmod{p}$$

Since $p \nmid a$, it follows that $k_1 + k_2 \equiv 0 \pmod{p}$, But this is impossible because $k_1 + k_2 \in [\![1, p-1]\!]$. The above shows that the integers $p - u_1, \cdots, p - u_s, v_1, \cdots, v_t$ are just a reordering of the integers $1, 2, \cdots, \frac{p-1}{2}$. Hence the relation (4.9.1) holds, and this proves the lemma. □

Now we can prove Prop. 4.8. i.e.,

$$\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}} .$$

Proof of Prop. 4.8. Let s be the number of least positive residues of the integers $1 \cdot 2, 2 \cdot 2, \cdots, \frac{p-1}{2} \cdot 2$ that are greater than $\frac{p}{2}$. Because these integers are all positive and less than p, they are the least positive residues mod p of themselves, Thus

$$s = \#\left\{k \in \left[\left[1, \frac{p-1}{2}\right]\right] \;\middle|\; 2k > \frac{p}{2}\right\} = \frac{p-1}{2} - \left[\frac{p}{4}\right] .$$

Since $(\frac{2}{p}) = (-1)^s$ by Gauss' lemma, it remains to show that

$$s = \frac{p-1}{2} - \left[\frac{p}{4}\right] \equiv \frac{p^2-1}{8} \pmod{2} .$$

By writing $p = 8m + k$ with $k \in \{1, 3, 5, 7\}$ (noticing that p is odd), we find that

$$s = \frac{p-1}{2} - \left[\frac{p}{4}\right] = 4m + \frac{k-1}{2} - \left(4m + \left[\frac{k}{4}\right]\right) = \frac{k-1}{2} - \left[\frac{k}{4}\right]$$

and

$$\frac{p^2-1}{8} = 8m^2 + 2mk + \frac{k^2-1}{8} \equiv \frac{k^2-1}{8} \pmod{2} .$$

So we can finish the proof by checking directly that

$$\frac{k-1}{2} - \left[\frac{k}{4}\right] \equiv \frac{k^2-1}{8} \pmod{2}$$

holds for all $k \in \{1, 3, 5, 7\}$. □

4.2 The law of quadratic reciprocity

Theorem 4.10 (Quadratic reciprocity 二次互反律)**.** *Let p and q be distinct odd primes. Then*

$$\left(\frac{p}{q}\right) \cdot \left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2} \cdot \frac{q-1}{2}}$$

Consequently

$$\left(\frac{p}{q}\right) = \begin{cases} \left(\frac{q}{p}\right) & \text{if } p \equiv 1 \pmod{4} \text{ or } q \equiv 1 \pmod{4} \\ -\left(\frac{q}{p}\right) & \text{if } p \equiv q \equiv 3 \pmod{4} \end{cases}$$

Prior to giving a proof, let us first look at how law can be used to evaluate legendre symbols.

Example 4.11.

(1) Since $17 \equiv 1 \pmod 4$, we have $\left(\frac{13}{17}\right) = \left(\frac{17}{13}\right) = \left(\frac{4}{13}\right) = \left(\frac{2}{13}\right)^2 = 1$.

(2)

$$\begin{aligned}\left(\frac{17}{19}\right) &= -\left(\frac{19}{7}\right) = (-1)\cdot\left(\frac{5}{7}\right) = (-1)\cdot\left(\frac{7}{5}\right)\\ &= (-1)\cdot\left(\frac{2}{5}\right) = (-1)\cdot(-1)^{\frac{5^2-1}{8}} = 1\end{aligned}$$

(3)

$$\begin{aligned}\left(\frac{713}{1009}\right) &= \left(\frac{23}{1009}\right)\cdot\left(\frac{31}{1009}\right) = \left(\frac{1009}{23}\right)\cdot\left(\frac{1009}{31}\right)\\ &= \left(\frac{20}{23}\right)\cdot\left(\frac{17}{31}\right)\\ &= \left(\frac{2}{23}\right)^2\cdot\left(\frac{5}{23}\right)\cdot\left(\frac{31}{17}\right) = \left(\frac{3}{5}\right)\cdot\left(\frac{-3}{17}\right)\\ &= \left(\frac{2}{3}\right)\cdot\left(\frac{-1}{17}\right)\cdot\left(\frac{17}{3}\right) = \left(\frac{-1}{17}\right)\cdot\left(\frac{2}{3}\right)^2\\ &= \left(\frac{-1}{17}\right) = 1 \quad \text{since } 17 \equiv 1 \pmod 4\end{aligned}$$

■

We now present a proof of the quadratic reciprocity law originally given by Max Eisenstein.

We begin with the following lemma, where a point $(x\,,\,y) \in \mathbb{R}^2$ is called a ***lattice point*** 格点 if x and y are both integers.

Lemma 4.12. *Let p be an odd prime and let $a \in \mathbb{Z}$ be an odd integer that is relatively prime to p. Then*

$$\left(\frac{a}{p}\right) = (-1)^{T(a,p)}$$

where

$$\begin{aligned}T(a,p) &= \sum_{j=1}^{\frac{p-1}{2}} \left[\frac{ja}{p}\right]\\ (\textit{if } a>0) &= \textit{the number of lattice points in the interior of the triangle}\\ &\left\{(x,\,y) \in \mathbb{R}^2 \mid 0 < x < \frac{p}{2},\ y > 0,\ y < \frac{a}{p}x\right\}.\end{aligned}$$

Proof. Consider the least positive residues of the integers $a, 2a, \cdots, \frac{p-1}{2}a$; let $u_1, \cdots, u_s$ be those greater than $\frac{p}{2}$ and let $v_1, \cdots, v_t$ be those less than $\frac{p}{2}$. Then for each $j \in \left[\!\left[1, \frac{p-1}{2}\right]\!\right]$, we have $ja = p\cdot\left[\frac{ja}{p}\right] + r_j$, with

$$\{r_1, r_2, \cdots, r_{\frac{p-1}{2}}\} = \{u_1, \cdots, u_s, v_1, \cdots, v_t\}.$$

Thus

$$a \cdot \sum_{j=1}^{\frac{p-1}{2}} j = \sum_{j=1}^{\frac{p-1}{2}} ja = p \cdot T(a,p) + \sum_{i=1}^{s} u_i + \sum_{k=1}^{t} v_k \tag{4.12.1}$$

As we have seen in the proof of Gauss' lemma (Lemma 4.9), $p-u_1, \cdots, p-u_s, v_1, \cdots, v_t$ are precisely the integers $1, 2, \cdots, \frac{p-1}{2}$. So we have

$$\sum_{j=1}^{\frac{p-1}{2}} j = ps - \sum_{i=1}^{s} u_i + \sum_{k=1}^{t} v_k = p \cdot s - 2 \cdot \sum_{i=1}^{s} u_i + \left(\sum_{i=1}^{s} u_i + \sum_{k=1}^{t} v_k \right) . \tag{4.12.2}$$

Subtracting (4.12.2) from (4.12.1) we get

$$(a-1) \cdot \sum_{j=1}^{\frac{p-1}{2}} j = pT(a,p) - ps + 2 \sum_{i=1}^{s} u_i$$

Notice that a and p are odd by assumption. So the above equality mod 2 yields

$$0 \equiv T(a,p) - s \pmod 2$$

Hence

$$(-1)^{T(a,p)} = (-1)^s = \left(\frac{a}{p} \right)$$

by Gauss' lemma (Lemma 4.9). □

Although Lemma 4.12 is used primarily as a tool in the proof of the quadratic reciprocity law, it can also be used to calculate Legendre symbols.

Example 4.13. To compute $(\frac{7}{11})$, we look at

$$\begin{aligned} \sum_{j=1}^{5} \left[\tfrac{7}{11} j \right] &= \left[\tfrac{7}{11} \right] + \left[\tfrac{14}{11} \right] + \left[\tfrac{21}{11} \right] + \left[\tfrac{28}{11} \right] + \left[\tfrac{35}{11} \right] \\ &= 0 + 1 + 1 + 2 + 3 = 7 . \end{aligned}$$

Hence

$$\left(\frac{7}{11} \right) = (-1)^7 = -1 .$$

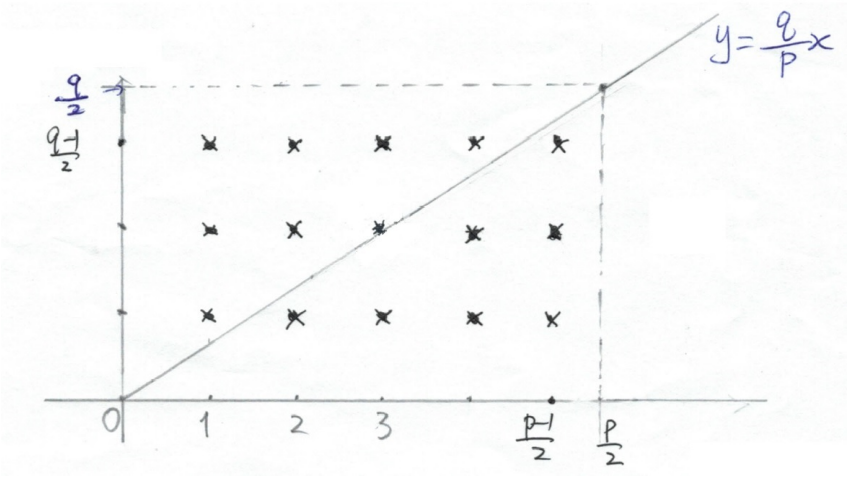
Proof of the quadratic reciprocity law. We consider the lattice points inside the rectangle

$$R := \left\{ (x,\, y) \in \mathbb{R}^2 \;\middle|\; 0 < x < \frac{p}{2}, 0 < y < \frac{q}{2} \right\} .$$

These points are exactly the elements of the Cartesian product $[\![1, \frac{p-1}{2}]\!] \times [\![1, \frac{q-1}{2}]\!]$. So there are exactly $\frac{p-1}{2} \cdot \frac{q-1}{2}$ of them.

Notice that on the diagonal of the rectangle R given by the equation $y = \frac{q}{p}x$ there exist no lattices points, because for

$$x, y \in \left[\!\left[1, \frac{p-1}{2} \right]\!\right] \times \left[\!\left[1, \frac{q-1}{2} \right]\!\right]$$



the equality $py = qx$ would imply $p \mid x$ and $q \mid y$, and these divisibility conditions cannot hold. Thus, the lattice points in R are partitioned into two sets: the set of lattice points in the upper and the lower triangles.

In the lower triangle $\{(x,\, y) \in \mathbb{R}^2 \mid 0 < x < \frac{p}{2},\ 0 < y < \frac{q}{p}x\}$, the number lattice points is $T(q,p)$. In the upper triangle $\{(x,\, y) \in \mathbb{R}^2 \mid 0 < y < \frac{q}{2},\ 0 < x < \frac{p}{q}y\}$ the number of lattice points is $T(p,q)$. By Lemma 4.12 we have

$$\left(\frac{p}{q}\right) \cdot \left(\frac{q}{p}\right) = (-1)^{T(p,q)} \cdot (-1)^{T(q,p)} = (-1)^{\frac{p-1}{2} \cdot \frac{q-1}{2}}$$

because

$$T(p,q) + T(q,p) = \text{ the number of lattice points in the rectangle } R = \frac{p-1}{2} \cdot \frac{q-1}{2}\,.$$

This completes the proof of the reciprocity law. □

As an application of the reciprocity law, we can prove the following primality test for Fermat numbers:

Theorem 4.14 (Pépin's test)**.** *The Fermat number $F_m = 2^{2^m} + 1$ is prime if and only if*

$$3^{\frac{F_m - 1}{2}} \equiv -1 \pmod{F_m}$$

Proof. First, if F_m is a prime number p, then by Euler's criterion,

$$3^{\frac{F_m-1}{2}} = 3^{\frac{p-1}{2}} \equiv \left(\frac{3}{p}\right) \pmod{p}\,.$$

Now using the reciprocity law, we find that

$$\left(\frac{3}{p}\right) = \left(\frac{3}{F_m}\right) = \left(\frac{F_m}{3}\right) = \left(\frac{2}{3}\right) = -1$$

noticing that we may assume $m \geq 1$ so that $F_m = 1 \pmod 4$ and $F_m = 2 \pmod 3$.

Conversely, if $3^{\frac{F_m-1}{2}} \equiv -1 \pmod{F_m}$, we may apply Proth's test (Thm. 3.27) to see that F_m is prime. (Here $n = F_m$ has the form $k \cdot 2^M + 1$ with $2^M := 2^{2^m} > k = 1$.) Altenatively, we may consider

a prime divisor p of F_m. The assumption $3^{\frac{F_m-1}{2}} \equiv -1 \pmod{F_m}$ implies $3^{\frac{F_m-1}{2}} \equiv -1 \pmod{p}$. Hence

$$\operatorname{ord}_p 3 \,|\, (F_m - 1) = 2^{2^m} =: 2^M$$

but

$$\operatorname{ord}_p 3 \nmid \frac{F_m - 1}{2} = 2^{2^m - 1} = 2^{M-1} .$$

This shows that

$$\operatorname{ord}_p 3 = 2^{2^m} = F_m - 1 \leq p - 1 .$$

Hence $F_m = p$ is a prime. □

Chapter 5

Arithmetic Functions and Dirichlet Series

5.1 Arithmetic functions

5.1.1 Multiplicative functions

Definition 5.1. An ***arithmetic*** (or ***arithmetical***) ***function*** (数论函数、算术函数) is any function $f : \mathbb{N}^* \to \mathbb{C}$. An arithmetic function $f : \mathbb{N}^* \to \mathbb{C}$ is called ***multiplicative*** (乘性的、积性的) if the following two conditions are satisfied:

1. f is not identically 0, i.e., $\exists n \in \mathbb{N}^*$ such that $f(n) \neq 0$;
2. For all m, $n \in \mathbb{N}^*$ with $gcd(m, n) = 1$, $f(mn) = f(m) \cdot f(n)$.

An arithmetic function f is called ***completely multiplicative*** (完全乘性的，完全积性的) if it satisfies the above condition (1) and the following $(2')$:

$2'$. For all m, $n \in \mathbb{N}^*$, $f(mn) = f(m)f(n)$. ■

(5.2) Notation. The following notation will be used frequently: Let $n \in \mathbb{N}^*$ and let $f : \mathbb{N}^* \to \mathbb{C}$ be an arithmetic function. Then

$$\sum_{d \mid n} f(d) := \text{the sum of the values } f(d) \text{ as } d \text{ runs over the positive divisors of } n\,,$$
$$\sum_{p \mid n} f(p) := \text{the sum of the values } f(p) \text{ as } p \text{ runs over the prime divisors of } n\,.$$

If $n = 1$ (so that n has no prime divisor), then we understand $\sum_{p \mid n} f(p)$ as 0.

Similarly,

$$\prod_{d \mid n} f(d) := \text{the product of the values } f(d) \text{ as } d \text{ runs over the positive divisors of } n\,,$$
$$\prod_{p \mid n} f(p) := \text{the product of the values } f(p) \text{ as } p \text{ runs over the prime divisors of } n\,.$$

If $n = 1$ (so that n has no prime divisor), then this last product is understood to be 1. □

The following lemma is easy to prove.

Lemma 5.3. *Let $f : \mathbb{N}^* \to \mathbb{C}$ be a multiplicative function.*

1. *We have $f(1) = 1$.*
2. *f is completely multiplicative if and only if for every prime number p and every $t \in \mathbb{N}$,*

$$f(p^t) = f(p)^t .$$

Proof. Since f is multiplicative by assumption, to check that f is completely multiplicative it suffices to check $f(p^t) = f(p)^t$ for all prime powers p^t. This proves (2). To see (1), note that by definition f is not identically 0. So we can choose an integer $n \in \mathbb{N}^*$ with $f(n) \neq 0$. The multiplicative property

$$f(n) = f(1 \cdot n) = f(1)f(n), \quad (\text{since } \gcd(1, n) = 1) ,$$

implies that $f(1) = 1$. □

Example 5.4. Let us list some most useful multiplicative functions:

1. Recall that Euler's phi-function $\phi : \mathbb{N}^* \to \mathbb{C}$ is defined by

$$\phi(n) = \#\{k \in [\![1, n]\!] \,|\, \gcd(k, n) = 1\} = \#(\mathbb{Z}/n\mathbb{Z})^* .$$

   We know that for every $n \in \mathbb{N}^*$,

$$\phi(n) = n \cdot \prod_{p \,|\, n} \left(1 - \frac{1}{p}\right) .$$

   Moreover, if $m, n \in \mathbb{N}^*$ are relatively prime, then $\phi(mn) = \phi(m)\phi(n)$ (Corollary 2.54), namely, ϕ is a multiplicative function.

   It is easy to see that ϕ is not completely multiplicative.

2. Fix $k \in \mathbb{R}$. The function $f(n) = n^k$ is obviously a completely multiplicative function.

3. We define the ***one-indicator*** function 验一函数

$$\mathbb{I} : \mathbb{N}^* \longrightarrow \mathbb{C}$$

   by

$$\mathbb{I}(n) = \left[\frac{1}{n}\right] = \begin{cases} 1 & \text{if } n = 1 , \\ 0 & \text{if } n > 1 . \end{cases}$$

   This is a completely multiplicative function.

4. The ***natural identity*** function 自然恒等函数

$$N : \mathbb{N}^* \longrightarrow \mathbb{C} ; \quad N(n) = n, \quad \forall\, n \in \mathbb{N}^*$$

   is a completely multiplicative function. (This is the case $k = 1$ of (2).)

5. The ***unit function*** 归一函数 $u : \mathbb{N}^* \longrightarrow \mathbb{C}$ is defined by

$$u(n) = 1, \quad \forall\, n \in \mathbb{N}^* .$$

It is completely multiplicative. (This is the case $k = 0$ of (2).) ■

Definition 5.5. The ***sum of divisors function*** 因子和函数 $\sigma : \mathbb{N} \to \mathbb{C}$ is defined by

$$\sigma(n) := \sum_{d \,|\, n} d\,, \quad n \in \mathbb{N}^* .$$

The ***number of divisors function*** 因子数函数 $\tau : \mathbb{N} \to \mathbb{C}$ is defined by

$$\tau(n) := \sum_{d \,|\, n} 1\,, \quad n \in \mathbb{N}^* .$$

More generally, for any $k \in \mathbb{R}$, we may define the function

$$\sigma_k \ : \ \mathbb{N}^* \longrightarrow \mathbb{C}\,; \quad \sigma_k(n) := \sum_{d \,|\, n} d^k\,, \quad n \in \mathbb{N}^* .$$

Note that $\sigma_0 = \tau$ and $\sigma_1 = \sigma$. ■

Definition 5.6. Let $f : \mathbb{N}^* \to \mathbb{C}$ be an arithmetic function. Its ***summatory function*** 因子取值和函数 is the function

$$F \ : \ \mathbb{N}^* \longrightarrow \mathbb{C}\,; \qquad F(n) := \sum_{d \,|\, n} f(d)\,.$$

■

Proposition 5.7. *If $f : \mathbb{N}^* \to \mathbb{C}$ is a multiplicative function, so is its summatory function $F(n) = \sum_{d \,|\, n} f(d)$.*

As an immediate corollary of this proposition, we get:

Corollary 5.8. *For every $k \in \mathbb{R}$, the function σ_k is a multiplicative function.*

In particular, the number of divisors function $\sigma_0 = \tau$ and the sum of divisors function $\sigma_1 = \sigma$ are multiplicative.

To prove Prop. 5.7, recall a lemma that was proved a long time ago.

Lemma 5.9 (= Lemma 1.42)**.** *Let m, $n \in \mathbb{N}^*$ be relatively prime. Write*

$$\begin{aligned} D &= \{ \text{ positive divisors of } mn \}\,, \\ D_1 &= \{ \text{ positive divisors of } m \}\,, \\ D_2 &= \{ \text{ positive divisors of } n \}\,. \end{aligned}$$

Then the map

$$D_1 \times D_2 \longrightarrow D\,; \quad (d_1\,,\, d_2) \mapsto d_1 \cdot d_2$$

is bijective.

In other words, every positive divisor d of mn can be written uniquely in the form

$$d = d_1 d_2 \ \text{ with } d_1 \,|\, m \text{ and } d_2 \,|\, n\,,\ d_i \in \mathbb{N}^* .$$

Proof of Prop. 5.7. Let $m, n \in \mathbb{N}^*$ be relatively prime. We need to show that

$$F(mn) = \sum_{d \,|\, mn} f(d)$$

is equal to the product of $F(m) = \sum_{d|m} f(d)$ and $F(n) = \sum_{d \,|\, n} f(d)$.

Note that when $d_1 \,|\, m$ and $d_2 \,|\, n$ we have $\gcd(d_1, d_2) = 1$ since $\gcd(m, n) = 1$. By virtue of Lemma 5.9, this is a rather straightforward computation:

$$\begin{aligned} F(mn) &= \sum_{d \in D} f(d) = \sum_{\substack{d_1 \in D_1 \\ d_2 \in D_2}} f(d_1 \cdot d_2) \\ \text{(since } f \text{ is multiplicative)} \ &= \sum_{\substack{d_1 \in D_1 \\ d_2 \in D_2}} f(d_1) f(d_2) = \sum_{d_1 \in D_1} f(d_1) \sum_{d_2 \in D_2} f(d_2) = F(m) \cdot F(n)\,. \end{aligned}$$

Here the sets D_1, D_2 and D are defined as in Lemma 5.9. □

Now we can provide formulas for $\sigma(n)$ and $\tau(n)$ for an arbitrary integer n.

Proposition 5.10. *Let $n \in \mathbb{N}^*$ have prime factorization $n = p_1^{a_1} \cdots p_s^{a_s}$ (with p_i distinct primes and $a_i \in \mathbb{N}$).*

Then

$$\sigma(n) = \prod_{i=1}^{s} \frac{p_i^{a_i+1} - 1}{p_i - 1}\,, \quad \tau(n) = \prod_{i=1}^{s} (a_i + 1)\,.$$

Proof. Since σ and τ are multiplicative, it suffices to consider the case $n = p^a$. Then

$$\{\text{positive divisors of } n\} = \{1, p, p^2\,, \cdots, p^a\}\,,$$

so the result follows immediately. □

Definition 5.11. A ***perfect number*** 完全数 is positive integer n such that $\sigma(n) = 2n$.

This definition is ancient, appearing as early as Euclid's *Elements* 几何原本. Euclid also proved that if q is a prime number of the form $q = 2^p - 1$, where p is a prime, then $n = \frac{q(q+1)}{2}$ is a perfect number. Note that since $p \geq 2$, we have $4 \,|\, (q + 1)$. So such a perfect number is even. Two millenia later, Euler proved that every even perfect number has this form.

Theorem 5.12 (Euclid–Euler)**.** *A positive integer n is an even perfect number if and only if*

$$n = \frac{q(q+1)}{2} = (2^p - 1)2^{p-1}$$

for some prime number q of the form $q = 2^p - 1$, where p is also prime.

It is worth mentioning the following definition:

Definition 5.13. For each $m \in \mathbb{N}^*$, the m-th ***Mersenne number*** 梅森数 is $M_m := 2^m - 1$. If a Mersenne number is a prime, it is called a ***Mersenne prime*** 梅森素数.

It is an easy to see that if $M_m = 2^m - 1$ is prime, then m must be a prime. ■

Proof of Thm. 5.12. First assume $n = (2^p - 1)2^{p-1}$, where $2^p - 1 = q$ is a Mersenne prime. Then

$$\sigma(n) = \sigma(2^p - 1)\sigma(2^{p-1}) = \sigma(q)\sigma(2^{p-1}) = (q+1)\frac{2^p - 1}{2 - 1} = 2^p \cdot (2^p - 1) = 2n\,,$$

so by definition, n is a perfect number.

Conversely, suppose n is an even perfect number. We write $n = 2^s \cdot t$ with $t \in \mathbb{N}^*$. Then

$$2^{s+1}t = 2n = \sigma(n) = \sigma(2^s)\sigma(t) = (2^{s+1} - 1)\sigma(t)$$

This implies $2^{s+1} \,|\, \sigma(t)$.

Let us write $\sigma(t) = 2^{s+1} \cdot r$ with $r \in \mathbb{N}^*$. The above equality gives

$$t = (2^{s+1} - 1)r\,, \quad \text{and hence } t + r = 2^{s+1}r = \sigma(t)\,. \tag{5.12.1}$$

We claim that $r = 1$. Indeed, if $r > 1$, then 1, r and t are 3 distinct positive divisors of t, Hence $\sigma(t) \geq 1 + r + t$, contradicting the equality $\sigma(t) = t + r$ in (5.12.1). Now from $r = 1$ and (5.12.1) we see that $t = 2^{s+1} - 1$ and that $\sigma(t) = t + 1$. This last equality means that t has no prime divisors other than 1 and itself, namely t is a prime. Hence $n = 2^s \cdot t = \frac{t(t+1)}{2}$ for some Mersenne prime t. This completes the proof. □

5.1.2 Dirichlet product and Möbius Inversion

We now study a less familiar, but very important, arithmetic function, called the Möbius function.

Definition 5.14. The ***Möbius function*** 莫比乌斯函数 $\mu : N^* \to \mathbb{C}$ is defined by

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1 \\ (-1)^r & \text{if } n = p_1 p_2 \cdots p_r\,, \text{ where } p_i \text{ are distinct primes} \\ 0 & \text{otherwise} \end{cases}$$

Thus, $\mu(n) = 0$ whenever n is divisible by the square of a prime. For $n > 1$, $\mu(n) \neq 0$ if and only if n is ***square-free*** 无平方因子 (i.e., not divisible by the square of a prime). ■

Example 5.15. From the definition of μ, we see that

$$\begin{array}{lll} \mu(1) = 1, & \mu(2) = \mu(3) = -1, & \mu(4) = 0 \\ \mu(5) = -1, & \mu(6) = (-1)^2 = 1, & \mu(7) = -1 \\ \mu(8) = 0, & \text{etc.} & \end{array}$$

and

$$\begin{array}{l} \mu(330) = \mu(3 \times 11 \times 2 \times 5) = (-1)^4 = 1; \\ \mu(660) = 0 \quad \text{since } 2^2 \,|\, 660\,. \end{array}$$

■

Proposition 5.16. *The Möbius function is multiplicative.*

Proof. An easy case-by-case verification. The details are left as an exercise. □

Theorem 5.17. *The summatory functon of the Möbius function is equal to the one-indicator function. That is, for all $n \in \mathbb{N}^*$,*

$$\sum_{d \mid n} \mu(d) = I(n) = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{if } n > 1 \end{cases}$$

Proof. Assume $n > 1$ and let $n = p_1^{a_1} \cdots p_k^{a_k}$ be its prime factorization. In the sum $\sum_{d \mid n} \mu(d)$ the only nonzero terms come from $d = 1$ and from those divisors d of n which are products of distinct primes. Hence

$$\begin{aligned} \sum_{d \mid n} \mu(d) &= \mu(1) + \big(\mu(p_1) + \cdots + \mu(p_k)\big) \\ &\quad + \big(\mu(p_1 p_2) + \cdots + \mu(p_{k-1} p_k)\big) \\ &\quad + \cdots\cdots \\ &\quad + \mu(p_1 p_2 \cdots p_k) \\ &= \mu(1) + \sum_{i=1}^{k} (-1)^i \cdot |D_i| \end{aligned}$$

where

$$D_i = \left\{ \begin{array}{c} \text{positive square-free divisor } d \text{ of } n \text{ that has exactly} \\ i \text{ distinct primes in its prime factorization} \end{array} \right\}$$

An element of D_i is uniquely determined by a subset with i elements of $\{p_1, \cdots, p_k\}$. Therefore $|D_i| = \binom{k}{i}$. So we obtain

$$\begin{aligned} \sum_{d \mid n} \mu(d) &= \mu(1) + \sum_{i=1}^{k} (-1)^i \, |D_i| \\ &= 1 + \sum_{i=1}^{k} (-1)^i \binom{k}{i} \\ &= \sum_{i=0}^{k} (-1)^i \binom{k}{i} = (1-1)^k = 0 \,. \end{aligned}$$

This proves the proposition. □

To fully understand the marvelous properties of the Möbius function, we need the following:

Definition 5.18. Let f and g be arithmetic function. Their ***Dirichlet product*** or ***Dirichlet convolution*** (Dirichlet 卷积) is defined to be the function $h : \mathbb{N}^* \to \mathbb{C}$ given by

$$h(n) := \sum_{d \mid n} f(d) g\left(\frac{n}{d}\right) .$$

We often write $f * g$ for this function h. ■

Lemma 5.19. *Dirichlet multiplication is commutative and associative. That is, for all arithmetic function f, g and h, we have*

$$f * g = g * f \quad \text{and} \quad (f * g) * h = f * (g * h) \,.$$

Proof. The first formula follows from the fact that the correspondence $d \mapsto d' := \frac{n}{d}$ establishes a

bijection from the set of positive divisors of n onto itself:

$$\begin{aligned}(f*g)(n) &= \sum_{d\,|\,n} f(d)g\left(\frac{n}{d}\right)\\ \text{(setting } d'=n/d) \ &= \sum_{d'\,|\,n} f(n/d')g(d') = \sum_{d'\,|\,n} g(d')f\left(\frac{n}{d'}\right) = (g*f)(n)\,.\end{aligned}$$

To prove the second formula, let us write $A = g*h$ and $B = f*g$. Then

$$\begin{aligned}(f*A)(n) &= \sum_{d\,|\,n} f(d)A\left(\frac{n}{d}\right)\\ &= \sum_{d\,|\,n} f(d)\sum_{c\,|\,\frac{n}{d}} g(c)h\left(\frac{n}{dc}\right)\\ &= \sum_{\substack{1\le d,\,c\le n\\ dc\,|\,n}} f(d)g(c)h\left(\frac{n}{dc}\right)\ .\end{aligned}\tag{5.19.1}$$

Here we use that

$$\big\{(d,\,c)\in[\![1,\,n]\!]^2\,|\,dc \text{ divides } n\big\} = \bigcup_{d\,|\,n}\left\{(d,c)\in[\![1,\,n]\!]^2\,\middle|\,c \text{ divides } \frac{n}{d}\right\}\,.$$

Similarly

$$\begin{aligned}(f*B)(n) &= \sum_{d\,|\,n} B(d)h\left(\frac{n}{d}\right) = \sum_{d'\,|\,n} h(d')B\left(\frac{n}{d'}\right)\\ &= \sum_{d'\,|\,n} h(d')\sum_{c'\,|\,\frac{n}{d'}} f(c')g\left(\frac{n}{d'c'}\right)\\ &= \sum_{\substack{1\le d',\,c'\le n\\ d'c'\,|\,n}} h(d')f(c')g\left(\frac{n}{d'c'}\right)\\ &= \sum_{\substack{1\le c',\,e'\le n\\ c'e'\,|\,n}} f(c')g(e')h\left(\frac{n}{c'e'}\right)\ .\end{aligned}\tag{5.19.2}$$

Comparison of the last summations in (5.19.1) and (5.19.2) show that $f*A = B*h$, as desired. □

Example 5.20. Let us use Dirichlet multiplication to interpret some results we have shown preciously.

1. For any arithmetic function f, its summatory function $F(n) = \sum_{d\,|\,n} f(d)$ may be viewed as the Dirichlet convolution $F = f*u$, where u is the unit function: $u(n) = 1$ for all $n\in\mathbb{N}^*$.

   Prop. 5.7 tells us that when f is multiplicative, so is the Dirichlet convolution $F = f*u$.

   In fact, one can prove that the Dirichlet product of any two multiplicative functions is multiplicative.

2. Recall that Euler's phi-function ϕ satisfies

$$\sum_{d\,|\,n}\phi(d) = n\ , \quad \text{for all } n\in\mathbb{N}^*\,.$$

Using the unit function u and the natural identity function $N(n) = n$, $n \in \mathbb{N}^*$, we may rewrite the above formula as

$$\phi * u = N \, .$$

3. Thm. 5.17 says that $\mu * u = \mathbb{I}$, where as before

   - μ is the Möbius function;
   - u is the unit function, i.e., $u(n) = 1$ for all $n \in \mathbb{N}^*$;
   - $\mathbb{I}$ is the one-indicator function, i.e., $\mathbb{I}(n) = \left[\frac{1}{n}\right]$ for all $n \in \mathbb{N}^*$. ■

The following result can be easily checked by direct calculation:

Lemma 5.21. *For every arithmetic function f, one has*

$$f * \mathbb{I} = f = \mathbb{I} * f \, .$$

The proof of the following remarkable theorem now becomes extremely easy:

Theorem 5.22 (Möbius Inversion formula 莫比乌斯反演公式)**.** *Let f and g be arithmetic functions. Then the following are equivalent:*

1. $f(n) = \sum\limits_{d \,|\, n} g(d)$ *for all $n \in \mathbb{N}^*$ (i.e., f is the summatory function of g).*
2. $g(n) = \sum\limits_{d \,|\, n} f(d)\mu\left(\frac{d}{n}\right)$ *for all $n \in \mathbb{N}^*$ (i.e., g is the Dirichlet convolution $f * \mu$).*

Proof. The assumption in (1) means that $f = g * u$. Thus,

$$f = g * u \Longrightarrow f * \mu = (g * u) * \mu = g * (u * \mu) = g * (\mu * u) = g * \mathbb{I} = g \, .$$

Here we have used the commutativity and the associativity of Dirichlet convolution (Lemma 5.19, the formula $\mu * u = \mathbb{I}$ given in Example 5.20 (3), and Lemma 5.21. Assertion (2) is the same as saying $f * \mu = g$. So we have (1)⇒(2).

Conversely, from $g = f * \mu$ we get

$$g * u = (f * \mu) * u = f * (\mu * u) = f * \mathbb{I} = f \, .$$

This shows (2)⇒(1). □

The Möbius inversion formula may be used to construct new identities that would be difficult to prove in another manner.

Example 5.23.

1. We have seen $\mu * u = \mathbb{I}$ and $\phi * u = N$ in Example 5.20.

   Thus, $\phi = \phi * \mathbb{I} = \phi * u * \mu = N * \mu$. In other words, by applying the Möbius inversion formula to the identity

   $$N(n) = \sum_{d \,|\, n} \phi(d) \, , \quad \forall \, n \in \mathbb{N}^*$$

   we obtain

   $$\phi(n) = \sum_{d \,|\, n} d{\cdot}\mu\left(\frac{n}{d}\right) = \sum_{d \,|\, n} \frac{n}{d}\mu(d) \, , \quad \forall \, n \in \mathbb{N}^* \, .$$

2. The number of divisors function $\tau(n) = \sum_{d\,|\,n} 1$ is nothing but the summatory function of the unit function u, i.e., $\tau = u * u$. The Möbius inversion yields $u = \tau * \mu$, i.e.,

$$1 = \sum_{d\,|\,n} \tau(d)\mu\left(\frac{n}{d}\right) , \quad \forall\, n \in \mathbb{N}^* .$$

3. The sum of divisors function $\sigma(n) = \sum_{d\,|\,n} d$ is the summatory function of the natural identity function N, i.e., $\sigma = N * u$. Therefore,

$$N = N * u * \mu = \sigma * \mu$$

i.e.,

$$n = \sum_{d\,|\,n} \sigma(d)\mu\left(\frac{n}{d}\right) , \quad \forall\, n \in \mathbb{N}^* .$$

■

5.2 Dirichlet series

5.2.1 Formal series and Euler products

(5.24) Let us recall some basics about infinite series, which should have been introduced in Analysis courses. A ***formal numberical series*** 形式数项级数 is a formal sum $\sum_{n\geq 1} a_n = \sum_{n=1}^{\infty} a_n$, where each $a_n \in \mathbb{C}$. The word "formal" here emphasizes that we only consider the expression $\sum_{n=1}^{\infty} a_n$ as a formal notation, and disregard the problem of convergence. Namely, we don't care about whether the formal sum $\sum_{n=1}^{\infty} a_n$ can really represent a number.

A ***formal function series*** 形式函数项级数 is a formal sum $\sum_{n=1}^{\infty} a_n(z)$, where each $a_n(z)$ is a $\mathbb{C}$-valued function in the variable z.
A most important class of functional series is that of the ***formal power series*** 形式幂级数

$$\sum_{n=1}^{\infty} a_n \cdot z^n , \qquad \text{with each } a_n \in \mathbb{C}$$

i.e., the function $a_n(z)$ is a scalar multiple of the n-th power function $z \mapsto z^n$ for each $n \in \mathbb{N}^*$. ■

(5.25) We should now discuss another kind of (formal) functional series that play a particularly important role in number theory.
Given an arithmetic function f, we define its ***formal Dirichlet series*** 形式 Dirichlet 级数 to be the series

$$D(f, s) := \sum_{n=1}^{\infty} f(n)n^{-s} = \sum_{n=1}^{\infty} \frac{f(n)}{n^s} .$$

Here, for historical reasons, the variable is denoted by the letter s. This is a functional series whose general term is a scalar multiple of the exponential function $s \mapsto n^{-s}$. The coefficients $f(n)$ are the values of the given function f.

If g is another arithmetic function, we define the addition $D(f,s)+D(g,s)$ in the obvious way:

$$D(f,s)+D(g,s) := D(f+g,s) \tag{5.25.1}$$

that is,

$$\sum_{n=1}^{\infty}\frac{f(n)}{n^s}+\sum_{n=1}^{\infty}\frac{g(n)}{n^s}=\sum_{n=1}^{\infty}\frac{f(n)+g(n)}{n^s}\,.$$

To define the multiplication $D(f,s)\cdot D(g,s)$, note that if we multiply in the usual may to obtain

$$\left(\sum_{n=1}^{\infty}\frac{f(n)}{n^s}\right)\cdot\left(\sum_{n=1}^{\infty}\frac{g(n)}{n^s}\right)=\sum_{n=1}^{\infty}\sum_{\substack{k,\,l\geq 1\\ k+l=n}}\frac{f(k)}{k^s}\frac{g(l)}{l^s}\,,$$

then it is not clear at all why the general term

$$\sum_{\substack{k,\,l\geq 1\\ k+l=n}}\frac{f(k)}{k^s}\frac{g(l)}{l^s}$$

takes the form $\frac{h(n)}{n^s}$ for a reasonably defined function h. However, if we compute formally in the way such that

$$\left(\sum_{n=1}^{\infty}\frac{f(n)}{n^s}\right)\cdot\left(\sum_{m=1}^{\infty}\frac{g(m)}{m^s}\right)=\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{f(n)g(m)}{(nm)^s}=\sum_{k=1}^{\infty}\left(\sum_{\substack{m,\,n\geq 1\\ mn=k}}f(m)g(n)\right)k^{-s}$$

then the result is naturally a Dirichlet series again: it is the Dirichlet series attached to the Dirichlet product $h=f*g$!

So, the correct definition of the product $D(f,s)\cdot D(g,s)$ should be

$$D(f,s)\cdot D(g,s) := D(f*g,s) \tag{5.25.2}$$

i.e.,

$$D(f,s)\cdot D(g,s)=\sum_{n=1}^{\infty}\left(\sum_{d\,|\,n}f(d)g\left(\frac{n}{d}\right)\right)n^{-s}=\sum_{n=1}^{\infty}\frac{(f*g)(n)}{n^s}\,.$$

Note that if s is a number such that the series

$$\sum_{n=1}^{\infty}\frac{f(n)}{n^s}\quad\text{and}\quad\sum_{n=1}^{\infty}\frac{g(n)}{n^s}$$

both converge absolutely, then in the calculation of the product $\left(\sum\frac{f(n)}{n^s}\right)\cdot\left(\sum\frac{g(n)}{n^s}\right)$ one can arrange the terms of the partial sums in any way he likes and always get a congruent series, whose limit is independent of the way he arranges the terms. In particular, if s takes such a value, then the numerical value of the product

$$D(f,s)\cdot D(g,s)=\sum_{n=1}^{\infty}\frac{(f*g)(n)}{n^s}$$

is the same as the value of

$$\sum_{n=1}^{\infty} \sum_{\substack{k,\, l\geq 1 \\ k+l=n}} \frac{f(k)}{k^s} \frac{g(l)}{l^s} \,,$$

which the is the usual product of the series $\sum_{n\geq 1} \frac{f(n)}{n^s}$ and $\sum_{n\geq 1} \frac{g(n)}{n^s}$.

Note that if $g = 0$ is the constant function with value 0, then $D(g, s) = 0$.

If $g = \mathbb{I}$ is the one-indicator function (验一函数) $n \mapsto \left[\frac{1}{n}\right]$, then the associated Dirichlet series

$$D(g, s) = D(\mathbb{I}, s) = \sum_{n=1}^{\infty} \frac{\mathbb{I}(n)}{n^s} = \frac{1}{1^s} + \frac{0}{2^s} + \frac{0}{3^s} + \cdots$$

may be identified with the constant 1.

In fact, if one prefers, he may say that the set of all Dirichlet series form a (commutative) ring, and he can even check that this ring is isomorphic to the ring of all arithmetic functions (with the usual addition and the Dirichlet product as the two operations in the definition of a ring). ■

Example 5.26. The Dirichlet series of the unit function (归一函数) $u(n) = 1\,, \forall\, n \in \mathbb{N}^*$ is denoted by

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} \,.$$

This series, or the function it represents, is the famous ***Riemann Zeta function*** 黎曼 Zeta 函数. ■

We will see that the most interesting Dirichlet series are those attached to multiplicative functions. This is to a large extent related to the usefulness of *Euler products*.

(5.27) Let f be an arithmetic function such that $f(1) = 1$. We define its ***Euler product*** 欧拉乘积 to be the formal infinite product

$$\mathcal{E}(f,\, s) = \prod_p \left(\sum_{m=0}^{\infty} \frac{f(p^m)}{p^{ms}} \right) = \prod_p \left(1 + f(p)p^{-s} + f(p^2)p^{-2s} + \cdots\right) \tag{5.27.1}$$

where the product is extended over all prime numbers p. We may rewrite $\mathcal{E}(f,\, s)$ in the form of a Dirichlet series

$$\mathcal{E}(f,\, s) = \sum_{n=1}^{\infty} \frac{\varepsilon_f(n)}{n^s} \tag{5.27.2}$$

in the following way:

For $n = 1$, we put $\varepsilon_f(1) = 1$.

For $n > 1$, we first write its prime factorization as $n = p_1^{a_1} \cdots p_r^{a_r}$. Then in the infinite product $\prod_p \left(\sum_{m=0}^{\infty} \frac{f(p^m)}{p^{ms}} \right)$ the coefficient $\varepsilon_f(n)$ of n^{-s} is only relevant to the partial product $\prod_{i=1}^{r} \left(\sum_{m=0}^{\infty} \frac{f(p_i^m)}{p_i^{ms}} \right)$.

Expanding this product into an infinite sum yields

$$\prod_{i=1}^{r}\left(\sum_{m=0}^{\infty}\frac{f(p_i^m)}{p_i^{ms}}\right)=\sum_{m_1\geq 0}^{\infty}\sum_{m_2\geq 0}^{\infty}\cdots\sum_{m_r\geq 0}^{\infty}\frac{f({p_1}^{m_1})\cdots f({p_r}^{m_r})}{({p_1}^{m_1}\cdots {p_r}^{m_r})^s}$$
$$=\sum_{k=1}^{\infty}\left(\sum_{\substack{m_1,\cdots,m_r\geq 0\\ p_1^{m_1}\cdots p_r^{m_r}=k}} f(p_1^{m_1})\cdots f(p_r^{m_r})\right)k^{-s}$$

The uniqueness statement in the fundamental theorem of arithmetic shows that for a given $k\geq 1$, the sum

$$\sum_{\substack{m_1,\cdots,m_r\geq 0\\ p_1^{m_1}\cdots p_r^{m_r}=k}} f(p_1^{m_1})\cdots f(p_r^{m_r})$$

is empty (hence equal to 0) if k has a prime divisor other than the p_i's, or is one single term

$$f(p_1^{m_1})\cdots f(p_r^{m_r})$$

if k has prime factorization $k=p_1^{m_1}\cdots p_r^{m_r}$. Therefore, the coefficient $\varepsilon_f(n)$ in the series (5.27.1) is given by

$$\varepsilon_f(n)=f(p_1^{a_1})\cdots f(p_r^{a_r}) \tag{5.27.3}$$

if $n=p_1^{a_1}\cdots p_r^{a_r}$ is the prime factorization of n. ■

Proposition 5.28. *Let f be an arithmetic function. Assume that $f(1)=1$.*

1. *f is multiplicative if and only if $D(f,s)=\mathcal{E}(f,s)$, i.e., the Dirichlet series of f coincides with Dirichlet series $\mathcal{E}(f,s)$ defined by its Euler product as in (5.27):*

$$\sum_{n=1}^{\infty}\frac{f(n)}{n^s}=\prod_p\left(\sum_{m=0}^{\infty}\frac{f(p^m)}{p^{ms}}\right).$$

2. *f is completely multiplicative if and only if*

$$\sum_{n=1}^{\infty}\frac{f(n)}{n^s}=\prod_p\left(\sum_{m=0}^{\infty}\frac{f(p)^m}{p^{ms}}\right)=\prod_p\left(1-f(p)p^{-s}\right)^{-1}.$$

Proof. We have already seen that as a Dirichlet series $\mathcal{E}(f,s)=\sum\frac{\varepsilon_f(n)}{n^s}$ has coefficients

$$\varepsilon_f(n)=f(p_1^{a_1})\cdots f(p_r^{a_r})\quad\text{ when }\ n=p_1^{a_1}\cdots p_r^{a_r}.$$

In order that $\varepsilon_f(n)=f(n)$ for all $n\geq 1$, it is sufficient and necessary that for all primes p_i and all $a_i\in\mathbb{N}$,

$$f(p_1^{a_1}\cdots p_r^{a_r})=f(p_1^{a_1})\cdots f(p_r^{a_r}).$$

This last condition means exactly that f ia a multiplicative function. This proves (1).

To finish the proof of (2), it suffices to observe that when f is known to be multiplicative, it is

completely multiplicative if and only if

$$f(p^m) = f(p)^m\,, \qquad \text{for all } m \in \mathbb{N} \text{ and all primes } p$$

(cf. Lemma 5.3). □

Example 5.29.

1. The unit function $u : n \mapsto 1$ is completely multiplicative. Therefore, for the Riemann zeta function $\zeta(s) = D(u,s) = \sum\limits_{n\geq 1} \frac{1}{n^s}$ one has the Euler product formula

$$\zeta(s) = \sum_{n\geq 1} \frac{1}{n^s} = \prod_p (1-p^{-s})^{-1}\,.$$

2. For any $k \in \mathbb{R}$, the function $f(n) = n^k$ is completely multiplicative, and we have clearly

$$D(f,s) = \sum_{n\geq 1} \frac{n^k}{n^s} = \sum_{n\geq 1} \frac{1}{n^{s-k}} = \zeta(s-k)$$

   Using Euler products we may write

$$\sum_{n\geq 1} \frac{n^k}{n^s} = \sum_{n\geq 1} \frac{1}{n^{s-k}} = \prod_p (1-p^{k-s})^{-1}\,.$$

3. Euler's ϕ-function is multiplicative, so we have the product formula

$$\begin{aligned}\sum_{n\geq 1} \frac{\phi(n)}{n^s} &= \prod_p \left(\sum_{m=0}^{\infty} \frac{\phi(p^m)}{p^{ms}}\right) = \prod_p \left(1 + \sum_{m=1}^{\infty} \frac{p^m(1-\frac{1}{p})}{p^{ms}}\right)\\ &= \prod_p \left(1 + \left(1-\frac{1}{p}\right)\frac{p^{1-s}}{1-p^{1-s}}\right) = \prod_p \frac{1-p^{-s}}{1-p^{1-s}}\\ &= \frac{\prod\limits_p (1-p^{-(s-1)})^{-1}}{\prod\limits_p (1-p^{-s})^{-1}} = \frac{\zeta(s-1)}{\zeta(s)}\,.\end{aligned}$$

   Equivalently, we get

$$D(\phi,s)\cdot D(u,s) = D(\phi,s)\cdot\zeta(s) = \zeta(s-1) = \sum_{n\geq 1}\frac{n}{n^s} = D(N,s)\,.$$

   Since $D(\phi,s)\cdot D(u,s) = D(\phi * u,\, s)$, we have re-discovered the formula

$$\phi * u = N$$

   (which is equivalent to $\phi = N * \mu$).

4. For the Möbius function μ we have

$$\begin{aligned} D(\mu, s) &= \sum_{n\geq 1} \frac{\mu(n)}{n^s} = \prod_p \left(\sum_{m\geq 0} \frac{\mu(p^m)}{p^{ms}} \right) \\ &= \prod_p (1 - p^{-s}) \quad \text{since } \mu(p^m) = \begin{cases} 1 & \text{if } m = 0 \\ -1 & \text{if } m = 1 \\ 0 & \text{if } m \geq 2 \end{cases} \\ &= \frac{1}{\prod_p (1 - p^{-s})^{-1}} = \frac{1}{\zeta(s)} \,. \end{aligned}$$

Written is an equivalent form, this gives

$$D(u, s) \cdot D(\mu, s) = D(u * \mu, s) = 1 = D(\mathbb{I},\, s)$$

so we can deduce that

$$u * \mu = \mathbb{I} \,.$$

5. For $k \in \mathbb{R}$, the function $\sigma_k(n) = \sum_{d \,|\, n} d^k$ is the Dirichlet product of the function $f(n) = n^k$ with the unit function $u(n) = 1$. One can check in at least 2 different ways that

$$D(\sigma_k,\, s) = \zeta(s) \cdot \zeta(s - k) \,.$$

The first method is to use the formulae

$$\begin{cases} \sigma_k = u * f \,, \\ D(u, s) = \zeta(s) \,, \quad \zeta(s - k) = D(f, s) \,. \end{cases}$$

The second way is to use the Euler product formula:

$$\begin{aligned} D(\sigma_k,\, s) &= \prod_p \left(\sum_{m\geq 0} \frac{\sigma_k(p^m)}{p^{ms}} \right) = \prod_p \left(\sum_{m\geq 0} \frac{p^{k(m+1)} - 1}{p^{ms}(p^k - 1)} \right) \\ &= \prod_p \left(\frac{p^k}{p^k - 1} \sum_{m\geq 0} (p^{k-s})^m - \frac{1}{p^k - 1} \sum_{m\geq 0} \frac{1}{p^{ms}} \right) \\ &= \prod_p \left(\frac{p^k}{p^k - 1} \cdot \frac{1}{1 - p^{k-s}} - \frac{1}{p^k - 1} \cdot \frac{1}{1 - p^{-s}} \right) \\ &= \prod_p \left(1 - p^{k-s} \right)^{-1} \cdot \left(1 - p^{-s} \right)^{-1} = \zeta(s) \cdot \zeta(s - k) \,. \end{aligned}$$

■

5.2.2 Dirichlet characters and L-functions

We now study another important class of multiplicative functions.
Throughout this subsection we fix an integer $k \geq 1$.

(5.30) A ***Dirichlet character*** (Dirichlet 特征) mod k is an arithmetic function $\chi : \mathbb{N}^* \to \mathbb{C}$ having the following properties:

1. $\chi(a) = 0$ if $\gcd(a, k) > 1$;
2. $\chi(1) = 1$;
3. For all $a, b \in \mathbb{N}^*$, one has $\chi(ab) = \chi(a)\chi(b)$. (Note that (2) and (3) mean exactly that χ is a completely multiplicative function)
4. $\chi(a + k) = \chi(a)$ for all $a \in \mathbb{N}^*$.

The function

$$\chi_1 : \mathbb{N}^* \longrightarrow \mathbb{C}$$

defined by

$$\chi_1(a) = \begin{cases} 1 & \text{if } \gcd(a, k) = 1 \\ 0 & \text{if } \gcd(a, k) \neq 1 \end{cases}$$

is obviously a Dirichlet character mod k. It is called the ***trivial character*** 平凡特征 or the ***principal character*** 主特征.

Note that if $k = 1$, then the situation $\gcd(a, k) > 1$ can never occur and from (2) and (4) we see that any character mod k must be identically equal to 1. This is of course not an interesting case. So we shall usually ignore this case and focus on the case with $k \geq 2$.

We denote by $\mathbf{X}(k)$ the set of all Dirichlet characters mod k. ■

(5.31) Let χ be a Dirichlet character mod k. We now show that there is a well defined function

$$\widetilde{\chi} : \mathbb{Z}/k\mathbb{Z} \longrightarrow \mathbb{C}$$

such that for any $\alpha \in \mathbb{Z}/k\mathbb{Z}$, if $a \in \mathbb{N}^*$ is a representative of α (i.e,$\alpha = [a]_k$), then $\widetilde{\chi}(a) = \chi(a)$. In fact, it is sufficient to show that if a and a' are two representatives of the congruence class α, both positive, then $\chi(a) = \chi(a')$.

We may assume $a' \geq a \geq 1$. Then the hypothesis $[a]_k = \alpha = [a']_k$ shows that $a' = a + mk$ for some $m \in \mathbb{N}$. By the periodicity of Dirichlet characters (cf. (5.30) (4)), we have $\chi(a') = \chi(a)$. The above mentioned map $\widetilde{\chi} : \mathbb{Z}/k\mathbb{Z} \to \mathbb{C}$ is thus well defined. From the definition of Dirichlet characters we can derive the following properties of the map $\widetilde{\chi}$:

1. $\widetilde{\chi}(a) = 0$ if $\alpha \notin (\mathbb{Z}/k\mathbb{Z})^*$;
2. $\widetilde{\chi}([1]_k) = 1$;
3. for all $\alpha, \beta \in (\mathbb{Z}/k\mathbb{Z})^*$, $\widetilde{\chi}(\alpha\beta) = \widetilde{\chi}(\alpha) \cdot \widetilde{\chi}(\beta)$;
4. for all $\alpha \in (\mathbb{Z}/k\mathbb{Z})^*$, $\widetilde{\chi}(\alpha)^{\phi(k)} = 1$. (In particular, $\widetilde{\chi}(\alpha) \neq 0$ for all $\alpha \in (\mathbb{Z}/k\mathbb{Z})^*$.)

Indeed, (1)–(3) are easily deduced from properties (1)–(3) of the character χ. To see (4), note that by Euler's theorem, $\alpha^{\phi(k)} = [1]_k$ for all $\alpha \in (\mathbb{Z}/k\mathbb{Z})^*$. Thus, by (2) and (3), we have

$$\widetilde{\chi}(\alpha)^{\phi(k)} = \widetilde{\chi}(\alpha^{\phi(k)}) = \widetilde{\chi}([1]_k) = \chi(1) = 1\,.$$

We may express property (4) by saying that all the nonzero values of $\widetilde{\chi}$ (or equivalently of χ) are $\phi(k)$-th roots of unity. (For $n \in \mathbb{N}^*$, an ***n-th root of unity*** (n 次単位根) is a complex number $z \in \mathbb{C}$ such that $z^n = 1$.) ■

Proposition 5.32. *Suppose that k has a primitive root.*

Then $|\mathbf{X}(k)| = \phi(k)$, i.e., there are precisely $\phi(k)$ distinct Dirichlet characters mod k.

Proof. To define a character χ mod k, it suffices to determine its values at integers that are relatively prime to k (since at the other integers its values must be 0). Let $r \in \mathbb{N}^*$ be a primitive root of k. Then for any $a \in \mathbb{N}^*$ with $\gcd(a, k) = 1$, there is a unique integer $i \in [\![1, \phi(k)]\!]$ such that

$$[a]_k = [r]_k^i = \left[r^i\right]_k \in (\mathbb{Z}/k\mathbb{Z})^* .$$

The integer i here is called the ***index*** of a mod k with respect to r. We denote it by $\mathrm{ind}_r[a]_k$.

Thus, for any $\chi \in \mathbf{X}(k)$, we have

$$\chi(a) = \chi(r^i) = \chi(r)^{\mathrm{ind}_r[a]_k} .$$

This shows that a character $\chi \in \mathbf{X}(k)$ is uniquely determined by the value $\chi(r) \in \mathbb{C}^*$.

As we have already seen in (5.31), $\chi(r)$ must be a $\phi(k)$-th root of unity. Thus, the only possible values of $\chi(r)$ are in the set

$$\left\{1, \theta, \theta^2, \cdots, \theta^{\phi(k)-1}\right\} = \{\text{all the } \phi(k)\text{-th roots of unity in } \mathbb{C}\} .$$

where $\theta = e^{2\pi\sqrt{-1}/\phi(k)} \in \mathbb{C}$. (Since the letter i is used above for another purpose, we temporarily write $\sqrt{-1}$ for the unit imaginary number that is usually denoted by $i \in \mathbb{C}$.) Therefore $|\mathbf{X}(k)| \leq \phi(k)$.

For any $j \in [\![0, \phi(k) - 1]\!]$, by setting $\chi_j(r) := \theta^j$ and

$$\chi_j(a) := \chi_j(r)^{\mathrm{ind}_r[a]_k} , \quad \text{for all } a \text{ with } \gcd(a, k) = 1 ,$$

we get a Dirichlet character $\chi_j \in \mathbf{X}(k)$. These characters are pairwise distinct, so we see that $|\mathbf{X}(k)| = \phi(k)$. This finishes the proof. □

Remark 5.33. With more group-theoretic analysis on the structure of $(\mathbb{Z}/k\mathbb{Z})^*$, one can prove that Prop. 5.32 is still true without assuming the existence of primitive roots. Namely, for every $k \in \mathbb{N}^*$, we have $|\mathbf{X}(k)| = \phi(k)$. We will admit this fact, and use in particular the finiteness of the set $\mathbf{X}(k)$. ■

Example 5.34. Let us determine all the Dirichlet characters mod k for $k = 4, 9, 10, 8$.

1. Any character mod 4 takes the value 0 at all integers that are not relatively prime to 4. So we need only determine its values at odd integers. These values depend only on the residue classes in $(\mathbb{Z}/4\mathbb{Z})^* = \{[1]_4, [3]_4\}$. For any $\chi \in \mathbf{X}(4)$, we have

$$\chi(a)^2 = \chi(a)^{\phi(4)} = 1 , \quad \text{for all } a \in \mathbb{N} \text{ such that } [a]_4 \in \{[1]_4 , [3]_4\} .$$

Since we must have $\chi(1) = 1$, the character χ is uniquely determined by its value $\chi(3) \in \{\pm 1\}$. So we find that there are exactly 2 Dirichlet characters mod 4. They are illustrated in the following table:

$a \pmod 4$	1	3	others
χ_1	1	1	0
χ_2	1	-1	0

2. Now suppose $k = 9$. It is sufficient to determine the values

$$\chi(2),\ \ \chi(4),\ \ \chi(5),\ \ \chi(7),\ \ \chi(8)$$

since

$$(\mathbb{Z}/9\mathbb{Z})^* = \{[1]_9\ ,\ [2]_9\ ,\ [4]_9\ ,\ [5]_9\ ,\ [7]_9\ ,\ [8]_9\}\,.$$

Note that $\phi(9) = 6$ and the only positive divisors of 6 are $1, 2, 3, 6$. Computation shows

$$2^2 \equiv 4\,,\ 2^3 \equiv 8 \equiv -1\,,\ 2^6 \equiv (-1) \equiv 1 \pmod 9\,.$$

So 2 is a primitive root of 9. Now since

$$\begin{aligned}
&\chi(4) = \chi(2)^2, \quad \chi(5) = \chi(2^5) = \chi(2)^5\\
&\chi(7) = \chi(-2) = \chi(2^4) = \chi(2)^4\\
&\chi(8) = \chi(2)^3
\end{aligned}$$

we only need to choose the value of $\chi(2)$, which is a sixth root of unity. The sixth roots of unity are

$$\begin{aligned}
&1\,,\ e^{\frac{\pi i}{6}} = \tfrac{1}{2} + \tfrac{\sqrt{3}}{2}i\,,\ e^{\frac{2\pi i}{3}} = -\tfrac{1}{2} + \tfrac{\sqrt{3}}{2}i\,,\ e^{\pi i} = -1\,,\\
&-e^{\frac{2\pi i}{6}} = -\tfrac{1}{2} - \tfrac{\sqrt{3}}{2}i\,,\ -e^{\frac{2\pi i}{3}} = \tfrac{1}{2} - \tfrac{\sqrt{3}}{2}i\,.
\end{aligned}$$

So we obtain the following table:

$a \mod 9$	1	2	$4 = 2^2$	$5 \equiv 2^5$	$7 \equiv 2^4$	$8 = 2^3$	others
χ_1	1	1	1	1	1	1	0
χ_2	1	$e^{\frac{2\pi i}{6}} = \frac{1}{2} + \frac{\sqrt{3}}{2}i$	$e^{\frac{2\pi i}{3}}$	$-e^{\frac{2\pi i}{3}}$	$-e^{\frac{2\pi i}{6}}$	-1	0
χ_3	1	$e^{\frac{2\pi i}{3}} = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$	$e^{\frac{4\pi i}{3}} = -e^{\frac{2\pi i}{6}}$	$-e^{\frac{2\pi i}{6}}$	$e^{\frac{2\pi i}{3}}$	1	0
χ_4	1	-1	1	-1	1	-1	0
χ_5	1	$-e^{\frac{2\pi i}{6}} = -\frac{1}{2} - \frac{\sqrt{3}}{2}i$	$e^{\frac{2\pi i}{3}}$	$e^{\frac{2\pi i}{3}}$	$-e^{\frac{2\pi i}{6}}$	1	0
χ_6	1	$-e^{\frac{2\pi i}{3}} = \frac{1}{2} - \frac{\sqrt{3}}{2}i$	$e^{\frac{4\pi i}{3}} = -e^{\frac{2\pi i}{6}}$	$e^{\frac{2\pi i}{6}}$	$e^{\frac{2\pi i}{3}}$	-1	0

3. We leave it to the reader to check the following table of Dirichlet characters mod 10:

$a \mod 10$	1	3	$7 \equiv 3^3$	$9 \equiv 3^2$	others
χ_1	1	1	1	1	0
χ_2	1	i	$-i$	-1	0
χ_3	1	-1	-1	1	0
χ_4	1	$-i$	i	-1	0

4. Now let $k = 8$. Notice that every element of $(\mathbb{Z}/8\mathbb{Z})^* = \{[1], [3], [5], [7]\}$ has order ≤ 2, and $[7] = [3] \cdot [5]$. The value $\chi(7)$ is thus uniquely determined by the values $\chi(3)$ and $\chi(5)$. Using $\chi(3), \chi(5) \in \{\pm 1\}$, we obtain:

$a \mod 8$	1	3	5	7	others
χ_1	1	1	1	1	0
χ_2	1	1	-1	-1	0
χ_3	1	-1	1	-1	0
χ_4	1	-1	-1	1	0

Careful readers may have already noticed that, in the above tables of Dirichlet characters,

- the sum of the entries in each row, except for the first one (i.e., the row for the trivial character χ_1), in equal to 0;
- the sum of the entries in each column, except for the first one (i.e., the column for $[a]_k = [1]_k$), is equal to 0. ■

We shall now prove that the above two properties actually hold in general.

Proposition 5.35 (The first orthogonality relation)**.** *Let $k \in \mathbb{N}^*$.*

1. *For any $\chi \in \mathbf{X}(k)$, we have*

$$\sum_{a \pmod k} \chi(a) = \begin{cases} \phi(k) & \text{if } \chi = \chi_1 \\ 0 & \text{otherwise} \end{cases}$$

Here the notation "$\sum_{a \pmod k}$" means that the sum is taken over all integers a in a complete system of residues mod k (e.g., all $a \in [\![1, k]\!]$).

2. *Set $r = \phi(k) = |\mathbf{X}(k)| = |(\mathbb{Z}/k\mathbb{Z})^*|$. Write*

$$\mathbf{X}(k) = \{\chi_1\,,\ \chi_2\,,\ \cdots,\ \chi_r\} \quad \text{and } (\mathbb{Z}/k\mathbb{Z})^* = \{[a_1]_k\,,\ \cdots,\ [a_r]_k\}\,.$$

Then for all i, $j \in [\![1, r]\!]$, we have

$$\frac{1}{r}\sum_{l=1}^{r} \chi_i(a_l)\overline{\chi_j(a_l)} = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Proof. (1) The formula is obvious if $\chi = \chi_1$. Now suppose $\chi \neq \chi_1$. There exists $[b]_k \in (\mathbb{Z}/k\mathbb{Z})^*$ such that $\chi(b) \neq 1$.

Putting $S := \sum_{a \pmod k} \chi(a) \in \mathbb{C}$, we have

$$\begin{aligned}\chi(b) \cdot S = \chi(b) \cdot \sum_{a \pmod k} \chi(a) &= \sum_{a \pmod k} \chi(ab) \\ &= \sum_{c \pmod k} \chi(c) = \sum_{a \pmod k} \chi(a) = S\,,\end{aligned}$$

which implies $S = 0$ since $\chi(b) \neq 1$. Here we have used the fact that when $[b]_k \in (\mathbb{Z}/k\mathbb{Z})^*$, the map

$$\mathbb{Z}/k\mathbb{Z} \longrightarrow \mathbb{Z}/k\mathbb{Z}\,; \quad [a]_k \longmapsto [ab]_k = [a]_k \cdot [b]_k$$

is a bijection.

(2) It is not difficult to check that for any $\chi\,,\, \chi' \in \mathbf{X}(k)$, the function

$$\chi''\ :\ \mathbb{N}^* \longrightarrow \mathbb{C}\,; \quad n \longmapsto \chi(n)\overline{\chi'(n)}$$

is again a Dirichlet character mod k, i.e., χ'' is completely multiplicative, periodic of period k, and satisfies $\chi''(n) = 0$ whenever $\gcd(n,k) > 1$.

Moreover, $\chi'' = \chi_1$ if and only if $\chi = \chi'$. (This follows from the fact that nonzero values of any Dirichlet character are all complex numbers of absolute value 1.)

Thus,

$$\begin{aligned}\frac{1}{r}\sum_{l=1}^{r} \chi_i(a_l)\overline{\chi}_j(a_l) = \frac{1}{r}\sum_{l=1}^{r}(\chi_i \cdot \overline{\chi}_j)(a_l) &= \frac{1}{r}\sum_{a \in (\mathbb{Z}/k\mathbb{Z})^*}(\chi_i \cdot \overline{\chi}_j)(a) \\ &= \frac{1}{r}\sum_{a \pmod k}^{r}(\chi_i \cdot \overline{\chi}_j)(a) \\ &= \begin{cases} 1 & \text{if } \chi_i \cdot \overline{\chi}_j = \chi_1 \\ 0 & \text{if } \chi_i \cdot \overline{\chi}_j \neq \chi_1 \end{cases} \\ &= \begin{cases} 1 & \text{if } \chi_i = \chi_j \\ 0 & \text{if } \chi_i \neq \chi_j \end{cases} \\ &= \delta_{ij}\end{aligned}$$

This finishes the proof. □

Remark 5.36. With notation as in Prop. 5.35 (2), if we put

$$A = \frac{1}{\sqrt{r}}\big(\chi_i(a_l)\big)_{1 \le i,\, l \le r} \in \mathbf{M}_r(\mathbb{C})\,,$$

the complex matrix with entries $A_{il} = \frac{1}{\sqrt{r}}\chi_i(a_l)$, then the proposition asserts that

$$A \cdot \overline{A}^t = I_r$$

where $\overline{A}^t$ denotes the conjugate transpose of A. That is, the matrix A is a unitary matrix. (Recall that a real unitary matrix is an orthogonal matrix.) ■

Proposition 5.37 (The second orthogonality relation)**.** *Let $k \in \mathbb{N}^*$.*

1. *Set $r = \phi(k) = |\mathbf{X}(k)| = |(\mathbb{Z}/k\mathbb{Z})^*|$ and write*

$$\mathbf{X}(k) = \{\chi_1\,,\,\chi_2\,,\,\cdots\,,\,\chi_r\} \qquad \text{and} \qquad (\mathbb{Z}/k\mathbb{Z})^* = \{[a_1]_k\,,\,\cdots\,,\,[a_r]_k\}\,.$$

   Then for all $i,\, j \in [\![1, r]\!]$, we have

$$\frac{1}{r}\sum_{l=1}^{r} \overline{\chi_l(a_i)} \cdot \chi_l(a_j) = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

2. *For any $a \in \mathbb{N}^*$, we have*

$$\sum_{\chi \,(\mathrm{mod}\, k)} \chi(a) = \begin{cases} \phi(k) & \text{if } a \equiv 1 \pmod{k} \\ 0 & \text{otherwise} \end{cases}$$

   Here the notation "$\sum_{\chi \,(\mathrm{mod}\, k)}$" means that the sum is taken over all Dirichlet characters χ modulo k, i.e., all $\chi \in \mathbf{X}(k)$.

Proof. (1) Define the matrix A as in Remark 5.36. We know from Prop. 5.35 (2) that $A \cdot \overline{A}^t = I_r$. This implies $\overline{A}^t \cdot A = I_r$. The desired formula is nothing but a re-interpretation of this equality.

(2). Using the notation of (1) we may assume $[a_1]_k = [1]_k$. Applying that formula with $i = 1$ we get

$$\frac{1}{\phi(k)}\sum_{l=1}^{r} \chi_l(a_j) = \begin{cases} 1 & \text{if } j = 1 \\ 0 & \text{otherwise} \end{cases}$$

i.e.,

$$\sum_{\chi \,(\mathrm{mod}\, k)} \chi(a) = \begin{cases} \phi(k) & \text{if } [a]_k = [1]_k \\ 0 & \text{otherwise} \end{cases}$$

This finishes the proof. □

(5.38) Let χ be a Dirichlet character mod k. The ***(Dirichlet) L-function*** or ***L-series*** (Dirichlet L-函数或 L-级数) associated to χ is the Dirichlet series

$$L(\chi, s) := D(\chi, s) = \sum_{n \geq 1} \frac{\chi(n)}{n^s}\,.$$

Since χ is completely multiplicative and vanishes at all integers that are not relatively prime to k, we have the Euler product formula

$$L(\chi, s) = \prod_{p} \left(1 - \chi(p)p^{-s}\right)^{-1} = \prod_{p \nmid k} \left(1 - \chi(p)p^{-s}\right)^{-1}\,. \tag{5.38.1}$$

If $\chi = \chi_1$ is the principal (i.e., trivial) character, then

$$\begin{aligned} L(\chi_1,\, s) &= \prod_{p \nmid k} \left(1 - \chi_1(p)p^{-s}\right)^{-1} = \prod_{p \nmid k} \left(1 - p^{-s}\right)^{-1} \\ &= \frac{\prod_p \left(1 - p^{-s}\right)^{-1}}{\prod_{p \mid k} \left(1 - p^{-s}\right)^{-1}} = \zeta(s) \cdot \prod_{p \mid k} \left(1 - p^{-s}\right) . \end{aligned} \tag{5.38.2}$$

■

5.3 Functions defined by Dirichlet series

5.3.1 Convergence of Dirichlet series

In this section, we shall study the convergence of Dirichlet series. We begin with a brief review of some basic facts in complex analysis. If the reader feels uncomfortable with functions in complex variables, he can restrict his attention to the case of real variables.

(5.39) Let $z_n = x_n + iy_n$, $n \geq 1$, with $x_n, y_n \in \mathbb{R}$ be a sequence of complex numbers. We say that the sequence (z_n) ***converges to*** a complex number $c = a + ib$ $(a, b \in \mathbb{R})$, if the following equivalent conditions hold:

1. $\lim\limits_{n\to\infty} x_n = a$ and $\lim\limits_{n\to\infty} y_n = b$.
2. $\lim\limits_{n\to\infty} |z_n - c| = 0$.

We also say that the sequence (z_n) has a limit equal to c and write $\lim\limits_{n\to\infty} z_n = c$.

Similar to the real case, a formal numerical sense $\sum\limits_{n\geq 1} z_n$ in $\mathbb{C}$ is said to be ***convergent*** if the limit $\lim\limits_{n\to\infty} \sum\limits_{k=1}^{n} z_k$ exists in $\mathbb{C}$. We say that the series $\sum\limits_{n\geq 1} z_n$ ***converges absolutely*** if the series $\sum\limits_{n\geq 1} |z_n|$ converges. Pointwise convergence and absolute convergence of functional series can be defined using the corresponding concept for numerical series. A formal functional series $\sum\limits_{n\geq 1} f(z_n)$ in the variable z is called ***convergent*** (resp. ***absolutely convergent***) at a point z_0 if the numerical series $\sum_{n\geq 1} f(z_0)$ converges (resp. converges absolutely). ■

(5.40) Recall that if $\sigma \in \mathbb{R}$, then the series $\sum_{n\geq 1} \frac{1}{n^\sigma}$ converges if and only if $\sigma > 1$. This can be shown by using ***Cauchy's integral test*** 柯西积分辨别法:

If f is a continuous real valued function from $[1, +\infty)$ to $[0, +\infty)$ and if f is monotone non-increasing, then the numerical series $\sum_{n\geq 1} f(n)$ converges if and only if the limit $\lim_{T\to+\infty} \int_1^T f(x)\mathrm{d}x$ exists in $\mathbb{R}$. ■

(5.41) Another useful fact is that if $\sigma \in \mathbb{R}$, then the series $\sum_{n\geq 1} \frac{(-1)^n}{n^\sigma}$ is convergent if and only if $\sigma > 0$. This is can be shown by using ***Leibniz's test*** 莱布尼茨判别法:

If (a_n) is a sequence of nonnegative real numbers that is monotone non-increasing and that converges to 0, then the series $\sum_{n\geq 1} (-1)^n a_n$ converges. ■

We recall some basic facts in complex analysis that are needed for the study of the convergence of Dirichlet series.

(5.42) First, the famous ***Euler formula*** says

$$e^{i\theta} = \cos\theta + i\sin\theta\,, \quad \text{for all } \theta \in \mathbb{R}\,.$$

Thus, for a complex number $s = \sigma + i\cdot t$ with $\sigma, t \in \mathbb{R}$, we have

$$e^s = e^{\sigma+it} = e^{\mathrm{Re}(s)}\,, \quad \text{where } \mathrm{Re}(s) = \text{ the real part of } s \in \mathbb{C}\,.$$

If $a \in \mathbb{R}$ and $a > 0$, we have

$$a^s = (e^{\log a})^s = e^{\sigma\log a + i\cdot t\log a}\,.$$

(Here we denote by "log" the natural logarithm, i.e., the logarithm with respect to the base e. Some others may prefer to write "ln" instead.) Thus

$$|a^s| = \left|e^{\sigma\log a + i\cdot t\log a}\right| = e^{\sigma\log a} = a^\sigma = a^{\mathrm{Re}(s)}\,.$$

If we want to study the absolute convergence of a Dirichlet series $D(f,s) = \sum\limits_{n\geq 1}\frac{f(n)}{n^s}$, we may reduce to considering the real variable series

$$\sum_{n\geq 1}\frac{|f(n)|}{|n^s|} = \sum_{n\geq 1}\frac{|f(n)|}{n^\sigma}\,, \quad \text{where } \sigma = \mathrm{Re}(s)\,.$$

■

A fundamental result in this direction is the following:

Theorem 5.43. *Let $D(f,s) = \sum\limits_{n\geq 1}\frac{f(n)}{n^s}$ be the Dirichlet series of an arithmetic function f.*

There exists $\sigma_a \in \mathbb{R}\cup\{\pm\infty\}$, called the **abscissa of absolute convergence** *(绝对收敛横坐标) of $D(f,s)$, such that*

- *for $\mathrm{Re}(s) > \sigma_a$, the series $D(f,s)$ converges absolutely;*
- *for $\mathrm{Re}(s) < \sigma_a$, the series $D(f,s)$ does not converge absolutely (but possibly converges conditionally) ("conditional convergence" means "convergence, but not absolute convergence").*

Here, $\sigma_a = -\infty$ means that $D(f,s)$ converges absolutely everywhere, and $\sigma_a = +\infty$ means that $D(f,s)$ never converges absolutely.

Proof. Let

$$A = \left\{\sigma\in\mathbb{R}\,\middle|\,\sum_{n\geq 1}\frac{|f(n)|}{n^\sigma}\ \text{converges}\right\}.$$

We may assume $A \neq \emptyset$ and $A \neq \mathbb{R}$, otherwise we can take $\sigma_a = +\infty$ or $\sigma_a = -\infty$ to conclude.

Note that it $\sigma \in \mathbb{R}$ belongs to A, then by the comparison test, the interval $[\sigma, +\infty)$ is contained in A. Since $A \neq \mathbb{R}$, the set A must have a lower bound. So the infimum $\sigma_a := \inf A \in \mathbb{R}$ exists. From what we have just said it follows that $(\sigma_a\,, +\infty) \subseteq A$ and $(-\infty, \sigma_a)\cap A = \emptyset$. This proves the theorem. □

Example 5.44. Here are some first examples of convergence of Dirichlet series.

1. The series $\zeta(s) = \sum\limits_{n\geq 1}\frac{1}{n^s}$ converges absolutely for $\mathrm{Re}(s) > 1$, and diverges for $s = 1$ (cf. (5.40)). Therefore, the abscissa of absolute convergence of $\zeta(s)$ is $\sigma_a = 1$.

2. If f is a bounded function, then $D(f,s) = \sum_{n\geq 1} \frac{f(n)}{n^s}$ converges absolutely for $\mathrm{Re}(s) > 1$, so we have $\sigma_a \leq 1$. (Since $(1,\infty) \subseteq A$, $\inf A \leq \inf(1,+\infty) = 1$.)

   If $k \in \mathbb{R}$ and $k > 0$, the bounded function $f_k(n) = \frac{1}{n^k}$ has Dirichlet series $D(f_k,s) = \zeta(s+k)$. Its abscissa of absolute convergence is therefore $\sigma_a = 1-k$.

3. If χ is a Dirichlet character modulo k for some $k \in \mathbb{N}^*$, then $|\chi(n)| \leq 1$ for all n. Thus, $D(\chi,s) = L(\chi,s) = \sum_{n\geq 1} \frac{\chi(n)}{n^s}$ has $\sigma_a \leq 1$. In other words, $L(\chi,s)$ converges absolutely for $\mathrm{Re}(s) > 1$.

4. If $f(n) = n^n$, the series $D(f,s) = \sum_{n\geq 1} \frac{n^n}{n^s}$ diverges for all s (since $\lim_{n\to\infty} |n^{n-s}| = +\infty$). Hence this series has $\sigma_a = +\infty$.

5. If $f(n) = n^{-n}$, the series $D(f,s) = \sum_{n\geq 1} \frac{1}{n^{n+s}}$ converges absolutely for all s, whence $\sigma_a = -\infty$.

   To see this, note that for all $\sigma \in \mathbb{R}$, $\frac{1}{n^{n+\sigma}} \leq \frac{1}{n^2}$ for all n large enough. So the convergence of $\sum_{n\geq 1} \frac{1}{n^2}$ implies that of $\sum_{n\geq 1} \frac{1}{n^{n+\sigma}}$. ■

(5.45) If a Dirichlet series $D(f,s) = \sum_{n\geq 1} \frac{f(n)}{n^s}$ has an abscissa of absolute convergence $\sigma_a = \sigma_a(f) < +\infty$, then $D(f,s)$ can be viewed as a function from the half-plane

$$H_{\sigma_a{}^+} := \{s \in \mathbb{C} \mid \mathrm{Re}(s) > \sigma_a\}$$

to $\mathbb{C}$. If g is another arithmetic function with abscissa of absolute convergence $\sigma_a(g) < +\infty$, then we have

$$D(f,s) \cdot D(g,s) = D(f * g, s)$$

as complex valued functions for all s with $\mathrm{Re}(s) > \max\{\sigma_a(f), \sigma_a(g)\}$. ■

(5.46) An infinite numerical product $\prod_{n=1}^{\infty} z_n$ with all $z_n \in \mathbb{C}$ is said to be ***convergent*** if

$$\text{the limit} \quad \lim_{N\to\infty} \prod_{n=1}^{N} z_n \quad \text{exists in } \mathbb{C}\,.$$

(This definition follows [SS, Chap. 5, § 3]. Some others (e.g. [Ahl, Chap. 5, § 2.2]) require further that $\lim_{N\to\infty} \prod_{n=1}^{N} z_n \neq 0$. For our purpose, the above definition is good enough.) ■

We admit the following result from analysis:

Proposition 5.47 (cf. [SS, Chap. 5, Prop. 3.1])**.** *Let $(z_n)_{n\geq 1}$ be a sequence of complex numbers such that the series $\sum_{n\geq 1} z_n$ converges absolutely.*

Then the infinite product $\prod_{n\geq 1} (1+z_n)$ converges and the product $\prod_{n\geq 1} (1+z_n)$ is nonzero unless one of its factors $1+z_n$ is 0.

Theorem 5.48. *Let $D(f,s) = \sum_{n\geq 1} \frac{f(n)}{n^s}$ be a Dirichlet series with abscissa of absolute convergence $\sigma_a < +\infty$.*

1. *If f is multiplicative, then the Euler product*

$$\mathcal{E}(f,s) = \prod_p \left(1 + \sum_{m\geq 1} \frac{f(p^m)}{p^{ms}}\right)$$

converges for $\mathrm{Re}(s) > \sigma_a$ *and one has*

$$\sum_{n\geq 1} \frac{f(n)}{n^s} = \prod_p \left(1 + \sum_{m\geq 1} \frac{f(p^m)}{p^{ms}}\right) \quad , \quad \text{for all } s \in \mathbb{C} \text{ with } \mathrm{Re}(s) > \sigma_a \, .$$

2. *If f is completely multiplicative, then*

$$\sum_{n\geq 1} \frac{f(n)}{n^s} = \prod_p \left(1 - f(p)p^{-s}\right)^{-1} , \quad \text{for } \mathrm{Re}(s) > \sigma_a \, .$$

Proof. For $\mathrm{Re}(s) > \sigma_a$, the series $\sum_{n\geq 1} \frac{|f(n)|}{|n^s|}$ converges, so the series $\sum_{m\geq 1} \frac{|f(p^m)|}{|p^{ms}|}$ converges as well. Thus $\sum_{m\geq 1} \frac{f(p^m)}{p^{ms}}$ converges to some complex number $z_p(s)$. Now for all $N \in \mathbb{N}^*$ we have

$$\sum_{p\leq N} |z_p(s)| = \sum_{p\leq N} \left| \sum_{m\geq 1} \frac{f(p^m)}{p^{ms}} \right| \leq \sum_{p\leq N} \sum_{m\geq 1} \left| \frac{f(p^m)}{p^{ms}} \right| \leq \sum_{n\geq 1} \left| \frac{f(n)}{n^s} \right| \, .$$

This proves the convergence of the series $\sum_p |z_p(s)|$. By Prop. 5.47, the product

$$\mathcal{E}(f,s) = \prod_p (1 + z_p(s)) = \prod_p \left(1 + \sum_{m\geq 1} \frac{f(p^m)}{p^{ms}}\right)$$

converges. The equality $D(f,s) = \mathcal{E}(f,s)$ for $\mathrm{Re}(s) > \sigma_a$ can be proved as in the case for formal series. This proves (1), and the assertion (2) then follows immediately. □

It is good to know the following result, whose proof we omit.

Theorem 5.49. *Let $D(f,s) = \sum_{n\geq 1} \frac{f(n)}{n^s}$ be a Dirichlet series.*

1. *There exists $\sigma_c \in \mathbb{R} \cup \{\pm\infty\}$, called the* **abscissa of convergence** *(收敛横坐标) of $D(f,s)$, such that*

   - *if $\mathrm{Re}(s) > \sigma_c$, then the series $\sum_{n\geq 1} \frac{f(n)}{n^s}$ converges;*
   - *if $\mathrm{Re}(s) < \sigma_c$, then the series $\sum_{n\geq 1} \frac{f(n)}{n^s}$ diverges.*

2. *Let $s \mapsto F(s) = \sum_{n\geq 1} \frac{f(n)}{n^s}$ be the function defined by the Dirichlet series on the half-plane*

$$H_{\sigma_c{}^+} := \{z \in \mathbb{C} \mid \mathrm{Re}(z) > \sigma_c\} \, .$$

Then F is infinitely differentiable on H_{σ_c+}, and its derivatives $F^{(k)}(s)$ can be obtained by termwise differentiating the series $\sum_{n\geq 1}\frac{f(n)}{n^s}$, that is,

$$F^{(k)}(s)=\sum_{n\geq 1}\frac{f(n)(-\log(n))^k}{n^s}\,,\qquad \textit{for } \mathrm{Re}(s)>\sigma_c\,.$$

5.3.2 Dirichlet L-functions and primes in arithmetic progressions

Our aim now is to study Dirichlet L-function and discuss as an application the main ideas of the proof of Dirichlet's theorem on primes in arithmetic progressions.

(5.50) First recall (cf. (5.38.2)) that for the principle character χ_1, we have

$$L(\chi_1,s)=\zeta(s)\cdot\prod_{p|k}(1-p^{-s})$$

where k is our fixed modulus of the characters under consideration. Thus, $L(\chi_1,s)$ differs form $\zeta(s)$ only by a finite product. It is clear that the series $\zeta(s)=\sum_{n\geq 1}\frac{1}{n^s}$ has abscissa of convergence $\sigma_c=\sigma_a=1$. Hence, $L(\chi_1,s)$ also has abscissa of convergence $\sigma_c=\sigma_a=1$. We will need more information about this function around $s=1$. ■

Proposition 5.51. *The function*

$$\psi(s):=\zeta(s)-\frac{1}{s-1}\,,\quad \textit{for } \mathrm{Re}(s)>1$$

can be extend to a **holomorphic function** 全纯函数 *on the half-plane*

$$H_{0+}:=\{z\in\mathbb{C}\mid \mathrm{Re}(z)>0\}\,.$$

(Here the word "holomorphic" means "infinitely differentiable" as a function in one complex variable. A holomorphic function is in particular continuous, differentiable, etc.)

Proof. We use the integral formula

$$\frac{1}{s-1}=\int_1^\infty t^{-s}dt=\sum_{n=1}^\infty\int_n^{n+1}\frac{1}{t^s}dt\,,\qquad \text{for } \mathrm{Re}(s)>1\,.$$

So we can write

$$\psi(s)=\zeta(s)-\frac{1}{s-1}=\sum_{n=1}^\infty\left(\frac{1}{n^s}-\int_n^{n+1}\frac{1}{t^s}dt\right)=\sum_{n=1}^\infty\int_n^{n+1}\left(\frac{1}{n^s}-\frac{1}{t^s}\right)dt\,.$$

Now put

$$\psi_n(s)=\int_n^{n+1}\left(\frac{1}{n^s}-\frac{1}{t^s}\right)dt=\frac{1}{n^s}-\int_n^{n+1}\frac{1}{t^s}dt\,.$$

Each $\psi_n(s)$ is clearly holomorphic on H_{0+}. It is sufficient to prove that the functional series $\sum_{n\geq 1}\psi_n(s)$ converges uniformly on every compact subset of H_{0+}. Notice that for fixed $n\geq 1$ and s, the function

$f_s(t) = t^{-s}$ has derivative $f_s'(t) = -st^{-(s+1)} = \frac{-s}{t^{s+1}}$. We have

$$|\psi_n(s)| \le \sup_{n\le t\le n+1} |f_s(n) - f_s(t)| \le \sup_{n\le t\le n+1} |f_s'(t)| \le \frac{|s|}{n^{1+\mathrm{Re}(s)}}\,.$$

If D is a compact subset of H_{0+}, then there exist positive real numbers $M > 0$ and $r > 0$, such that

$$\frac{|s|}{n^{1+\mathrm{Re}(s)}} \le \frac{M}{n^{1+r}}\,, \quad \text{for all } s \in D\,,\ n \ge 1\,.$$

By a criterion of Weierstrass, this implies that $\sum \psi_n(s)$ converges uniformly on D. □

Corollary 5.52. *We have*

$$\lim_{s\to 1^+} \zeta(s)\cdot(s-1) = 1\,.$$

In the language of complex analysis, this means that $\zeta(s)$ has a **simple pole** 单极点 *at $s = 1$ and its* **residue** 留数 *at $s = 1$ is 1. Here the notation "$\lim_{s\to 1^+}$" means the limit as s tends to 1 from the half-plane $H_{1+} = \{z \in \mathbb{C} \,|\, \mathrm{Re}(z) > 1\}$.*

Proof. By Prop. 5.51 we have

$$\zeta(s)\cdot(s-1) = 1 + (s-1)\psi(s)$$

where $\psi(s)$ is a continuous function on H_{0+}. In particular $\lim_{s\to 1} \psi(s) = \psi(1)$ exists in $\mathbb{C}$. The corollary thus follows immediately. □

The proof of the following corollary is left to the reader as an exercise.

Corollary 5.53. *Let χ_1 be the principal Dirichlet character mod k. Then*

$$\lim_{s\to 1^+} (s-1)L(\chi_1, s) = \frac{\phi(k)}{k}\,.$$

Next we study $L(\chi, s)$ for $\chi \ne \chi_1$.

Proposition 5.54. *Let χ be a Dirichlet character* mod k *and suppose $\chi \ne \chi_1$.*

1. *For every $N \in \mathbb{N}^*$, $\left|\sum_{n=1}^{N} \chi(n)\right| \le \phi(k)$.*
2. *The series $L(\chi, s) = \sum_{n\ge 1} \frac{\chi(n)}{n^s}$ has abscissa of convergence $\sigma_c = 0$.*

Proof. (1) By the first orthogonality relation (cf. Prop. 5.35). We have

$$\sum_{n=mk+1}^{mk+k} \chi(n) = 0\,, \quad \text{for all } m \in \mathbb{N}\,.$$

Thus, if we write $N = qk + r$ with $0 \le r \le k-1$, then

$$\begin{aligned}\sum_{n=1}^{N} \chi(n) &= \sum_{m=0}^{q-1} \sum_{n=mk+1}^{mk+k} \chi(n) + \sum_{n=qk+1}^{qk+r} \chi(n) = \sum_{n=qk+1}^{qk+r} \chi(n)\\ \text{(since } k \text{ is a period of } \chi) \quad &= \sum_{n=1}^{r} \chi(n) = \sum_{\substack{1\le n\le r\\ \gcd(n,k)=1}} \chi(n)\,.\end{aligned}$$

It then follows easily that

$$\left|\sum_{n=1}^{N}\chi(n)\right| = \left|\sum_{\substack{1\le n\le r\\ \gcd(n,\,k)=1}}\chi(n)\right| \le \sum_{\substack{1\le n\le r\\ \gcd(n,\,k)=1}}|\chi(n)| \le \phi(k)\,.$$

(2) First notice that if $\mathrm{Re}(s) < 0$, the series $\sum_{n\ge1}\frac{\chi(n)}{n^s}$ diverges, because the sequence $\frac{|\chi(n)|}{|n^s|} = \frac{|\chi(n)|}{n^{\mathrm{Re}(s)}}$ is unbounded when $\mathrm{Re}(s) < 0$. This implies that $\sigma_c \ge 0$.

To show $\sigma_c = 0$, it is sufficient to show that for all real number $\sigma > 0$, the series $\sum_{n\ge1}\frac{\chi(n)}{n^\sigma}$ converges.

Let us write $\chi(n) = A(n) + iB(n)$ with $A(n), B(n) \in \mathbb{R}$. We need only to show that the real series

$$\sum_{n\ge1}\frac{A(n)}{n^\sigma} \quad \text{and} \quad \sum_{n\ge1}\frac{B(n)}{n^\sigma} \quad \text{converges.}$$

Notice that by (1), the partial sums

$$\sum_{n=1}^{N}A(n) \quad \text{and} \quad \sum_{n=1}^{N}B(n) \quad \text{are bounded as } N\to\infty\,.$$

The sequence $(\frac{1}{n^\sigma})_{n\ge1}$ converges decreasingly to 0. So the desired convergence follows from a well known criterion of Dirichlet:

Dirichlet's convergence test: Let $(a_n)_{n\ge1}$ and $(b_n)_{n\ge1}$ be sequences of real numbers such that

- the sequence $\left(\sum_{n=1}^{m}b_n\right)_{m\ge1}$ is bounded;
- the sequence $(a_n)_{n>1}$ tends to 0 monotonously.

Then the series $\sum_{n\ge1}a_nb_n$ converges. □

The proof of the next theorem relies on more advanced techniques in analysis, so we omit its proof.

Theorem 5.55. *For any character $\chi \ne \chi_1$ mod k, one has $L(\chi, 1) \ne 0$.*

Corollary 5.56. *Let $a, k \in \mathbb{N}^*$ be such that $\gcd(a\,,\,k) = 1$. Then*

$$\lim_{s\to1^+}(s-1)\cdot\prod_{\chi\,(\mathrm{mod}\,k)}L(\chi,s)^{\overline{\chi(a)}}$$

exits and is nonzero.

Proof. There are only finitely many characters mod k and for $\chi \ne \chi_1$, we have

$$\lim_{s\to1}L(\chi,s) = L(\chi,1) \ne 0$$

by Thm. 5.55. This combined with Corollary 5.53 yields the desired result. □

(5.57) The significance of the above corollary lies in its application to Dirichlet's theorem on primes in arithmetic progressions, a remarkable theorem we will soon state. In the proof of Dirichlet's theorem, a ***logarithm function*** in complex variables is needed. As in the real case, the series $\sum_{n\ge1}\frac{(-1)^{n+1}}{n}z^n$

converges uniformly on any compact subset of the open disc $\{z \in \mathbb{C} : |z| < 1\}$. We may thus define the function

$$D_1 := \{z \in \mathbb{C} : |z| < 1\} \longrightarrow \mathbb{C}$$
$$z \longmapsto \log(1+z) := \sum_{n \geq 1} \frac{(-1)^{n+1}}{n} z^n$$

and thus we have a logarithm function

$$\log : \quad \{z \in \mathbb{C} : |z-1| < 1\} \longrightarrow \mathbb{C}$$
$$z \longmapsto \log(z) := \sum_{n \geq 1} \frac{(-1)^{n+1}}{n} (z-1)^n \,.$$

For any $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 1$, we have $|p^{-s}| < p^{-1} \leq \frac{1}{2}$ and

$$\left| \frac{1}{1-p^{-s}} - 1 \right| = \frac{|p^{-s}|}{|1-p^{-s}|} < 1 \,.$$

(It is an easy exercise to show that $|z| < \frac{1}{2}$ implies $\left|\frac{z}{1-z}\right| < 1$.)

Hence the logarithm values $\log(1-p^{-s})$ and $\log(\frac{1}{1-p^{-s}})$ are both well defined, and we have

$$\log\left(\frac{1}{1-p^{-s}}\right) = -\log(1-p^{-s}) \quad \text{when} \quad \mathrm{Re}(s) > 1 \,.$$

Similarly, for any character χ mod k, the logarithm $\log\left(1-\chi(p)p^{-s}\right)^{-1}$ is well defined for $\mathrm{Re}(s) > 1$. Using the Euler product formula, we get

$$\begin{aligned}
\log L(\chi, s) &= \sum_p \log\left(\frac{1}{1-\chi(p)p^{-s}}\right) = \sum_p (-1) \cdot \log\left(1-\chi(p)p^{-s}\right) \\
&= \sum_p (-1) \cdot \sum_{n \geq 1} \frac{(-1)^{n+1} \cdot (-p^{-s})^n \cdot \chi(p)^n}{n} \\
&= \sum_p \sum_{n \geq 1} \frac{\chi(p)^n}{np^{ns}} = \sum_p \frac{\chi(p)}{p^s} + \sum_p \sum_{n \geq 2} \frac{\chi(p)^n}{np^{ns}}
\end{aligned}$$

for $\mathrm{Re}(s) > 1$, and all $\chi \in \mathbf{X}(k)$. Thus, for $\mathrm{Re}(s) > 1$ we have

$$\begin{aligned}
\log\left(\prod_{\chi \,(\mathrm{mod}\, k)} L(\chi, s)^{\overline{\chi(a)}}\right) &= \sum_{\chi \,(\mathrm{mod}\, k)} \overline{\chi(a)} \log L(\chi, s) \\
&= \sum_{\chi \,(\mathrm{mod}\, k)} \sum_p \frac{\overline{\chi(a)}\chi(p)}{p^s} + \sum_{\chi \,(\mathrm{mod}\, k)} \sum_{n \geq 2,\, p} \frac{\chi(p^n) \cdot \overline{\chi(a)}}{np^{ns}} \\
&= \sum_p \left(\sum_\chi \overline{\chi(a)}\chi(p)\right) p^{-s} + \sum_p \sum_{n \geq 2} \frac{1}{np^{ns}} \left(\sum_\chi \overline{\chi(a)}\chi(p^n)\right) \,.
\end{aligned}$$

Using the second orthogonality relation we can further simplify the last two terms in the above formula:

$$\begin{aligned}&\sum_p \left(\sum_\chi \overline{\chi(a)}\chi(p)\right) p^{-s} + \sum_p \sum_{n\geq 2} \frac{1}{np^{ns}} \left(\sum_\chi \overline{\chi(a)}\chi(p^n)\right)\\ &= \sum_{p\equiv a \pmod k} \frac{\phi(k)}{p^s} + \sum_{\substack{n\geq 2\\ p^n\equiv a \pmod k}} \frac{\phi(k)}{np^{ns}}\end{aligned}$$

So we obtain

$$\log\left(\prod_{\chi \pmod k} L(\chi,s)^{\overline{\chi(a)}}\right) = \phi(k)\cdot\left(\sum_{p\equiv a \pmod k} \frac{1}{p^s} + \sum_{\substack{n\geq 2\\ p^n\equiv a \pmod k}} \frac{1}{np^{ns}}\right) \tag{5.57.1}$$

for $\mathrm{Re}(s)>1$. Notice that when $\sigma=\mathrm{Re}(s)>1$, we have

$$\begin{aligned}\left|\sum_{\substack{n\geq 2\\ p^n\equiv a \pmod k}} \frac{1}{np^{ns}}\right| &\leq \sum_{\substack{n\geq 2\\ p^n\equiv a \pmod k}} \frac{1}{np^{n\sigma}} \leq \sum_{n\geq 2,\, p} \frac{1}{np^n}\\ &\leq \sum_p \sum_{n\geq 2} \frac{1}{p^n} = \sum_p \frac{1}{p^2-p} = \sum_p \frac{1}{p(p-1)} = \sum_{n\geq 1} \frac{1}{(p_n-1)p_n}\end{aligned}$$

where $p_1<p_2<\cdots$ denotes the sequence of prime numbers. Since $p_j\geq j+1$ for all j,

$$\sum_{n\geq 1} \frac{1}{(p_n-1)p_n} \leq \sum_{n\geq 1} \frac{1}{n(n+1)} < +\infty\,.$$

Thus, form (5.57.1) we deduce that

$$\prod_{\chi \pmod k} L(\chi,s)^{\overline{\chi(a)}} = e^{\phi(k)\cdot\left(\sum_{p\equiv a \pmod k} \frac{1}{p^s}\right)}\cdot\rho(s) \tag{5.57.2}$$

where the function $\rho(s)$ is bounded for $\mathrm{Re}(s)>1$. ∎

Now we can easily prove the following:

Theorem 5.58 (Dirichlet)**.** *Let $a,k\in\mathbb{N}^*$ be such that $\gcd(a,k)=1$. Then*

$$\lim_{s\to 1^+} \sum_{p\equiv a \pmod k} \frac{1}{p^s} = \infty\,.$$

In particular, there are infinitely many primes in the arithmetic progression $\{km+a\,|\,m\in\mathbb{N}\}$.

Proof. If $\sum_{p\equiv a \pmod k} \frac{1}{p^s}$ is bounded as $s\to 1^+$, then from (5.57.2) we will get

$$\lim_{s\to 1^+} (s-1) \prod_{\chi \pmod k} L(\chi,s)^{\overline{\chi(a)}} = 0\,.$$

But this contradicts Corollary (5.56). □

5.3.3 Complements on the Riemann zeta function and the Riemann hypothesis

In this introductory subsection, we provide some more information about the Riemann zeta function $\zeta(s)$ that could be interesting to know.

(5.59) Our first remarks is that the values of $\zeta(s)$ at positive even integers can be expressed in terms of the Bernoulli numbers. These numbers are first studied by Jakob Bernoulli. They are define by the identify

$$\frac{x}{e^x - 1} = \sum_{n\geq 0} \frac{B_n}{n!} x^n \,. \tag{5.59.1}$$

Using the Taylor expression $e^x = \sum\limits_{n\geq 0} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \cdots$ one can compute the Bernoulli numbers inductively. For example, $B_0 = 1$, $B_1 = -\frac{1}{2}$, $B_2 = \frac{1}{6}$.

The zeta values $\zeta(2n)$ for $n \geq 1$ are given by

$$\zeta(2n) = (-1)^{n+1} \frac{(2\pi)^{2n}}{2 \cdot (2n)!} B_{2n}, \forall n \geq 1 \tag{5.59.2}$$

In particular $\zeta(2) = \sum\limits_{n\geq 1} \frac{1}{n^2} = \frac{\pi^2}{6}$. ■

(5.60) Although it is far beyond the scope of this course, a fact we would like to mention is that the function

$$s \in \{z \in \mathbb{C} \,|\, \mathrm{Re}(z) > 1\} \longmapsto \sum_{n\geq 1} \frac{1}{n^s}$$

can be extended to a holomorphic function that is defined on $\mathbb{C}^* = \mathbb{C} - \{0\}$. That function is unique, and is still denoted by $\zeta(s)$. (What people call ***Riemann zeta function*** usually means this extended function.)

In the language of complex analysis, we can say that $\zeta(s)$ has an ***analytic continuation*** 解析延拓 to the whole complex plane $\mathbb{C}$, where it has only 1 pole at $s = 1$ and $s = 1$ is a simple pole of $\zeta(s)$. For the extended Riemann zeta function, the following statements have been shown:

(1) The negative even integers $-2, -4, -6, \cdots$ are all simple zeros of $\zeta(s)$. These are called ***trivial zeros*** of $\zeta(s)$.

(2) There are no zeros of $\zeta(s)$ outside the strip $\{z \in \mathbb{C} : 0 < \mathrm{Re}(z) < 1\}$, which is often called the ***critical strip*** 临界带.

(3) Inside the critical strip $\zeta(s)$ has no real zeros, and the zeros in this strip are symmetric about the line $\mathrm{Re}(s) = \frac{1}{2}$, which is called the ***critical line*** 临界直线.

(4) As a complement to (5.59.2) we have the following special values of the zeta function

$$\begin{cases} \zeta(0) = -\frac{1}{2}\,, \\ \zeta(-2n) = 0\,, \quad \text{for } n \geq 1\,, \\ \zeta(-m) = (-1)^m \frac{B_{m+1}}{m+1}\,, \quad \text{for } m \geq 1\,. \end{cases}$$

■

The famous Riemann Hypothesis is the following conjecture.

Conjecture 5.61 (Riemann Hypothesis)**.** *All the nontrivial zeros of the Riemann zeta function* $\zeta(s)$ *lie on the critical line* $\mathrm{Re}(s) = \frac{1}{2}$.

Chapter 6

Lattices and Minkowski's Theorem

6.1 Lattice points and Minkowski's theorem

Let $n \in \mathbb{N}^*$. we shall write elements of $\mathbb{R}^n$ as row vectors and also call them ***points***. Since $\mathbb{R}^n$ has a natural structure of $\mathbb{R}$-vector space, addition, subtraction and scalar multiplication are well defined operations on points of $\mathbb{R}^n$.

A point $x = (x_1\,, \cdots, x_n) \in \mathbb{R}^n$ is called an ***integral point*** 整点 if all the x_i are integers. A lot of questions in number theory can be thought of as questions about integral points satisfying a collection of constraint conditions. It is thus interesting to know which subsets of $\mathbb{R}^n$ must contain integral points. Naturally, such a subset should not have an irregular shape.

(6.1) A nonempty subset Ω of $\mathbb{R}^n$ is called ***convex*** 凸的 if for all points $P, Q \in \mathbb{R}^n$, the line segment

$$[PQ] := \{(1-t)P + tQ \mid 0 \le t \le 1\}$$

is contained in Ω.

Here are some simple observations and examples:

- A subset of $\mathbb{R}$ is convex if and only if it is an interval.
- A regular polygon with its interior is a convex subset of $\mathbb{R}^2$.
- For any real number $r > 0$, the open ball

$$B(r) := \left\{ x \in R^n \;\middle|\; \sum_{i=1}^{n} x_i^2 < r^2 \right\}$$

is a convex subset. ■

We record a purely technical result, without proof.

Lemma 6.2. *A bounded convex subset $\Omega \subseteq \mathbb{R}^n$ is Jordan measurable, that is, the characteristic function*

$$1_\Omega \,:\, \mathbb{R}^n \longrightarrow \mathbb{R}\,; \quad x \longmapsto \begin{cases} 1 & \text{if } x \in \Omega \\ 0 & \text{if } x \notin \Omega \end{cases}$$

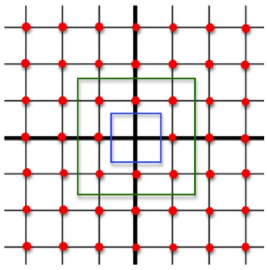
is Riemann integrable. Therefore we can define the **volume** 体积 *of* Ω *as*

$$\mathrm{Vol}(\Omega) := \int_{\mathbb{R}^n} 1_\Omega dx_1 \cdots dx_n$$

Definition 6.3. A nonempty subset $\Omega \subseteq \mathbb{R}^n$ is ***centrally symmetric*** 中心对称的 if for all $P \in \Omega$, one has $-P \in \Omega$. A ***convex body*** 凸体 in $\mathbb{R}^n$ is a nonempty, bounded, centrally symmetric convex subset. ■

The following is the first fundamental theorem in the theory of geometry of numbers.

Theorem 6.4 (Minkowski's lattice point theorem, first form)**.** *If* $\Omega \subseteq \mathbb{R}^n$ *is a convex body with* $vol(\Omega) > 2^n$*, then* Ω *must contain an integral point other than the origin.*

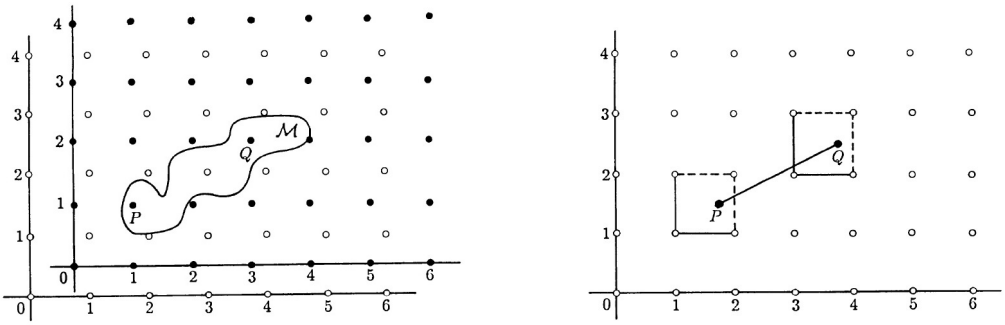


We shall prove this theorem following the approach of Blichfeldt.

Lemma 6.5. *Let* $M \subseteq \mathbb{R}^n$ *be a bounded subset with a volume* $\mathrm{Vol}(M) > 1$.

Then M *contains two distinct points* x *and* y *such that* $x - y$ *is an integral point (not necessarily in* M*).*

Geometrically, there are two ways to interpret the result of Lemma 6.5. First, the lemma says that we can translate the lattice of all integral points in $\mathbb{R}^n$, meanwhile leaving M unmoved, so that M will contain at least two points of the new lattice. Alternatively, if we don't move any point but simply look at the relative position of each point of M with respect to the unit cube with integral points as its vertices and surrounding that point, then there are two distinct points of M for which the relative positions are the same.



Proof of Lemma 6.5. Let 1_M denote the characteristic function of M: $1_M(x) = \begin{cases} 1 & \text{if } x \in M \\ 0 & \text{if } x \notin M \end{cases}$. Con-

sider the function

$$\psi : \mathbb{R}^n \longrightarrow \mathbb{R} ; \quad \psi(x) := \sum_{g \in \mathbb{Z}^n} 1_M(x+g)$$

where the sum is over all integral points $g \in \mathbb{Z}^n$. Since M is bounded, the sum $\sum_{g\in\mathbb{Z}^n} 1_M(x+g)$ is in fact a finite sum. So the function is well defined, and clearly it takes only nonnegative integer values.

Now let ε be the unit cube defined by

$$\mathcal{E} := \{x = (x_1 , \cdots , x_n) \in \mathbb{R}^n \mid 0 \le x_i < 1 , \text{ for all } i\} .$$

The whole space $\mathbb{R}^n$ is clearly covered by all the translations of $\mathcal{E}$, i.e.,

$$\mathbb{R}^n = \bigcup_{g\in\mathbb{Z}^n} (\mathcal{E} + g) .$$

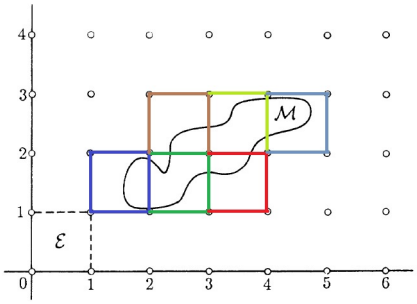
Now consider the integral

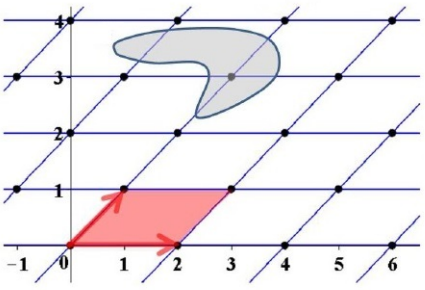
$$I := \int_{\mathcal{E}} \psi(x)dx = \sum_g \int_{\mathcal{E}} 1_M(x+g)dx = \sum_g \int_{\mathcal{E}+g} 1_M(y)dy = \int_{\mathbb{R}^n} 1_M(y)dy = \mathrm{Vol}(M) > 1 .$$

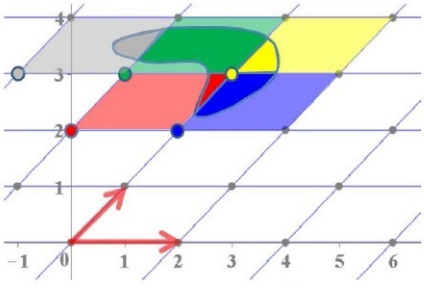
From its definition, $\psi(x)$ is a nonnegative, bounded integer for all $x \in \mathcal{E}$, therefore it must attain a maximum, say m, in $\mathcal{E}$. Thus

$$m \ge I = \int_{\mathcal{E}} \psi(x)dx = \mathrm{Vol}(M) > 1 .$$

Since $m \in \mathbb{Z}$, we have $m \ge 2$. Let $x_0 \in \mathcal{E}$ be a point such that $\psi(x_0) = m \ge 2$. Then there are at least two distinct points $g, g' \in \mathbb{Z}^n$ such that $1_M(x_0 + g) = 1_M(x_0 + g') = 1$. Taking $x = x_0 + g$ and $y = x_0 + g'$ finishes the proof. □







Proof of Theorem 6.4. Let $M := \frac{1}{2}\Omega = \left\{\frac{1}{2}x | x \in \Omega\right\}$. Then

$$\mathrm{Vol}(M) = \frac{1}{2^n}\mathrm{Vol}(\Omega) > 1\,.$$

By Lemma 6.5, there exist two distinct points $x, y \in M$ such that $g := x - y$ is an integral point. Now $2x$ and $2y$ are points in Ω, and since Ω is centrally symmetric, $-2y \in \Omega$. The convexity of Ω then implies

$$g = \frac{1}{2}\cdot(2x) + \frac{1}{2}\cdot(-2y) \in \Omega\,.$$

This proves the theorem. □

In practice, it turns out more convenient if we consider not only the integral points themselves but also their images under linear automorphisms of $\mathbb{R}^n$. We therefore introduce the following definition.

Definition 6.6. A (***complete***) ***lattice*** (完备) 格 in $\mathbb{R}^n$ is a subset Λ of the form

$$\Lambda = \mathbb{Z}v_1 \oplus \cdots \oplus \mathbb{Z}v_n := \{a_1v_1 + \cdots + a_nv_n \mid a_i \in \mathbb{Z}\}$$

where $v_1\,, \cdots, v_n$ is a basis of $\mathbb{R}^n$ as a vector space. Equivalently, $\Lambda = T(\mathbb{Z}^n)$ is the image of the lattice $\mathbb{Z}^n$ of all integral points under a linear automorphism T of $\mathbb{R}^n$.

If $A = \mathcal{M}(T)$ is the matrix of T with respect to a basis of $\mathbb{R}^n$, then the absolute value $|\det(T)| = |\det(A)|$ is called the ***determinant*** 行列式 of the lattice Λ, denoted by $\det(\Lambda)$. We may choose A to be the matrix whole columns are the coordinates of the basis vectors $v_1, \cdots, v_n$. The determinant $\det(\Lambda)$ is equal to $|\det(T)| = |\det(A)|$. The determinant $\det(A) = \det(T)$ is independent of the choice of the basis, so the same is true for $\det(\Lambda)$.

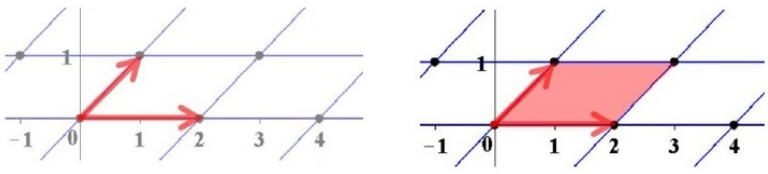
If $\mathcal{E}$ denote the unit cube

$$\mathcal{E} = \{x = (x_1\,, \cdots, x_n) \in R^n \mid 0 \le x_i < 1\,, \text{ for all } i\}$$

then

$$\det(\Lambda) = \mathrm{Vol}(T(\mathcal{E})) = |\det(T)| \cdot \mathrm{Vol}(\mathcal{E}) = |\det(T)|\,.$$

The image $g := T(\mathcal{E})$ is called the ***fundamental parallelepiped*** 基本平行六面体 of the lattice Λ. ■



We leave it to the reader to prove the following easy lemma.

Lemma 6.7. *Let T be a linear automorphism of $\mathbb{R}^n$ and let $\Omega \subseteq \mathbb{R}^n$ be a nonempty subset.*

Let $\mathcal{P}$ be one of the following properties:

(a) bounded;

(b) convex;

(c) centrally symmetric;

(d) Jordan measurable.

Then Ω has the property $\mathcal{P}$ if and only if so does $T(\Omega)$.
When Ω is measurable, one has

$$\mathrm{Vol}(T(\Omega)) = |\det(T)| \cdot \mathrm{Vol}(\Omega)\,.$$

Now we can state the second form of Minkowski's lattice point theorem.

Theorem 6.8 (Minkowski's lattice point theorem, second form)**.** *Let Λ be a (complete) lattice in $\mathbb{R}^n$ and let $\Omega \subseteq \mathbb{R}^n$ be a convex body such that $\mathrm{Vol}(\Omega) > 2^n \cdot \det(\Lambda)$.*
Then Ω must contain a nonzero lattice point of Λ.

Proof. Suppose $\Lambda = T(\mathbb{Z}^n)$ for some linear automorphism $T : \mathbb{R}^n \to \mathbb{R}^n$. According to Lemma 6.7, we may apply Thm. 6.4 to $T^{-1}(\Omega) \subseteq \mathbb{R}^n$. □

Exercise 6.9. Let K be a circle of diameter 26.2 m centered at the origin. Trees of diameter 0.16 m grow at each lattice point within K except for the origin, which is where Shrek is standing.
Prove: Shrek can't see outside this mini forest.

6.2 Applications of Minkowski's theorem

6.2.1 Sums of squares

Now we use Minkowski's theorem to solve a number theoretic problem.

Lemma 6.10. *If p is a prime such that $p \equiv 1 \pmod 4$, then there exist integers $x, y \in \mathbb{Z}$ such that $p = x^2 + y^2$.*

Proof. We consider the convex body $\Omega \subseteq \mathbb{R}^2$ given by

$$\Omega := \left\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 2p\right\}\,.$$

It is sufficient to find a lattice Λ satisfying the following conditions

(1) $\Lambda \subseteq \mathbb{Z}^2$;

(2) For all $(x, y) \in \Lambda$, $\|(x, y)\|^2 := x^2 + y^2$ is divisible by p;

(3) Ω contains a nonzero point of the lattice Λ.

The third condition is satisfied as soon as

$$2^2 \cdot |\det(\Lambda)| < \mathrm{Vol}(\Omega) = 2\pi p\,. \tag{6.10.1}$$

by Minkowski's theorem (Thm. 6.8). Condition (1) holds if we take $\Lambda = T(\mathbb{Z}^2)$, where $T : \mathbb{R}^2 \to \mathbb{R}^2$ is defined by

$$\begin{pmatrix} s \\ t \end{pmatrix} \mapsto A \begin{pmatrix} s \\ t \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} s \\ t \end{pmatrix}$$

for some matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with integer entries $a, b, c, d \in \mathbb{Z}$. For such a lattice, condition (2) means that

$$(as + bt)^2 + (cs + dt)^2 \equiv 0 \pmod{p}, \quad \text{for all } s, t \in \mathbb{Z}$$

i.e.,

$$(a^2 + c^2)s^2 + 2(ab + cd)st + (b^2 + d^2)t^2 \equiv 0, \quad \text{for all } s, t \in \mathbb{Z}.$$

This is clearly equivalent to

$$a^2 + c^2 \equiv ab + cd \equiv b^2 + d^2 \equiv 0 \pmod{p}. \tag{6.10.2}$$

Combining (6.10.1) and (6.10.2) we obtain

$$\begin{cases} a^2 + c^2 \equiv ab + cd \equiv b^2 + d^2 \equiv 0 \pmod{p} \\ |ad - bc| < \frac{\pi p}{2} \end{cases} \tag{6.10.3}$$

It remains to find integers $a, b, c, d \in \mathbb{Z}$ satisfying (6.10.3). Notice that taking $b = 0$ and $d = p$ simplifies the condition to the following:

$$\begin{cases} p \cdot |a| < \frac{\pi p}{2} \\ a^2 + c^2 \equiv 0 \pmod{p} \end{cases}$$

The inequality is clearly true when $a = 1$. Then the existence of $c \in \mathbb{Z}$ satisfying $1 + c^2 \equiv 0 \pmod{p}$ is guaranteed by the fact that $\left(\frac{-1}{p}\right) = 1$. (Here we use the assumption $p \equiv 1 \pmod{4}$.) □

Theorem 6.11 (Sum of 2 squares theorem)**.** *A positive integer n is a sum of two squares of integers if and only if for every prime p such that $p \equiv 3 \pmod{4}$, $v_p(n)$ is even.*

In other words, if n has prime factorization

$$n = 2^a p_1^{b_1} \cdots p_r^{b_r} q_1^{c_1} \cdots q_s^{c_s}, \quad \text{with } a, b_i, c_j \in \mathbb{N}$$

where p_i are distinct primes congruent to $1 \mod 4$ and and q_j are distinct primes congruent to $3 \mod 4$, then the equation $n = x^2 + y^2$ has integer solutions if and only if all c_j are even.

Proof. We first prove that if n has the form described in the theorem, then it is a sum of 2 squares (of integers). Indeed, if all c_i are even, then $m := {q_1}^{c_1} \cdots {q_s}^{c_s}$ is a square, hence $m = x^2 + y^2$ has integer solutions. Each p_i is a sum of 2 integer squares by Lemma 6.10. Also, it is clear that $2 = 1^2 + 1^2$ is a sum of 2 squares. Now it suffices to notice that the set of integers that are expressible as the sum of 2 integer squares is closed under multiplication, thanks to the following identify:

$$(a^2 + b^2)(c^2 + d^2) = (ac - bd)^2 + (ad + bc)^2.$$

(This identify is related to the formula for modulus of complex numbers:

$$|z_1|^2 \cdot |z_2|^2 = |z_1 z_2|^2$$

if $z_1 = a + bi$ and $z_2 = c + di$.)

It remains to show that if $n = x^2 + y^2$ has integer solutions, then $v_p(n)$ is even for all primes p such that $p \equiv 3 \pmod{4}$.

We prove this statement by a contradiction.

Suppose that n has a prime divisor p which is congruent to 3 mod 4, and suppose that there exist $x, y \in \mathbb{Z}$ such that $n = x^2 + y^2$. Assume that $v_p(n)$ is odd. Then we must have $xy \neq 0$ (otherwise $v_p(n) = 2v_p(x)$ or $2v_p(y)$, which is even). Moreover, by properties of the p-adic valuation, we have $v_p(x) = v_p(y)$. For otherwise, $v_p(x^2) = 2v_p(x) \neq 2v_p(y) = v_p(y^2)$, and this implies that

$$v_p(x^2 + y^2) = \min(v_p(x^2), v_p(y^2)) \equiv 0 \pmod 2 .$$

But this contradicts our assumption that $v_p(n)$ is odd.

Thus, we have $v_p(x) = v_p(y) = t$ for some $t \in \mathbb{N}$. Putting $x_1 = x/p^t$, $y_1 = y/p^t$ and $m = n/p^{2t}$, we get $m = x_1^2 + y_1^2$. (Note that $m \in \mathbb{N}$ because $v_p(n) = v_p(x^2 + y^2) \geq \min(v_p(x^2), v_p(y^2)) = 2t$.) Since $v_p(m) = v_p(n) - 2t$ is odd and ≥ 0. We have $m \equiv 0 \pmod p$. Hence

$$x_1^2 + y_1^2 \equiv 0 \pmod p .$$

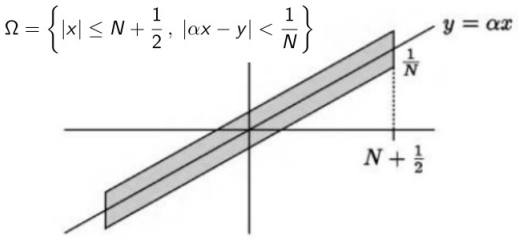
Since $v_p(x_1) = v_p(x) - t = 0$, x_1 is invertible mod p. Letting $u \in \mathbb{Z}$ be an inverse of x_1 mod p we see that $1 + (uy_1)^2 \equiv 0 \pmod p$. This leads to a contradiction to the fact $\left(\frac{-1}{p}\right) = -1$ (recall that we have assumed $p \equiv 3 \pmod 4$). We have thus finished the proof. □

Remark 6.12. Just for information, let us mention that around 1770, Lagrange proved that every positive integer can be written as a sum of four integer squares (the 4 squares theorem). This result also has a proof that is based on Minkowski's lattice point theorem. (We indicate such a proof in the problem set.) Another fact is that in 1798, Legendre proved that a positive integer n is a sum of 3 integer squares if and only if it is ***not*** of the form $n = 4^a(8b + 7)$, with $a, b \in \mathbb{N}$. ■

6.2.2 Dirichlet's approximation theorem

The following theorem was first proved by Dirichlet in 1834, using the pigeonhole principle.

Theorem 6.13 (Dirichlet's approximations theorem)**.** *Let $\alpha \in \mathbb{R}$ and let N be a positive integer. Then there exists a pair of integers m, n such that $1 \leq n \leq N$ and $|n\alpha - m| < \frac{1}{N}$.*



Proof. Consider the convex body

$$\Omega := \left\{ (x, y) \in \mathbb{R}^2 \;\middle|\; |x| \leq N + \frac{1}{2},\ |\alpha x - y| < \frac{1}{N} \right\} .$$

Then $\mathrm{Vol}(\Omega) = 2(N + \frac{1}{2}) \cdot \frac{2}{N} > 4$. Hence Ω contains a nonzero lattice point $(n, m) \in \mathbb{Z}^2$. By symmetry, we may assume $n > 0$, and the definition of Ω gives $n \leq N$ and $|n\alpha - m| < \frac{1}{N}$. This completes the proof. □

Bibliography

[Ahl] Lars V. Ahlfors, *Complex Analysis (third edition)*, McGraw-Hill Inc., 1979.

[Ros] Kenneth H. Rosen, *Elementary Number Theory and its applications (6th edition)*, Pearson, 2011. 国内有机械工业出版社的影印版.

[SS] Elias M., Stein and Rami, Shakarchi, Rami, *Complex analysis*, Princeton University Press, Princeton, NJ, 2003.

[Tao] Terence Tao, *Analysis I, (Third edition)*, Springer, 2016. 中文翻译版书名《陶哲轩实分析 (第 3 版)》, 李馨 译, 人民邮电出版社, 2018.