A Senior Seminar On Quadratic Reciprocity

Angel V. Kumchev

Contents

Chapter 1. Elementary Number Theory: A Review 1

1.1. Divisibility 1

1.2. The greatest common divisor of two integers 2

1.3. Prime numbers and the fundamental theorem of arithmetic 5

1.4. Congruence modulo *m* 5

1.5. Complete and reduced systems modulo *m* 8

1.6. Exercises 9

Chapter 2. The Law Of Quadratic Reciprocity 11

2.1. Polynomial congruences 11

2.2. Linear congruences 12

2.3. Quadratic congruences. Quadratic residues and nonresidues 13

2.4. The Legendre symbol 14

2.5. Quadratic reciprocity 16

2.6. Exercises 17

Chapter 3. Complex Numbers 19

3.1. Definition and algebraic properties 19

3.2. Geometric interpretation, moduli and conjugates 20

3.3. Exponential form 22

3.4. Roots of complex numbers 23

3.5. The complex exponential 23

3.6. Trigonometric and hyperbolic functions of a complex argument 24

3.7. The complex logarithm and power functions 26

3.8. Gauss sums 28

3.9. Exercises 30

Chapter 4. Algebra Over The Complex Numbers 33

4.1. Roots of polynomials with complex coefficients 33

4.2. Linear algebra over the complex numbers 35

4.3. Review of determinants 40

4.4. The resultant of two polynomials 44

4.5. Symmetric polynomials in several variables 45

4.6. Exercises 46

Chapter 5. Fourier Series 51

5.1. Definition 51

5.2. Calculus of complex-valued functions 51

5.3. An example 53

5.4. Representing functions by Fourier series 54

5.5. Poisson’s summation formula 56
5.6. Exercises 57

Chapter 6. Algebraic Numbers And Algebraic Integers 61
6.1. Definition 61
6.2. The ring of algebraic integers 61
6.3. Congruences between algebraic integers 63
6.4. A proof of Theorem 2.17 64
6.5. Exercises 65

Chapter 7. Proofs Of The Law Of Quadratic Reciprocity 67
7.1. S.Y. Kim’s elementary proof 67
7.2. Eisenstein’s lemma and counting lattice points 68
7.3. Another proof of Eisenstein’s 69
7.4. An algebraic proof: Gauss sums are algebraic integers 69
7.5. A proof using the formula for the Gauss sum 70
7.6. An evaluation of the Gauss sum using Fourier series 71
7.7. An evaluation of the Gauss sum using matrices 72
7.8. A proof using polynomials and resultants 73

CHAPTER 1

Elementary Number Theory: A Review

As we said already in the Introduction, the focus of this course is a theorem from number theory known as the *law of quadratic reciprocity*. Thus, it should not come as a surprise that we shall need certain facts from elementary number theory. In this chapter, we review some of the basics of arithmetic: the definition of divisibility, the greatest common divisor of two integers, the definition and basic properties of prime numbers, and properties of congruences. It is very likely that you are familiar with much of the material in the chapter from earlier courses, so we aim mainly to refresh your knowledge, fill any potential gaps in it, and set up the notation and terminology for future reference. The pace of the exposition is thus rather brisk, and several proofs are left as exercises.

1.1. Divisibility

Number theory studies the properties of the integers, and that study usually starts with the notion of divisibility of one integer by another.

Definition 1.1. If $a, b \in \mathbb{Z}$ and $b \neq 0$, we say that b *divides* a and write $b \mid a$, if $a = bq$ for some $q \in \mathbb{Z}$. If b does not divide a, we write $b \nmid a$. When $b \mid a$, we may also say that b is a *divisor* of a, b is a *factor* of a, or a is a *multiple* of b.

Example 1.2. We have $3 \mid 15$ and $6 \mid 324$, but $100 \nmid 2010$.

Given two integers a and $b \neq 0$, we sometimes want to divide a by b even though b may not be a divisor of a: say, we may want to divide 37 candy bars among 4 children. This problem leads to the concept of "division with a quotient and a remainder." It is summarized by the following theorem, known as the *quotient-remainder theorem* or the *division algorithm*.

Theorem 1.3 (Quotient-remainder theorem)**.** *If $a, b \in \mathbb{Z}$ and $b > 0$, then there exist unique integers q and r such that*

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

Proof. "Existence." Let q be the largest integer m with $m \leq a/b$ and let $r = a - bq$. Then

$$q \leq a/b \quad \Longrightarrow \quad 0 \leq a - bq = r.$$

Furthermore, since q is the largest integer not exceeding a/b, we have

$$a/b < q + 1 \quad \Longrightarrow \quad a < bq + b \quad \Longrightarrow \quad r = a - bq < b.$$

"Uniqueness." Suppose that q_1, q_2, r_1 and r_2 are integers such that

$$a = bq_1 + r_1 = bq_2 + r_2, \quad 0 \leq r_1, r_2 < b.$$

Then

$$bq_1 + r_1 = bq_2 + r_2 \quad \Longrightarrow \quad r_1 - r_2 = b(q_2 - q_1).$$

On the other hand,

$$0 \leq r_1, r_2 < b \quad \Longrightarrow \quad -b < r_1 - r_2 < b.$$

Hence,

$$-b < b(q_2 - q_1) < b \quad \Longrightarrow \quad -1 < q_2 - q_1 < 1.$$

Since $q_2 - q_1$ is an integer, we conclude that $q_2 - q_1 = 0$. Thus, $q_1 = q_2$ and

$$bq_1 + r_1 = bq_2 + r_2 \quad \Longrightarrow \quad r_1 = r_2. \qquad \square$$

1.2. The greatest common divisor of two integers

Definition 1.4. Let a and b be integers, not both 0. The *greatest common divisor of a and b*, denoted $\gcd(a, b)$ or (a, b), is the largest natural number d such that $d \mid a$ and $d \mid b$.

Example 1.5. The positive divisors of 2 are 1 and 2, and the positive divisors of 6 are 1, 2, 3 and 6. Thus, $(2, 6) = 2$.

The positive divisors of 4 are 1, 2 and 4, and the positive divisors of 6 are 1, 2, 3 and 6. Thus, $(4, 6) = 2$.

The positive divisors of -4 are 1, 2 and 4, and the positive divisors of 0 are all natural numbers. Thus, $(-4, 0) = 4$.

If $a \neq 0$, then any positive divisor of a is $\leq |a|$. Moreover, $|a|$ (which is either a or $-a$) is a divisor of a, so $|a|$ is the largest positive divisor of a. Thus, $(a, 0) = |a|$. □

In the above example, we computed $(a, 0)$ for all $a \neq 0$. We want to find a method for computing the gcd of any two numbers. The next lemma reduces that problem to the case when a and b are natural numbers.

Lemma 1.6. *If a and b are integers and $b \neq 0$, then $(a, b) = (a, |b|)$.*

Lemma 1.7. *Let b, q, and r be integers, and suppose that b and r are not both 0. Then*

$$(bq + r, b) = (r, b).$$

Proof. Let $a = bq + r$, $d_1 = (a, b)$, and $d_2 = (r, b)$. By the definition of (r, b), $d_2 \mid r$ and $d_2 \mid b$, so $b = d_2 m$ and $r = d_2 n$ for some $m, n \in \mathbb{Z}$. Hence,

$$a = bq + r = (d_2 m)q + d_2 n = d_2(mq + n).$$

Since $mq + n$ is an integer, it follows that $d_2 \mid a$. Since d_2 is a positive divisor of b, we conclude that d_2 is a natural number that divides both a and b. Then d_2 does not exceed the largest natural number that divides both a and b: that is, $d_2 \leq d_1$.

Next, by the definition of (a, b), $d_1 \mid a$ and $d_1 \mid b$, so $a = d_1 n$ and $b = d_1 m$ for some $n, m \in \mathbb{Z}$. Hence,

$$r = (bq + r) - bq = a - bq = d_1 n - (d_1 m)q = d_1(n - mq).$$

Since $n - mq$ is an integer, it follows that $d_1 \mid r$. Since d_1 is a positive divisor of b, we conclude that d_1 is a natural number that divides both r and b. Then d_1 does not exceed the largest natural number that divides both r and b: that is, $d_1 \leq d_2$.

We proved that $d_1 \leq d_2$ and $d_2 \leq d_1$. Therefore, $d_1 = d_2$. □

We can use Lemma 1.7 to calculate the gcd of any two positive integers. We illustrate the main idea by an example, and then prove a theorem that describes the method in general.

Example 1.8. Let us use Lemma 1.7 to calculate $(216, 51)$. We apply Lemma 1.7 repeatedly as follows:

$$(216, 51) = (4 \cdot 51 + 12, 51) = (12, 51) = (12, 4 \cdot 12 + 3) = (12, 3) = 3.$$

It seems plausible that we should be able to proceed similarly to Example 1.8 to compute (a, b) for general a and b. The following theorem, known as the *Euclidean algorithm*, establishes that this is indeed the case.

Theorem 1.9 (Euclidean algorithm)**.** *Let a and b be integers, with $0 < b < a$.*

i) *If $b \mid a$, then $(a, b) = b$.*

ii) *If $b \nmid a$, then there exist integers $q_1, r_1, q_2, r_2, \dots, r_m, q_{m+1}$ such that*

$$\begin{aligned} a &= bq_1 + r_1, && 0 < r_1 < b; \\ b &= r_1 q_2 + r_2, && 0 \le r_2 < r_1; \\ r_1 &= r_2 q_3 + r_3, && 0 \le r_3 < r_2; \\ &\vdots \\ r_{m-2} &= r_{m-1} q_m + r_m, && 0 \le r_m < r_{m-1}; \\ r_{m-1} &= r_m q_{m+1}; \end{aligned}$$

and $(a, b) = r_m$.

Proof. i) This part is an easy consequence of the definition of greatest common divisor.

ii) We use induction on a to prove the statement:

> *If b is an integer such that $0 < b < a$ and $b \nmid a$, then there exist integers $q_1, r_1, q_2, r_2, \dots, r_m, q_{m+1}$ such that*

$$\begin{aligned} a &= bq_1 + r_1, && 0 < r_1 < b; \\ b &= r_1 q_2 + r_2, && 0 \le r_2 < r_1; \\ r_1 &= r_2 q_3 + r_3, && 0 \le r_3 < r_2; \\ &\vdots \\ r_{m-2} &= r_{m-1} q_m + r_m, && 0 \le r_m < r_{m-1}; \\ r_{m-1} &= r_m q_{m+1}; && r_m = (a, b). \end{aligned}$$

The base case of the induction is $a = 3$. When $a = 3$, the only admissible value of b is $b = 2$, and the above statement holds with $m = 1$, $q_1 = r_1 = 1$, $q_2 = 2$:

$$3 = 2 \cdot 1 + 1, \ 0 < 1 < 2; \quad 2 = 1 \cdot 2; \quad 1 = (3, 2).$$

Now, suppose that the above statement holds for all integers a with $2 < a < n$, and consider the case $a = n$. Let b be an integer with $0 < b < n$ and $b \nmid n$. By Theorem 1.3, there exist integers q, r such that

$$n = bq + r, \quad 0 \le r < b.$$

Furthermore, since $b \nmid n$, the remainder r cannot be 0, so we can strengthen the above inequality to get

$$n = bq + r, \quad 0 < r < b. \tag{1.1}$$

If $r \mid b$, then $b = rq'$ for some integer q', and Lemma 1.7 and part i) give

$$(n, b) = (bq + r, b) = (r, b) = r.$$

Hence, the above statement for the pair n, b holds with $m = 1$, $q_1 = q$, $r_1 = r$, and $q_2 = q'$. Next, suppose that $r \nmid b$. Since $0 < r < b < n$ and $r \nmid b$, we can apply the inductive hypothesis to the pair b, r: there exist integers $q_1, r_1, q_2, r_2, \dots, r_m, q_{m+1}$ such that

$$\begin{aligned} b &= rq_1 + r_1, & 0 &< r_1 < r; \\ r &= r_1q_2 + r_2, & 0 &\le r_2 < r_1; \\ r_1 &= r_2q_3 + r_3, & 0 &\le r_3 < r_2; \\ &\vdots \\ r_{m-2} &= r_{m-1}q_m + r_m, & 0 &\le r_m < r_{m-1}; \\ r_{m-1} &= r_mq_{m+1}; & r_m &= (b, r). \end{aligned}$$

Combining these conditions with (1.1), we find that the integers $q, r, q_1, r_1, q_2, \dots, r_m, q_{m+1}$ have all the desired properties relative to the pair n, b, with the possible exception of $(n, b) = r_m$. To verify the latter property, we use Lemma 1.7 and the identity $(b, r) = r_m$ above:

$$(n, b) = (bq + r, b) = (r, b) = r_m. \qquad \square$$

The Euclidean algorithm is very effective for computational purposes. However, it does not relate (a, b) directly to a and b. The next theorem provides exactly such a relation.

Theorem 1.10. *Let a and b be integers, not both 0, and let $d = (a, b)$. Then d is the least positive integer in the set*

$$\mathcal{S} = \left\{ ax + by \mid x, y \in \mathbb{Z} \right\}.$$

In particular, there exist integers x, y such that $d = ax + by$.

There are two standard ways to prove Theorem 1.10: one based on the well-ordering axiom of the natural numbers, the other based on the Euclidean algorithm. Both proofs are sketched in the exercises (cf. Exercises 1.4 and 1.5). We now state several important consequences of Theorem 1.10, which we will use numerous times throughout the course.

Theorem 1.11. *If a, b, c are integers such that $(a, b) = (a, c) = 1$, then $(a, bc) = 1$.*

Proof. Suppose that a, b and c are integers such that $(a, b) = (a, c) = 1$. Then by Theorem 1.10, there exist integers x, y, u, v such that

$$ax + by = (a, b) = 1, \quad au + cv = (a, c) = 1.$$

Hence,

$$1 = ax + by = ax + by(1) = ax + by(au + cv) = a(x + byu) + (bc)(yv).$$

Since $x + byu$ and yv are integers, this shows that 1 is a positive integer from the set

$$\mathcal{S} = \left\{ am + (bc)n \mid m, n \in \mathbb{Z} \right\}.$$

Since there is no positive integer less than 1, it follows that 1 is the least positive integer in $\mathcal{S}$. By Theorem 1.10, the least positive integer in $\mathcal{S}$ is (a, bc), so we conclude that $(a, bc) = 1$. $\square$

Corollary 1.12. *If $k \ge 2$ and $a, b_1, \dots, b_k$ are integers such that*

$$(a, b_1) = (a, b_2) = \cdots = (a, b_k) = 1,$$

then $(a, b_1b_2 \cdots b_k) = 1$.

Theorem 1.13. *If a, b, c are integers such that $a \mid bc$ and $(a, b) = 1$, then $a \mid c$.*

Theorem 1.14. *If a, b, c are integers such that $a \mid c$, $b \mid c$ and $(a, b) = 1$, then $ab \mid c$.*

1.3. Prime numbers and the fundamental theorem of arithmetic

We now recall the definition of prime and composite numbers.

Definition 1.15. An integer $p > 1$ is called a *prime number*, or simply a *prime*, if its only positive divisors are 1 and p. An integer $n > 1$ which is not prime is called a *composite number*. The numbers 0 and 1 are neither prime, nor composite.

Example 1.16. The integers 2, 3, 13 and 89 are prime; 4, 6, 51 and 837 are composite.

The next theorem provides a necessary and sufficient condition for primality that is often more useful in number-theoretic proofs than the definition of a prime.

Theorem 1.17. *A positive integer $p > 1$ is prime if and only if p has the property:*

$$(\forall a, b \in \mathbb{Z})(p \mid ab \implies p \mid a \text{ or } p \mid b). \tag{1.2}$$

Proof. "⇒". Suppose that p is a prime number and a and b are integers such that $p \mid ab$. If $p \mid a$, statement (1.2) is true. Now, suppose that $p \nmid a$. By the definition of a prime, the only positive divisors of p are 1 and p. Hence, (p, a) is 1 or p. However, $(p, a) \neq p$, because $p \nmid a$. Hence, $(p, a) = 1$. Since $p \mid ab$ and $(p, a) = 1$, it follows from Theorem 1.13 that $p \mid b$. Therefore, (1.2) is true again.

"⇐". Suppose that $p > 1$ has property (1.2). We must show that p is prime. Suppose that p is composite. Then p has a positive divisor other than 1 and p: $p = ab$, where $a, b \in \mathbb{N}$ and $1 < a < p$. In particular, since $a < p$, we conclude that $p \nmid a$. Note also that

$$p = ab,\ a > 1 \quad \implies \quad b < p \quad \implies \quad p \nmid b.$$

However, since $p \mid p = ab$, it follows from property (1.2) that $p \mid a$ or $p \mid b$; a contradiction. Therefore, p is prime. □

The importance of prime numbers comes from the next theorem, which says roughly that one can use multiplication to build any integer $n > 1$ from primes, and that there is essentially a unique way to do so. In other words, one can view the primes as the basic building blocks of the integers under multiplication.

Theorem 1.18 (Fundamental theorem of arithmetic)**.** *Let $n > 1$ be an integer. Then n has a unique factorization of the form*

$$n = p_1 p_2 \cdots p_k, \quad p_1 \le p_2 \le \cdots \le p_k,$$

where $p_1, p_2, \dots, p_k$ are prime numbers.

1.4. Congruence modulo m

Definition 1.19. Let $m \in \mathbb{N}$, with $m > 1$, and let $a, b \in \mathbb{Z}$. We say that *a is congruent to b modulo m*, and write $a \equiv b \pmod{m}$, if $m \mid (a - b)$.

Theorem 1.20. *Let $a, b, m \in \mathbb{Z}$ and $m > 1$. Then $a \equiv b \pmod{m}$ if and only if there exists an integer k such that $a = b + mk$.*

Proof. By the definitions of congruence and divisibility,

$$\begin{aligned} a \equiv b \pmod{m} \quad &\iff \quad m \mid (a - b) \\ &\iff \quad a - b = mk \\ &\iff \quad a = b + mk, \end{aligned}$$

where $k \in \mathbb{Z}$. □

The next theorem summarizes some basic properties of congruences. Properties i)–iii) demonstrate that congruence modulo m is an equivalence relation on the integers. Properties iv)–vi) introduce the basic arithmetic operations with congruences.

Theorem 1.21. *Let $m \in \mathbb{N}$ and $a, b, c, d \in \mathbb{Z}$. Then:*

i) $a \equiv a \pmod{m}$*;*
ii) *if* $a \equiv b \pmod{m}$*, then* $b \equiv a \pmod{m}$*;*
iii) *if* $a \equiv b \pmod{m}$ *and* $b \equiv c \pmod{m}$*, then* $a \equiv c \pmod{m}$*;*
iv) *if* $a \equiv b \pmod{m}$ *and* $c \equiv d \pmod{m}$*, then* $a + c \equiv b + d \pmod{m}$*;*
v) *if* $a \equiv b \pmod{m}$ *and* $c \equiv d \pmod{m}$*, then* $ac \equiv bd \pmod{m}$*;*
vi) *if* $ac \equiv bc \pmod{m}$ *and* $(c, m) = d$*, then* $a \equiv b \pmod{m/d}$*.*

Proof. i) Since $m \mid 0 = (a - a)$, we have $a \equiv a \pmod{m}$.

ii) By Theorem 1.20, $a = b + km$ for some integer k, whence

$$b = a - km = a + (-k)m.$$

Since $-k$ is also an integer, it follows from Theorem 1.20 that $b \equiv a \pmod{m}$.

iii) By Theorem 1.20,

$$a \equiv b \pmod{m} \implies a = b + km, \quad b \equiv c \pmod{m} \implies b = c + lm,$$

where $k, l \in \mathbb{Z}$. Hence,

$$a = b + km = (c + lm) + km = c + (l + k)m = c + nm,$$

where $n = l + k$ is also an integer. Thus, $a \equiv c \pmod{m}$, by Theorem 1.20.

iv) By Theorem 1.20,

$$a \equiv b \pmod{m} \implies a = b + km, \quad c \equiv d \pmod{m} \implies c = d + lm,$$

where $k, l \in \mathbb{Z}$. Hence,

$$a + c = (b + km) + (d + lm) = (b + d) + (k + l)m = (b + d) + nm,$$

where $n = k + l$ is also an integer. Thus, $a + c \equiv b + d \pmod{m}$, again by Theorem 1.20.

v) By Theorem 1.20,

$$a \equiv b \pmod{m} \implies a = b + km, \quad c \equiv d \pmod{m} \implies c = d + lm,$$

where $k, l \in \mathbb{Z}$. Hence,

$$ac = (b + km)(d + lm) = (bd) + (kd + lb + klm)m = (bd) + nm,$$

where $n = kd + lb + klm$ is also an integer. Thus, by Theorem 1.20, $ac \equiv bd \pmod{m}$.

vi) By Theorem 1.20,

$$ac \equiv bc \pmod{m} \implies ac = bc + km,$$

where $k \in \mathbb{Z}$. Moreover, because $(c, m) = d$, there exist $u, v \in \mathbb{Z}$ such that $cu + mv = d$. Combining these two identities, we find that

$$\begin{aligned} ac = bc + km &\implies acu = bcu + kmu \\ &\implies a(d - mv) = b(d - mv) + kmu \\ &\implies ad = bd - bmv + kmu + amv = bd + nm \\ &\implies a = b + n(m/d), \end{aligned}$$

where $n = -bv + ku + av$ is also an integer. Thus, yet another appeal to Theorem 1.20 gives $a \equiv b \pmod{m/d}$. □

Since congruence modulo m is an equivalence relation on $\mathbb{Z}$, it partitions the integers into equivalence classes called *residue* (or *congruence*) *classes modulo m*. The equivalence class $[a]$ of an integer a under congruence modulo m is denoted $[a]_m$, that is,

$$[a]_m = \big\{n \in \mathbb{Z} \mid n \equiv a \pmod{m}\big\}.$$

We denote the set of all residue classes modulo m by $\mathbb{Z}_m$. When the modulus m is clear from the context, we often write the congruence class $[a]_m$ just as $[a]$. We make use of this convention in the next two examples.

Example 1.22. When $m = 2$, there are only two residue classes: the set $\{0, \pm 2, \pm 4, \dots\}$ of the even integers is the residue class $[k]_2$ of any even integer k; and the set $\{\pm 1, \pm 3, \dots\}$ of the odd integers is the residue class $[k]_2$ of any odd integer k. Thus, $\mathbb{Z}_2 = \{[0], [1]\}$.

Example 1.23. When $m = 3$, there are three residue classes:

$$\begin{aligned} [0] = [3] = [6] = \cdots &= \{\dots, -9, -6, -3, 0, 3, 6, 9, \dots\},\\ [-2] = [1] = [4] = \cdots &= \{\dots, -8, -5, -2, 1, 4, 7, 10, \dots\},\\ [-1] = [2] = [5] = \cdots &= \{\dots, -7, -4, -1, 2, 5, 8, 11, \dots\}. \end{aligned}$$

Thus, $\mathbb{Z}_3 = \{[0], [1], [2]\}$.

The next theorem generalizes the last two examples.

Theorem 1.24. *Let $m \in \mathbb{N}$ and $m > 1$. Then $[0], [1], \dots, [m-1]$ is a complete list of distinct residue classes modulo m.*

Proof. First, we must show that the congruence class $[n]_m$ of every integer n appears in the above list. By Theorem 1.3 with $a = n$ and $b = m$, there exist integers q and r such that $n = mq + r$ and $0 \le r < m$. By Theorem 1.20 and the definition of congruence class,

$$n = mq + r \quad\Longrightarrow\quad n \equiv r \pmod{m} \quad\Longrightarrow\quad n \in [r].$$

This proves that the congruence classes $[n]$ and $[r]$ are not disjoint (they both contain n). Thus, $[n] = [r]$. Since r is one of the numbers $0, 1, \dots, m-1$, this proves that $[n]$ is one of the residue classes in the above list.

Next, suppose that $[a] = [b]$ for some integers $a, b \in \{0, 1, \dots, m-1\}$. Then $a \in [b]$, so

$$a \equiv b \pmod{m} \quad\Longrightarrow\quad m \mid (a-b) \quad\Longrightarrow\quad a - b = mq,$$

for some $q \in \mathbb{Z}$. On the other hand,

$$0 \le a, b < m \quad\Longrightarrow\quad -m < a - b < m \quad\Longrightarrow\quad -1 < q < 1.$$

Hence, $q = 0$. We conclude that $a - b = 0$, which proves that $a = b$. Hence, the residue classes $[0], [1], \dots, [m-1]$ are distinct. □

Remark. You may be familiar with the concept of a "binary operation" on a set. Properties iv) and v) in Theorem 1.21 can be used to justify the definition of the following two operations on $\mathbb{Z}_m$: given residue classes $[a]$ and $[b]$ modulo m, we define the sum $[a] + [b]$ and the product $[a][b]$ by

$$[a] + [b] = [a + b], \qquad [a][b] = [ab]. \tag{1.3}$$

1.5. Complete and reduced systems modulo m

Definition 1.25. A *complete system of residues modulo m* is a system of integers $a_1, \dots, a_s$ such that $\{[a_1], \dots, [a_s]\} = \mathbb{Z}_m$ and $a_i \not\equiv a_j \pmod{m}$ when $i \neq j$.

Example 1.26. By Theorem 1.24, the numbers $0, 1, \dots, m-1$ form a complete residue system modulo m. Another such system is $1, 2, \dots, m$. When m is an odd integer, the numbers $0, \pm 1, \dots, \pm\frac{1}{2}(m-1)$ also form a complete system of residues modulo m.

By Theorem 1.24, $\mathbb{Z}_m$ contains m residue classes, so any complete residue system modulo m must contain exactly m integers, which must be pairwise incongruent modulo m. The next theorem demonstrates that the converse is also true: any m incongruent integers modulo m form a complete residue system modulo m.

Theorem 1.27. *If $a_1, \dots, a_m$ are integers such that $a_i \not\equiv a_j \pmod{m}$ when $i \neq j$, then $a_1, \dots, a_m$ is a complete system of residues modulo m.*

Proof. For $j = 1, \dots, m$, we write

$$a_j = mq_j + r_j, \qquad 0 \le r_j < m.$$

We claim that $r_1, \dots, r_m$ are a permutation of $0, 1, \dots, m-1$. Indeed, $r_1, \dots, r_m$ are m integers from the set $\{0, 1, \dots, m-1\}$. Thus, either each number from this set appears among $r_1, \dots, r_m$, or some number appears twice: say, $r_i = r_j = r$ for some $i \neq j$. But in the latter case, we have

$$a_i - a_j = (mq_i + r) - (mq_j + r) = m(q_i - q_j) \quad \Longrightarrow \quad a_i \equiv a_j \pmod{m},$$

which contradicts our assumption about the a_i's. This establishes our claim. Now, since $a_j = mq_j + r_j$, we have

$$a_j \equiv r_j \pmod{m} \quad \Longrightarrow \quad [a_j] = [r_j],$$

and we conclude that

$$\big\{[a_1], \dots, [a_m]\big\} = \big\{[r_1], \dots, [r_m]\big\}.$$

Since $[r_1], \dots, [r_m]$ are a permutation of $[0], [1], \dots, [m-1]$, this proves the theorem. □

Corollary 1.28. *If a, b are integers, with $(a, m) = 1$, and $r_1, \dots, r_m$ is a complete system modulo m, then*

$$ar_1 + b, ar_2 + b, \dots, ar_m + b$$

is also a complete system modulo m.

Definition 1.29. A *reduced system of residues modulo m* is a system of integers $a_1, \dots, a_s$ such that

$$\big\{[a_1], \dots, [a_s]\big\} = \big\{[x]_m \mid (x, m) = 1\big\}$$

and $a_i \not\equiv a_j \pmod{m}$ when $i \neq j$. That is, the residue class $[x]$ of every integer x, with $(x, m) = 1$, appears among $[a_1], \dots, [a_s]$ exactly once.

Example 1.30. If p is a prime number, the integers $1, 2, \dots, p-1$ form a reduced residue system modulo p. Indeed, if $(x, p) = 1$, then by Theorem 1.24, $[x] = [r]$ for some r with $0 \le r \le p - 1$. Furthermore, if $r = 0$, we get

$$x \equiv 0 \pmod{p} \quad \Longrightarrow \quad p \mid x,$$

which contradicts our assumption that $(x, p) = 1$. Hence, $r > 0$, that is r is one of the integers $1, 2, \dots, p-1$.

Similarly, it can be shown that if $p > 2$ is prime, the integers $\pm 1, \pm 2, \dots, \pm\frac{1}{2}(p-1)$ form a reduced system of residues modulo p.

Theorem 1.31. *If $r_1, \dots, r_s$ is a reduced system modulo m and $(a, m) = 1$, then $ar_1, \dots, ar_s$ is also a reduced system modulo m.*

Theorem 1.32 (Fermat's little theorem)**.** *If p is a prime number and $(a, p) = 1$, then*

$$a^{p-1} \equiv 1 \pmod{p}.$$

Proof. By Example 1.30 and Theorem 1.31, the integers $a, 2a, \dots, (p-1)a$ form a reduced residue system modulo p. Hence, the classes $[a], [2a], \dots, [(p-1)a]$ modulo p form a permutation of the classes $[1], [2], \dots, [p-1]$. In particular, the product of the classes in the first list equals the product of the classes in the second:

$$[a][2a] \cdots [(p-1)a] = [1][2] \cdots [p-1].$$

Using (1.3), we deduce

$$[a(2a) \cdots ((p-1)a)] = [1 \cdot 2 \cdots (p-1)] \quad \Longrightarrow \quad [a^{p-1}(p-1)!] = [(p-1)!].$$

Rewriting the last identity as a congruence modulo p, we get

$$a^{p-1}(p-1)! \equiv (p-1)! \pmod{p}. \tag{1.4}$$

By Corollary 1.12 with $a = p$, $k = p - 1$ and $b_j = j$, we get $(p, (p-1)!) = 1$, so the result follows from (1.4) and part vi) of Theorem 1.21. □

1.6. Exercises

Exercise 1.1. Use the Euclidean algorithm to find the gcd's: $(102, 222)$; $(981, 1234)$; $(20785, 44350)$.

Exercise 1.2. Prove Lemma 1.6.

Exercise 1.3. Let $(a, b) = d$, $a = da_1$ and $b = db_1$. Prove that $(a_1, b_1) = 1$.

Exercise 1.4. The purpose of this exercise is to deduce Theorem 1.10 from the well-ordering axiom of the natural numbers: *Every non-empty subset of $\mathbb{N}$ has a unique least element.*

(a) Let a and b be integers, not both zero. Show that the set

$$\mathcal{S} = \left\{ ax + by \mid x, y \in \mathbb{Z},\ ax + by > 0 \right\}$$

is a non-empty subset of $\mathbb{N}$. Deduce that $\mathcal{S}$ has a least element c.

(b) Let $d = (a, b)$. Show that d divides all the elements of $\mathcal{S}$. Deduce that $d \mid c$.

(c) By the quotient-remainder theorem, there exist integers q and r such that $a = cq + r$ and $0 \le r < c$. Show that r can be expressed in the form $r = au + bv$, with $u, v \in \mathbb{Z}$. Deduce that $r = 0$, and so $c \mid a$.

(d) Argue similarly to part (c) to show that $c \mid b$. Deduce that $c \mid d$.

(e) Combine parts (b) and (d) to show that $c = d$.

Exercise 1.5. The purpose of this exercise is to deduce Theorem 1.10 from the Euclidean algorithm.

(a) Let $S \subset \mathbb{N}$, and let $d \in S$ divides all the elements of S. Show that d is the least element of S.

(b) Let $a, b \in \mathbb{N}$, with $b < a$, and let $d = (a, b)$. Show that d divides all the elements of

$$\mathcal{S} = \left\{ ax + by \mid x, y \in \mathbb{Z},\ ax + by > 0 \right\}.$$

(c) Let $b \mid a$. Use part (a) to show that b is the least element of the set $\mathcal{S}$ in part (b).

(d) Let $b \nmid a$. Use induction on m to show that each of the remainders $r_1, \dots, r_m$ in the Euclidean algorithm is an element of $\mathcal{S}$. In particular, $r_m = d$ is an element of $\mathcal{S}$. Use parts (a) and (b) to deduce that d is the least element of $\mathcal{S}$.

Exercise 1.6. Use mathematical induction to prove Corollary 1.12.

Exercise 1.7. Prove Theorem 1.13.

Exercise 1.8. Prove Theorem 1.14.

Exercise 1.9. If $a, b \in \mathbb{Z}$ and $b > 0$, prove that there exist unique integers q and r such that

$$a = bq + r, \quad 1 \le r \le b.$$

Exercise 1.10. Let p be a prime number and k be an integer with $1 \le k \le p-1$. Prove that p divides the binomial coefficient $\binom{p}{k} = \frac{p!}{k!(p-k)!}$.

Exercise 1.11. Prove Theorem 1.31.

Exercise 1.12. Use mathematical induction to show that if n is a positive integer, then $4^n \equiv 1 + 3n \pmod{9}$.

Exercise 1.13. Find the least positive residue modulo 47 of: 2^{10}; 2^{47}; 2^{2010}. [Hint. 47 is prime, so Fermat's little theorem may be of help.]

Exercise 1.14. Let $(m, n) = 1$. Use Theorem 1.27 to prove that if x runs through a complete residue system modulo m and y runs through a complete residue system modulo n, the sums $nx + my$ run through a complete system of residues modulo mn.

Exercise 1.15. Let $k \ge 2$ and $m_1, m_2, \dots, m_k$ be pairwise relatively prime integers (i.e., $(m_i, m_j) = 1$ whenever $i \ne j$). Define $M = m_1 m_2 \cdots m_k$ and $M_j = M/m_j = \prod_{i \ne j} m_i$ for $j = 1, 2, \dots, k$. Use mathematical induction to prove that if $x_1, x_2, \dots, x_k$ run through complete systems of residues modulo $m_1, m_2, \dots, m_k$, respectively, then the sums $M_1 x_1 + M_2 x_2 + \cdots + M_k x_k$ run through a complete system of residues modulo M. [Hint. The previous exercise provides the base of the induction and should also help with the inductive step.]

The next couple of exercises review some of the divisibility tests and their proofs. When $a_0, a_1, \dots, a_k$ are digits (i.e., integers between 0 and 9), the integer with those decimal digits is denoted $(a_k a_{k-1} \dots a_1 a_0)_{10}$, i.e.,

$$(a_k a_{k-1} \dots a_1 a_0)_{10} = a_k 10^k + a_{k-1} 10^{k-1} + \cdots + a_1 10 + a_0. \tag{1.5}$$

Most divisibility tests use congruences to simplify the right side of (1.5).

Exercise 1.16. In this exercise, we prove the tests for divisibility by 2, 5, 3, 9 and powers of 2 and 5. Let $n = (a_k \dots a_1 a_0)_{10}$.

- **(a)** Compare the two sides of (1.5) modulo 2 to show that: $2 \mid n$ if and only if $2 \mid a_0$, i.e., if 2 divides the units digit of n.
- **(b)** Compare the two sides of (1.5) modulo 5 to show that: $5 \mid n$ if and only if $5 \mid a_0$, i.e., if 5 divides the units digit of n.
- **(c)** Compare the two sides of (1.5) modulo 9 to show that: $9 \mid n$ if and only if $9 \mid (a_0 + a_1 + \cdots + a_k)$, i.e., if 9 divides the sum of the digits of n. [Hint. $10^j \equiv 1 \pmod{9}$ for all $j \in \mathbb{N}$.]
- **(d)** Use part (c) to show that: $3 \mid n$ if and only if 3 divides the sum of the digits of n.
- **(e)** Compare the two sides of (1.5) modulo 4 to show that: $4 \mid n$ if and only if $4 \mid (a_1 a_0)_{10}$, i.e., if 4 divides the two-digit number formed from the last two digits of n.
- **(f)** Compare the two sides of (1.5) modulo 8 to show that: $8 \mid n$ if and only if $8 \mid (a_2 a_1 a_0)_{10}$, i.e., if 8 divides the three-digit number formed from the last three digits of n.
- **(g)** Compare the two sides of (1.5) modulo 25 to show that: $25 \mid n$ if and only if $25 \mid (a_1 a_0)_{10}$, i.e., if 25 divides the two-digit number formed from the last two digits of n.
- **(h)** Generalize parts (e)–(g) to state and prove divisibility tests by powers of 2 and 5.

Exercise 1.17. In this exercise, we prove the tests for divisibility by 7, 11, 13, 27 and 37. Let $n = (a_k \dots a_1 a_0)_{10}$.

- **(a)** Compare the two sides of (1.5) modulo 11 to show that: $11 \mid n$ if and only if $9 \mid (a_0 - a_1 + a_2 - \cdots)$, i.e., if 11 divides the alternating sum of the digits of n. [Hint. $10^j \equiv (-1)^j \pmod{11}$ for all $j \in \mathbb{N}$.]
- **(b)** Compare the two sides of (1.5) modulo 999 to show that: $999 \mid n$ if and only if 999 divides the sum
$$(a_0 a_1 a_2)_{10} + (a_3 a_4 a_5)_{10} + (a_6 a_7 a_8)_{10} + \cdots .$$
- **(c)** Note that $999 = 27 \cdot 37$. Use this and part (b) to obtain divisibility tests modulo 27 and 37.
- **(d)** Compare the two sides of (1.5) modulo 1001 to show that: $1001 \mid n$ if and only if 1001 divides the alternating sum
$$(a_0 a_1 a_2)_{10} - (a_3 a_4 a_5)_{10} + (a_6 a_7 a_8)_{10} - \cdots .$$
- **(e)** Note that $1001 = 7 \cdot 11 \cdot 13$. Use this and part (e) to obtain divisibility tests modulo 7 and 13.

CHAPTER 2

The Law Of Quadratic Reciprocity

In this chapter, we state the law of quadratic reciprocity and some related theorems and describe their importance for the solution of quadratic congruences in one unknown.

2.1. Polynomial congruences

A *polynomial congruence (in one variable)* is a congruence of the form

$$a_nx^n + \cdots + a_1x + a_0 \equiv 0 \pmod{m}, \tag{2.1}$$

where $a_0, a_1, \dots, a_n$ are given integers and x is an integer unknown. If x_0 is a solution of (2.1) and $x_1 \equiv x_0 \pmod{m}$, then by parts iv) and v) of Theorem 1.21,

$$a_nx_1^n + \cdots + a_1x_1 + a_0 \equiv a_nx_0^n + \cdots + a_1x_0 + a_0 \equiv 0 \pmod{m},$$

that is, x_1 is also a solution of (2.1). Consequently, if an integer x_0 is a solution of (2.1), all the integers in the congruence class $[x_0]$ are also solutions. Thus, we usually describe the solutions of a polynomial congruence modulo m as a collection of residue classes modulo m.

Example 2.1. Consider the congruence

$$2x^4 - 3x^2 + x + 7 \equiv 0 \pmod{5}.$$

Since there are only five congruence classes modulo 5 and each of them either is a solution of the above congruence, or is not, we may simply perform an exhaustive search for the solutions:

$$\begin{aligned}
2 \cdot 0^4 - 3 \cdot 0^2 + 0 + 7 &= 7 \not\equiv 0 \pmod{5};\\
2 \cdot 1^4 - 3 \cdot 1^2 + 1 + 7 &= 7 \not\equiv 0 \pmod{5};\\
2(-1)^4 - 3(-1)^2 + (-1) + 7 &= 5 \equiv 0 \pmod{5};\\
2 \cdot 2^4 - 3 \cdot 2^2 + 2 + 7 &= 29 \not\equiv 0 \pmod{5};\\
2(-2)^4 - 3(-2)^2 + (-2) + 7 &= 25 \equiv 0 \pmod{5}.
\end{aligned}$$

Hence, the solutions of this congruence are the integers x in the congruence classes

$$x \equiv -1 \pmod{5} \qquad \text{and} \qquad x \equiv -2 \pmod{5}.$$

Clearly, the method used to solve the above example can be applied to any explicit congruence, and it will eventually find all the solutions of the congruence. However, this method has two significant weaknesses. First, it can be extremely inefficient—just try to apply it to the congruence

$$2010x^{2010} + \cdots + 3x^3 + 2x^2 + x + 31415926 \equiv 0 \pmod{10^{10}}.$$

Furthermore, even when the method can be applied in reasonable time (as in Example 2.1), it still provides no insight into the structure of the solutions. These shortcomings of the brute force method from Example 2.1 have led to the development of more sophisticated techniques for solving polynomial congruences. The ensuing theory of polynomial congruences shares many features with the theory of polynomial equations over the real numbers that you are familiar with. For example, you should be familiar with the following result from algebra:

> *If $f(x)$ is a polynomial with real coefficients of degree n, then the equation $f(x) = 0$ has at most n distinct roots.*

There is a similar result for polynomial congruences modulo a prime modulus p. We state it here for future reference. The proof is sketched in the exercises.

Theorem 2.2 (Lagrange)**.** *Let p be a prime number and $f(x) = a_nx^n + \cdots + a_1x + a_0$ be a polynomial with integer coefficients such that $p \nmid a_n$. Then the congruence $f(x) \equiv 0 \pmod{p}$ has at most n solutions modulo p.*

2.2. Linear congruences

A *linear congruence (in one variable)* is a congruence of the form

$$ax \equiv b \pmod{m}, \tag{2.2}$$

where x is an integer unknown. It turns out that it is almost as easy to describe the solutions of a linear congruence as it is to describe the solutions of a linear equation. First, we consider the case when $(a, m) = 1$.

Theorem 2.3. *Let $a, b \in \mathbb{Z}$ and $m \in \mathbb{N}$, and let $(a, m) = 1$. Then the linear congruence* (2.2) *has a unique solution modulo m.*

Proof. By Theorem 1.10, there exist $u, v \in \mathbb{Z}$ such that $au + mv = 1$. Multiplying both sides of this identity by b, we obtain

$$aub + mvb = b \quad\Longrightarrow\quad a(ub) = b + m(-vb) \quad\Longrightarrow\quad a(ub) \equiv b \pmod{m}.$$

That is, ub is a solution of (2.2).

Moreover, if x and y are two solutions of (2.2), then by part vi) of Theorem 1.21,

$$ax \equiv b \equiv ay \pmod{m} \quad\Longrightarrow\quad x \equiv y \pmod{m},$$

that is, any two solutions of (2.2) must belong to the same congruence class modulo m. □

Let $(a, m) = 1$. In the special case $b = 1$, the unique solution modulo m of the congruence $ax \equiv 1 \pmod{m}$ is called the *inverse of a modulo m* and is denoted by a^*, i.e., a^* denotes any integer such that $aa^* \equiv 1 \pmod{m}$. By the above theorem, a^* is determined up to congruence modulo m.

The next theorem extends Theorem 2.3 to linear congruences (2.2), where a and m are not necessarily coprime.

Theorem 2.4. *Let $a, b \in \mathbb{Z}$ and $m \in \mathbb{N}$, and let $(a, m) = d$. If $d \nmid b$, then the congruence* (2.2) *has no solutions. If $d \mid b$, then* (2.2) *has exactly d incongruent solutions modulo m,*

$$x \equiv x_0 + r(m/d) \pmod{m} \qquad (r = 0, 1, \dots, d-1). \tag{2.3}$$

Proof. Let $a = da_1$ and $m = dm_1$. We note that then $(a_1, m_1) = 1$, by Exercise 1.3. By Theorem 1.20, for each solution x of (2.2), we can find a $k \in \mathbb{Z}$ such that

$$ax = b + km \quad\Longrightarrow\quad b = ax - km = d(a_1x - km_1) \quad\Longrightarrow\quad d \mid b.$$

Therefore, the condition $d \mid b$ is necessary for the existence of solutions of (2.2). This establishes the first part of the theorem.

Now, suppose that $d \mid b$ and $b_1 = b/d$. By part vi) of Theorem 1.21, (2.2) is equivalent to the congruence

$$a_1 x \equiv b_1 \pmod{m_1}.$$

Since $(a_1, m_1) = 1$, the last congruence has a unique solution modulo m_1 by Theorem 2.3. Hence, the solutions of (2.2) are the integers in some congruence class $x \equiv x_0 \pmod{m_1}$. Let $x_0 + km_1$, $k \in \mathbb{Z}$, be one of these solutions. By the quotient-remainder theorem, there exist integers q, r such that $k = dq + r$. Hence,

$$x_0 + km_1 = x_0 + rm_1 + q(dm_1) \equiv x_0 + rm_1 \pmod{m}.$$

Therefore, any solution of (2.2) belongs to one of the congruence classes (2.3). Conversely, any integer x that belongs to one of the congruence classes (2.3) satisfies $x \equiv x_0 \pmod{m_1}$, and therefore, x is a solution of (2.2). □

Example 2.5. Consider the congruences $10x \equiv 5 \pmod{12}$ and $10x \equiv 6 \pmod{12}$. The first congruence has no solution, because $(10, 12) = 2$ and $2 \nmid 5$. The second congruence has solutions. By parts v) and vi) of Theorem 1.21,

$$10x \equiv 6 \pmod{12} \quad \Longrightarrow \quad 5x \equiv 3 \pmod{6},$$

and

$$5x \equiv 3 \pmod{6} \quad \Longrightarrow \quad -5x \equiv -3 \pmod{6} \quad \Longrightarrow \quad x \equiv 3 \pmod{6}.$$

Thus, the second congruence has two incongruent solutions modulo 12:

$$x \equiv 3 \pmod{12}, \qquad x \equiv 9 \pmod{12}.$$

2.3. Quadratic congruences. Quadratic residues and nonresidues

Next, it is natural to consider the quadratic congruence

$$ax^2 + bx + c \equiv 0 \pmod{m},$$

where $(a, m) = 1$. A variation of the classic trick of completing of the square can be used to reduce any quadratic congruence to congruence of the form $x^2 \equiv d \pmod{m}$ (see Exercise 2.11), so we shall focus on such congruences.

Definition 2.6. Let $a, m \in \mathbb{Z}$, with $m > 1$ and $(a, m) = 1$. We say that a is a *quadratic residue modulo m* if the congruence $x^2 \equiv a \pmod{m}$ has a solution; we say that a is a *quadratic nonresidue modulo m* if it is not a quadratic residue.

Example 2.7. Consider the case $m = 7$: 2 is a quadratic residue modulo 7, but 3 and 5 are not. In fact, the square of any integer is congruent modulo 7 to the square of one of the numbers $0, \pm1, \pm2, \pm3$. Hence, for any $x \in \mathbb{Z}$,

$$x^2 \equiv 0, 1, 2 \text{ or } 4 \pmod{7}.$$

Therefore, the quadratic residues modulo 7 are 1, 2 and 4, and the quadratic nonresidues are 3, 5 and 6; 0 is neither residue nor nonresidue.

Example 2.8. The only quadratic residue modulo 8 is 1. Indeed, the square of any integer is congruent modulo 8 to the square of one of the numbers $0, \pm1, \pm2, \pm3$ or 4. Hence, for any $x \in \mathbb{Z}$,

$$x^2 \equiv 0, 1 \text{ or } 4 \pmod{8}.$$

Therefore, the quadratic residue modulo 8 is 1, and the quadratic nonresidues are 3, 5 and 7; 0, 2, 4 and 6 are neither residues nor nonresidues.

Lemma 2.9. *Let $p > 2$ be a prime and a be a quadratic residue modulo p. Then the congruence*

$$x^2 \equiv a \pmod{p} \tag{2.4}$$

has exactly two solutions modulo p.

Proof. Since a is a quadratic residue, there is an integer b such that $b^2 \equiv a \pmod{p}$. Furthermore, we have $(-b)^2 = b^2 \equiv a \pmod{p}$, so $-b$ is also a solution of (2.4). These two solutions are distinct modulo p. Indeed, if they were not, we would have

$$b \equiv -b \pmod{p} \quad\Longrightarrow\quad 2b \equiv 0 \pmod{p} \quad\Longrightarrow\quad b \equiv 0 \pmod{p},$$

by part vi) of Theorem 1.21 (note that $(2, p) = 1$). Hence $a \equiv b^2 \equiv 0 \pmod{p}$, which contradicts our assumption that a is a quadratic residue. Therefore, congruence (2.4) has at least two distinct solutions modulo p. On the other hand, by Theorem 2.2, (2.4) has at most two distinct solutions modulo p. It follows that (2.4) has exactly two distinct solutions modulo p. □

Compare the above lemma with Example 2.8: the congruence $x^2 \equiv 1 \pmod{8}$ has not two but four solutions modulo 8. This demonstrates that the primality of p is crucial for the conclusion of Lemma 2.9.

Theorem 2.10. *Let p be an odd prime. There are $\frac{1}{2}(p-1)$ quadratic residues and $\frac{1}{2}(p-1)$ quadratic nonresidues modulo p.*

Proof. If a is a quadratic residue, then (2.4) has solutions $x \equiv \pm b \pmod{p}$ for some integer $b \in \left\{1, 2, \dots, \frac{1}{2}(p-1)\right\}$. Hence, every quadratic residue is congruent to one of the integers

$$1^2, 2^2, \dots, \tfrac{1}{4}(p-1)^2. \tag{2.5}$$

These integers are clearly quadratic residues modulo p. Moreover, the integers (2.5) are pairwise distinct modulo p. Indeed, if $i^2 \equiv j^2 \pmod{p}$, with $1 \le i < j \le \frac{1}{2}(p-1)$, then $i, -i, j$ and $-j$ are all solutions of (2.4) with $a = i^2$; since those four numbers are incongruent modulo p, this contradicts Theorem 2.2. Therefore, the integers (2.5) are exactly the quadratic residues modulo p. So, there are $\frac{1}{2}(p-1)$ quadratic residues modulo p. Since there are $p-1$ reduced residue classes modulo p, we deduce that

$$\begin{aligned}\#(\text{quadratic nonresidues}) &= (p-1) - \#(\text{quadratic residues})\\ &= (p-1) - \tfrac{1}{2}(p-1) = \tfrac{1}{2}(p-1).\end{aligned}$$

□

2.4. The Legendre symbol

Definition 2.11. Let p be an odd prime and $(a, p) = 1$. The *Legendre symbol* $\left(\frac{a}{p}\right)$ is defined by

$$\left(\frac{a}{p}\right) = \begin{cases} +1 & \text{if } a \text{ is a quadratic residue modulo } p,\\ -1 & \text{if } a \text{ is a quadratic nonresidue modulo } p.\end{cases}$$

Note that the sign of the Legendre symbol $\left(\frac{a}{p}\right)$ tells us whether the congruence (2.4) has two solutions or no solution at all, just as the sign of a real number a tells us whether the equation $x^2 = a$ has two real solutions or no solution at all. Thus, we now proceed to study Legendre symbols. Our main objective will be to find an efficient way of computing the Legendre symbol $\left(\frac{a}{p}\right)$ for a given odd prime p and an integer a.

Theorem 2.12 (Euler's criterion)**.** *Let p be an odd prime and $(a, p) = 1$. Then*

$$\left(\frac{a}{p}\right) \equiv a^{(p-1)/2} \pmod{p}.$$

Proof. Suppose first that a is a quadratic residue and $a \equiv b^2 \pmod{p}$, with $(b, p) = 1$. Then

$$a^{(p-1)/2} \equiv \left(b^2\right)^{(p-1)/2} = b^{p-1} \equiv 1 \pmod{p},$$

by Fermat's little theorem. Therefore, $a^{(p-1)/2} \equiv \left(\frac{a}{p}\right) \pmod{p}$.

Now, suppose that a is a quadratic nonresidue. We just showed that every quadratic residue is a solution of the polynomial congruence

$$x^{(p-1)/2} - 1 \equiv 0 \pmod{p}. \tag{2.6}$$

The degree of this congruence is $(p-1)/2$, so by Theorem 2.2, it has at most $(p-1)/2$ solutions modulo p. Since there are exactly $(p-1)/2$ quadratic residues modulo p, we conclude that the solutions of (2.6) are exactly the quadratic residues modulo p. In particular, a is not a solution of (2.6): that is, $p \nmid \left(a^{(p-1)/2} - 1\right)$. On the other hand, by Fermat's little theorem,

$$a^{p-1} - 1 \equiv 0 \pmod{p} \quad \Longrightarrow \quad \left(a^{(p-1)/2} - 1\right)\left(a^{(p-1)/2} + 1\right) \equiv 0 \pmod{p}.$$

Since $p \nmid \left(a^{(p-1)/2} - 1\right)$, we deduce that

$$a^{(p-1)/2} + 1 \equiv 0 \pmod{p} \quad \Longrightarrow \quad a^{(p-1)/2} \equiv -1 = \left(\frac{a}{p}\right) \pmod{p}. \qquad \square$$

Corollary 2.13. *Let p be an odd prime. Then*

$$\left(\frac{-1}{p}\right) = \begin{cases} +1 & \text{if } p \equiv 1 \pmod{4}, \\ -1 & \text{if } p \equiv 3 \pmod{4}. \end{cases}$$

Theorem 2.14. *Let p be an odd prime and $(a, p) = (b, p) = 1$. Then:*

i) *if $a \equiv b \pmod{p}$, then $\left(\frac{a}{p}\right) = \left(\frac{b}{p}\right)$;*
ii) $\left(\frac{a}{p}\right)\left(\frac{b}{p}\right) = \left(\frac{ab}{p}\right)$.

Proof. i) Using Theorem 2.12 and the hypothesis $a \equiv b \pmod{p}$, we get

$$\left(\frac{a}{p}\right) \equiv a^{(p-1)/2} \equiv b^{(p-1)/2} \equiv \left(\frac{b}{p}\right) \pmod{p}.$$

Since $\left(\frac{a}{p}\right) = \pm 1$ and $\left(\frac{b}{p}\right) = \pm 1$,

$$\left(\frac{a}{p}\right) - \left(\frac{b}{p}\right) = 2, 0 \text{ or } -2.$$

Since $p \geq 3$ must divide this difference, we deduce that

$$\left(\frac{a}{p}\right) - \left(\frac{b}{p}\right) = 0. \qquad \square$$

Corollary 2.15. *Let p be an odd prime. The product of two quadratic residues or of two quadratic non-residues (modulo p) is a quadratic residue; the product of a quadratic residue and a quadratic non-residue is a quadratic non-residue.*

2.5. Quadratic reciprocity

We can now state the main theorem of this course. Gauss, who was the first to give a complete proof of the quadratic reciprocity law (actually, he found at least eight different proofs), called it *Theorema Aureum* (i.e., golden theorem). To this day, it is considered one of the crown jewels of elementary number theory.

Theorem 2.16 (Quadratic reciprocity law)**.** *Let p and q be distinct odd primes. Then*

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{1}{2}(p-1)\cdot\frac{1}{2}(q-1)} = \begin{cases} +1 & \text{if } p \text{ or } q \equiv 1 \pmod 4, \\ -1 & \text{if } p \equiv q \equiv 3 \pmod 4. \end{cases}$$

The law of quadratic reciprocity is usually complemented with a formula for the Legendre symbol $\left(\frac{2}{p}\right)$, which we state in the next theorem.

Theorem 2.17. *If p is an odd prime, then*

$$\left(\frac{2}{p}\right) = (-1)^{(p^2-1)/8} = \begin{cases} +1 & \text{if } p \equiv \pm 1 \pmod 8, \\ -1 & \text{if } p \equiv \pm 3 \pmod 8. \end{cases}$$

Combining Theorems 2.16 and 2.17, Corollary 2.13 and the two properties of Legendre symbols in Theorem 2.14, we can easily compute Legendre symbols. The next example illustrates the usual strategy.

Example 2.18. The number 9973 is prime. Let us determine whether 19920 is a quadratic residue modulo 9973. We need to compute $\left(\frac{19920}{9973}\right)$. First, we use Theorem 2.14 to replace the number 19920 in the top position of the Legendre symbol by a smaller number:

$$\left(\frac{19920}{9973}\right) = \left(\frac{19920-9973}{9973}\right) = \left(\frac{9947}{9973}\right) = \left(\frac{9947-9973}{9973}\right) = \left(\frac{-26}{9973}\right).$$

Using Corollary 2.13 and Theorem 2.17, we get

$$\left(\frac{-1}{9973}\right) = +1, \quad \text{because } 9973 \equiv 1 \pmod 4;$$
$$\left(\frac{2}{9973}\right) = -1, \quad \text{because } 9973 \equiv 5 \pmod 8.$$

Thus, by part ii) of Theorem 2.14,

$$\left(\frac{-26}{9973}\right) = \left(\frac{-1}{9973}\right)\left(\frac{2}{9973}\right)\left(\frac{13}{9973}\right) = (+1)(-1)\left(\frac{13}{9973}\right) = -\left(\frac{13}{9973}\right).$$

Now, since $13 \equiv 1 \pmod 4$, the quadratic reciprocity law yields

$$\left(\frac{13}{9973}\right)\left(\frac{9973}{13}\right) = 1 \quad \Longrightarrow \quad \left(\frac{13}{9973}\right) = \left(\frac{9973}{13}\right).$$

Again, we use part i) of Theorem 2.14 to replace 9973 by a smaller number:

$$\left(\frac{9973}{13}\right) = \left(\frac{873}{13}\right) = \left(\frac{93}{13}\right) = \left(\frac{2}{13}\right) = -1,$$

because of Theorem 2.17. Therefore,

$$\left(\frac{19920}{9973}\right) = \left(\frac{-26}{9973}\right) = -\left(\frac{13}{9973}\right) = -\left(\frac{9973}{13}\right) = 1,$$

and 19920 is a quadratic residue modulo 9973. □

The purpose of the remainder of this course is to explain several different proofs of the quadratic reciprocity law. There are many such proofs. In fact, they are so numerous that Dr. Franz Lemmermeyer has set up a special website:

`http://www.rzuser.uni-hd.de/~hb3/frchrono.html`,

with a list of all known (to him) proofs of Theorem 2.16. At the time of this writing (February, 2010), the website lists 233 proofs. Not all those proofs are completely different—and some are just minor modifications of others, but there is still a great variety of methods. Our discussion will focus on two main groups of proofs, which rely on Theorems 2.19 and 2.21 below.

Theorem 2.19 (Gauss' lemma)**.** *Let p be an odd prime and $(a,p)=1$. Let μ be the number of least positive residues of the integers $a, 2a, 3a, \dots, \frac{1}{2}(p-1)a$ modulo p that are greater than $p/2$. Then $\left(\frac{a}{p}\right) = (-1)^{\mu}$.*

Example 2.20. Let $p = 13$ and $a = 7$. Then the integers in the statement of Gauss' lemma are $7, 14, 21, 28, 35$ and 42. Their least positive residues modulo 13 are $7, 1, 8, 2, 9$ and 3, of which three are greater than $13/2$. Hence, $\mu = 3$ and Gauss' lemma states that $\left(\frac{7}{13}\right) = (-1)^3 = -1$.

In the next theorem, $\mathrm{i} = \sqrt{-1}$ is the imaginary unit. We shall review complex numbers and the meaning of the complex exponential $e^{2\pi \mathrm{i} n^2/m}$ in the next lecture.

Theorem 2.21 (Gauss sum formula)**.** *Let $m \in \mathbb{N}$. Then*

$$\sum_{n=1}^{m} e^{2\pi \mathrm{i} n^2/m} = \frac{1+\mathrm{i}^{-m}}{1+\mathrm{i}^{-1}}\sqrt{m}.$$

2.6. Exercises

Exercise 2.1. Solve the linear congruences: $102x \equiv 5 \pmod{22}$; $7x \equiv 3 \pmod{5}$; $24x \equiv 6 \pmod{39}$.

Exercise 2.2. The purpose of this exercise is to establish Theorem 2.2. Let $f(x) = a_n x^n + \cdots + a_1 x + a_0$ be a polynomial with integer coefficients of degree $n \geq 1$.

(a) Let r be an integer. Synthetic division gives $f(x) = (x-r)g_r(x) + f(r)$, where $g_r(x)$ is a polynomial of degree $n-1$. Prove that $g_r(x)$ also has integer coefficients and that its leading coefficient is a_n.

(b) Suppose that p is a prime and r and s are integers such that $r \not\equiv s \pmod{p}$ and $f(r) \equiv f(s) \equiv 0 \pmod{p}$. Let $g_r(x)$ be the polynomial with integer coefficients from part (a). Prove that $g_r(s) \equiv 0 \pmod{p}$.

(c) Use mathematical induction on n to prove Theorem 2.2.
[Hint. Theorem 2.3 provides the base of the induction. Your inductive hypothesis should be that the theorem is true for all polynomials of degree $n-1$, where $n \geq 2$. Then the inductive step should establish the theorem for a generic $f(x)$ of degree n. Let $r, r_2, \dots, r_k$ be the solutions of $f(x) \equiv 0 \pmod{p}$. Use part (b) to show that $r_2, \dots, r_k$ are solutions of $g_r(x) \equiv 0 \pmod{p}$, where $g_r(x)$ is the polynomial from part (a); then apply the inductive hypothesis to $g_r(x)$.]

Exercise 2.3. Evaluate the Legendre symbols: $\left(\frac{3}{53}\right)$; $\left(\frac{15}{101}\right)$; $\left(\frac{105}{1009}\right)$; $\left(\frac{1973}{2011}\right)$.

Exercise 2.4. Prove Corollary 2.13.

Exercise 2.5. Prove part ii) of Theorem 2.14.

Exercise 2.6. Use Gauss' lemma with $a = 2$ to prove Theorem 2.17.

Exercise 2.7. Let p and q be two distinct primes and $(a, pq) = 1$.

(a) Prove that if a is a quadratic residue modulo pq, then a is a quadratic residue modulo p and modulo q.

(b) Prove that if $y^2 \equiv a \pmod{p}$ and $y^2 \equiv a \pmod{q}$, then $y^2 \equiv a \pmod{pq}$.

(c) Show that there exist integers u and v such that $pu + qv = 1$.

(d) Let $y = psu + qrv$, where u and v are the integers from part (c). Prove that $y \equiv r \pmod{p}$ and $y \equiv s \pmod{q}$.

(e) Prove that if a is a quadratic residue modulo p and modulo q, then a is a quadratic residue modulo pq.
[Hint. Let r be a solution of $x^2 \equiv a \pmod{p}$ and s be a solution of $x^2 \equiv a \pmod{q}$. Use part (d) to construct an integer y such that $y^2 \equiv a \pmod{p}$ and $y^2 \equiv a \pmod{q}$. Then apply part (b).]

(f) Prove that if a is a quadratic residue modulo pq, then the congruence $x^2 \equiv a \pmod{pq}$ has four solutions.
[Hint. Every solution y of $x^2 \equiv a \pmod{pq}$ arises from a solution r of $x^2 \equiv a \pmod{p}$ and a solution s of $x^2 \equiv a \pmod{q}$ via the procedure outlined in the hint to part (e). Since the latter congruences have two solutions each, the congruence modulo pq has four solutions.]

Exercise 2.8. The purpose of this exercise is to establish Gauss' lemma. Let p be an odd prime and $(a, p) = 1$, and let μ be the number from the statement of Gauss' lemma.

(a) For each $j = 1, 2, \dots, \frac{1}{2}(p-1)$, ja is congruent to one of the numbers $\pm1, \pm2, \dots, \pm\frac{1}{2}(p-1)$: that is, $ja \equiv \varepsilon_j r_j \pmod{p}$, where $1 \le r_j \le \frac{1}{2}(p-1)$ and $\varepsilon_j = \pm 1$. Prove that $(-1)^{\mu} = \varepsilon_1 \varepsilon_2 \cdots \varepsilon_{(p-1)/2}$.

(b) Prove that $r_1, r_2, \dots, r_{(p-1)/2}$ are pairwise distinct.

(c) Prove that $r_1 r_2 \cdots r_{(p-1)/2} = \big(\frac{1}{2}(p-1)\big)!$.

(d) Use parts (a) and (c) to prove that $a^{(p-1)/2}\big(\frac{1}{2}(p-1)\big)! \equiv (-1)^{\mu}\big(\frac{1}{2}(p-1)\big)! \pmod{p}$.

(e) Deduce Gauss' lemma from part (d) and Euler's criterion. [Hint. $p \nmid \big(\frac{1}{2}(p-1)\big)!$.]

Exercise 2.9. This exercise establishes *Wilson's theorem:* If p is a prime, then $(p-1)! \equiv -1 \pmod{p}$.

(a) Use Theorems 2.2 and 2.3 to show that the numbers $2, 3, \dots, p-2$ can be partitioned into $\frac{1}{2}(p-3)$ pairs a, a^* such that $aa^* \equiv 1 \pmod{p}$.

(b) Use part (a) to deduce that $(p-1)! \equiv 1^{(p-3)/2}(p-1) \equiv -1 \pmod{p}$.

Exercise 2.10. Use the quadratic reciprocity law to prove that if p is an odd prime, then

$$\left(\frac{3}{p}\right) = \begin{cases} +1 & \text{if } p \equiv \pm 1 \pmod{12}, \\ -1 & \text{if } p \equiv \pm 5 \pmod{12}. \end{cases}$$

Exercise 2.11. Consider the congruence $ax^2 + bx + c \equiv 0 \pmod{p}$, where p is a prime and a, b and c are integers with $p \nmid a$.

(a) Let $p = 2$. Determine which quadratic congruences modulo 2 have solutions. [Hint. There are only four congruences to worry about.]

(b) Let p be an odd prime and let $d = b^2 - 4ac$. Show that the congruence $ax^2 + bx + c \equiv 0 \pmod{p}$ is equivalent to the congruence $y^2 \equiv d \pmod{p}$, where $y = 2ax + b$.

(c) Use part (b) to show that if $p \mid d$, then the congruence $ax^2 + bx + c \equiv 0 \pmod{p}$ has only one solution modulo p.

(d) Use part (b) to show that if d is a quadratic residue modulo p, then the congruence $ax^2 + bx + c \equiv 0 \pmod{p}$ has two incongruent solutions modulo p.

(e) Use part (b) to show that if d is a quadratic nonresidue modulo p, then the congruence $ax^2 + bx + c \equiv 0 \pmod{p}$ has no solution modulo p.

Exercise 2.12. Let $m > 1$ be an odd integer, and let $m = p_1 p_2 \cdots p_k$ be its unique factorization as a product of primes (cf. Theorem 1.18). The *Jacobi symbol* $\left(\frac{\cdot}{m}\right)$ modulo m is defined for all integers a such that $(a, m) = 1$ by

$$\left(\frac{a}{m}\right) = \left(\frac{a}{p_1}\right)\left(\frac{a}{p_2}\right)\cdots\left(\frac{a}{p_k}\right),$$

where the factors $\left(\frac{a}{p_j}\right)$ on the right side are Legendre symbols.

(a) Use Corollary 2.13 to show that $\left(\frac{-1}{m}\right) = (-1)^{(m-1)/2}$.

(b) Use Theorem 2.14 to show that $\left(\frac{a}{m}\right)\left(\frac{b}{m}\right) = \left(\frac{ab}{m}\right)$ and that $\left(\frac{a}{m}\right) = \left(\frac{b}{m}\right)$ whenever $a \equiv b \pmod{m}$.

(c) Use the quadratic reciprocity law to show that if $m, n \in \mathbb{N}$ and $(m, n) = 1$, then

$$\left(\frac{n}{m}\right)\left(\frac{m}{n}\right) = (-1)^{\frac{1}{2}(m-1)\cdot\frac{1}{2}(n-1)}.$$

CHAPTER 3

Complex Numbers

In this chapter, we review the definition and the properties of the complex numbers. We also extend the definitions of some of some elementary functions from real to complex argument and give a brief overview of some of the properties of the extended functions. Most likely, some of the material in this chapter will be familiar from earlier classes and some will be new (although those who have taken a course in complex analysis will likely find nothing new).

3.1. Definition and algebraic properties

We start with the definition of a complex number. In lower-level courses, one usually encounters complex numbers as expressions of the form $a + \mathrm{i}b$, where $a, b \in \mathbb{R}$ and i is a non-real "number" such that $\mathrm{i}^2 = -1$ (hence, we often say "...where $\mathrm{i} = \sqrt{-1}$"). What one usually does not encounter is any explanation where the number i comes from and why it exists in the first place. Here, we will give a more systematic treatment of the definition of a complex number. Our definition may seem strange at first, but eventually it will lead us to the same concept and also will reveal the mystery behind the "imaginary unit" i.

A *complex number* z is an ordered pair (x, y), with $x, y \in \mathbb{R}$. The set of all complex numbers is denoted by $\mathbb{C}$, that is,

$$\mathbb{C} = \left\{(x, y) \mid x, y \in \mathbb{R}\right\}.$$

The components x and y of the complex number $z = (x, y)$ are called *real* and *imaginary parts* of z and are denoted $x = \operatorname{Re} z$ and $y = \operatorname{Im} z$, respectively.

We define the algebraic operations *addition*, +, and *multiplication*, × (also denoted '·' or not at all), of two complex numbers $z_1 = (x_1, y_1)$ and $z_2 = (x_2, y_2)$ as follows:

$$z_1 + z_2 := (x_1 + x_2, y_1 + y_2), \quad z_1 \cdot z_2 := (x_1x_2 - y_1y_2, x_1y_2 + x_2y_1). \tag{3.1}$$

We call the complex number $(0, 1)$ the *imaginary unit* and denote it by i. Further, for $x \in \mathbb{R}$, we identify x and the complex number $(x, 0)$, that is, we write $x = (x, 0)$. With these conventions, we can use (3.1) to express a complex number $z = (x, y)$ in the form

$$z = (x, y) = (x, 0) + (0, 1) \cdot (y, 0) = x + \mathrm{i}y.$$

Henceforth, we will always write the complex number $z = (x, y)$ as $z = x + \mathrm{i}y$. That is,

$$\mathbb{C} = \left\{x + \mathrm{i}y \mid x, y \in \mathbb{R}\right\},$$

with addition and multiplication given by

$$z_1 + z_2 := (x_1 + x_2) + \mathrm{i}(y_1 + y_2), \quad z_1z_2 := (x_1x_2 - y_1y_2) + \mathrm{i}(x_1y_2 + x_2y_1). \tag{3.2}$$

Note that with these definitions, the imaginary unit i satisfies the relation $\mathrm{i}^2 = -1$. This is where the phrase "...where $\mathrm{i} = \sqrt{-1}$" comes from.

Combining the two definitions (3.2) with the algebraic properties of real numbers, we can extend many of those properties to complex numbers. Proposition 3.1 below lists several such properties. We note that the properties of complex numbers listed in the proposition establish that $\mathbb{C}$ with the two operations defined above is a field.

Proposition 3.1. *Let $z = x + \mathrm{i}y$, $z_1 = x_1 + \mathrm{i}y_1$, $z_2 = x_2 + \mathrm{i}y_2$ and $z_3 = x_3 + \mathrm{i}y_3$ be complex numbers. Then:*

i) $z_1 + z_2 = z_2 + z_1$;
ii) $(z_1 + z_2) + z_3 = z_1 + (z_2 + z_3)$;
iii) *the number* zero, $0 = 0 + \mathrm{i}0$, *satisfies* $z + 0 = z$;
iv) *the number* $(-z) = (-x) + \mathrm{i}(-y)$ *satisfies* $z + (-z) = 0$;
v) $z_1z_2 = z_2z_1$;
vi) $(z_1z_2)z_3 = z_1(z_2z_3)$;
vii) *the number* one, $1 = 1 + \mathrm{i}0$, *satisfies* $z \cdot 1 = z$;
viii) *if* $z \neq 0$, *there is a number* z^{-1} *satisfying* $zz^{-1} = 1$;
ix) $(z_1 + z_2)z = z_1z + z_2z$.

Proof. The proofs of these properties use the properties of the real numbers. For example, consider v). We have

$$\begin{aligned} z_1z_2 &= (x_1x_2 - y_1y_2) + \mathrm{i}(x_1y_2 + x_2y_1) && \text{by (3.2)} \\ &= (x_2x_1 - y_2y_1) + \mathrm{i}(x_2y_1 + x_1y_2) && \text{by the properties of } \mathbb{R} \\ &= z_2z_1 && \text{by (3.2).} \end{aligned}$$

We also remark that the number z^{-1} in viii) is given by

$$z^{-1} = \frac{x}{x^2 + y^2} + \mathrm{i}\frac{-y}{x^2 + y^2}. \qquad \square$$

3.2. Geometric interpretation, moduli and conjugates

It is natural to associate with every complex number $z = x + \mathrm{i}y$ the point (x, y) in the xy-plane. When the plane is used to represent geometrically the complex numbers in this fashion, it is usually referred to as the *complex plane*. Oftentimes, it is also useful to visualize the number $z = x + \mathrm{i}y$ as the vector from the origin to the point (x, y). For example, the sum of $z_1 + z_2$ is represented this way by the sum of the vectors representing z_1 and z_2 (see Figure 3.1).

The vector interpretation is also helpful in extending the notion of absolute value of a real number to $\mathbb{C}$. We define the *modulus*, $|z|$, of a complex number $z = x + \mathrm{i}y$ by

$$|z| = \sqrt{x^2 + y^2}.$$

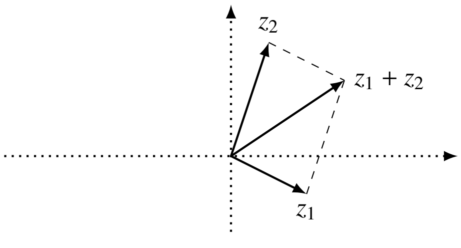

Figure 3.1. Geometric interpretation of the addition

The *complex conjugate of* z is the number $\bar{z} = x - \mathrm{i}y$. Sometimes, we will refer to $|z_1 - z_2|$ as the *distance between* z_1 *and* z_2 (since it represents the length of the line segment with endpoints z_1 and z_2).

Example 3.2. The distance between $z_1 = 1 - \mathrm{i}$ and $z_2 = 3 + 2\mathrm{i}$ is

$$|z_1 - z_2| = |-2 - 3\mathrm{i}| = \sqrt{(-2)^2 + (-3)^2} = \sqrt{13}.$$

Proposition 3.3. *If* z, z_1 *and* z_2 *are complex numbers, then:*

i) $|z|^2 = z\bar{z}$, $\operatorname{Re} z = \frac{1}{2}(z + \bar{z})$, $\operatorname{Im} z = \frac{1}{2i}(z - \bar{z})$, $\bar{\bar{z}} = z$;
ii) $\overline{z_1 \pm z_2} = \bar{z}_1 \pm \bar{z}_2$, $\overline{z_1 z_2} = \bar{z}_1\bar{z}_2$, $\overline{z_1/z_2} = \bar{z}_1/\bar{z}_2$;
iii) $|z_1 z_2| = |z_1| \cdot |z_2|$, $|z_1/z_2| = |z_1|/|z_2|$;
iv) $z = \bar{z}$ *if and only if* $z \in \mathbb{R}$.

Theorem 3.4 (Triangle inequality)**.** *For any pair of complex numbers* z_1 *and* z_2*, we have*

$$|z_1 + z_2| \le |z_1| + |z_2|. \tag{3.3}$$

Two alternative formulations of the triangle inequality are

$$|z_1 - z_2| \le |z_1| + |z_2|, \quad |z_1 - z_2| \ge \big||z_1| - |z_2|\big|. \tag{3.4}$$

Each of the three inequalities in (3.3) and (3.4) can be used to prove the other two (see the exercises).

Proof. Since both sides of (3.3) are non-negative, it suffices to show that

$$|z_1 + z_2|^2 \le (|z_1| + |z_2|)^2. \tag{3.5}$$

Using the properties in Proposition 3.3, we can represent the left side of (3.5) as

$$\begin{aligned}|z_1 + z_2|^2 &= (z_1 + z_2)\overline{(z_1 + z_2)} = (z_1 + z_2)(\bar{z}_1 + \bar{z}_2) \\ &= z_1\bar{z}_1 + z_1\bar{z}_2 + z_2\bar{z}_1 + z_2\bar{z}_2 = |z_1|^2 + 2\operatorname{Re}(z_1\bar{z}_2) + |z_2|^2,\end{aligned}$$

while the right side of (3.5) equals

$$|z_1|^2 + 2|z_1 z_2| + |z_2|^2 = |z_1|^2 + 2|z_1\bar{z}_2| + |z_2|^2.$$

Hence, (3.5) is equivalent to the inequality

$$|z_1|^2 + 2\operatorname{Re}(z_1\bar{z}_2) + |z_2|^2 \le |z_1|^2 + 2|z_1\bar{z}_2| + |z_2|^2 \quad \Longleftrightarrow \quad \operatorname{Re}(z_1\bar{z}_2) \le |z_1\bar{z}_2|.$$

The last inequality follows from the observation that

$$\operatorname{Re} z \le |\operatorname{Re} z| \le |z| \quad \forall z \in \mathbb{C}. \qquad \square$$

Example 3.5. Suppose that $|z| \le 1$ and bound the expression $|z^2 + 3|$ from above and below.

Solution. By the triangle inequality with $z_1 = z^2$ and $z_2 = 3$,

$$|z^2 + 3| \le |z|^2 + |3| \le 1 + 3 = 4,$$

while by the second inequality in (3.4),

$$|z^2 + 3| \ge \big|3 - |z|^2\big| \ge 3 - |z|^2 \ge 3 - 1 = 2. \qquad \square$$

3.3. Exponential form

Let $z = x + \mathrm{i}y$ be a nonzero complex number and let (r, θ) be the polar coordinates of the point (x, y). Since $x = r\cos\theta$ and $y = r\sin\theta$, we can express z in the form

$$z = r(\cos\theta + \mathrm{i}\sin\theta). \tag{3.6}$$

Clearly, $r = |z|$. It is also clear that the number θ is determined up to a shift by a multiple of 2π, that is, if θ satisfies (3.6), so does any number of the form $\theta + 2n\pi$, $n \in \mathbb{Z}$. Any such number θ is called an *argument* of z and the set of all the arguments of z is denoted $\arg z$: that is,

$$\arg z = \left\{\theta + 2n\pi \mid n \in \mathbb{Z}\right\}.$$

We also define the *principal argument* of z, $\operatorname{Arg} z$, as the unique argument of z lying in the interval $-\pi < \theta \leq \pi$.

Definition 3.6. If $\theta \in \mathbb{R}$, we define

$$e^{\mathrm{i}\theta} = \exp(\mathrm{i}\theta) = \cos\theta + \mathrm{i}\sin\theta.$$

In particular, we can write (3.6) in *exponential form*, $z = re^{\mathrm{i}\theta}$.

Example 3.7. If $z = -1 + \mathrm{i}$, we have

$$\operatorname{Arg} z = \tfrac{3\pi}{4}, \quad \arg z = \left\{\tfrac{3\pi}{4} + 2n\pi \mid n \in \mathbb{Z}\right\}, \quad z = \sqrt{2}e^{3\pi\mathrm{i}/4}.$$

Proposition 3.8. *If $z_1 = r_1e^{\mathrm{i}\theta_1}$ and $z_2 = r_2e^{\mathrm{i}\theta_2}$ are complex numbers in exponential form, then:*

i) *$z_1 = z_2$ if and only if $r_1 = r_2$ and $\theta_1 = \theta_2 + 2n\pi$;*
ii) *$z_1z_2 = r_1r_2e^{\mathrm{i}(\theta_1+\theta_2)}$;*
iii) *$z_1/z_2 = (r_1/r_2)e^{\mathrm{i}(\theta_1-\theta_2)}$;*
iv) *$z_1^n = r_1^ne^{\mathrm{i}n\theta_1}$ for all $n \in \mathbb{Z}$.*

Proof. ii) It suffices to consider the case $r_1 = r_2 = 1$. We have

$$\begin{aligned} e^{\mathrm{i}\theta_1}e^{\mathrm{i}\theta_2} &= (\cos\theta_1 + \mathrm{i}\sin\theta_1)(\cos\theta_2 + \mathrm{i}\sin\theta_2) \\ &= (\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2) + \mathrm{i}(\cos\theta_1\sin\theta_2 + \cos\theta_2\sin\theta_1) \\ &= \cos(\theta_1 + \theta_2) + \mathrm{i}\sin(\theta_1 + \theta_2) = e^{\mathrm{i}(\theta_1+\theta_2)}. \end{aligned}$$ □

Example 3.9 (De Moivre's formula)**.** By part iv) of Proposition 3.8,

$$(\cos\theta + \mathrm{i}\sin\theta)^n = \cos n\theta + \mathrm{i}\sin n\theta.$$

Comparing this identity and the expansion for the left side that follows from the binomial formula, we can derive formulas for $\cos n\theta$ and $\sin n\theta$. For example, when $n = 2$, we have

$$\cos 2\theta + \mathrm{i}\sin 2\theta = (\cos\theta + \mathrm{i}\sin\theta)^2 = (\cos^2\theta - \sin^2\theta) + \mathrm{i}(2\sin\theta\cos\theta),$$

whence

$$\cos 2\theta = \cos^2\theta - \sin^2\theta, \quad \sin 2\theta = 2\sin\theta\cos\theta.$$

Example 3.10. We can use the exponential form of $\sqrt{3} - \mathrm{i}$ to compute quickly $\left(\sqrt{3} - \mathrm{i}\right)^{-6}$. We have $\sqrt{3} - \mathrm{i} = 2e^{-\pi\mathrm{i}/6}$, so

$$\left(\sqrt{3} - \mathrm{i}\right)^{-6} = \left(2e^{-\pi\mathrm{i}/6}\right)^{-6} = 2^{-6}e^{\pi\mathrm{i}} = -\tfrac{1}{64}.$$

3.4. Roots of complex numbers

Definition 3.11. A complex number w is called an *nth root* of $z \in \mathbb{C}$ if $w^n = z$. We will write $z^{1/n}$ for the set of all nth roots of a complex number z: that is,

$$z^{1/n} = \{ w \in \mathbb{C} \mid w^n = z \}.$$

The notation $\sqrt[n]{r}$ will be reserved for nth roots of non-negative real numbers, that is, if $r \geq 0$, $\sqrt[n]{r}$ is the unique real number $x \geq 0$ satisfying the equation $x^n = r$.

The number 0 has only one nth root—itself. If $z = re^{\mathrm{i}\theta} \neq 0$, a number $w = \rho e^{\mathrm{i}\phi}$ is an nth root of z if

$$\begin{aligned} \rho^n e^{\mathrm{i}n\phi} = re^{\mathrm{i}\theta} \quad &\iff \quad \rho^n = r \ \text{ and } \ n\phi = \theta + 2k\pi \ \text{ for some } \ k \in \mathbb{Z} \\ &\iff \quad \rho = \sqrt[n]{r} \ \text{ and } \ \phi = (\theta + 2k\pi)/n \ \text{ for some } \ k \in \mathbb{Z}. \end{aligned}$$

We observe that among the values of $e^{\mathrm{i}\phi}$ there are exactly n distinct, which we can obtain by letting k run through n consecutive integers, say, $k = 1, 2, \dots, n$ or $k = 0, 1, \dots, n-1$. Therefore, z has exactly n nth roots:

$$\sqrt[n]{r}\, e^{\mathrm{i}\theta/n}, \ \sqrt[n]{r}\, e^{\mathrm{i}(\theta+2\pi)/n}, \ \sqrt[n]{r}\, e^{\mathrm{i}(\theta+4\pi)/n}, \dots, \ \sqrt[n]{r}\, e^{\mathrm{i}(\theta+2(n-1)\pi)/n}.$$

Example 3.12. We have $(-2)^{1/4} = \sqrt[4]{2} \exp\big(\mathrm{i}(\pi/8 + k\pi/2)\big)$, $k = 1, 2, 3, 4$.

Example 3.13. We have

$$4^{1/2} = \left\{ \sqrt{4} e^{\mathrm{i}(\pi + k\pi)} \mid k = 1, 2 \right\} = \left\{ 2e^{2\pi\mathrm{i}}, 2e^{3\pi\mathrm{i}} \right\} = \{2, -2\}.$$

3.5. The complex exponential

Definition 3.14. We define the *(complex) exponential function*, denoted e^z or $\exp z$, by

$$e^{x+\mathrm{i}y} = e^x e^{\mathrm{i}y} = e^x(\cos y + \mathrm{i} \sin y).$$

Notice that the exponential function defined above agrees with the real exponential function for real arguments: $\exp(x + \mathrm{i}0) = e^x$, where the right side represents the real exponential function from calculus.

Proposition 3.15. *The complex exponential function has the following properties:*

i) $e^{z_1} e^{z_2} = e^{z_1+z_2}$ *and* $e^{z_1}/e^{z_2} = e^{z_1-z_2}$ *for all* $z_1, z_2 \in \mathbb{C}$.
ii) $|e^z| = e^{\mathrm{Re}\, z}$ *for all* $z \in \mathbb{C}$.
iii) $e^{z+2\pi\mathrm{i}} = e^z$ *for all* $z \in \mathbb{C}$*, that is, the exponential function is* $2\pi\mathrm{i}$*-periodic.*
iv) *The equation* $e^z = c$ *has infinitely many solutions for every* $c \neq 0$.
v) *If* $m \in \mathbb{N}$ *and* $a, b \in \mathbb{Z}$ *satisfy* $a \equiv b \pmod{m}$*, then* $e^{2\pi\mathrm{i}a/m} = e^{2\pi\mathrm{i}b/m}$.
vi) *If a and m are integers and* $(a, m) = 1$*, then the list* $e^{2\pi\mathrm{i}a/m}, e^{2\pi\mathrm{i}2a/m}, \dots, e^{2\pi\mathrm{i}ma/m}$ *is a permutation of the list* $e^{2\pi\mathrm{i}/m}, e^{2\pi\mathrm{i}2/m}, \dots, e^{2\pi\mathrm{i}m/m} = 1$.

Proof. i)–iii). These follow from the definition, the properties of the real exponential function (from calculus), and the properties of $e^{\mathrm{i}y}$ in Proposition 3.8.

iv) Writing $z = x + \mathrm{i}y$ and $c = re^{\mathrm{i}\theta}$, we can express the equation $e^z = c$ in the form

$$e^x e^{\mathrm{i}y} = re^{\mathrm{i}\theta}.$$

By part i) of Proposition 3.8, the latter equation is equivalent to

$$e^x = r, \ y = \theta + 2n\pi \quad (n \in \mathbb{Z}) \quad \iff \quad x = \ln r, \ y = \theta + 2n\pi \quad (n \in \mathbb{Z}).$$

That is, the given equation has infinitely many solutions

$$z_n = \ln |c| + \mathrm{i}(\mathrm{Arg}\, c + 2n\pi) \quad (n \in \mathbb{Z}).$$

vi) By Corollary 1.28, when k runs through the integers $1, 2, \ldots m$, the product ak runs through a complete system of residues modulo m. That is, the numbers in the list $a, 2a, \ldots, ma$ can be rearranged as $n_1, n_2, \ldots, n_m$ so that $n_k \equiv k \pmod{m}$. Then, by part v), $e^{2\pi i n_k/m} = e^{2\pi i k/m}$, so rearranging the numbers in the list $e^{2\pi i a/m}, e^{2\pi i 2a/m}, \ldots, e^{2\pi i ma/m}$ as $e^{2\pi i n_1/m}, e^{2\pi i n_2/m}, \ldots, e^{2\pi i n_m/m}$ yields the desired result. □

3.6. Trigonometric and hyperbolic functions of a complex argument

We define the *(complex) sine and cosine functions*, denoted $\sin z$ and $\cos z$, by

$$\sin z = \frac{e^{iz} - e^{-iz}}{2i} \quad \text{and} \quad \cos z = \frac{e^{iz} + e^{-iz}}{2}.$$

We also define the functions $\tan z$, $\cot z$, $\sec z$, $\csc z$ in the usual way:

$$\tan z = \frac{\sin z}{\cos z}, \quad \cot z = \frac{\cos z}{\sin z}, \quad \sec z = \frac{1}{\cos z}, \quad \csc z = \frac{1}{\sin z}.$$

We define the *(complex) hyperbolic sine and hyperbolic cosine functions*, denoted $\sinh z$ and $\cosh z$, by

$$\sinh z = \frac{e^{z} - e^{-z}}{2} \quad \text{and} \quad \cosh z = \frac{e^{z} + e^{-z}}{2}.$$

We also define the functions $\tanh z$, $\coth z$, $\operatorname{sech} z$, $\operatorname{csch} z$ by

$$\tanh z = \frac{\sinh z}{\cosh z}, \quad \coth z = \frac{\cosh z}{\sinh z}, \quad \operatorname{sech} z = \frac{1}{\cosh z}, \quad \operatorname{csch} z = \frac{1}{\sinh z}.$$

We note that, with the exception of $\sin z$ and $\cos z$, all these functions are defined (at least formally) by extending their known definitions from calculus to the complex case. The definitions of $\sin z$ and $\cos z$, on the other hand, extend Euler's formulas

$$\sin x = \frac{e^{ix} - e^{-ix}}{2i} \quad \text{and} \quad \cos x = \frac{e^{ix} + e^{-ix}}{2}$$

from the real to the complex case.

Proposition 3.16. *The following properties hold true:*

i) $\sin(z_1 + z_2) = \sin z_1 \cos z_2 + \cos z_1 \sin z_2$, $\cos(z_1 + z_2) = \cos z_1 \cos z_2 - \sin z_1 \sin z_2$;
ii) $\sin^2 z + \cos^2 z = 1$;
iii) *the trigonometric functions are* 2π*-periodic; the functions* $\tan z$ *and* $\cot z$ *are* π*-periodic;*
iv) $\sinh z = -i \sin(iz)$, $\cosh z = \cos(iz)$;
v) $\sinh(z_1 + z_2) = \sinh z_1 \cosh z_2 + \cosh z_1 \sinh z_2$, $\cosh(z_1 + z_2) = \cosh z_1 \cosh z_2 + \sinh z_1 \sinh z_2$;
vi) $\cosh^2 z - \sinh^2 z = 1$;
vii) *the hyperbolic functions are* $2\pi i$*-periodic; the functions* $\tanh z$ *and* $\coth z$ *are* πi*-periodic.*

Proof. i) By part i) of Proposition 3.15, we have

$$
\begin{aligned}
&\cos z_1 \cos z_2 - \sin z_1 \sin z_2\\
&= \left(\frac{e^{\mathrm{i}z_1} + e^{-\mathrm{i}z_1}}{2}\right)\left(\frac{e^{\mathrm{i}z_2} + e^{-\mathrm{i}z_2}}{2}\right) - \left(\frac{e^{\mathrm{i}z_1} - e^{-\mathrm{i}z_1}}{2\mathrm{i}}\right)\left(\frac{e^{\mathrm{i}z_2} - e^{-\mathrm{i}z_2}}{2\mathrm{i}}\right)\\
&= \frac{e^{\mathrm{i}(z_1+z_2)} + e^{\mathrm{i}(z_1-z_2)} + e^{\mathrm{i}(z_2-z_1)} + e^{-\mathrm{i}(z_1+z_2)}}{4} + \frac{e^{\mathrm{i}(z_1+z_2)} - e^{\mathrm{i}(z_1-z_2)} - e^{\mathrm{i}(z_2-z_1)} + e^{-\mathrm{i}(z_1+z_2)}}{4}\\
&= \frac{e^{\mathrm{i}(z_1+z_2)} + e^{-\mathrm{i}(z_1+z_2)}}{2} = \cos(z_1 + z_2).
\end{aligned}
$$

iii) and vii). The 2π-periodicity of the trigonometric functions and the $2\pi\mathrm{i}$-periodicity of the hyperbolic functions are corollaries of their definitions and the $2\pi\mathrm{i}$-periodicity of $\exp z$. For example,

$$\sec(z + 2\pi) = \frac{2}{\exp(\mathrm{i}z + 2\pi\mathrm{i}) + \exp(-\mathrm{i}z - 2\pi\mathrm{i})} = \frac{2}{\exp(\mathrm{i}z) + \exp(-\mathrm{i}z)} = \sec z.$$

iv) These are straightforward from the definitions of $\sin z$, $\cos z$, $\sinh z$, and $\cosh z$. For example,

$$-\mathrm{i}\sin(\mathrm{i}z) = (-\mathrm{i})\frac{e^{\mathrm{i}(\mathrm{i}z)} - e^{-\mathrm{i}(\mathrm{i}z)}}{2\mathrm{i}} = -\frac{e^{-z} - e^{z}}{2} = \sinh z.$$

v and vi). These (and many other) identities can be deduced from the corresponding identities for trigonometric functions by means of the identities iv). For example,

$$\cosh^2 z - \sinh^2 z = [\cos(\mathrm{i}z)] - [(-\mathrm{i})\sin(\mathrm{i}z)]^2 = \cos^2(\mathrm{i}z) + \sin^2(\mathrm{i}z) = 1, \quad \text{by ii)}. \qquad \square$$

If $f(z)$ is a trigonometric or a hyperbolic function, the equation $f(z) = c$ has infinitely many solutions for all but at most two exceptional complex numbers c. The exceptional values of c exist when $f(z)$ is one of the functions $\tan z$, $\cot z$, $\sec z$, $\csc z$, $\tanh z$, $\coth z$, $\operatorname{sech} z$ or $\operatorname{csch} z$; there are no exceptional values for $\sin z$, $\cos z$, $\sinh z$ and $\cosh z$. In the next example, we illustrate this in the case $f(z) = \tan z$. Note that unlike the real case (the equation $\tan x = y$ has solutions for every value of y), in the complex case the equation $\tan z = c$ can fail to have a solution.

Example 3.17. The equation $\tan z = c$, where $c \in \mathbb{C}$, has infinitely solutions if and only if $c \neq \pm\mathrm{i}$.

Solution. Since

$$\tan z = \frac{e^{2\mathrm{i}z} - 1}{\mathrm{i}(e^{2\mathrm{i}z} + 1)},$$

the equation $\tan z = c$ is equivalent to $e^{2\mathrm{i}z} = w$, where w is a solution of

$$\frac{w-1}{w+1} = c\mathrm{i}. \tag{3.7}$$

When $c \neq -\mathrm{i}$, the last equation has a unique solution

$$w = \frac{1 + c\mathrm{i}}{1 - c\mathrm{i}},$$

whereas when $c = -\mathrm{i}$, (3.7) has no solution. When $c = \mathrm{i}$, the solution of (3.7) is $w = 0$ and the resulting equation for z, $e^{2\mathrm{i}z} = 0$, has no solution. When $c \neq \pm\mathrm{i}$, we have $w \neq 0$. We write w in exponential form, $w = re^{\mathrm{i}\theta}$ and find

$$\begin{aligned} e^{2\mathrm{i}z} = re^{\mathrm{i}\theta} \quad &\Longleftrightarrow \quad e^{-2y+2\mathrm{i}x} = re^{\mathrm{i}\theta} \\ &\Longleftrightarrow \quad e^{-2y} = r, \quad 2x = \theta + 2n\pi \qquad (n \in \mathbb{Z}) \\ &\Longleftrightarrow \quad y = -\tfrac{1}{2}\ln r, \quad x = \tfrac{1}{2}\theta + n\pi \qquad (n \in \mathbb{Z}) \\ &\Longleftrightarrow \quad z = \tfrac{1}{2}\theta + n\pi - \tfrac{\mathrm{i}}{2}\ln r \qquad (n \in \mathbb{Z}). \end{aligned}$$ □

The next example illustrates that over the complex numbers we can solve the equation $\sin z = c$ for any complex number c. You should compare this with the equation $\sin x = c$, where $c \in \mathbb{R}$ and x is a real variable—the latter has no solutions when $|c| > 1$.

Example 3.18. We now solve the equation $\sin z = 4$. By the substitution $w = e^{\mathrm{i}z}$, this equation reduces to

$$\frac{w - w^{-1}}{2\mathrm{i}} = 4 \quad \Longleftrightarrow \quad w^2 - 8\mathrm{i}w - 1 = 0.$$

This quadratic equation has two complex roots

$$w_{1,2} = \frac{8\mathrm{i} + \left((-8\mathrm{i})^2 + 4\right)^{1/2}}{2} = \left(4 \pm \sqrt{15}\right)\mathrm{i} = \left(4 \pm \sqrt{15}\right)e^{\pi\mathrm{i}/2}.$$

Hence, by the proof of part iv) of Proposition 3.15, we obtain two infinite families of solutions of $\sin z = 4$:

$$\begin{aligned} e^{\mathrm{i}z} = w_1 \quad &\Longleftrightarrow \quad \mathrm{i}z = \ln\left(4 + \sqrt{15}\right) + \mathrm{i}(\pi/2 + 2n\pi) \quad (n \in \mathbb{Z}) \\ &\Longleftrightarrow \quad z = \pi/2 + 2n\pi - \mathrm{i}\ln\left(4 + \sqrt{15}\right) \quad (n \in \mathbb{Z}); \\ e^{\mathrm{i}z} = w_2 \quad &\Longleftrightarrow \quad \mathrm{i}z = \ln\left(4 - \sqrt{15}\right) + \mathrm{i}(\pi/2 + 2n\pi) \quad (n \in \mathbb{Z}) \\ &\Longleftrightarrow \quad z = \pi/2 + 2n\pi - \mathrm{i}\ln\left(4 - \sqrt{15}\right) \quad (n \in \mathbb{Z}). \end{aligned}$$

Notice that similarly to the real case, we obtain the solutions in our two infinite series differ one from another by a multiple of 2π, the period of the sine function. □

3.7. The complex logarithm and power functions

Definition 3.19. We define the *logarithm* $\log z$ for all nonzero complex numbers z as the set of solutions w of the equation $e^w = z$, that is,

$$\log z = \ln|z| + \mathrm{i}\arg z,$$

where $\ln x$ is the natural logarithm, defined for $x > 0$ as $\ln x = \int_1^x t^{-1}dt$.

Note that $\log z$ is not a function but a relation—it is the inverse relation of the exponential function viewed as a relation on $\mathbb{C}$. In order to deal with the complex logarithm as a function, we introduce the notion of a "branch". If D is a region in the complex plane, a function $F : D \to \mathbb{C}$ is called a *branch of the logarithm* if for each $z \in D$, $F(z)$ is one of the values of $\log z$.

Example 3.20. The simplest example of a branch of the logarithm is the function

$$\operatorname{Log} z = \ln|z| + \mathrm{i}\operatorname{Arg} z \quad (-\pi < \operatorname{Arg} z < \pi),$$

known as the *principal branch of the logarithm.*

Proposition 3.21. *The logarithms* $\log z$ *and* $\operatorname{Log} z$ *have the following properties:*

i) *For every* $z \neq 0$, *we have* $e^{\log z} = z$, *but not necessarily* $\log(e^z) = z$.

ii) *If* $|\operatorname{Im} z| < \pi$*, then* $\operatorname{Log}(e^z) = z$*.*

iii) *If* $z_1, z_2 \neq 0$*, then*

$$\log(z_1 z_2) = \log z_1 + \log z_2, \quad \log\left(\frac{z_1}{z_2}\right) = \log z_1 - \log z_2.$$

iv) *If* $-\pi < \arg z_1, \arg z_2 < \pi$*, then the identities*

$$\operatorname{Log}(z_1 z_2) = \operatorname{Log} z_1 + \operatorname{Log} z_2 \quad \textit{and} \quad \operatorname{Log}\left(\frac{z_1}{z_2}\right) = \operatorname{Log} z_1 - \operatorname{Log} z_2$$

may fail.

Note that together parts i) and ii) establish that $\operatorname{Log} z$ is the inverse function of the restriction of $\exp z$ to the horizontal strip $-\pi < \operatorname{Im} z < \pi$.

Proof. i) The first part is the definition of $\log z$. The second part is obvious, since the function $\log(e^z)$ is an infinite set, whereas the function $w = z$ has a single value.

ii) This follows from the definition of $\operatorname{Log} z$.

iii) The left-hand side of the identity is

$$\log(z_1 z_2) = \ln|z_1 z_2| + \mathrm{i}\arg(z_1 z_2) = (\ln|z_1| + \ln|z_2|) + \mathrm{i}\arg(z_1 z_2),$$

and the right-hand side is

$$\begin{aligned}\log z_1 + \log z_2 &= (\ln|z_1| + \mathrm{i}\arg z_1) + (\ln|z_2| + \mathrm{i}\arg z_2)\\ &= (\ln|z_1| + \ln|z_2|) + \mathrm{i}(\arg z_1 + \arg z_2).\end{aligned}$$

Thus, it suffices to show that

$$\arg(z_1 z_2) = \arg z_1 + \arg z_2. \tag{3.8}$$

Suppose that $z_1 = r_1 e^{\mathrm{i}\theta_1}$ and $z_2 = r_2 e^{\mathrm{i}\theta_2}$. Then $z_1 z_2 = r_1 r_2 e^{\mathrm{i}(\theta_1+\theta_2)}$, so the left side of (3.8) is the set

$$\{\theta_1 + \theta_2 + 2n\pi \mid n \in \mathbb{Z}\}. \tag{3.9}$$

The right side of (3.8), on the other hand, is the set of all possible sums of a value of $\arg z_1$ and a value of $\arg z_2$. That is, the right side of (3.8) is

$$\{(\theta_1 + 2k\pi) + (\theta_2 + 2m\pi) \mid k, m \in \mathbb{Z}\}. \tag{3.10}$$

Therefore, (3.8) claims that the sets (3.9) and (3.10) are equal. Now, if $\phi = \theta_1 + \theta_2 + 2n\pi$ is an element of (3.9), then ϕ is also an element of (3.10): take $k = 0$ and $m = n$. Conversely, if $\phi = (\theta_1 + 2k\pi) + (\theta_2 + 2m\pi)$ is an element of (3.10), then ϕ is also an element of (3.9): simply take $n = k + m$. This proves that the sets (3.9) and (3.10) are equal.

iv) Consider, for example, $z_1 = z_2 = \mathrm{i}$. Then $\operatorname{Log} z_1 = \operatorname{Log} z_2 = \frac{\pi \mathrm{i}}{2}$, but $\operatorname{Log}(z_1 z_2)$ is not even defined. □

Definition 3.22. If $z \neq 0$ and $a \in \mathbb{C}$, we define z^a by

$$z^a = \exp(a \log z).$$

Note that z^a may have more than one value, because $\log z$ has infinitely many values. Whether these infinitely many values yield a single value, a finite number of values, or an infinite number of values of z^a depends on a. Writing $\log z$ in the form

$$\log z = \ln|z| + \mathrm{i}(\operatorname{Arg} z + 2n\pi) \quad (n \in \mathbb{Z}),$$

we find that

$$z^a = \exp\left[a(\ln|z| + \mathrm{i}(\operatorname{Arg} z + 2n\pi))\right] = \exp\left[a(\ln|z| + \mathrm{i}\operatorname{Arg} z)\right]\exp(2\pi\mathrm{i}na) \quad (n \in \mathbb{Z}).$$

Hence, the number of distinct values of z^a is equal to the number of distinct values of $\exp(2\pi \mathrm{i} na)$ as n runs through $\mathbb{Z}$. Four different cases can occur:

1) If $a \in \mathbb{Z}$, then $e^{2\pi \mathrm{i} na} = 1$ for all $n \in \mathbb{Z}$, so z^a has a single value. In fact, if $a = k$, the power function z^k equals the monomial z^k.
2) If $a \in \mathbb{Q} - \mathbb{Z}$ (i.e., $a = p/q$, with $q > 1$ and $(p,q) = 1$), then z^a has q distinct values. Indeed, by part vi) of Proposition 3.15, the numbers $e^{2\pi \mathrm{i} p/q}, e^{2\pi \mathrm{i} 2p/q}, \dots, e^{2\pi \mathrm{i} qp/q}$ can be rearranged as $e^{2\pi \mathrm{i}/q}, e^{2\pi \mathrm{i} 2/q}, \dots, e^{2\pi \mathrm{i} q/q} = 1$. The latter are q distinct values. On the other hand, if n is any other integer, we can write it in the form $n = kq + r$, where $k, r \in \mathbb{Z}$ and $1 \le r \le q$ (see Exercise 1.9). Hence,
$$e^{2\pi \mathrm{i} na} = e^{2\pi \mathrm{i}(kq+r)p/q} = e^{2\pi \mathrm{i} rp/q},$$
and the latter number is among the q numbers found earlier.
3) If a is irrational, all the values of z^a are distinct. Indeed, if
$$e^{2\pi \mathrm{i} na} = e^{2\pi \mathrm{i} ma} \quad \text{for some } m \neq n,$$
then
$$2\pi \mathrm{i} na = 2\pi \mathrm{i} ma + 2\pi \mathrm{i} k \quad \text{for some } k \in \mathbb{Z} \quad \Longrightarrow \quad a = \frac{k}{n-m} \in \mathbb{Q}.$$
4) If a is not real, all the values of z^a are again distinct, since each number $e^{2\pi \mathrm{i} na}$ has a different modulus.

Example 3.23. Let us evaluate the expressions $(1-\mathrm{i})^{\mathrm{i}}$ and $(2\mathrm{i})^{1/3}$. We have
$$(1-\mathrm{i})^{\mathrm{i}} = \exp\big(\mathrm{i}\log(1-\mathrm{i})\big) = \exp\big(\mathrm{i}(\ln|1-\mathrm{i}| + \mathrm{i}\arg(1-\mathrm{i}))\big).$$
Since $|1-\mathrm{i}| = \sqrt{2}$ and $\arg(1-\mathrm{i}) = -\frac{\pi}{4} + 2k\pi$, $k \in \mathbb{Z}$, we obtain
$$(1-\mathrm{i})^{\mathrm{i}} = \exp\Big(\mathrm{i}\big(\ln\sqrt{2} + \mathrm{i}(-\tfrac{\pi}{4} + 2k\pi)\big)\Big) = e^{\frac{\pi}{4} - 2k\pi + \mathrm{i}\ln\sqrt{2}} \quad (k \in \mathbb{Z}).$$

Next, we have
$$(2\mathrm{i})^{1/3} = \exp\big(\tfrac{1}{3}\log(2\mathrm{i})\big) = \exp\big(\tfrac{1}{3}(\ln|2\mathrm{i}| + \mathrm{i}\arg(2\mathrm{i}))\big).$$
Since $|2\mathrm{i}| = 2$ and $\arg(2\mathrm{i}) = \frac{\pi}{2} + 2k\pi$, $k \in \mathbb{Z}$, we obtain
$$(2\mathrm{i})^{1/3} = \exp\Big(\tfrac{1}{3}\big(\ln 2 + \mathrm{i}(\tfrac{\pi}{2} + 2k\pi)\big)\Big) = e^{\frac{1}{3}\ln 2} e^{\mathrm{i}(\frac{\pi}{6} + \frac{2}{3}k\pi)} = \sqrt[3]{2}\, e^{\mathrm{i}(\frac{\pi}{6} + \frac{2}{3}k\pi)} \quad (k \in \mathbb{Z}).$$
Finally, note that the exponential $e^{\mathrm{i}(\frac{\pi}{6} + \frac{2}{3}k\pi)}$ takes on only three distinct values: $e^{\pi\mathrm{i}/6}, e^{5\pi\mathrm{i}/6}$ and $e^{3\pi\mathrm{i}/2}$. We conclude that
$$(2\mathrm{i})^{1/3} = \sqrt[3]{2}\, e^{\mathrm{i}(\frac{\pi}{6} + \frac{2}{3}k\pi)} \quad (k = 0, 1, 2).$$

3.8. Gauss sums

In this section, we use the properties of the complex exponentials covered earlier in the chapter to compute the modulus of the Gauss sum. The main result is the following theorem.

Theorem 3.24. *Let $m \in \mathbb{N}$ be odd and let $(a,m) = 1$. Define*
$$G(m,a) = \sum_{x=1}^{m} e^{2\pi \mathrm{i} ax^2/m}.$$
Then $|G(m,a)|^2 = m$.

The proof of the theorem uses two lemmas.

Lemma 3.25. *Let $m \in \mathbb{N}$ and $a \in \mathbb{Z}$. Then*

$$\sum_{x=1}^{m} e^{2\pi \mathrm{i} ax/m} = \begin{cases} m & \text{if } m \mid a, \\ 0 & \text{if } m \nmid a. \end{cases}$$

Lemma 3.26. *Let $f(x) = a_n x^n + \cdots + a_1 x + a_0$ be a polynomial with integer coefficients and let $r_1, r_2, \ldots, r_m$ be a complete system of residues modulo m. Then*

$$\sum_{k=1}^{m} e^{2\pi \mathrm{i} f(r_k)/m} = \sum_{k=1}^{m} e^{2\pi \mathrm{i} f(k)/m}.$$

Proof. Let $s_1, s_2, \ldots, s_m$ be a permutation of $r_1, r_2, \ldots, r_m$ such that $s_k \equiv k \pmod{m}$. Then

$$\sum_{k=1}^{m} e^{2\pi \mathrm{i} f(r_k)/m} = \sum_{k=1}^{m} e^{2\pi \mathrm{i} f(s_k)/m},$$

because the sum on the right is simply a rearrangement of the sum on the left. Furthermore,

$$s_k \equiv k \pmod{m} \quad \Longrightarrow \quad a_n s_k^n + \cdots + a_1 s_k + a_0 \equiv a_n k^n + \cdots + a_1 k + a_0 \pmod{m}.$$

Thus, part v) of Proposition 3.15 gives $e^{2\pi \mathrm{i} f(s_k)/m} = e^{2\pi \mathrm{i} f(k)/m}$, whence

$$\sum_{k=1}^{m} e^{2\pi \mathrm{i} f(s_k)/m} = \sum_{k=1}^{m} e^{2\pi \mathrm{i} f(k)/m}.$$

□

Proof of Theorem 3.24. We have

$$\begin{aligned} |G(m,a)|^2 &= \left(\sum_{x=1}^{m} e^{2\pi \mathrm{i} ax^2/m}\right) \overline{\left(\sum_{y=1}^{m} e^{2\pi \mathrm{i} ay^2/m}\right)} && \text{(by Proposition 3.3)} \\ &= \left(\sum_{x=1}^{m} e^{2\pi \mathrm{i} ax^2/m}\right) \left(\sum_{y=1}^{m} \overline{e^{2\pi \mathrm{i} ay^2/m}}\right) && \text{(by Exercise 3.15)} \\ &= \left(\sum_{x=1}^{m} e^{2\pi \mathrm{i} ax^2/m}\right) \left(\sum_{y=1}^{m} e^{-2\pi \mathrm{i} ay^2/m}\right) && \text{(by Exercise 3.20)} \\ &= \sum_{x=1}^{m} \sum_{y=1}^{m} e^{2\pi \mathrm{i} ax^2/m} e^{-2\pi \mathrm{i} ay^2/m} = \sum_{x=1}^{m} \sum_{y=1}^{m} e^{2\pi \mathrm{i} a(x^2-y^2)/m}. \end{aligned}$$

Next, in the sum over y, we change the summation variable to $z = x + y$. When y runs through the numbers $1, 2, \ldots, m$, z runs through the numbers $x+1, x+2, \ldots, x+m$. Also,

$$x^2 - y^2 = (x-y)(x+y) = (2x - (x+y))(x+y) = 2xz - z^2.$$

Hence,

$$|G(m,a)|^2 = \sum_{x=1}^{m} \sum_{z=x+1}^{x+m} e^{2\pi \mathrm{i} a(2xz - z^2)/m}. \tag{3.11}$$

Now, since $x+1, x+2, \ldots, x+m$ is a complete residue system modulo m, we can use Lemma 3.26 to get

$$\sum_{z=x+1}^{x+m} e^{2\pi \mathrm{i} a(2xz-z^2)/m} = \sum_{z=1}^{m} e^{2\pi \mathrm{i} a(2xz-z^2)/m}.$$

Substituting this in the right side of (3.11), we obtain

$$|G(m,a)|^2 = \sum_{x=1}^{m}\sum_{z=1}^{m} e^{2\pi \mathrm{i}a(2xz-z^2)/m} = \sum_{z=1}^{m}\sum_{x=1}^{m} e^{2\pi \mathrm{i}a(2xz-z^2)/m}$$
$$= \sum_{z=1}^{m} e^{-2\pi \mathrm{i}az^2/m} \sum_{x=1}^{m} e^{2\pi \mathrm{i}(2az)x/m}.$$

By Lemma 3.25 with $2az$ in place of a, the sum over x equals m when $m \mid 2az$ and 0 otherwise. Since $(m,a) = 1$, it follows from Theorem 1.13 that if $m \mid 2az$, then $m \mid 2z$. Moreover, since m is odd, if $m \mid 2z$, then $m \mid z$. In particular, the only number among $2az$, $1 \le z \le m$, that is divisible by m is $2am$. Therefore, the sum over x vanishes when $z = 1, 2, \dots, m-1$ and equals m when $z = m$. We conclude that

$$|G(m,a)|^2 = me^{-2\pi \mathrm{i}am^2/m} + \sum_{z=1}^{m-1} 0e^{-2\pi \mathrm{i}az^2/m} = me^{-2\pi \mathrm{i}am} = m.$$

□

3.9. Exercises

Exercise 3.1. Evaluate the expressions: $(2+3\mathrm{i})(5-2\mathrm{i})$; $\dfrac{3-2\mathrm{i}}{2+3\mathrm{i}}$; $\dfrac{3+6\mathrm{i}}{2-\mathrm{i}}$.

Exercise 3.2. Find the complex numbers z such that $z^2 = 2+\mathrm{i}$. [Hint. Let $z = x + \mathrm{i}y$. The equation $z^2 = 2+\mathrm{i}$ is equivalent to a system of two equations in the unknowns x and y. Solve that system to find z.]

Exercise 3.3. Solve the quadratic equation $z^2 + (2+\mathrm{i})z + (2+\mathrm{i}) = 0$.

Exercise 3.4. Prove parts vi), viii) and ix) of Proposition 3.1.

Exercise 3.5. Prove Proposition 3.3.

Exercise 3.6. **(a)** Use the triangle inequality to prove the two inequalities in (3.4).

(b) Use the first inequality in (3.4) to prove the triangle inequality.

(c) Use the second inequality in (3.4) to prove the triangle inequality.

Exercise 3.7. Suppose that $|z| \le 2$. Use the triangle inequality to bound the expression $2z^3 - 4z - 3 - \mathrm{i}$ from above.

Exercise 3.8. Suppose that $|z| \le \frac{1}{2}$. Use the second inequality in (3.4) to show that

$$\left|2z^3 - 4z - 3 - \mathrm{i}\right| \ge \sqrt{10} - \tfrac{5}{2}.$$

Exercise 3.9. Write the given complex numbers in exponential form: $3\mathrm{i}$; -2; $\sqrt{3} - \mathrm{i}$; $2 + 2\mathrm{i}$; $-4 + 3\mathrm{i}$.

Exercise 3.10. Use De Moivre's formula to derive formulas for: $\sin 3\theta$; $\cos 3\theta$; $\sin 4\theta$; $\cos 4\theta$.

Exercise 3.11. Let $\theta \in \mathbb{R}$. Prove *Euler's formulas*:

$$\cos\theta = \frac{e^{\mathrm{i}\theta} + e^{-\mathrm{i}\theta}}{2}, \quad \sin\theta = \frac{e^{\mathrm{i}\theta} - e^{-\mathrm{i}\theta}}{2\mathrm{i}}.$$

Exercise 3.12. Describe geometrically the following sets in the complex plane:

$$\{z \in \mathbb{C} \mid |z - \mathrm{i}| = 3\}; \quad \{z \in \mathbb{C} \mid |z| \ge 1\}; \quad \{z \in \mathbb{C} \mid 1 < |z-2| < 2\}.$$

Exercise 3.13. Describe geometrically the following sets in the complex plane:

$$\{z \in \mathbb{C} \mid \operatorname{Arg} z = \pi/4\}; \quad \{z \in \mathbb{C} \mid \pi/3 < \operatorname{Arg} z \le 2\pi/3\}; \quad \{z \in \mathbb{C} \mid 0 < \operatorname{Arg}(z - 2 + \mathrm{i}) < \tfrac{3}{4}\pi\}.$$

Exercise 3.14. Describe geometrically the following sets in the complex plane:

$$\{z \in \mathbb{C} \mid \operatorname{Im}(z - \mathrm{i}) < 2\}; \quad \{z \in \mathbb{C} \mid \operatorname{Re}(z - 2 + \mathrm{i}) > 1\};$$
$$\{z \in \mathbb{C} \mid \operatorname{Re}(z - 1 - \mathrm{i}) < 2 < \operatorname{Im}(z - 1 - \mathrm{i})\}.$$

Exercise 3.15. Use Proposition 3.3 and mathematical induction to prove that for all $n \ge 1$,

$$\overline{z_1 + z_2 + \cdots + z_n} = \bar{z}_1 + \bar{z}_2 + \cdots + \bar{z}_n, \quad \overline{z_1 z_2 \cdots z_n} = \bar{z}_1 \bar{z}_2 \cdots \bar{z}_n.$$

Exercise 3.16. Solve the equations: $\sin z = 2$; $\sinh z = -2\mathrm{i}$; $e^{2z} = -1 + \mathrm{i}$.

Exercise 3.17. Let $a \in \mathbb{R}$. Describe geometrically the sets

$$\left\{e^z \mid \operatorname{Re} z = a\right\}; \quad \left\{e^z \mid \operatorname{Im} z = a\right\}; \quad \left\{e^z \mid |z| = a\right\}.$$

Exercise 3.18. Prove parts i)–iii) and v) of Proposition 3.15.

Exercise 3.19. Prove parts ii), iv), v) and vii) of Proposition 3.16.

Exercise 3.20. Prove the identities:

$$\overline{e^z} = e^{\bar{z}}, \quad \overline{\sin z} = \sin \bar{z}, \quad \overline{\cos z} = \cos \bar{z}, \quad \overline{\tan z} = \tan \bar{z}, \quad \overline{\sinh z} = \sinh \bar{z}, \quad \overline{\cosh z} = \cosh \bar{z}.$$

Exercise 3.21. Evaluate the expressions: $(1-\mathrm{i})^{4\mathrm{i}}$; $(-1+\mathrm{i}\sqrt{3})^{3/2}$; $\operatorname{Log}(1-\mathrm{i})$.

Exercise 3.22. Find all the solutions of the equations: $\sin z = \cosh 3$; $\cosh z = \mathrm{i}$.

Exercise 3.23. Prove Lemma 3.25.

Exercise 3.24. Let $m \in \mathbb{N}$ be even and $(a, m) = 1$, and let $G(m, a)$ be the sum in Theorem 3.24. Modify the proof of Theorem 3.24 to prove that

$$|G(m,a)|^2 = \begin{cases} 2m & \text{if } m \equiv 0 \pmod 4, \\ 0 & \text{if } m \equiv 2 \pmod 4. \end{cases}$$

Exercise 3.25. The purpose of this exercise is to provide an easy proof of the trigonometric identities

$$\sin\theta + \sin 2\theta + \cdots + \sin n\theta = \frac{\sin\left(\frac{1}{2}n\theta\right)\sin\left(\frac{1}{2}(n+1)\theta\right)}{\sin\left(\frac{1}{2}\theta\right)} \quad (\theta \neq 2k\pi), \tag{$*$}$$

$$\cos\theta + \cos 2\theta + \cdots + \cos n\theta = \frac{\sin\left((n+\frac{1}{2})\theta\right)}{2\sin\left(\frac{1}{2}\theta\right)} - \frac{1}{2} \quad (\theta \neq 2k\pi). \tag{$**$}$$

(a) Prove that $1 + e^{\mathrm{i}\theta} + e^{\mathrm{i}2\theta} + \cdots + e^{\mathrm{i}n\theta} = \dfrac{e^{\mathrm{i}(n+1)\theta} - 1}{e^{\mathrm{i}\theta} - 1}$.

(b) Use Euler's formula for the sine function to show that $\dfrac{e^{\mathrm{i}(n+1)\theta} - 1}{e^{\mathrm{i}\theta} - 1} = \dfrac{e^{\mathrm{i}n\theta/2}\sin\left(\frac{1}{2}(n+1)\theta\right)}{\sin\left(\frac{1}{2}\theta\right)}$.

(c) Use parts (a) and (b) to show that

$$e^{\mathrm{i}\theta} + e^{\mathrm{i}2\theta} + \cdots + e^{\mathrm{i}n\theta} = \frac{e^{\mathrm{i}n\theta/2}\sin\left(\frac{1}{2}(n+1)\theta\right)}{\sin\left(\frac{1}{2}\theta\right)} - 1.$$

(d) Prove the identities ($*$) and ($**$) by comparing the real and imaginary parts of the two sides of the identity in part (c).

Exercise 3.26. Prove the identities:

(a) $\arcsin x = -\mathrm{i}\operatorname{Log}\left(\mathrm{i}x + \sqrt{1-x^2}\right)$, where $-1 \leq x \leq 1$;

(b) $\arctan x = \dfrac{\mathrm{i}}{2}\operatorname{Log}\left(\dfrac{\mathrm{i}+x}{\mathrm{i}-x}\right)$, where $x \in \mathbb{R}$.

Exercise 3.27. Let $n \in \mathbb{N}$. An *nth root of unity* is a solution of the equation $z^n = 1$, i.e., one of the n values of $1^{1/n}$. A *primitive nth root of unity* is an nth root of unity which is not an mth root of unity for any $m < n$. For example, -1 is a primitive second root of unity, but 1 is not, since it is also a first root of unity. Also, i and $-$i are primitive fourth roots of unity, but 1 and -1 are not, since they are also second roots of unity.

(a) Prove that ζ is an nth root of unity if and only if $\zeta = e^{2\pi\mathrm{i}k/n}$ for some integer k.

(b) Prove that if ζ is a non-primitive nth root of unity, then $\zeta^m = 1$ for some positive integer m with $m \mid n$. [Hint. Let m be the least positive integer such that ζ is an mth root of unity. Write n in the form $n = mq + r$, with $0 \leq r < m$, and show that $r = 0$. Deduce that $m \mid n$.]

(c) Prove that ζ is a primitive nth root of unity if and only if $\zeta = e^{2\pi\mathrm{i}k/n}$ for some integer k with $(k, n) = 1$.

(d) Prove that if $(a, n) = 1$, the numbers $1, e^{2\pi\mathrm{i}a/n}, e^{2\pi\mathrm{i}2a/n}, \ldots, e^{2\pi\mathrm{i}(n-1)a/n}$ are a complete list of nth roots of unity.

CHAPTER 4

Algebra Over The Complex Numbers

4.1. Roots of polynomials with complex coefficients

Recall that a *zero* (or a *root*) of the polynomial $f(z) = a_n z^n + \cdots + a_1 z + a_0$ is a number α such that $f(\alpha) = 0$. If α is a root of a polynomial $f(z)$ of degree n, we can use long division to express $f(z)$ in the form

$$f(z) = (z - \alpha)g(z), \tag{4.1}$$

where $g(z)$ is a polynomial of degree $n - 1$.

Example 4.1. Let $f(z) = z^6 + 3z^2 + 4$. Then $f(\mathrm{i}) = \mathrm{i}^6 + 3\mathrm{i}^2 + 4 = 0$, and long division of $f(z)$ by $z - \mathrm{i}$ gives

$$f(z) = (z - \mathrm{i})\big(z^5 + \mathrm{i}z^4 - z^3 - \mathrm{i}z^2 + 4z + 4\mathrm{i}\big).$$

If we have $g(\alpha) \neq 0$ in (4.1), we say that α is a *simple root* of $f(z)$. On the other hand, if $g(\alpha) = 0$, we can apply (4.1) to $g(z)$ to obtain $g(z) = (z - \alpha)h(z)$, where $h(z)$ is a polynomial of degree $n - 2$. Substituting this expression for $g(z)$, we obtain

$$f(z) = (z - \alpha)^2 h(z).$$

If $h(\alpha) \neq 0$, we say that α is a *double root* of $f(z)$. On the other hand, if $h(\alpha) = 0$, we can repeat the above procedure yet again. In general, if α is a root of $f(z)$, we can find an integer m such that

$$f(z) = (z - \alpha)^m g(z),$$

where $g(z)$ is some polynomial of degree $n - m$ such that $g(\alpha) \neq 0$. The integer m is called the *multiplicity* of α, and we say that α is a *zero of multiplicity m*.

Where roots are concerned, there is a significant difference between polynomials over the real numbers and polynomials over the complex numbers. Recall that a polynomial with real coefficients may have no real zeros. For example, $f(z) = z^2 + 1$ has no real zeros, because its value is at least 1 for all real values of the variable z. However, if we allow z to take on complex values, then we have the following remarkable result.

Theorem 4.2 (Fundamental theorem of algebra). *Let $f(z)$ be a polynomial (with complex coefficients) of degree n. Then $f(z)$ has exactly n complex roots, counting multiplicities.*

The phrase "counting multiplicities" means that if a polynomial of degree five has a zero at $z = 2$ that is a zero of multiplicity 3 and another zero at $z = -1$ that has multiplicity 2, then the five zeros of $f(z)$ are $-1, -1, 2, 2, 2$.

Note that Theorem 4.2 says that a polynomial of degree n has exactly n complex roots, but it gives no clue how to find those. You know from high-school algebra the quadratic formula: the roots of $f(z) = az^2 + bz + c$ are

$$\alpha_{1,2} = \frac{-b + (b^2 - 4ac)^{1/2}}{2a}.$$

There are similar formulas for polynomials of degrees 3 or 4. Those formulas resemble the quadratic formula, except that they are much messier. For example, one of the solutions (the "nicest" of them) of the cubic equation

$$z^3 + pz + q = 0$$

is given by the formula

$$\alpha_1 = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}.$$

When the degree is 5 or higher, the situation is even worse. Then we have a theorem that says that, in fact, such a general formula does not exist. That is not to say that we cannot solve any polynomial equation of degree 5, just that there is no way to express the solutions in terms of the coefficients using only radicals and the basic arithmetic operations in $\mathbb{C}$.

Although the fundamental theorem of algebra does not yield an algorithm for solving polynomial equations, it does provide a great deal of valuable information about polynomials.

Corollary 4.3. *If $f(z) = a_nz^n + \cdots + a_1z + a_0$ and $g(z) = b_nz^n + \cdots + b_1z + b_0$ are polynomials of degree n such that $f(\alpha_j) = g(\alpha_j)$ for $n+1$ distinct numbers $\alpha_1, \alpha_2, \dots, \alpha_{n+1}$, then $f(z) = g(z)$: that is, $a_j = b_j$ for $j = 0, 1, \dots, n$.*

Corollary 4.4. *If $f(z) = a_nz^n + \cdots + a_1z + a_0$ is a polynomial of degree n and $\alpha_1, \alpha_2, \dots, \alpha_n$ are its roots, listed according to their multiplicities, then*

$$f(z) = a_n(z - \alpha_1)(z - \alpha_2)\cdots(z - \alpha_n).$$

Corollary 4.5. *Let $f(z) = a_nz^n + \cdots$ and $g(z) = a_nz^n + \cdots$ be two polynomials of degree n with the same leading coefficient and with the same roots, then $f(z) = g(z)$.*

Corollary 4.6. *Let $f(z) = a_nz^n + \cdots$ and $g(z) = a_nz^n + \cdots$ be two polynomials of degree n with the same leading coefficient. If $f(\alpha_j) = g(\alpha_j)$ for n distinct numbers $\alpha_1, \alpha_2, \dots, \alpha_n$, then $f(z) = g(z)$.*

Theorem 4.7. *Let $f(z)$ be a polynomial with real coefficients and suppose that α is a complex root of $f(z)$. Then $\bar{\alpha}$ is also a root of $f(z)$.*

Proof. Let $f(z) = a_nz^n + \cdots + a_1z + a_0$, where $a_0, a_1, \dots, a_n \in \mathbb{R}$. Note that since the a_j's are real, we have $a_j = \bar{a}_j$ for all $j = 0, 1, \dots, n$. Then, on using the identities in Exercise 3.15,

$$\begin{aligned} f(\bar{\alpha}) &= a_n(\bar{\alpha})^n + a_{n-1}(\bar{\alpha})^{n-1} + \cdots + a_1\bar{\alpha} + a_0 \\ &= \bar{a}_n(\bar{\alpha})^n + \bar{a}_{n-1}(\bar{\alpha})^{n-1} + \cdots + \bar{a}_1\bar{\alpha} + \bar{a}_0 \\ &= \overline{a_n\alpha^n} + \overline{a_{n-1}\alpha^{n-1}} + \cdots + \overline{a_1\alpha} + \bar{a}_0 \\ &= \overline{a_n\alpha^n + a_{n-1}\alpha^{n-1} + \cdots + a_1\alpha + a_0} = \overline{f(\alpha)} = \bar{0} = 0. \end{aligned}$$

That is, $\bar{\alpha}$ is also a root of $f(z)$. □

In particular, the non-real roots of a polynomial with real coefficients come in pairs $a \pm b$i.

Example 4.8. Let $f(z) = z^5 - 2z^3 + 2z^2 - 3z + 2$. Then $f(\mathrm{i}) = 0$, so it follows from Theorem 4.7 that $-\mathrm{i}$ is also a root of $f(z)$. Hence, by Corollary 4.4, we have

$$f(z) = (z - \mathrm{i})(z + \mathrm{i})(z - \alpha_3)(z - \alpha_4)(z - \alpha_5) = (z^2 + 1)g(z), \quad \text{say},$$

where α_3, α_4 and α_5 are the remaining three complex roots of $f(z)$. Using long division, we find that $g(z) = z^3 - 3z + 2$. Since $g(1) = 0$, we have $g(z) = (z-1)h(z)$, and long division gives

$$h(z) = z^2 + z - 2 = (z+2)(z-1).$$

Hence,

$$f(z) = (z^2+1)(z-1)^2(z+2).$$

The idea behind the above example can be used to give an inductive proof of the following result.

Theorem 4.9. *Let $f(x) = a_n x^n + \cdots + a_1 x + a_0$ be a polynomial with real coefficients of degree n. Then $f(x)$ can be expressed in the form*

$$f(x) = a_n(x - r_1)\cdots(x - r_k)(x^2 + p_1 x + q_1)\cdots(x^2 + p_m x + q_m),$$

where $r_1, \dots, r_k$ are the real roots of $f(x)$ listed according to their multiplicities and $p_1, q_1, p_2, q_2, \dots, p_m, q_m$ are real numbers such that $p_j^2 - 4q_j < 0$ for all $j = 1, \dots, m$. Also, $k + 2m = n$.

4.2. Linear algebra over the complex numbers

In linear algebra too, working with complex vectors and matrices sometimes has advantages. Recall that a (complex) number λ is called an *eigenvalue* of an $n \times n$ matrix A if there exists a nonzero n-dimensional vector $\mathbf{x}$ such that $A\mathbf{x} = \lambda\mathbf{x}$; any such vector $\mathbf{x}$ is called an *eigenvector* for λ. We recall the following theorem from linear algebra.

Theorem 4.10. *Let A be an $n \times n$ matrix. A complex number λ is an eigenvalue of A if and only if* $\det(A - \lambda I) = 0$, *where I is the $n \times n$ identity matrix. If λ is an eigenvalue of A, the eigenvectors of A for λ are the nonzero solutions of the linear system* $(A - \lambda I)\mathbf{x} = \mathbf{0}$.

The determinant $\det(A - zI)$ is called the *characteristic polynomial* of A. Expanding this determinant yields

$$\det(A - zI) = (-1)^n z^n + b_1 z^{n-1} + \cdots + b_n, \tag{4.2}$$

where

$$b_1 = (-1)^{n-1}\sum_{k=1}^{n} a_{kk}, \quad b_n = \det A, \tag{4.3}$$

a_{ij} being the (i, j)-th entry of the matrix A. In particular, A has exactly n complex eigenvalues, counting multiplicities, because the polynomial (4.2) has exactly n roots, by the fundamental theorem of algebra. Let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the eigenvalues of A. By Corollary 4.4,

$$\det(A - zI) = (-1)^n (z - \lambda_1)(z - \lambda_2)\cdots(z - \lambda_n).$$

Multiplying out the right side of this identity, we obtain alternative expressions for the coefficients b_1 and b_n in (4.2):

$$b_1 = (-1)^{n+1}\sum_{k=1}^{n} \lambda_k, \quad b_n = (-1)^{2n}\prod_{k=1}^{n} \lambda_k.$$

Comparing these expressions with (4.3), we obtain

$$\operatorname{tr} A = \sum_{k=1}^{n} a_{kk} = \sum_{k=1}^{n} \lambda_k, \quad \det A = \prod_{k=1}^{n} \lambda_k. \tag{4.4}$$

(The sum of the diagonal entries of a square matrix A is called the *trace* of A and is denoted $\operatorname{tr} A$.)

Example 4.11. Compute the eigenvalues and eigenvectors of the matrix $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

Solution. The characteristic polynomial of A is

$$\begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1,$$

so the eigenvalues are i and −i. Since

$$\begin{bmatrix} -\mathrm{i} & -1 & 0 \\ 1 & -\mathrm{i} & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -\mathrm{i} & 0 \\ -\mathrm{i} & -1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -\mathrm{i} & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

the eigenspace for $\lambda = \mathrm{i}$ contains the solutions $\mathbf{x}$ of

$$x_1 - \mathrm{i}x_2 = 0 \quad \Longleftrightarrow \quad \mathbf{x} = x_2 \begin{bmatrix} \mathrm{i} \\ 1 \end{bmatrix} \quad \Longleftrightarrow \quad \mathbf{x} \in \operatorname{Span} \left\{ \begin{bmatrix} \mathrm{i} \\ 1 \end{bmatrix} \right\}.$$

Similarly, the eigenspace for $\lambda = -\mathrm{i}$ is $\operatorname{Span} \left\{ \begin{bmatrix} -\mathrm{i} \\ 1 \end{bmatrix} \right\}$.

Notice that

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \mathrm{i} \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ \mathrm{i} \end{bmatrix} = \mathrm{i} \begin{bmatrix} \mathrm{i} \\ 1 \end{bmatrix}$$

and

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -\mathrm{i} \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -\mathrm{i} \end{bmatrix} = (-\mathrm{i}) \begin{bmatrix} -\mathrm{i} \\ 1 \end{bmatrix},$$

so the vectors $\begin{bmatrix} \mathrm{i} \\ 1 \end{bmatrix}$ and $\begin{bmatrix} -\mathrm{i} \\ 1 \end{bmatrix}$ are indeed eigenvectors for $\lambda = \mathrm{i}$ and $\lambda = -\mathrm{i}$. □

Recall from linear algebra that if A is an $n \times n$ real matrix with n real eigenvalues $\lambda_1, \dots, \lambda_n$ (listed according to their multiplicities) and n linearly independent eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ such that $A\mathbf{x}_j = \lambda_j \mathbf{x}_j$, then A is *diagonalizable*, i.e., there exist an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$. Moreover,

$$P = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{bmatrix}, \quad D = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}.$$

The same is true for complex matrices, except the eigenvalues, eigenvectors and the matrices P and D are allowed to be complex numbers, vectors and matrices.

Example 4.12. Diagonalize (over $\mathbb{C}$) the matrix A from Example 4.11.

Solution. We have $A = PDP^{-1}$, where

$$P = \begin{bmatrix} \mathrm{i} & -\mathrm{i} \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{bmatrix}.$$

□

Recall that an $n \times n$ real matrix A may fail to be diagonalizable over $\mathbb{R}$ for two reasons: it may have fewer than n real eigenvalues (even counting multiplicities); or it may have n eigenvalues, but fewer than n linearly independent eigenvectors. Because of the fundamental theorem of algebra, the former scenario never occurs for complex matrices: every $n \times n$ matrix has n complex eigenvalues. In particular, a real matrix that has fewer than n

real eigenvalues must have some nonreal eigenvalues. Since the characteristic polynomial of A has real coefficients, by Theorem 4.7, the nonreal eigenvalues of A come in complex-conjugate pairs $a \pm \mathrm{i}b$. It turns out that the same holds for the complex eigenvectors of A. We extend the definitions of complex conjugate, real part and imaginary part of a complex number to vectors in $\mathbb{C}^n$ (i.e., n-dimensional vectors with complex entries):

$$\text{for } \mathbf{z} = \begin{bmatrix} x_1 + \mathrm{i}y_1 \\ x_2 + \mathrm{i}y_2 \\ \vdots \\ x_n + \mathrm{i}y_n \end{bmatrix}: \qquad \bar{\mathbf{z}} = \begin{bmatrix} x_1 - \mathrm{i}y_1 \\ x_2 - \mathrm{i}y_2 \\ \vdots \\ x_n - \mathrm{i}y_n \end{bmatrix}, \quad \operatorname{Re}\mathbf{z} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \operatorname{Im}\mathbf{z} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

With these definitions, we have the following result.

Theorem 4.13. *Let A be an $n \times n$ matrix with real entries. If λ and $\bar{\lambda}$ are a pair of complex-conjugate eigenvalues of A and $\mathbf{z}$ is an eigenvector for λ, then $\bar{\mathbf{z}}$ is an eigenvector for $\bar{\lambda}$.*

Notice that this is consistent with Example 4.11: we found the complex conjugate eigenvalues ±i, whose respective eigenvectors were also complex conjugate:

$$\overline{\begin{bmatrix} \mathrm{i} \\ 1 \end{bmatrix}} = \begin{bmatrix} -\mathrm{i} \\ 1 \end{bmatrix}.$$

Consider a real matrix A that is diagonalizable over $\mathbb{C}$, but not over $\mathbb{R}$. That is, there exist an invertible matrix P and a diagonal matrix D with complex entries such that $A = PDP^{-1}$, but no such matrices with real entries. Is it possible to represent A as $A = QCQ^{-1}$, where Q and C are real matrices and C, though not diagonal, is still pretty simple to work with? Not only is the answer to this question "yes", but the matrices Q and C in this representation can be easily derived from the matrices P and D in the diagonalization of A over the complex numbers. Since A is diagonalizable over $\mathbb{C}$, we can find real eigenvalues $\lambda_1, \dots, \lambda_k$ and pairs of complex conjugate eigenvalues $\mu_1, \bar{\mu}_1, \dots, \mu_m, \bar{\mu}_m$, altogether $n = k + 2m$ of them. We can also find k linearly independent real eigenvectors $\mathbf{x}_1, \dots, \mathbf{x}_k$ corresponding to the eigenvalues $\lambda_1, \dots, \lambda_k$ and m linearly independent pairs of complex eigenvectors $\mathbf{z}_1, \bar{\mathbf{z}}_1, \dots, \mathbf{z}_m, \bar{\mathbf{z}}_m$ corresponding to the pairs of complex eigenvalues $\mu_1, \bar{\mu}_1, \dots, \mu_m, \bar{\mu}_m$. Then the matrix Q in the desired representation has columns $\mathbf{x}_1, \dots, \mathbf{x}_k$, $\operatorname{Re}\mathbf{z}_1, \operatorname{Im}\mathbf{z}_1, \dots, \operatorname{Re}\mathbf{z}_m, \operatorname{Im}\mathbf{z}_m$, and the matrix C is a block matrix of the form (a *block-diagonal matrix*)

$$\begin{bmatrix} \lambda_1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & \lambda_k & 0 & \cdots & 0 \\ 0 & \cdots & 0 & C_1 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & C_m \end{bmatrix}, \tag{4.5}$$

where $C_1, \dots, C_m$ are 2×2 blocks of the form

$$C_j = \begin{bmatrix} \operatorname{Re}\mu_j & \operatorname{Im}\mu_j \\ -\operatorname{Im}\mu_j & \operatorname{Re}\mu_j \end{bmatrix}.$$

Example 4.14. If possible, represent the matrix

$$A = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 2 & 0 & 1 & 2 & -1 \\ 1 & 3 & 0 & 1 & 2 \end{bmatrix}$$

as $A = QCQ^{-1}$, where Q and C are real matrices and C is either diagonal or block-diagonal of the form (4.5) above.

Solution. The characteristic polynomial of A is

$$\begin{vmatrix} 2-\lambda & 0 & 0 & 0 & 0 \\ 0 & 2-\lambda & 0 & 0 & 0 \\ 0 & 1 & 1-\lambda & 0 & 0 \\ 2 & 0 & 1 & 2-\lambda & -1 \\ 1 & 3 & 0 & 1 & 2-\lambda \end{vmatrix} = (2-\lambda)^2(1-\lambda)\begin{vmatrix} 2-\lambda & -1 \\ 1 & 2-\lambda \end{vmatrix}$$

$$= (2-\lambda)^2(1-\lambda)\left[(2-\lambda)^2+1\right],$$

so the eigenvalues of A are $1, 2, 2$, and the two roots of $(2-\lambda)^2+1=0$:

$$(2-\lambda)^2 = -1 \iff 2-\lambda = \pm \mathrm{i} \iff \lambda = 2 \pm \mathrm{i}.$$

To find the eigenvectors for $\lambda = 2-\mathrm{i}$, we solve $(A-(2-\mathrm{i})I)\mathbf{x} = \mathbf{0}$:

$$\begin{bmatrix} A-(2-\mathrm{i})I & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathrm{i} & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathrm{i} & 0 & 0 & 0 & 0 \\ 0 & 1 & -1+\mathrm{i} & 0 & 0 & 0 \\ 2 & 0 & 1 & \mathrm{i} & -1 & 0 \\ 1 & 3 & 0 & 1 & \mathrm{i} & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathrm{i} & 0 & 0 & 0 & 0 \\ 0 & 1 & -1+\mathrm{i} & 0 & 0 & 0 \\ 2 & 0 & 1 & \mathrm{i} & -1 & 0 \\ 1 & 3 & 0 & 1 & \mathrm{i} & 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathrm{i} & 0 & 0 & 0 & 0 \\ 0 & 1 & -1+\mathrm{i} & 0 & 0 & 0 \\ 0 & 0 & 1 & \mathrm{i} & -1 & 0 \\ 0 & 3 & 0 & 1 & \mathrm{i} & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1+\mathrm{i} & 0 & 0 & 0 \\ 0 & 0 & 1 & \mathrm{i} & -1 & 0 \\ 0 & 3 & 0 & 1 & \mathrm{i} & 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1+\mathrm{i} & 0 & 0 & 0 \\ 0 & 0 & 1 & \mathrm{i} & -1 & 0 \\ 0 & 0 & 0 & 1 & \mathrm{i} & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & \mathrm{i} & -1 & 0 \\ 0 & 0 & 0 & 1 & \mathrm{i} & 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathrm{i} & -1 & 0 \\ 0 & 0 & 0 & 1 & \mathrm{i} & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \mathrm{i} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Hence, the vector form of the solution of $(A - (2 - \mathrm{i})I)\mathbf{x} = \mathbf{0}$ and a basis for the eigenspace for $\lambda = 2 - \mathrm{i}$ are

$$\mathbf{x} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -\mathrm{i}x_5 \\ x_5 \end{bmatrix} = x_5 \begin{bmatrix} 0 \\ 0 \\ 0 \\ -\mathrm{i} \\ 1 \end{bmatrix}, \qquad \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \\ -\mathrm{i} \\ 1 \end{bmatrix} \right\}.$$

By Theorem 4.13, a basis for this eigenspace $\lambda = 2 + \mathrm{i}$ is

$$\left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \\ \mathrm{i} \\ 1 \end{bmatrix} \right\}.$$

Furthermore, computations similar to the above with $\lambda = 2$ and $\lambda = 1$ yield the bases

$$\left\{ \begin{bmatrix} 1 \\ -2 \\ -2 \\ 5 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ -1 \\ -1 \\ 0 \\ 5 \end{bmatrix} \right\} \quad \text{and} \quad \left\{ \begin{bmatrix} 0 \\ 0 \\ 2 \\ -1 \\ 1 \end{bmatrix} \right\}$$

for the eigenspaces for $\lambda = 2$ and $\lambda = 1$, respectively.

It follows that A is diagonalizable (over $\mathbb{C}$) and $A = PDP^{-1}$, where

$$P = \begin{bmatrix} 1 & 3 & 0 & 0 & 0 \\ -2 & -1 & 0 & 0 & 0 \\ -2 & -1 & 2 & 0 & 0 \\ 5 & 0 & -1 & -\mathrm{i} & \mathrm{i} \\ 0 & 5 & 1 & 1 & 1 \end{bmatrix}, \qquad D = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 - \mathrm{i} & 0 \\ 0 & 0 & 0 & 0 & 2 + \mathrm{i} \end{bmatrix}.$$

Since A has nonreal eigenvalues, it is not diagonalizable over $\mathbb{R}$, but it can be represented as $A = QCQ^{-1}$, where

$$Q = \begin{bmatrix} 1 & 3 & 0 & 0 & 0 \\ -2 & -1 & 0 & 0 & 0 \\ -2 & -1 & 2 & 0 & 0 \\ 5 & 0 & -1 & 0 & -1 \\ 0 & 5 & 1 & 1 & 0 \end{bmatrix}, \qquad C = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix}.$$

□

Theorem 4.15. *If $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of the $n \times n$ matrix A, then the eigenvalues of A^2 are the numbers $\lambda_1^2, \lambda_2^2, \ldots, \lambda_n^2$.*

Proof. Let $f(z)$ denote the characteristic polynomial of A, and let $g(z)$ denote the characteristic polynomial of A^2. By Corollary 4.4,

$$f(z) = (-1)^n(z - \lambda_1)(z - \lambda_2) \cdots (z - \lambda_n) = \prod_{j=1}^{n} (\lambda_j - z).$$

We shall show that

$$g(z) = \prod_{j=1}^{n} \left(\lambda_j^2 - z\right). \tag{4.6}$$

Let m be an integer. Since

$$A^2 - m^2 I = (A - mI)(A + mI),$$

we have

$$\begin{aligned} g(m^2) &= \det(A^2 - m^2 I) = \det(A - mI)\det(A + mI) = f(m)f(-m) \\ &= \prod_{j=1}^{n}(\lambda_j - m)\prod_{j=1}^{n}(\lambda_j + m) = \prod_{j=1}^{n}(\lambda_j^2 - m^2). \end{aligned}$$

This establishes that the two sides of (4.6) are equal when z is replaced by the square of any integer m. Since the two sides of (4.6) are polynomials of degree n, the identity follows from Corollary 4.3. □

4.3. Review of determinants

You should be familiar with the definition and properties of determinants from linear algebra. However, the focus of linear algebra courses is rarely on the computation of determinants, and hence, it is likely that you have a limited experience with proofs of formulas for special $n \times n$ determinants.

Definition 4.16. Let A be an $n \times n$ matrix, whose (i, j)th entry is denoted by a_{ij}. The *determinant of* A, denoted $\det A$, is defined by the following recursive procedure:

1. If $n = 1$ and $A = [a_{11}]$, then $\det A = a_{11}$.
2. If $n \geq 2$, for each i and j, we introduce the matrix A_{ij}, which is the $(n-1)\times(n-1)$ matrix obtained from A by deleting its ith row and jth column. Then

$$\det A = a_{11}\det A_{11} - a_{12}\det A_{12} + \cdots + (-1)^{1+n}a_{1n}\det A_{1n}. \tag{4.7}$$

Formula (4.7) is known as the *expansion of* $\det A$ *along the first row*.

Example 4.17. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Then

$$a_{11} = a, \quad a_{12} = b, \quad a_{21} = c, \quad a_{22} = d, \quad A_{11} = \begin{bmatrix} d \end{bmatrix}, \quad A_{12} = \begin{bmatrix} c \end{bmatrix},$$

so (4.7) with $n = 2$ gives

$$\det A = a_{11}\det A_{11} - a_{12}\det A_{12} = ad - bc.$$

Notice that this is the usual expression for the determinant of a 2×2 matrix.

Recall from linear algebra that (4.7) is a special case of a more general formula known as the *cofactor expansion of a determinant*. Given an $n \times n$ matrix A with entries a_{ij}, the (i, j)*th cofactor of* A is the number

$$C_{ij} = (-1)^{i+j}\det A_{ij}.$$

That is, up to a sign, the cofactor is the determinant $\det A_{ij}$. The sign $(-1)^{i+j}$ depends on the position of the entry a_{ij} in the following way:

$$\begin{bmatrix} + & - & + & \cdots \\ - & + & - & \cdots \\ + & - & + & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

In particular, the right side of (4.7) is just the sum of the products of the numbers in the first row of A and their respective cofactors:

$$\det A = a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n}.$$

It turns out that we can replace the first row by any row or column of A.

Theorem 4.18. *The determinant of an $n \times n$ matrix A can be computed by a cofactor expansion across any row or down any column of A. More precisely, for any i, $1 \le i \le n$, we have*

$$\det A = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}, \tag{4.8}$$

and for any j, $1 \le j \le n$, we have

$$\det A = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}. \tag{4.9}$$

Formulas (4.8) and (4.9) are called the *cofactor expansion across the ith row* and the *cofactor expansion down the jth column*, respectively. We now show how to use these formulas to take advantage of possible zero entries.

Corollary 4.19. *If A is a triangular matrix, then* $\det A$ *is the product of the entries on the main diagonal.*

Proof. We will prove the corollary for upper triangular matrices. The proof for lower triangular matrices is similar. We argue by induction on the size of A.

If $A = [a_{11}]$, then $\det A = a_{11}$, which is the product of the diagonal entries of A.

Suppose now that $n \ge 2$ and the corollary holds for $(n-1) \times (n-1)$ upper diagonal matrices. Consider an $n \times n$ diagonal matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}.$$

Expanding $\det A$ across the last row, we get

$$\det A = (-1)^{2n} a_{nn} \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} \\ 0 & a_{22} & \cdots & a_{2,n-1} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{n-1,n-1} \end{vmatrix}.$$

Since the last determinant is of order $n-1$, it equals $a_{11}a_{22}\cdots a_{n-1,n-1}$, by the inductive hypothesis. Hence, $\det A = a_{11}a_{22}\cdots a_{n-1,n-1}a_{nn}$, which completes the induction. □

Having various cofactor expansions to play with is only useful if the matrix contains many zeros. If there are no zero entries, no matter which cofactor expansion we use, we will end up performing tons of arithmetic operations. We can avoid this by using the properties of determinants to replace the given determinant by an equal one that does contain many zeros. Here is a list of properties, which are useful in this context.

Proposition 4.20. *Let A and B be square matrices. Then:*

i) *If B is obtained from A by a row replacement, then* $\det B = \det A$.
ii) *If B is obtained from A by the interchange of two rows, then* $\det B = -\det A$.
iii) *If B is obtained from A by multiplying one of its rows by a number k, then* $\det B = k \det A$.
iv) $\det A^t = \det A$.

v) $\det(AB) = (\det A)(\det B)$.
vi) *A is invertible if and only if* $\det A \neq 0$.

Corollary 4.21. *If A is a square matrix that has two equal rows or two equal columns, then* $\det A = 0$.

Next, we shall use the properties of determinants stated in Proposition 4.20 to evaluate a special determinant. Let $n \geq 2$ and let $z_1, z_2, \dots, z_n$ be complex numbers. The $n \times n$ determinant

$$V(z_1, z_2, \dots, z_n) = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ z_1 & z_2 & \cdots & z_n \\ z_1^2 & z_2^2 & \cdots & z_n^2 \\ \vdots & \vdots & & \vdots \\ z_1^{n-1} & z_2^{n-1} & \cdots & z_n^{n-1} \end{vmatrix}$$

is called *Vandermonde's determinant* of the variables $z_1, z_2, \dots, z_n$.

Theorem 4.22. *Let* $V(z_1, \dots, z_n)$ *be Vandermonde's determinant of the variables* $z_1, \dots, z_n$. *Then*

$$V(z_1, \dots, z_n) = \prod_{1 \leq i < j \leq n} (z_j - z_i). \tag{4.10}$$

First proof of Theorem 4.22. If two of the numbers $z_1, z_2, \dots, z_n$ are equal, Vandermonde's determinant is zero by Corollary 4.21. The right side of (4.10) is also zero, because some of the differences in the product is zero.

Next, we consider the case when all n numbers are distinct and argue by induction on n. When $n = 2$, we have

$$V(z_1, z_2) = \begin{vmatrix} 1 & 1 \\ z_1 & z_2 \end{vmatrix} = z_2 - z_1 = \prod_{1 \leq i < j \leq 2} (z_j - z_i).$$

Suppose that $n \geq 3$ and (4.10) holds for $V(z_1, \dots, z_{n-1})$. Expanding $V(z_1, \dots, z_n)$ down its last column, we get

$$V(z_1, \dots, z_n) = V(z_1, \dots, z_{n-1}) z_n^{n-1} + D_1 z_n^{n-2} + \cdots + D_{n-2} z_n + D_{n-1},$$

where $D_1, D_2, \dots, D_{n-1}$ are determinants with entries that depend only on $z_1, \dots, z_{n-1}$. Therefore, as a function of z_n, $V(z_1, \dots, z_n)$ is a polynomial of degree $n - 1$ with leading coefficient

$$V(z_1, \dots, z_{n-1}) = \prod_{1 \leq i < j \leq n-1} (z_j - z_i).$$

Further, by Corollary 4.21, we have

$$V(z_1, \dots, z_{n-1}, z_1) = 0, \quad V(z_1, \dots, z_{n-1}, z_2) = 0, \quad \dots, \quad V(z_1, \dots, z_{n-1}, z_{n-1}) = 0,$$

because in each of these determinants the last column equals one of the other columns. These equations show that the numbers $z_1, z_2, \dots, z_{n-1}$ are roots of $V(z_1, \dots, z_n)$ (as a polynomial in z_n). That is, $V(z_1, \dots, z_n)$ is a polynomial of degree $n - 1$ with roots $z_1, \dots, z_{n-1}$ and leading coefficient $V(z_1, \dots, z_{n-1})$. On the other hand, the right side of (4.10) is

$$\prod_{1 \leq i < j \leq n-1} (z_j - z_i)(z_n - z_1) \cdots (z_n - z_{n-1}),$$

that is, the right side of (4.10) is also a polynomial of degree $n - 1$ with roots $z_1, \dots, z_{n-1}$ and leading coefficient $V(z_1, \dots, z_{n-1})$. By Corollary 4.5, these two polynomials are equal. This completes the induction and establishes (4.10). □

Second proof of Theorem 4.22. We argue by induction on n. When $n = 2$, we have

$$V(z_1, z_2) = \begin{vmatrix} 1 & 1 \\ z_1 & z_2 \end{vmatrix} = z_2 - z_1 = \prod_{1 \le i < j \le 2} (z_j - z_i).$$

Suppose that $n \geq 3$ and (4.10) holds for $V(z_2, \ldots, z_n)$. By parts i) and iv) of Proposition 4.20, we can subtract the first column of Vandermonde's determinant from each of the remaining $n-1$ columns to get

$$V(z_1, \ldots, z_n) = \begin{vmatrix} 1 & 0 & \cdots & 0 \\ z_1 & z_2 - z_1 & \cdots & z_n - z_1 \\ z_1^2 & z_2^2 - z_1^2 & \cdots & z_n^2 - z_1^2 \\ \vdots & \vdots & & \vdots \\ z_1^{n-1} & z_2^{n-1} - z_1^{n-1} & \cdots & z_n^{n-1} - z_1^{n-1} \end{vmatrix}.$$

Expanding the last determinant across the first row, we obtain

$$V(z_1, \ldots, z_n) = \begin{vmatrix} z_2 - z_1 & z_3 - z_1 & \cdots & z_n - z_1 \\ z_2^2 - z_1^2 & z_3^2 - z_1^2 & \cdots & z_n^2 - z_1^2 \\ \vdots & \vdots & & \vdots \\ z_2^{n-1} - z_1^{n-1} & z_3^{n-1} - z_1^{n-1} & \cdots & z_n^{n-1} - z_1^{n-1} \end{vmatrix}. \tag{4.11}$$

Define $S_0(x, y) = 1$ and

$$S_k(x, y) = x^k + x^{k-1}y + \cdots + xy^{k-1} + y^k \quad (k \geq 1).$$

Using the identity $x^k - y^k = (x - y)S_{k-1}(x, y)$, we can rewrite the right side of (4.11) as

$$V(z_1, \ldots, z_n) = \begin{vmatrix} z_2 - z_1 & z_3 - z_1 & \cdots & z_n - z_1 \\ (z_2 - z_1)S_1(z_1, z_2) & (z_3 - z_1)S_1(z_1, z_3) & \cdots & (z_n - z_1)S_1(z_1, z_n) \\ \vdots & \vdots & & \vdots \\ (z_2 - z_1)S_{n-2}(z_1, z_2) & (z_3 - z_1)S_{n-2}(z_1, z_3) & \cdots & (z_n - z_1)S_{n-2}(z_1, z_n) \end{vmatrix}.$$

We factor out $z_2 - z_1$ from the first column of the latter determinant, $z_3 - z_1$ from its second column, etc. to get

$$V(z_1, \ldots, z_n) = \prod_{j=2}^{n} (z_j - z_1) \times \begin{vmatrix} 1 & 1 & \cdots & 1 \\ S_1(z_1, z_2) & S_1(z_1, z_3) & \cdots & S_1(z_1, z_n) \\ \vdots & \vdots & & \vdots \\ S_{n-2}(z_1, z_2) & S_{n-2}(z_1, z_3) & \cdots & S_{n-2}(z_1, z_n) \end{vmatrix}. \tag{4.12}$$

Note that

$$S_k(x, z) - xS_{k-1}(x, z) = x^k + x^{k-1}z + \cdots + z^k - x\left(x^{k-1} + x^{k-2}z + \cdots + z^{k-1}\right) = z^k. \tag{4.13}$$

We now subtract (in this order):

z_1 times the second-to-last row of the determinant in (4.12) from its last row;
z_1 times the third-to-last row of the determinant in (4.12) from its second-to-last row;

$$\vdots$$

z_1 times the second row of the determinant in (4.12) from its third row;
z_1 times the first row of the determinant in (4.12) from its second row.

After multiple applications of (4.13), we obtain

$$V(z_1,\dots,z_n)=\prod_{j=2}^{n}(z_j-z_1)\times\begin{vmatrix}1&1&\cdots&1\\ z_2&z_3&\cdots&z_n\\ \vdots&\vdots&&\vdots\\ z_2^{n-2}&z_3^{n-2}&\cdots&z_n^{n-2}\end{vmatrix}=\prod_{j=2}^{n}(z_j-z_1)\times V(z_2,\dots,z_n).$$

Hence, (4.10) follows by applying the inductive hypothesis to $V(z_2,\dots,z_n)$. □

4.4. The resultant of two polynomials

Sometimes, one wants to determine whether two polynomials have a common root without actually finding all the roots of the two polynomials. It turns out that this question can be answered by calculating a special quantity, the "resultant" of the two polynomials, which can be expressed solely in terms of the coefficients of the polynomials.

Definition 4.23. Let $f(z)$ and $g(z)$ be two polynomials with complex coefficients, and suppose that they can be factored over $\mathbb{C}$ as

$$f(z)=a(z-\alpha_1)(z-\alpha_2)\cdots(z-\alpha_n),\quad g(z)=b(z-\beta_1)(z-\beta_2)\cdots(z-\beta_m),\tag{4.14}$$

respectively. The *resultant* of f and g, denoted $\operatorname{res}(f,g)$, is the number

$$\operatorname{res}(f,g)=a^m b^n\prod_{i=1}^{n}\prod_{j=1}^{m}(\alpha_i-\beta_j).$$

Corollary 4.24. *Let $f(z)$ and $g(z)$ be two polynomials with complex coefficients. Then $f(z)$ and $g(z)$ have a common root if and only if* $\operatorname{res}(f,g)=0$.

Proposition 4.25. *Let $f(z), g(z)$ and $h(z)$ be polynomials, and suppose that $f(z)$ and $g(z)$ are of the form* (4.14). *Then:*

i) $\operatorname{res}(f,g)=a^m g(\alpha_1)g(\alpha_2)\cdots g(\alpha_n)$.
ii) $\operatorname{res}(g,f)=(-1)^{mn}\operatorname{res}(f,g)$.
iii) $\operatorname{res}(fg,h)=\operatorname{res}(f,h)\operatorname{res}(g,h)$.
iv) $\operatorname{res}(f,fh+g)=\operatorname{res}(f,g)$.

The definition of the resultant ties it nicely to the roots of the two polynomials, but it is not particularly convenient for evaluation purposes. The next result expresses the resultant of two polynomials solely in terms of the coefficients of the polynomials.

Theorem 4.26. *Let*

$$f(z)=a_nz^n+\cdots+a_1z+a_0,\quad g(z)=b_mz^m+\cdots+b_1z+b_0$$

be two polynomials with complex coefficients. Then $\operatorname{res}(f,g)$ *equals the* $(n+m)\times(n+m)$ *determinant where:*

- *if $1\le j\le m$, the jth row is* $\begin{bmatrix}\underbrace{0\,0\cdots0}_{j-1}\ a_n\ a_{n-1}\cdots a_1\ a_0\ 0\cdots0\end{bmatrix}$;
- *if $m+1\le j\le m+n$, the jth row is* $\begin{bmatrix}\underbrace{0\,0\cdots0}_{j-m-1}\ b_m\ b_{m-1}\cdots b_1\ b_0\ 0\cdots0\end{bmatrix}$.

For example, when $f(z) = a_2z^2+a_1z+a_0$ and $g(z) = b_3z^3+b_2z^2+b_1z+b_0$, Theorem 4.26 says that

$$\operatorname{res}(f,g) = \begin{vmatrix} a_2 & a_1 & a_0 & 0 & 0 \\ 0 & a_2 & a_1 & a_0 & 0 \\ 0 & 0 & a_2 & a_1 & a_0 \\ b_3 & b_2 & b_1 & b_0 & 0 \\ 0 & b_3 & b_2 & b_1 & b_0 \end{vmatrix}.$$

Corollary 4.27. *If $f(z)$ and $g(z)$ have integer (or rational, or real) coefficients, then the resultant* $\operatorname{res}(f,g)$ *is an integer (or a rational, or a real number).*

4.5. Symmetric polynomials in several variables

Definition 4.28. Let $n \geq 2$. A polynomial $f(z_1,\dots,z_n)$ of n variables with complex coefficients is said to be *symmetric* if for any permutation $z_{j_1}, z_{j_2},\dots,z_{j_n}$ of the variables $z_1, z_2,\dots,z_n$, one has

$$f(z_{j_1}, z_{j_2},\dots,z_{j_n}) = f(z_1, z_2,\dots,z_n).$$

Example 4.29. Let $n = 2$ and denote the variables by x and y instead of z_1 and z_2. Then the polynomials

$$\sigma_1 = x+y \quad \text{and} \quad \sigma_2 = xy$$

are symmetric. Indeed, in this case, the only possible permutations of the variables are the trivial permutation x, y and the transposition y, x, and neither of those changes σ_1 and σ_2.

Example 4.30. Let $n = 3$ and denote the variables by x, y and z. Then the polynomials

$$\sigma_1 = x+y+z, \quad \sigma_2 = xy+yz+zx, \quad \text{and} \quad \sigma_3 = xyz$$

are symmetric, because a permutation of x, y, z simply changes the order of terms and/or factors in the σ's.

Example 4.31. Let $n \geq 2$ and $1 \leq k \leq n$. Then the polynomial

$$\sigma_k = \sum_{1\leq i_1<i_2<\cdots<i_k\leq n} z_{i_1}z_{i_2}\cdots z_{i_k} \tag{4.15}$$

is symmetric. Here, the summation is over all k-element subsets of $\{1,2,\dots,n\}$, so we are adding all the possible product of k distinct variables. Thus,

$$\sigma_1 = \sum_{i=1}^{n} z_i, \quad \sigma_2 = \sum_{1\leq i<j\leq n} z_iz_j, \quad \text{and} \quad \sigma_n = z_1z_2\cdots z_n.$$

Note that Examples 4.29 and 4.30 are just the special cases $n = 2$ and $n = 3$ of the present example.

Definition 4.32. Let $n \geq 2$. The symmetric polynomials $\sigma_1,\dots,\sigma_n$ defined by (4.15) are called the *elementary symmetric polynomials* in the variables $z_1,\dots,z_n$.

It turns out that every symmetric polynomial in $z_1,\dots,z_n$ can be rewritten (via substitutions) as a polynomial in $\sigma_1,\dots,\sigma_n$, (treated as independent variables). More precisely, we have the following result.

Theorem 4.33 (Fundamental theorem for symmetric polynomials)**.** *Let $f(z_1,\dots,z_n)$ be a symmetric polynomial. Then there exists a polynomial $g(w_1,\dots,w_n)$ such that*

$$f(z_1,\dots,z_n) = g(\sigma_1,\dots,\sigma_n),$$

where $\sigma_1, \dots, \sigma_n$ are the elementary symmetric polynomials (4.15). *Furthermore, if f has integer (or rational, or real) coefficients, then so does g.*

The proof of this theorem goes beyond the scope of these notes, but we can illustrate the idea of the proof by an example.

Example 4.34. Let $n = 2$ and consider the polynomial $f(x, y) = x^5 + y^5$, which is clearly a symmetric polynomial in x and y. Then

$$f_1(x, y) = f(x, y) - (x + y)^5 = f(x, y) - \sigma_1^5$$

is also symmetric in x and y, and (by the binomial theorem)

$$f_1(x, y) = -5x^4y - 10x^3y^2 - 10x^2y^3 - 5xy^4.$$

Note that the highest power of x in $f_1(x, y)$ is smaller than that in $f(x, y)$. Next, consider

$$f_2(x, y) = f_1(x, y) + 5xy(x + y)^3 = f_1(x, y) + 5\sigma_2\sigma_1^3.$$

This polynomial is also symmetric, and we have

$$f_2(x, y) = 5x^3y^2 + 5x^2y^3,$$

so the highest power of x in $f_2(x, y)$ is smaller than that in $f_1(x, y)$ (which, in turn, was smaller than that in $f(x, y)$). Next, we consider

$$f_3(x, y) = f_2(x, y) - 5x^2y^2(x + y) = f_2(x, y) - 5\sigma_2^2\sigma_1.$$

Since $f_3(x, y) = 0$, we have $f_2(x, y) = 5\sigma_2^2\sigma_1$. Hence,

$$\begin{aligned} f(x, y) &= \sigma_1^5 + f_1(x, y) \\ &= \sigma_1^5 - 5\sigma_2\sigma_1^3 + f_2(x, y) \\ &= \sigma_1^5 - 5\sigma_2\sigma_1^3 + 5\sigma_2^2\sigma_1 \\ &= g(\sigma_1, \sigma_2), \quad \text{where } g(u, v) = u^5 - 5u^3v + 5uv^2. \end{aligned}$$

4.6. Exercises

Exercise 4.1. Observe that 2i is a root of the equation $z^{10} + 4z^8 + 4z^6 + 16z^4 + 4z^2 + 16 = 0$ and use this information to find its remaining nine roots.

Exercise 4.2. Write the polynomial $x^{10} + 4x^8 + 4x^6 + 16x^4 + 4x^2 + 16$ as a product of the form stated in Theorem 4.9.

Exercise 4.3. Prove Corollary 4.3.

Exercise 4.4. Prove Corollary 4.4.

Exercise 4.5. Prove Corollary 4.5.

Exercise 4.6. Prove Corollary 4.6.

Exercise 4.7. Prove Theorem 4.13.

Exercise 4.8. If possible, diagonalize the given matrices. If a real matrix is diagonalizable over $\mathbb{C}$ but not over $\mathbb{R}$, then represent it as QCQ^{-1}, where Q and C are real matrices and C is block-diagonal.

$$A = \begin{bmatrix} 4 & 1 & 0 & 1 \\ 2 & 3 & 0 & 1 \\ -2 & 1 & 2 & -3 \\ 2 & -1 & 0 & 5 \end{bmatrix}; \quad B = \begin{bmatrix} 1 & 1 & \frac{1}{2} & -1 \\ -2 & 1 & -1 & 0 \\ 0 & 2 & 0 & 2 \\ 2 & 1 & -\frac{3}{2} & 2 \end{bmatrix}.$$

Exercise 4.9. The matrix A below has a double real eigenvalue and a pair of complex eigenvalues. Furthermore, it is known that the real eigenvalue is an integer. Represent A in the form $A = QCQ^{-1}$, where Q and C are real matrices and C is block-diagonal.

$$A = \begin{bmatrix} 18.5 & 10 & -2.5 & -45 \\ -6.5 & -3 & 1.5 & 17 \\ 8.5 & 6 & -2.5 & -23 \\ 5.5 & 3 & -0.5 & -13 \end{bmatrix}.$$

Exercise 4.10. Let A be a 7×7 real matrix with eigenvalues $1, 1, -2, 1 + \mathrm{i}, 1 + \mathrm{i}, 1 - \mathrm{i}, 1 - \mathrm{i}$, and respective eigenvectors

$$\begin{bmatrix} 1 \\ 0 \\ 2 \\ 0 \\ 0 \\ 1 \\ 3 \end{bmatrix}, \quad \begin{bmatrix} -1 \\ 0 \\ 0 \\ -2 \\ -1 \\ 0 \\ -3 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 4 \\ 2 \end{bmatrix}, \quad \begin{bmatrix} 2 \\ 1+\mathrm{i} \\ 0 \\ 1 \\ 0 \\ -1 \\ 2 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \\ 2\mathrm{i} \\ 3-\mathrm{i} \\ 0 \\ 0 \\ -5 \end{bmatrix}, \quad \begin{bmatrix} 2 \\ 1-\mathrm{i} \\ 0 \\ 1 \\ 0 \\ -1 \\ 2 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \\ -2\mathrm{i} \\ 3+\mathrm{i} \\ 0 \\ 0 \\ -5 \end{bmatrix}.$$

Represent A in the form $A = QCQ^{-1}$, where Q and C are real matrices and C is block-diagonal.

Exercise 4.11. Prove Theorem 4.9.

Exercise 4.12. Evaluate the determinants:

$$\begin{vmatrix} 2 & 1 & 0 & 2 \\ 4 & 2 & 7 & -2 \\ 0 & 2 & 1 & 1 \\ 0 & 2 & 3 & 0 \end{vmatrix}, \quad \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 \\ 1 & 1 & 3 & 1 \\ 1 & 4 & 1 & 1 \end{vmatrix}, \quad \begin{vmatrix} 4 & 8 & 8 & 8 & 5 \\ 0 & 1 & 0 & 0 & 0 \\ 6 & 8 & 8 & 8 & 7 \\ 0 & 8 & 8 & 3 & 0 \\ 0 & 8 & 2 & 0 & 0 \end{vmatrix}.$$

Exercise 4.13. Use induction on n to prove that

$$\begin{vmatrix} 0 & \cdots & 0 & c_1 \\ 0 & \cdots & c_2 & 0 \\ \vdots & & \vdots & \vdots \\ c_n & \cdots & 0 & 0 \end{vmatrix} = (-1)^{n(n-1)/2} c_1 c_2 \cdots c_n.$$

Exercise 4.14. Let D_n be the $n \times n$ determinant

$$D_n = \begin{vmatrix} 2 & -1 & 0 & \cdots & 0 & 0 \\ -1 & 2 & -1 & \cdots & 0 & 0 \\ 0 & -1 & 2 & \cdots & 0 & 0 \\ \vdots & \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & -1 \\ 0 & 0 & 0 & \cdots & -1 & 2 \end{vmatrix}.$$

That is, the diagonal entries of D_n are all equal to 2, the entries one position off the diagonal are all equal to -1, and all the remaining entries are zeros.

(a) Use expansions across the first row to prove that $D_n = 2D_{n-1} - D_{n-2}$.
(b) Use part (a) and induction on n to prove that $D_n = n + 1$.

Exercise 4.15. Let

$$R(\lambda; a_0, a_1, \ldots, a_{n-1}) = \begin{vmatrix} \lambda & 0 & 0 & \cdots & 0 & a_0 \\ -1 & \lambda & 0 & \cdots & 0 & a_1 \\ 0 & -1 & \lambda & \cdots & 0 & a_2 \\ \vdots & \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & a_{n-2} \\ 0 & 0 & 0 & \cdots & -1 & \lambda + a_{n-1} \end{vmatrix}.$$

(a) Use an expansion across the first row to prove that $R(\lambda; a_0, a_1, \ldots, a_{n-1}) = a_0 + \lambda R(\lambda; a_1, a_2, \ldots, a_{n-1})$.
(b) Use part (a) and induction on n to prove that $R(\lambda; a_0, a_1, \ldots, a_{n-1}) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$.

Exercise 4.16. Let

$$D_n = \begin{vmatrix} a_1+b_1 & b_1 & b_1 & \cdots & b_1 \\ b_2 & a_2+b_2 & b_2 & \cdots & b_2 \\ b_3 & b_3 & a_3+b_3 & \cdots & b_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_n & b_n & b_n & \cdots & a_n+b_n \end{vmatrix}.$$

(a) Use the properties of determinants to prove that

$$D_n = \begin{vmatrix} a_1+b_1 & -a_1 & -a_1 & \cdots & -a_1 \\ b_2 & a_2 & 0 & \cdots & 0 \\ b_3 & 0 & a_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_n & 0 & 0 & \cdots & a_n \end{vmatrix}.$$

(b) Expand the determinant in part (a) across the last row to show that $D_n = a_nD_{n-1} + (-1)^{n+1}b_nE_{n-1}$, where E_{n-1} is an $(n-1)\times(n-1)$ determinant.

(c) Show that in part (b), $E_{n-1} = (-1)^{n+1}a_1a_2\cdots a_{n-1}$.

(d) Use parts (b) and (c) and induction on n to compute D_n.

Exercise 4.17. A square matrix $A = [a_{ij}]$ is called *antisymmetric* if $a_{ij} + a_{ji} = 0$ for all pairs of indices i, j. Prove that if n is odd and A is an $n \times n$ antisymmetric matrix, then $\det A = 0$. [Hint. Use the properties of determinants to show that $\det A^t = (-1)^n \det A$.]

Exercise 4.18. Let n be an even integer n and a, b be complex numbers. Define

$$D_n = \begin{vmatrix} a & 0 & 0 & \cdots & 0 & b \\ 0 & a & 0 & \cdots & b & 0 \\ 0 & 0 & a & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & b & 0 & \cdots & a & 0 \\ b & 0 & 0 & \cdots & 0 & a \end{vmatrix}.$$

(a) Expand D_n across the first row to show that $D_n = aE'_n - bE''_n$, where E'_n and E''_n are $(n-1)\times(n-1)$ determinants.

(b) Show that in part (a), $E'_n = aD_{n-2}$ and $E''_n = bD_{n-2}$.

(d) Use parts (a) and (b) and induction on n to show that $D_n = (a^2 - b^2)^{n/2}$.

Exercise 4.19. Let $f(z) = a_nz^n + \cdots + a_1z + a_0$ and let $\alpha_1, \alpha_2, \ldots, \alpha_{n+1}$ be $n+1$ distinct complex numbers.

(a) Consider the equations

$$f(\alpha_1) = 0, \quad f(\alpha_2) = 0, \quad \ldots, \quad f(\alpha_{n+1}) = 0$$

as a homogeneous linear system for the coefficients $a_0, a_1, \ldots, a_n$. Prove that this system has a unique solution. [Hint. This has something to do with Vandermonde's determinant.]

(b) Use part (a) to prove that if $f(\alpha_j) = 0$ for $j = 1, 2, \ldots, n+1$, then $f(z) = 0$.

Exercise 4.20. Let $f(z) = a_kz^k + \cdots + a_1z + a_0$ and A be an $n\times n$ matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Generalize the proof of Theorem 4.15 given above to establish that the matrix $f(A) = a_kA^k + \cdots + a_1A + a_0$ has eigenvalues $f(\lambda_1), f(\lambda_2), \ldots, f(\lambda_n)$.

Exercise 4.21. Prove Corollary 4.24.

Exercise 4.22. Prove Proposition 4.25.

Exercise 4.23. Compute $\operatorname{res}(z^n, f(z))$, where $f(z)$ is a polynomial of degree m.

Exercise 4.24. Prove Corollary 4.27.

Exercise 4.25. Show that if $f(z_1, \ldots, z_n)$ and $g(z_1, \ldots, z_n)$ are symmetric polynomials, then so are

$$af(z_1, \ldots, z_n) + bg(z_1, \ldots, z_n) \quad \text{and} f(z_1, \ldots, z_n)g(z_1, \ldots, z_n).$$

Exercise 4.26. Which of the following polynomials are symmetric polynomials in x and y: $x^3 - y^3$, $x^2y + xy^2$, $x^4 + xy + y^4$, $x + 2y + 3xy^2$?

Exercise 4.27. Which of the following polynomials are symmetric polynomials in x, y and z:

$$x^3 + y^3 + z^3, \quad x^2y + y^2z + z^2x, \quad x^3 + y^3?$$

Exercise 4.28. Express the following symmetric polynomials as polynomials of the elementary symmetric polynomials $\sigma_1, \sigma_2, \ldots$:

$$x^2 + y^2 + z^2, \quad x^2y^2 + y^2z^2 + x^2z^2, \quad x^4 + y^4 + z^4, \quad z_1^2 + z_2^2 + \cdots + z_n^2.$$

Exercise 4.29. A polynomial $f(x, y)$ of the form

$$f(x, y) = a_nx^n + a_{n-1}x^{n-1}y + \cdots + a_1xy^{n-1} + a_0y^n,$$

where at least one of the coefficients $a_0, a_1, \ldots, a_n$ is nonzero is called a *homogeneous polynomial of degree n*.

(a) Show that if $f(x, y)$ and $g(x, y)$ are two homogeneous polynomials of degree n, then $f(x, y) + g(x, y)$ is either a homogeneous polynomial of degree n, or identically 0.

(b) Show that if $f(x, y)$ and $g(x, y)$ are two homogeneous polynomials of degrees n and m, respectively, then $f(x, y)g(x, y)$ is a homogeneous polynomial of degree $n + m$.

(c) Show that if $f(x, y)$ is a homogeneous polynomial of degree n with real coefficients, then either the equation $f(x, y) = 0$ has the unique solution $x = y = 0$, or it has infinitely many solutions, given by the points on k lines through the origin, where $1 \leq k \leq n$. [Hint. The solutions of $f(x, y) = 0$ are related to the solutions of the equation $g(z) = 0$, where $g(z) = f(z, 1)$.]

CHAPTER 5

Fourier Series

In this chapter, we provide a brief introduction to Fourier series, a topic which may be new to many of you. Moreover, even those who have encountered Fourier series in earlier courses are likely to be familiar mainly with Fourier sine and cosine series, whereas we shall focus on complex Fourier series, in which the sine and cosine functions are replaced by complex exponentials.

5.1. Definition

The usual definition of Fourier series starts with a p-periodic real function f (often $p = 2\pi$ for convenience). The *Fourier series* of f then is

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} \big(a_n \cos(2\pi n x/p) + b_n \sin(2\pi n x/p) \big), \tag{5.1}$$

where the coefficients $a_0, a_1, b_1, \dots, a_n, b_n, \dots$ are given by the formulas

$$a_n = \frac{2}{p} \int_0^p f(t) \cos(2\pi n t/p)\, dt, \qquad b_n = \frac{2}{p} \int_0^p f(t) \sin(2\pi n t/p)\, dt. \tag{5.2}$$

In these notes, we shall give preference to an alternative definition that uses complex exponentials instead of sines and cosines. Also, we shall allow our functions to be complex-valued.

Definition 5.1. Let $f : \mathbb{R} \to \mathbb{C}$ be a 1-periodic function. The *Fourier series* of f is

$$S(f; x) = \sum_{n=-\infty}^{\infty} c_n e^{2\pi \mathrm{i} n x} = \lim_{N \to \infty} \sum_{n=-N}^{N} c_n e^{2\pi \mathrm{i} n x}, \tag{5.3}$$

where the coefficients c_n are given by the formula

$$c_n = \int_0^1 f(t) e^{-2\pi \mathrm{i} n t}\, dt \qquad (n \in \mathbb{Z}). \tag{5.4}$$

5.2. Calculus of complex-valued functions

In order to make sense of the formulas in (5.3) and (5.4), we need to extend some definitions from calculus to complex-valued functions. Let $\{z_n\}_{n=1}^{\infty}$, $z_n = x_n + \mathrm{i} y_n$, be a sequence of complex numbers. We say that this sequence *converges* to the complex number $c = a + \mathrm{i}b$, and write $\lim\limits_{n\to\infty} z_n = c$, if the two real sequences $\{x_n\}_{n=1}^{\infty}$ and $\{y_n\}_{n=1}^{\infty}$ converge to a and b, respectively:

$$\lim_{n\to\infty} (x_n + \mathrm{i} y_n) = a + \mathrm{i} b \quad \iff \quad \lim_{n\to\infty} x_n = a, \; \lim_{n\to\infty} y_n = b.$$

Similarly, the series $\sum_{n=1}^{\infty} z_n$ if the two real series $\sum_{n=1}^{\infty} x_n$ and $\sum_{n=1}^{\infty} y_n$ converge, and

$$\sum_{n=1}^{\infty} z_n = \sum_{n=1}^{\infty} x_n + \mathrm{i} \sum_{n=1}^{\infty} y_n.$$

Furthermore, if $f : D \to \mathbb{C}$ is a complex-valued function of a real argument, then it can be expressed in the form $f(t) = u(t) + \mathrm{i}v(t)$, where u and v are real functions defined in the same domain D as the function f. We say that f is *differentiable* (resp., *integrable*) if both u and v are differentiable (resp., integrable), and define the derivative $f'(t)$ and the integral $\int_a^b f(t)\,dt$ of f by

$$f'(t) = u'(t) + \mathrm{i}v'(t), \qquad \int_a^b f(t)\,dt = \int_a^b u(t)\,dt + \mathrm{i} \int_a^b v(t)\,dt.$$

With these definitions, most integration and differentiation rules from calculus remain true for complex-valued functions. Below, we list some integration formulas for complex-valued functions, which will be useful later.

Theorem 5.2 (Newton–Leibnitz). *Let $a, b \in \mathbb{R}$ and $f : [a, b] \to \mathbb{C}$ have a continuous derivative on $[a, b]$. Then*

$$\int_a^b f'(t)\,dt = f(b) - f(a).$$

Proof. Let $f(t) = u(t) + \mathrm{i}v(t)$. Then

$$\int_a^b f'(t)\,dt = \int_a^b (u'(t) + \mathrm{i}v'(t))\,dt = \int_a^b u'(t)\,dt + \mathrm{i} \int_a^b v'(t)\,dt.$$

Since u and v are real-valued functions, we can evaluate the two integrals on the right side of the above identity by the fundamental theorem of calculus. We get

$$\int_a^b f'(t)\,dt = \big(u(b) - u(a)\big) + \mathrm{i}\big(v(b) - v(a)\big) = f(b) - f(a). \qquad \square$$

Theorem 5.3. *Suppose that $a, b \in \mathbb{R}$, $\alpha \in \mathbb{C}$ and $f, g : [a, b] \to \mathbb{C}$ are continuous on $[a, b]$. Then:*

i) $\int_a^b \alpha f(t)\,dt = \alpha \int_a^b f(t)\,dt;$
ii) $\int_a^b \big(f(t) + g(t)\big)\,dt = \int_a^b f(t)\,dt + \int_a^b g(t)\,dt;$
iii) *if $a \le c \le b$, then $\int_a^b f(t)\,dt = \int_a^c f(t)\,dt + \int_c^b f(t)\,dt;$*
iv) *if f is p-periodic and $b - a = p$, then $\int_a^b f(t)\,dt = \int_0^p f(t)\,dt.$*

Theorem 5.4 (Integration by parts). *Suppose that $a, b \in \mathbb{R}$ and $f, g : [a, b] \to \mathbb{C}$ have continuous derivatives on $[a, b]$. Then*

$$\int_a^b f(t)g'(t)\,dt = f(b)g(b) - f(a)g(a) - \int_a^b f'(t)g(t)\,dt.$$

Example 5.5. Let $\alpha = p + \mathrm{i}q$ be a nonzero complex number and $f(t) = e^{\alpha t}$. Then

$$\begin{aligned} f'(t) &= \big(e^{pt}\cos qt + \mathrm{i}e^{pt}\sin qt\big)' = \big(e^{pt}\cos qt\big)' + \mathrm{i}\big(e^{pt}\sin qt\big)' \\ &= \big(pe^{pt}\cos qt - qe^{pt}\sin qt\big) + \mathrm{i}\big(pe^{pt}\sin qt + qe^{pt}\cos qt\big) \\ &= (p + \mathrm{i}q)e^{pt}\cos qt + (-q + \mathrm{i}p)e^{pt}\sin qt \\ &= (p + \mathrm{i}q)e^{pt}\cos qt + \mathrm{i}(p + \mathrm{i}q)e^{pt}\sin qt \\ &= (p + \mathrm{i}q)e^{pt}(\cos qt + \mathrm{i}\sin qt) = \alpha e^{pt}e^{\mathrm{i}qt} = \alpha e^{\alpha t}. \end{aligned}$$

Hence, by Theorem 5.2 and part i) of Theorem 5.3,

$$\int_a^b e^{\alpha t}\,dt = \alpha^{-1}\int_a^b f'(t)\,dt = \alpha^{-1}\big(f(b)-f(a)\big) = \alpha^{-1}\big(e^{\alpha b}-e^{\alpha a}\big). \tag{5.5}$$

Taking $\alpha = 2\pi i m$, $a = 0$ and $b = 1$, we obtain the following important formula.

Lemma 5.6. *Let $m \in \mathbb{Z}$. Then*

$$\int_0^1 e^{2\pi i m t}\,dt = \begin{cases} 1 & \text{if } m = 0, \\ 0 & \text{if } m \neq 0. \end{cases}$$

5.3. An example

Example 5.7. Let $f(x)$ be the 1-periodic function (see Figure 5.1) defined by

$$f(x) = x^2 \quad (|x| \le 1/2), \qquad f(x+1) = f(x).$$

Let us compute the Fourier series of this function. When $n \neq 0$, we have

$$\begin{aligned} c_n &= \int_0^1 f(t)e^{-2\pi i n t}\,dt = \int_{-1/2}^{1/2} f(t)e^{-2\pi i n t}\,dt \\ &= \int_{-1/2}^{1/2} t^2 e^{-2\pi i n t}\,dt = \frac{-1}{2\pi i n}\int_{-1/2}^{1/2} t^2\,d\big(e^{-2\pi i n t}\big) \\ &= \frac{-1}{2\pi i n}\big[t^2 e^{-2\pi i n t}\big]_{-1/2}^{1/2} + \frac{1}{2\pi i n}\int_{-1/2}^{1/2} e^{-2\pi i n t}\,d\big(t^2\big) \\ &= \frac{\frac{1}{4}e^{\pi i n} - \frac{1}{4}e^{-\pi i n}}{2\pi i n} + \frac{1}{\pi i n}\int_{-1/2}^{1/2} t e^{-2\pi i n t}\,dt \\ &= \frac{1}{\pi i n}\int_{-1/2}^{1/2} t e^{-2\pi i n t}\,dt = \frac{-1}{2(\pi i n)^2}\int_{-1/2}^{1/2} t\,d\big(e^{-2\pi i n t}\big) \\ &= \frac{1}{2\pi^2 n^2}\big[t e^{-2\pi i n t}\big]_{-1/2}^{1/2} - \frac{1}{2\pi^2 n^2}\int_{-1/2}^{1/2} e^{-2\pi i n t}\,dt \\ &= \frac{\frac{1}{2}e^{-\pi i n} + \frac{1}{2}e^{\pi i n}}{2\pi^2 n^2} - \frac{1}{2\pi^2 n^2}\int_0^1 e^{-2\pi i n t}\,dt = \frac{(-1)^n}{2\pi^2 n^2}. \end{aligned}$$

Also,

$$c_0 = \int_0^1 f(t)\,dt = \int_{-1/2}^{1/2} f(t)\,dt = \int_{-1/2}^{1/2} t^2\,dt = \frac{1}{12}.$$

Hence,

$$S(f;x) = \frac{1}{12} + \frac{1}{2\pi^2}\sum_{n\neq 0}\frac{(-1)^n e^{2\pi i n x}}{n^2}.$$

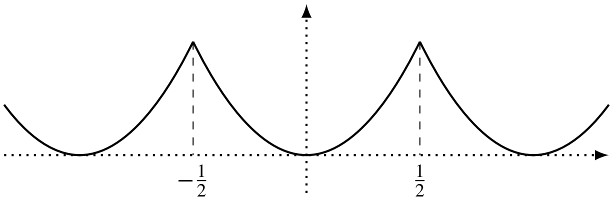

FIGURE 5.1. 1-periodic extension of $f(x) = x^2$, $|x| \le 1/2$

Note that the above series is absolutely convergent: the series whose terms are the moduli of the terms of $S(f;x)$ is

$$\frac{1}{12} + \lim_{N\to\infty} \frac{1}{2\pi^2} \sum_{n=1}^{N} \left(\frac{1}{n^2} + \frac{1}{(-n)^2} \right) = \frac{1}{12} + \frac{1}{\pi^2} \sum_{n=1}^{\infty} \frac{1}{n^2},$$

which is a convergent p-series ($p = 2$).

5.4. Representing functions by Fourier series

The question now arises: Is there a relation between $f(x)$ and the sum $S(f;x)$ of its Fourier series? The next theorem establishes the connection between them. However, before we state the theorem, we need to introduce some notation. We shall write

$$f(a-0) = \lim_{x\to a^-} f(x) \quad \text{and} \quad f(a+0) = \lim_{x\to a^+} f(x)$$

for the left and right limits of $f(x)$ as $x \to a$. Also, we say that f *has at most a jump discontinuity at* $x = a$, if both of the above limits exist. Note that this means that: either f is continuous at $x = a$ (then $f(a-0) = f(a+0) = f(a)$); or f has a removable discontinuity at $x = a$ (then $f(a-0) = f(a+0) \neq f(a)$); or f has a jump discontinuity at $x = a$ (then $f(a-0) \neq f(a+0)$). If $f : \mathbb{R} \to \mathbb{C}$ is a function such that every real number is at most a jump discontinuity of f, then we define the function $\tilde{f} : \mathbb{R} \to \mathbb{C}$ by

$$\tilde{f}(x) = \frac{f(x-0) + f(x+0)}{2}. \tag{5.6}$$

When f is continuous at $x = a$, $\tilde{f}$ is also continuous at $x = a$ and $\tilde{f}(a) = f(a)$. When f has a removable discontinuity at $x = a$, $\tilde{f}$ is continuous at $x = a$ and $\tilde{f}(a)$ equals the common value of $f(a+0)$ and $f(a-0)$. And when f has a jump discontinuity at $x = a$, $\tilde{f}$ also has a jump discontinuity at $x = a$ and $\tilde{f}(a)$ is the midpoint between $f(a+0)$ and $f(a-0)$.

Theorem 5.8. *Suppose that $f : \mathbb{R} \to \mathbb{C}$ is a 1-periodic function that is continuous on* $[0,1]$*, with the possible exception of a finite number of points $a_1, a_2, \dots, a_n$ at which f has at most jump discontinuities. Then the Fourier series of f converges for all $x \in \mathbb{R}$ and $S(f;x) = \tilde{f}(x)$, where $\tilde{f}$ is the function defined by* (5.6)*.*

Example 5.9. Returning to the Fourier series from Example 5.7, we can now say that

$$x^2 = \frac{1}{12} + \frac{1}{2\pi^2} \sum_{n\neq 0} \frac{(-1)^n e^{2\pi i n x}}{n^2} \qquad (|x| \leq 1/2). \tag{5.7}$$

Indeed, because the function f in Example 5.7 is continuous at all real numbers, we have $\tilde{f}(x) = f(x)$ for all $x \in \mathbb{R}$. Hence, Theorem 5.8 gives $S(f;x) = f(x)$ for all $x \in \mathbb{R}$.

When $x = \frac{1}{2}$, equation (5.7) gives

$$\frac{1}{4} = \frac{1}{12} + \frac{1}{2\pi^2} \sum_{n\neq 0} \frac{(-1)^n e^{\pi i n}}{n^2} = \frac{1}{12} + \frac{1}{2\pi^2} \sum_{n\neq 0} \frac{1}{n^2} = \frac{1}{12} + \frac{1}{\pi^2} \sum_{n=1}^{\infty} \frac{1}{n^2},$$

whence

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

This formula was first proved by L. Euler in 1735. At the time, the question of computing $\sum_{n=1}^{\infty} n^{-2}$ was an open problem known as the Basel problem. Before Euler, many mathematicians had tried and failed to compute the sum of this series. Not only did Euler succeed, but he was in fact able to find a method for computing the sums $\sum_{n=1}^{\infty} n^{-2k}$ for all

$k \in \mathbb{N}$. His solution of the Basel problem brought Euler instant fame and established him as the premier mathematician of the day.

Example 5.10. Let $f(x)$ be the 1-periodic function (see Figure 5.2) defined by

$$f(x) = x \quad (-1/2 \le x < 1/2), \qquad f(x+1) = f(x).$$

When $n \neq 0$, we have

$$\begin{aligned} c_n &= \int_0^1 f(t)e^{-2\pi \mathrm{i} nt}\,dt = \int_{-1/2}^{1/2} f(t)e^{-2\pi \mathrm{i} nt}\,dt \\ &= \int_{-1/2}^{1/2} te^{-2\pi \mathrm{i} nt}\,dt = \frac{-1}{2\pi \mathrm{i} n}\int_{-1/2}^{1/2} t\,d\big(e^{-2\pi \mathrm{i} nt}\big) \\ &= \frac{-1}{2\pi \mathrm{i} n}\big[te^{-2\pi \mathrm{i} nt}\big]_{-1/2}^{1/2} + \frac{1}{2\pi \mathrm{i} n}\int_{-1/2}^{1/2} e^{-2\pi \mathrm{i} nt}\,dt \\ &= \frac{\frac{1}{2}e^{-\pi \mathrm{i} n} + \frac{1}{2}e^{\pi \mathrm{i} n}}{-2\pi \mathrm{i} n} + \frac{1}{2\pi \mathrm{i} n}\int_0^1 e^{-2\pi \mathrm{i} nt}\,dt = \frac{(-1)^{n+1}}{2\pi \mathrm{i} n}. \end{aligned}$$

Also,

$$c_0 = \int_0^1 f(t)\,dt = \int_{-1/2}^{1/2} f(t)\,dt = \int_{-1/2}^{1/2} t\,dt = 0.$$

Hence, the Fourier series of f is

$$S(f;x) = \frac{-1}{2\pi \mathrm{i}}\sum_{n\neq 0} \frac{(-1)^n e^{2\pi \mathrm{i} nx}}{n}.$$

Note that the above series is not absolutely convergent: the series whose terms are the moduli of the terms of $S(f;x)$ is essentially the harmonic series, which is divergent. Nonetheless, the series $S(f;x)$ converges and $S(f;x) = f(x)$ for all real numbers except for $\pm\frac{1}{2}, \pm\frac{3}{2}, \pm\frac{5}{2}, \dots$. When x is one of these values, we have

$$S(f;\tfrac{1}{2}+k) = \lim_{N\to\infty} \frac{-1}{2\pi \mathrm{i}} \sum_{\substack{n=-N\\ n\neq 0}}^{N} \frac{(-1)^n e^{\pi \mathrm{i} n}}{n} = \lim_{N\to\infty} \frac{-1}{2\pi \mathrm{i}} \sum_{\substack{n=-N\\ n\neq 0}}^{N} \frac{1}{n} = 0,$$

because the terms with opposite values of n in the last sum cancel each other. Notice that this is consistent with Theorem 5.8: at the points of discontinuity of f, the Fourier series converges not to $f(x)$ (which equals $-\frac{1}{2}$) but to $\tilde{f}(x)$ (which equals 0).

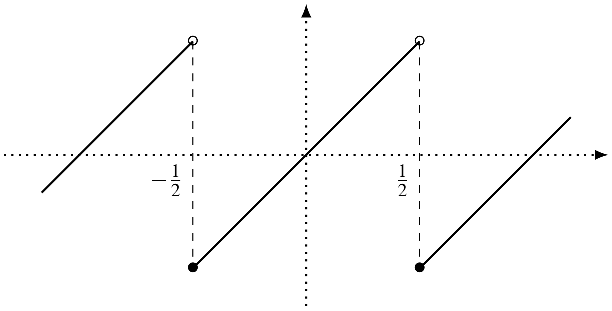

FIGURE 5.2. The saw-tooth function

5.5. Poisson's summation formula

We start this section by defining the Fourier transform of a function.

Definition 5.11. Let $f : \mathbb{R} \to \mathbb{C}$ be a continuous function such that the improper integral $\int_{-\infty}^{\infty} |f(x)|\,dx$ is convergent. The *Fourier transform* of f, denoted $\hat{f}(y)$, is defined by

$$\hat{f}(y) = \int_{-\infty}^{\infty} f(x) e^{-2\pi \mathrm{i} xy}\,dx.$$

The main result of the section is the following theorem.

Theorem 5.12 (Poisson's summation formula)**.** *Let the function $f : \mathbb{R} \to \mathbb{C}$ be continuous everywhere, except possibly for a finite number of points $a_1, a_2, \dots, a_n$ at which f has at most jump discontinuities. Suppose also that at the points $a_1, a_2, \dots, a_n$, we have*

$$f(a_j) = \tfrac{1}{2}\big(f(a_j - 0) + f(a_j + 0)\big),$$

and that the improper integral $\int_{-\infty}^{\infty} |f(x)|\,dx$ converges. Then

$$\sum_{n=-\infty}^{\infty} f(n) = \sum_{n=-\infty}^{\infty} \hat{f}(n),$$

where $\hat{f}(y)$ is the Fourier transform of f.

We do not prove Theorem 5.12 in these notes, but we do sketch the main idea of its proof in a special case. Our next example is essentially the proof of the theorem when $f(x) = e^{-\pi x^2 t}$ for some $t > 0$.

Example 5.13. Let $t > 0$ and define the function

$$\theta(x) = \theta(x;t) = \sum_{m=-\infty}^{\infty} e^{-\pi(m+x)^2 t} \qquad (-\infty < x < \infty).$$

By Exercise 5.8, the above series converges. Moreover, we have $\theta(x+1) = \theta(x)$, because the mth term in the sum defining $\theta(x+1)$ is the same as the $(m+1)$th term in the sum defining $\theta(x)$. Let us compute the nth Fourier coefficient of $\theta(x)$.

Suppose that $|x| \le \frac{1}{2}$. If $M \in \mathbb{N}$, we have

$$\theta(x) = \theta_1(x;M) + \theta_2(x;M),$$

where

$$\theta_1(x;M) = \sum_{m=-M}^{M} e^{-\pi(m+x)^2 t}, \quad \theta_2(x;M) = \sum_{|m|>M} e^{-\pi(m+x)^2 t}.$$

The nth Fourier coefficient of $\theta(x)$ is

$$c_n = \int_{-1/2}^{1/2} \theta(x) e^{-2\pi \mathrm{i} nx}\,dx = I_1(n,M) + I_2(n,M), \tag{5.8}$$

where

$$I_1(n,M) = \int_{-1/2}^{1/2} \theta_1(x;M) e^{-2\pi \mathrm{i} nx}\,dx, \quad I_2(n,M) = \int_{-1/2}^{1/2} \theta_2(x;M) e^{-2\pi \mathrm{i} nx}\,dx.$$

We find that

$$\begin{aligned}\theta_2(x; M) &= \sum_{m=M+1}^{\infty} \left(e^{-\pi(m+x)^2 t} + e^{-\pi(x-m)^2 t}\right) \leq 2 \sum_{m=M+1}^{\infty} e^{-\pi(m-1/2)^2 t} \\ &= 2\sum_{k=M}^{\infty} e^{-\pi(k+1/2)^2 t} \leq 2\sum_{k=M}^{\infty} e^{-\pi k t} = \varepsilon_M, \quad \text{say.}\end{aligned}$$

Hence, Exercise 5.7 gives

$$|I_2(n; M)| \leq \int_{-1/2}^{1/2} \left|\theta_2(x; M) e^{-2\pi i n x}\right| dx \leq \int_{-1/2}^{1/2} \varepsilon_M \, dx = \varepsilon_M. \tag{5.9}$$

On the other hand,

$$\begin{aligned} I_1(n, M) &= \sum_{m=-M}^{M} \int_{-1/2}^{1/2} e^{-\pi(m+x)^2 t} e^{-2\pi i n x} \, dx \\ &= \sum_{m=-M}^{M} \int_{m-1/2}^{m+1/2} e^{-\pi u^2 t} e^{-2\pi i n(u-m)} \, du \qquad (u = x + m) \\ &= \sum_{m=-M}^{M} \int_{m-1/2}^{m+1/2} e^{-\pi u^2 t} e^{-2\pi i n u} \underbrace{e^{2\pi i n m}}_{=1} \, du \\ &= \int_{-M-1/2}^{M+1/2} e^{-\pi u^2 t} e^{-2\pi i n u} \, du. \end{aligned} \tag{5.10}$$

Combining (5.9) and (5.10), we conclude that

$$\lim_{M\to\infty} \left(I_1(n, M) + I_2(n, M)\right) = \int_{-\infty}^{\infty} e^{-\pi u^2 t} e^{-2\pi i n u} \, du = \hat{\phi}(n),$$

where $\hat{\phi}$ is the Fourier transform of $\phi(u) = \phi(u; t) = e^{-\pi u^2 t}$.

Since the left side of (5.8) is independent of M, by letting $M \to \infty$ in (5.8), we deduce that $c_n = \hat{\phi}(n)$. Therefore, the Fourier series of $\theta(x)$ is

$$S(\theta; x) = \sum_{n=-\infty}^{\infty} \hat{\phi}(n) e^{2\pi i n x}.$$

As it can be shown (see Exercise 5.9) that $\theta(x)$ is continuous at all real x, it follows that

$$\theta(x; t) = \sum_{n=-\infty}^{\infty} \phi(n + x; t) = \sum_{n=-\infty}^{\infty} \hat{\phi}(n; t) e^{-2\pi i n x}, \quad \phi(u; t) = e^{-\pi u^2 t}.$$

5.6. Exercises

Exercise 5.1. Let $f(x)$ be the 1-periodic function defined by

$$f(x) = x^4 \quad (|x| \leq 1/2), \qquad f(x + 1) = f(x).$$

(a) Find the Fourier series of $f(x)$.
(b) Use part (a) to prove that $\sum_{n=1}^{\infty} n^{-4} = \frac{1}{90}\pi^4$.

Exercise 5.2. Let $f(x)$ be the 1-periodic function defined by

$$f(x) = x^6 \quad (|x| \leq 1/2), \qquad f(x + 1) = f(x).$$

Use the Fourier series of $f(x)$ to obtain a formula for $\sum_{n=1}^{\infty} n^{-6}$.

Exercise 5.3. Let $f_k(x)$ be the 1-periodic function defined by

$$f_k(x) = x^{2k} \quad (|x| \le 1/2), \qquad f_k(x+1) = f_k(x).$$

Use the Fourier series of $f_k(x)$ and mathematical induction to prove that $\sum_{n=1}^{\infty} n^{-2k} = A_k \pi^{2k}$, where A_k is a rational number.

Exercise 5.4. Prove Theorem 5.3.

Exercise 5.5. Prove Theorem 5.4.

Exercise 5.6. Use equation (5.5) to prove the formulas

$$\int_a^b e^{pt} \cos qt\, dt = \frac{e^{pb}(q \sin qb + p \cos qb) - e^{pa}(q \sin qa + p \cos qa)}{p^2 + q^2},$$
$$\int_a^b e^{pt} \sin qt\, dt = \frac{e^{pb}(p \sin qb - q \cos qb) - e^{pa}(p \sin qa - q \cos qa)}{p^2 + q^2}.$$

[Hint. Compare the real and imaginary parts of the two sides of (5.5).]

Remark. Let $g : [a,b] \to \mathbb{R}$. For each $n \in \mathbb{N}$, define a sequence of Riemann sums

$$S_n(g) = \sum_{k=1}^{n} g(a + k\Delta_n)\Delta_n, \quad \Delta_n = (b-a)/n.$$

Recall from calculus that if g is continuous, then

$$\int_a^b g(x)\, dx = \lim_{n\to\infty} S_n(g).$$

Exercise 5.7. Let $a, b \in \mathbb{R}$ and $f : [a,b] \to \mathbb{C}$ be continuous. Prove that

$$\left| \int_a^b f(x)\, dx \right| \le \int_a^b |f(x)|\, dx.$$

[Hint. Let $f(x) = u(x) + \mathrm{i}v(x)$. Use the above remark to write the left side of the inequality as $|\lim_{n\to\infty}(S_n(u) + \mathrm{i}S_n(v))|$. Then use the triangle inequality to show that the last limit is $\le \lim_{n\to\infty} S_n(|f|)$. Complete the proof by using the remark one more time.]

Exercise 5.8. Let $t > 0$ and $x \in \mathbb{R}$. Prove that the series $\sum_{m=1}^{\infty} e^{-\pi(m+x)^2 t}$ and the series $\sum_{m=1}^{\infty} e^{-\pi(x-m)^2 t}$ both converge. Deduce that the series used to define the function $\theta(x;t)$ in Example 5.13 also converges.

Exercise 5.9. The purpose of this exercise is to establish that the function $\theta(x)$ in Example 5.13 is continuous at all real numbers. Let a be an arbitrary real number with $|a| \le 1$. Since $\theta(x)$ is a 1-periodic function, it suffices to show that it is continuous at any such a.

(a) Fix an $\varepsilon > 0$. Prove that there exists an integer N such that for all x in the range $a - 1 < x < a + 1$, we have

$$\left| \theta(x) - \sum_{n=-N}^{N} e^{-\pi(n+x)^2 t} \right| < \frac{\varepsilon}{3}.$$

(b) Prove that the function $\sum_{n=-N}^{N} e^{-\pi(n+x)^2 t}$ is continuous at $x = a$.

(c) Use parts (a) and (b) to prove that there exists a number $\delta > 0$ such that

$$|x - a| < \delta \quad \Longrightarrow \quad |\theta(x) - \theta(a)| < \varepsilon.$$

Deduce that $\theta(x)$ is continuous at $x = a$.

Exercise 5.10. The purpose of this exercise is to establish the connection between the Fourier sine and cosine series of a function and its complex Fourier series. Let $f : \mathbb{R} \to \mathbb{R}$ be a 1-periodic function. For $n \in \mathbb{Z}$, we define the sequences

$$a_n = 2\int_0^1 f(t)\cos(2\pi nt)\, dt, \quad b_n = 2\int_0^1 f(t)\sin(2\pi nt)\, dt, \quad c_n = \int_0^1 f(t) e^{-2\pi \mathrm{i} nt}\, dt.$$

(a) Show that $a_0 = 2c_0$ and that, for $n \ge 1$, $a_n = c_n + c_{-n}$ and $b_n = \mathrm{i}(c_n - c_{-n})$.

(b) Use Euler's formulas and part (a) to show that

$$\frac{a_0}{2} + \sum_{n=1}^{N} \big(a_n \cos(2\pi nx) + b_n \sin(2\pi nx)\big) = \sum_{n=-N}^{N} c_n e^{2\pi \mathrm{i} nx}.$$

Exercise 5.11. Let $f : \mathbb{R} \to \mathbb{C}$ be a continuous function such that $\int_{-\infty}^{\infty} x^2 |f(x)|\, dx$ converges.

(a) Let $x, y \in \mathbb{R}$. Show that

$$\frac{e^{\mathrm{i}xy} - 1}{y} - \mathrm{i}x = -yx^2 \left(\frac{1 - \cos(xy)}{(xy)^2} + \mathrm{i}\frac{xy - \sin(xy)}{(xy)^2} \right).$$

Observe that the functions $(1 - \cos t)/t^2$ and $(t - \sin t)/t^2$ are bounded to deduce that

$$\frac{e^{\mathrm{i}xy} - 1}{y} = \mathrm{i}x + yx^2 g(x, y),$$

where $g(x, y)$ is a complex-valued function such that $|g(x, y)| \le C$ for some constant $C > 0$.

(b) Let $\hat{f}(y)$ be the Fourier transform of f, and let $h \in \mathbb{R}$. Use part (a) to show that

$$\frac{\hat{f}(y + h) - \hat{f}(y)}{h} = \int_{-\infty}^{\infty} f(x) e^{-2\pi \mathrm{i} yx} \big((-2\pi \mathrm{i} x) + 4\pi^2 h x^2 g(-2\pi x, h) \big)\, dx,$$

where g is the function from part (a).

(c) Use part (b) and the bound for g from part (a) to prove that $\hat{f}(y)$ is differentiable and

$$\hat{f}'(y) = -2\pi \mathrm{i} \int_{-\infty}^{\infty} x f(x) e^{-2\pi \mathrm{i} yx}\, dx.$$

Exercise 5.12. The purpose of this exercise is to compute the Fourier transform of the Gaussian density function $\phi(x) = e^{-x^2/2}$.

(a) Show that $\int_{-\infty}^{\infty} x^2 e^{-x^2/2}\, dx$ converges.

(b) Use part (a) and Exercise 5.11 to show that

$$\hat{\phi}'(y) = -2\pi \mathrm{i} \int_{-\infty}^{\infty} x e^{-x^2/2} e^{-2\pi \mathrm{i} yx}\, dx.$$

(c) Note that $x e^{-x^2/2}\, dx = -d\big(e^{-x^2/2}\big)$ and use integration by parts to show that

$$\int_{-\infty}^{\infty} x e^{-x^2/2} e^{-2\pi \mathrm{i} yx}\, dx = -2\pi \mathrm{i} y \int_{-\infty}^{\infty} e^{-x^2/2} e^{-2\pi \mathrm{i} yx}\, dx.$$

(d) Combine parts (b) and (c) to show that $\hat{\phi}(y)$ satisfies the differential equation $\hat{\phi}' = -4\pi^2 y \hat{\phi}$. Deduce that $\hat{\phi}(y) = c e^{-2\pi^2 y^2}$ for some constant $c > 0$.

(e) By part (d), $\hat{\phi}(0) = c$. On the other hand, it is known from calculus and probability theory that

$$\hat{\phi}(0) = \int_{-\infty}^{\infty} e^{-x^2/2}\, dx = \sqrt{2\pi}.$$

Deduce that $\hat{\phi}(y) = \sqrt{2\pi} e^{-2\pi^2 y^2}$.

Exercise 5.13. Let $\phi(x; t) = e^{-\pi x^2 t}$. In Example 5.13, we showed that

$$\sum_{n \in \mathbb{Z}} \phi(n; t) = \sum_{n \in \mathbb{Z}} \hat{\phi}(n; t). \qquad (*)$$

(a) Use Exercise 5.12 and a change of variables to show that $\hat{\phi}(y; t) = e^{-\pi y^2/t} / \sqrt{t}$.

(b) Use part (a) and $(*)$ above to show that the function

$$\theta(t) = \sum_{n \in \mathbb{Z}} e^{-\pi n^2 t} \qquad (t > 0)$$

satisfies the identity $\theta(1/t) = \sqrt{t}\theta(t)$. The function $\theta(t)$ is a special case of *Jacobi's theta-function*; the identity is known as its *transformation formula*.

Exercise 5.14. Recall the Maclaurin series expansions of the sine and cosine functions: for all real x,

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + \cdots, \qquad \cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots + (-1)^n \frac{x^{2n}}{(2n)!} + \cdots.$$

(a) Use the definition of $e^{\mathrm{i}x}$ and the above Maclaurin expansions to prove that

$$e^{\mathrm{i}x} = \sum_{n=0}^{\infty} \frac{(\mathrm{i}x)^n}{n!} = 1 + (\mathrm{i}x) + \frac{(\mathrm{i}x)^2}{2!} + \cdots \qquad (x \in \mathbb{R}).$$

This shows that the Maclaurin expansion of the real exponential e^x can be extended to the the case of complex exponentials $e^{\mathrm{i}x}$.

(b) Recall that if $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$ are absolutely convergent series with real terms, then

$$\left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right) = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} a_k b_{n-k}\right).$$

Show that this identity remain true for the series with complex terms.

(c) Use parts (a) and (b), the Maclaurin expansion of e^x and the binomial formula to prove that

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!} = 1 + z + \frac{z^2}{2!} + \cdots \qquad (z \in \mathbb{C}).$$

CHAPTER 6

Algebraic Numbers And Algebraic Integers

6.1. Definition

Definition 6.1. An *algebraic number* is a complex number α that is a root of a polynomial $f(x) = a_nx^n + \cdots + a_1x + a_0$, where $a_0, a_1, \dots, a_n \in \mathbb{Q}$. An *algebraic integer* is a complex number ω that is a root of a polynomial $f(x) = x^n + b_{n-1}x^{n-1} + \cdots + b_1x + b_0$, where $b_0, b_1, \dots, b_{n-1} \in \mathbb{Z}$. In this context, we refer to the elements of $\mathbb{Z}$ as *rational integers*.

Example 6.2. Every rational number m/n is algebraic, because it is the root of $f(x) = mx - n$.

Every rational integer n is an algebraic integer, because it is root of $f(x) = x - n$.

Every root of unity $\zeta = e^{2\pi ik/n}$, where $k, n \in \mathbb{Z}$, is an algebraic integer, because it is root of $f(x) = x^n - 1$.

Theorem 6.3. *If a rational number $r \in \mathbb{Q}$ is an algebraic integer, then $r \in \mathbb{Z}$.*

Proof. Let $r = p/q$, where $p, q \in \mathbb{Z}$ and $(p, q) = 1$. Since r is an algebraic integer, there exist integers $b_0, b_1, \dots, b_{n-1}$ such that

$$\begin{aligned} & (p/q)^n + b_{n-1}(p/q)^{n-1} + \cdots + b_1(p/q) + b_0 = 0 \\ \Longrightarrow \quad & p^n + b_{n-1}p^{n-1}q + \cdots + b_1pq^{n-1} + b_0q^n = 0 \\ \Longrightarrow \quad & p^n = q\big(- b_{n-1}p^{n-1} - \cdots - b_1pq^{n-2} - b_0q^{n-1}\big). \end{aligned}$$

It follows that $q \mid p^n$, whence $(q, p^n) = |q|$. On the other hand, since $(p, q) = 1$, Corollary 1.12 gives $(q, p^n) = 1$. We conclude that $q = \pm 1$, and hence, $r \in \mathbb{Z}$. □

6.2. The ring of algebraic integers

In this section, we shall show that the set of all algebraic complex numbers is a field and the set of algebraic integers is a commutative ring.

Lemma 6.4. *Let $\theta_1, \theta_2, \dots, \theta_n$ be complex numbers, not all equal to 0, and define the set*

$$M = M(\theta_1, \theta_2, \dots, \theta_n) = \big\{c_1\theta_1 + c_2\theta_2 + \cdots + c_n\theta_n \mid c_1, c_2, \dots, c_n \in \mathbb{Q}\big\}.$$

If $\alpha \in \mathbb{C}$ is such that $\alpha\theta_j \in M$ for $j = 1, 2, \dots, n$, then α is an algebraic number.

Proof. Since $\alpha\theta_i \in M$, we have

$$\alpha\theta_i = c_{i1}\theta_1 + c_{i2}\theta_2 + \cdots + c_{in}\theta_n \quad \text{for some } c_{ij} \in \mathbb{Q}.$$

Combining these identities, we find that $\theta_1, \theta_2, \dots, \theta_n$ is a nontrivial solution of the homogeneous linear system

$$A\mathbf{x} = \alpha\mathbf{x} \quad \Longleftrightarrow \quad (A - \alpha I)\mathbf{x} = \mathbf{0},$$

where A is the square matrix $A = [c_{ij}]$. The above system has nontrivial solutions if and only if

$$\begin{vmatrix} c_{11} - \alpha & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} - \alpha & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} - \alpha \end{vmatrix} = 0.$$

The value of the determinant is an expression of the form

$$(-1)^n \alpha^n + a_{n-1}\alpha^{n-1} + \cdots + a_1\alpha + a_0,$$

where each a_k is a rational number, because it is a finite sum of products of c_{ij}'s. Therefore, α is a root of a polynomial with rational coefficients. □

Theorem 6.5. *Let* α *and* β *be nonzero algebraic numbers. Then so are* $\alpha \pm \beta$*,* $\alpha\beta$*,* β^{-1} *and* α/β.

Of course, if α or β is zero, $\alpha \pm \beta$ and $\alpha\beta$ are also algebraic.

Proof. Suppose that

$$a_n\alpha^n + \cdots + a_1\alpha + a_0 = 0, \qquad b_m\beta^m + \cdots + b_1\beta + b_0 = 0,$$

where $a_0, a_1, \ldots, a_n, b_0, b_1, \ldots, b_m \in \mathbb{Q}$, $a_n \neq 0$, and $b_m \neq 0$. It follows that

$$\alpha^n = a'_{n-1}\alpha^{n-1} + \cdots + a'_1\alpha + a'_0, \qquad \beta^m = b'_{m-1}\beta^{m-1} + \cdots + b'_1\beta + b'_0, \tag{6.1}$$

where $a'_k = -a_k/a_n$ and $b'_k = -b_k/b_m$. Let M be the set of all sums of the form

$$\sum_{i=0}^{n-1}\sum_{j=0}^{m-1} c_{ij}\alpha^i\beta^j, \tag{6.2}$$

with coefficients $c_{ij} \in \mathbb{Q}$. Note that M is a set of the type considered in Lemma 6.4.

To prove that $\alpha + \beta$ is algebraic, we show that $(\alpha+\beta)\alpha^i\beta^j \in M$ for all $i = 0, 1, \ldots, n-1$ and $j = 0, 1, \ldots, m-1$, and then we apply the lemma. We have

$$(\alpha + \beta)\alpha^i\beta^j = \alpha^{i+1}\beta^j + \alpha^i\beta^{j+1}.$$

When $i < n-1$ and $j < m-1$, the sum on the right is of the form (6.2) with $c_{i+1,j} = c_{i,j+1} = 1$ and all the other coefficients equal to 0. When $i = n-1$ and $j < m-1$, we use (6.1) to obtain

$$\alpha^n\beta^j + \alpha^{n-1}\beta^{j+1} = a'_{n-1}\alpha^{n-1}\beta^j + \cdots + a'_1\alpha\beta^j + a'_0\beta^j + \alpha^i\beta^{j+1},$$

and the sum on the right is again of the form (6.2). Similarly, when $i < n-1$ and $j = m-1$ or $i = n-1$ and $j = m-1$, we can use (6.1) to express the sum $\alpha^n\beta^j + \alpha^{n-1}\beta^{j+1}$ in the form (6.2). This proves that $\alpha + \beta$ is algebraic.

The proofs that $\alpha - \beta$ and $\alpha\beta$ are algebraic are similar; β^{-1} is algebraic, because it is a root of the polynomial $b_0x^n + b_1x^{n-1} + \cdots + b_n$, which has rational coefficients. Finally, to show that α/β is algebraic, we observe that it is the product of α and β^{-1}, both of which are algebraic. □

The proofs of the next two results are similar to those of Lemma 6.4 and Theorem 6.5, so we leave them as exercises.

Lemma 6.6. *Let* $\theta_1, \theta_2, \ldots, \theta_n$ *be complex numbers, not all equal to* 0*, and define the set*

$$M = M(\theta_1, \theta_2, \ldots, \theta_n) = \left\{c_1\theta_1 + c_2\theta_2 + \cdots + c_n\theta_n \mid c_1, c_2, \ldots, c_n \in \mathbb{Z}\right\}.$$

If $\alpha \in \mathbb{C}$ *is such that* $\alpha\theta_j \in M$ *for* $j = 1, 2, \ldots, n$*, then* α *is an algebraic integer.*

Theorem 6.7. *Let* α *and* β *be algebraic integers. Then so are* $\alpha \pm \beta$ *and* $\alpha\beta$.

Remark. In the terminology of abstract algebra, Theorems 6.5 and 6.7 establish that the set of algebraic numbers and the set of algebraic integers are, respectively, a subfield and a subring of the field of complex numbers.

6.3. Congruences between algebraic integers

Let Ω denote the ring of all algebraic integers.

Definition 6.8. Let $\omega \in \Omega$, with $\omega \neq 0$, and let $\alpha, \beta \in \Omega$. We say that α *is congruent to* β *modulo* ω, and write $\alpha \equiv \beta \pmod{\omega}$, if $\alpha - \beta = \omega\eta$ for some $\eta \in \Omega$.

Remark. Suppose that $a, b \in \mathbb{Z}$ and $m \in \mathbb{N}$. Then a, b and m belong also to Ω, so the notation $a \equiv b \pmod{m}$ now has a double meaning: according to Theorem 1.20, it means that $a - b = mk$ for some $k \in \mathbb{Z}$; and according to the above definition, it means that $a - b = m\eta$ for some $\eta \in \Omega$. The ambiguity is, however, superficial. Indeed, if $a - b = m\eta$, with $a, b, m \in \mathbb{Z}$, $m \neq 0$ and $\eta \in \Omega$, then

$$\eta = (a - b)/m \in \mathbb{Q} \cap \Omega.$$

Therefore, it follows from Theorem 6.3 that if $a - b = m\eta$, with $\eta \in \Omega$, then in fact, $\eta \in \mathbb{Z}$.

The next theorem summarizes some basic properties of congruences between algebraic integers. Note the similarities between the properties below and the properties of congruence modulo a rational integer m listed in Theorem 1.21. The proof of the theorem is also similar to the proof of Theorem 1.21, so we leave it as an exercise.

Theorem 6.9. *Let* $\omega \in \Omega$, *with* $\omega \neq 0$, *and* $\alpha, \beta, \gamma, \delta \in \Omega$. *Then:*

i) $\alpha \equiv \alpha \pmod{\omega}$*;*
ii) *if* $\alpha \equiv \beta \pmod{\omega}$*, then* $\beta \equiv \alpha \pmod{\omega}$*;*
iii) *if* $\alpha \equiv \beta \pmod{\omega}$ *and* $\beta \equiv \gamma \pmod{\omega}$*, then* $\alpha \equiv \gamma \pmod{\omega}$*;*
iv) *if* $\alpha \equiv \beta \pmod{\omega}$ *and* $\gamma \equiv \delta \pmod{\omega}$*, then* $\alpha + \gamma \equiv \beta + \delta \pmod{\omega}$*;*
v) *if* $\alpha \equiv \beta \pmod{\omega}$ *and* $\gamma \equiv \delta \pmod{\omega}$*, then* $\alpha\gamma \equiv \beta\delta \pmod{\omega}$.

Lemma 6.10. *Let p be a prime number and let* $\alpha, \beta \in \Omega$. *Then*

$$(\alpha + \beta)^p \equiv \alpha^p + \beta^p \pmod{p}.$$

Proof. By the binomial theorem,

$$(\alpha + \beta)^p = \alpha^p + \beta^p + \sum_{k=1}^{p-1} \binom{p}{k} \alpha^k \beta^{p-k}.$$

By Exercise 1.10, $\binom{p}{k} \equiv 0 \pmod{p}$ for all $k = 1, 2, \dots, p - 1$, so multiple applications of parts iv) and v) of Theorem 6.9 show that

$$\sum_{k=1}^{p-1} \binom{p}{k} \alpha^k \beta^{p-k} \equiv 0 \pmod{p} \implies (\alpha + \beta)^p \equiv \alpha^p + \beta^p \pmod{p}.$$

□

6.4. A proof of Theorem 2.17

Let $\zeta = e^{2\pi \mathrm{i}/8}$ and $\tau = \zeta + \zeta^{-1}$. Note that ζ and ζ^{-1} are algebraic integers, since both are roots of the polynomial $x^4 + 1$. Also, $\tau \in \Omega$ by Theorem 6.7. Furthermore, we have

$$\tau^2 = \zeta^2 + 2 + \zeta^{-2} = e^{\pi \mathrm{i}/2} + 2 + e^{-\pi \mathrm{i}/2} = \mathrm{i} + 2 - \mathrm{i} = 2.$$

Raising both sides of the above identity to the $(p-1)/2$ power and using Euler's criterion, we obtain

$$\tau^{p-1} = (\tau^2)^{(p-1)/2} = 2^{(p-1)/2} \equiv \left(\frac{2}{p}\right) \pmod{p},$$

whence

$$\tau^{p+1} \equiv \left(\frac{2}{p}\right)\tau^2 = \left(\frac{2}{p}\right)2 \pmod{p}. \tag{6.3}$$

Next, we make use of congruences between algebraic integers. By Lemma 6.10,

$$\tau^p = \left(\zeta + \zeta^{-1}\right)^p \equiv \zeta^p + \zeta^{-p} \pmod{p},$$

whence,

$$\tau^{p+1} \equiv \left(\zeta^p + \zeta^{-p}\right)\tau \pmod{p}.$$

Note that the value of the sum $\zeta^p + \zeta^{-p}$ depends only on the residue class of p modulo 8:

$$\zeta^p + \zeta^{-p} = \begin{cases} \zeta + \zeta^{-1} & \text{if } p \equiv \pm 1 \pmod{8}, \\ \zeta^3 + \zeta^{-3} & \text{if } p \equiv \pm 3 \pmod{8}. \end{cases}$$

Thus, when $p \equiv \pm 1 \pmod{8}$, we have

$$\tau^{p+1} \equiv \tau^2 = 2 \pmod{p};$$

and when $p \equiv \pm 3 \pmod{8}$, we have

$$\tau^{p+1} \equiv \left(\zeta^3 + \zeta^{-3}\right)\left(\zeta + \zeta^{-1}\right) = \zeta^4 + \zeta^2 + \zeta^{-2} + \zeta^{-4} = -2 \pmod{p}.$$

Since the expression $(p^2 - 1)/8$ is even when $p \equiv \pm 1 \pmod{8}$ and odd when $p \equiv \pm 3 \pmod{8}$, we can summarize the above computation as

$$\tau^{p+1} \equiv (-1)^{(p^2-1)/8}2 \pmod{p}.$$

Comparing this congruence with (6.3), we deduce that

$$\left(\frac{2}{p}\right)2 \equiv (-1)^{(p^2-1)/8}2 \pmod{p}.$$

We proved this congruence as a congruence between algebraic integers, but since both sides and the modulus are in $\mathbb{Z}$, our earlier remark shows that the same congruence holds modulo p. Since $(2, p) = 1$, we can now cancel the 2's from the last congruence to obtain

$$\left(\frac{2}{p}\right) \equiv (-1)^{(p^2-1)/8} \pmod{p}.$$

Since both sides of this congruence are equal to 1 or -1, their difference is ± 2 or 0. Since that difference must be divisible by $p > 2$, it can only equal 0 and Theorem 2.17 follows. □

6.5. Exercises

Exercise 6.1. Let $m, n \in \mathbb{N}$. Show that $\sqrt[n]{m}$ is an algebraic integer.

Exercise 6.2. Show that $\sqrt[3]{2} + \sqrt[3]{7}$ is an algebraic integer without using Theorem 6.7.

Exercise 6.3. Show that $\sqrt{6 + \sqrt[3]{4}} - \sqrt{6 - \sqrt[3]{2}}$ is an algebraic integer without using Theorem 6.7.

Exercise 6.4. Provide the details of the proof of Theorem 6.5 for $\alpha\beta$.

Exercise 6.5. Prove Lemma 6.6.

Exercise 6.6. Prove Theorem 6.7.

Exercise 6.7. Prove Theorem 6.9.

Recall that a set X is called *countable* if either X is finite, or there exists a bijection from X onto $\mathbb{N}$, the set of natural numbers. Recall also the following fact.

If $X_1, X_2, \dots, X_n, \dots$ are countable sets, then their union $\cup_{n=1}^{\infty} X_n$ is also countable.

Exercise 6.8. The purpose of this exercise is to show that the set Ξ of all algebraic numbers is countable.

(a) Let Ξ_n be the set of algebraic numbers α that satisfy polynomial equations over $\mathbb{Q}$ of degrees $\leq n$,

$$a_n\alpha^n + \cdots + a_1\alpha + a_0 = 0 \qquad (a_i \in \mathbb{Q}).$$

Show that Ξ_n is countable.

(b) Use part (a) and the above fact to show that Ξ is countable.

(c) Use part (b) to show that there exist complex numbers that are not algebraic (such numbers are called *transcendental*).

Exercise 6.9. Let $\alpha \in \mathbb{C}$ be algebraic, and let $p(x) \in \mathbb{Q}[x]$ be a non-zero polynomial with rational coefficients that has the least degree among all polynomials that have α as a root. Then any polynomial $f(x) \in \mathbb{Q}[x]$ such that $f(\alpha) = 0$ is divisible by $p(x)$.

Exercise 6.10. Consider $\mathbb{C}$ with the operation of addition of complex numbers and the operation of multiplication of complex numbers by rationals.

(a) Show that $\mathbb{C}$, with the above operations, is a linear space over $\mathbb{Q}$.

(b) Show that the set Ξ of all algebraic numbers is a linear subspace of the linear space from part (a). The purpose of the remainder of the exercise is to show that Ξ is an infinite-dimensional subspace.

(c) Show if the linear subspace Ξ from part (b) is n-dimensional, then every algebraic number α satisfies a polynomial equation with rational coefficients and of degree at most n.

(d) Let $\zeta = e^{2\pi \mathrm{i}/m}$, $m \in \mathbb{N}$, and let $p(x) = x^k + a_{k-1}x^{k-1} + \cdots + a_1 x + a_0 \in \mathbb{Q}[x]$ be the monic polynomial of least degree that has ζ as its root. Use Exercise 6.9 to show that $p(x)$ divides $x^m - 1$. Deduce that all the roots of $p(x)$ are of the form ζ^l, $l \in \mathbb{N}$.

(e) Use part (d) to show that $|a_j| \leq \binom{k}{j}$ for all $j = 0, 1, \dots, k-1$.

(f) Use parts (c) and (e) to show that if Ξ is an n-dimensional subspace of $\mathbb{C}$, then there are only a finite number of distinct numbers of the form $\zeta = e^{2\pi \mathrm{i}/m}$, $m \in \mathbb{N}$. Observe that the numbers of this form are, in fact, pairwise distinct, and deduce that Ξ must be an infinite-dimensional subspace of $\mathbb{C}$.

CHAPTER 7

Proofs Of The Law Of Quadratic Reciprocity

This chapter consists of a series of exercises that develop six proofs of the law of quadratic reciprocity.

7.1. S.Y. Kim's elementary proof

Our first proof is a relatively recent one: it was published by S.Y. Kim in 2004[1]. It uses only elementary properties of congruences and some basic combinatorics. It also uses the results of Exercises 2.7 and 2.9.

Exercise 7.1 (First proof of the law of quadratic reciprocity)**.** Let p and q be distinct odd primes and define

$$R = \left\{a \in \mathbb{N} \mid 1 \le a \le \tfrac{1}{2}(pq-1),\ (a,pq)=1\right\}, \quad \Pi = \prod_{a\in R} a.$$

(a) Prove that $T \subseteq S$ and that $R = S - T$, where

$$S = \left\{a \in \mathbb{N} \mid 1 \le a \le \tfrac{1}{2}(pq-1),\ (a,p)=1\right\}, \quad T = \left\{q, 2q, \dots, \tfrac{1}{2}(p-1)q\right\}.$$

(b) Use part (a), Wilson's theorem (Exercise 2.9) and Euler's criterion to prove that $\Pi \equiv (-1)^{(q-1)/2}\left(\frac{q}{p}\right) \pmod{p}$.

(c) By switching the roles of p and q in part (b), show that $\Pi \equiv (-1)^{(p-1)/2}\left(\frac{p}{q}\right) \pmod{q}$.

(d) Use parts (b) and (c) to prove that $(-1)^{(q-1)/2}\left(\frac{q}{p}\right) = (-1)^{(p-1)/2}\left(\frac{p}{q}\right)$ if and only if $\Pi \equiv \pm 1 \pmod{pq}$.

(e) Let $U = \left\{a \in R \mid a^2 \equiv \pm 1 \pmod{pq}\right\}$. Prove that $\Pi \equiv \pm \prod_{a\in U} a \pmod{pq}$.

[Hint. For any $a \in R$, exactly one of $\bar{a}$ or $pq - \bar{a}$ is also an element of R. Denote that element of R by a^*. Then $aa^* \equiv \pm 1 \pmod{pq}$ for all $a \in R$. Pair the elements of $R - U$ into pairs a, a^*.]

(f) By Exercise 2.7, the congruence $x^2 \equiv 1 \pmod{pq}$ has four solutions: ± 1 and $\pm\alpha$ for some integer α. Show that only two of these four solutions are congruent to integers in U: 1 and one of α and $-\alpha$.

(g) Use Exercise 2.7 and Corollary 2.13 to prove that when $p \equiv 3 \pmod 4$ or $q \equiv 3 \pmod 4$, the congruence $x^2 \equiv -1 \pmod{pq}$ has no solution.

(h) Use Exercise 2.7 and Corollary 2.13 to prove that when $p \equiv q \equiv 1 \pmod 4$, the congruence $x^2 \equiv -1 \pmod{pq}$ has four solutions.

(i) Let $p \equiv q \equiv 1 \pmod 4$ and let β be one of the four solutions of $x^2 \equiv -1 \pmod{pq}$. Prove that the other three solutions are $-\beta$ and $\pm\alpha\beta$, where α is the integer from part (f). Show that only two of these four solutions are congruent to integers in U: one of β and $-\beta$ and one of $\alpha\beta$ or $-\alpha\beta$.

[1] An Elementary Proof of the Quadratic Reciprocity Law, *Amer. Math. Monthly* **111** (2004), 48–50.

(j) Use parts (e)–(i) to prove that $\Pi \equiv \pm 1 \pmod{pq}$ if and only if $p \equiv q \equiv 1 \pmod 4$.
[Hint. The set U is either $\{1, *\alpha\}$ or $\{1, *\alpha, *\beta, *\alpha\beta\}$, where each $*$ denotes either $+$ or $-$ sign.]

(k) Use parts (d) and (j) to prove that $(-1)^{(q-1)/2}\left(\frac{q}{p}\right) = (-1)^{(p-1)/2}\left(\frac{p}{q}\right)$ if and only if $p \equiv q \equiv 1 \pmod 4$.

(l) Deduce the law of quadratic reciprocity from part (k).

7.2. Eisenstein's lemma and counting lattice points

Our second proof of the law of quadratic reciprocity is a variant of one of Gauss's proofs. It uses a lemma of the German mathematician Eisenstein and a counting argument for lattice points in the plane.

Exercise 7.2 (Eisenstein's lemma)**.** Let p be an odd prime and a an odd integer with $(a, p) = 1$, and define

$$T(a,p) = \sum_{j=1}^{(p-1)/2} [ja/p].$$

(a) For $j = 1, 2, \ldots, \frac{1}{2}(p-1)$, let $ja = pq_j + r_j$ with $0 \le r_j < p$. Show that $q_j = [ja/p]$ and $r_j \ne 0$.

(b) Define

$$u_j = \begin{cases} r_j & \text{if } 0 < r_j < p/2, \\ p - r_j & \text{if } p/2 < r_j < p. \end{cases}$$

Prove that $u_1, u_2, \ldots, u_{(p-1)/2}$ are pairwise distinct. Deduce that

$$u_1 + u_2 + \cdots + u_{(p-1)/2} = 1 + 2 + \cdots + \tfrac{1}{2}(p-1).$$

(c) Let μ be the number from the statement of Gauss' lemma. Prove that

$$r_1 + r_2 + \cdots + r_{(p-1)/2} \equiv u_1 + u_2 + \cdots + u_{(p-1)/2} + \mu \pmod 2.$$

[Hint. $p - u_j \equiv u_j + 1 \pmod 2$.]

(d) Use parts (a), (b) and (c) to prove that

$$a \sum_{j=1}^{(p-1)/2} j \equiv pT(a,p) + \sum_{j=1}^{(p-1)/2} j + \mu \pmod 2.$$

Deduce that $T(a,p) \equiv \mu \pmod 2$.

(e) Use part (d) and Gauss' lemma to prove that $\left(\frac{a}{p}\right) = (-1)^{T(a,p)}$.

Exercise 7.3 (Second proof of the law of quadratic reciprocity)**.** Let p and q be distinct odd primes and let $\mathbf{R}$ be the rectangle with vertices $O(0,0)$, $A(p/2,0)$, $B(0,q/2)$ and $C(p/2,q/2)$.

(a) Prove that the number of lattice points (that is, points (x,y) with integer coordinates) inside $\mathbf{R}$ is $\frac{1}{2}(p-1)\cdot\frac{1}{2}(q-1)$.

(b) Prove that there are no lattice points on the diagonal OC of $\mathbf{R}$.

(c) Let $T(a,p)$ be the quantity defined in Exercise 7.2. Prove that the number of lattice points inside the triangle $\mathbf{T}_1$ with vertices O, A and C is $T(q,p)$.

(d) Prove that the number of lattice points inside the triangle $\mathbf{T}_2$ with vertices O, B and C is $T(p,q)$.

(e) Use parts (a)–(d) to prove that $T(p,q) + T(q,p) = \frac{1}{2}(p-1)\cdot\frac{1}{2}(q-1)$.

(f) Deduce the law of quadratic reciprocity from part (e) and Exercise 7.2(e).

7.3. Another proof of Eisenstein's

In this section, we present a proof published by Eisenstein in 1845[2].

Exercise 7.4 (Third proof of the law of quadratic reciprocity)**.** Let p and q be distinct odd primes. Define the function $f(z) = e^{2\pi \mathrm{i} z} - e^{-2\pi \mathrm{i} z} = 2\mathrm{i}\sin(2\pi z)$, where $\mathrm{i} = \sqrt{-1}$.

(a) Prove that if $n \in \mathbb{N}$ is odd, then $x^n - y^n = \prod_{k=0}^{n-1} \left(\zeta^k x - \zeta^{-k} y\right)$, where $\zeta = e^{2\pi \mathrm{i}/n}$. [Hint. Consider both sides of the identity as polynomials in x. Show that those polynomials have the same roots and the same leading coefficients.]

(b) Use part (a) to prove that if $n \in \mathbb{N}$ is odd, then

$$x^n - y^n = (x-y) \prod_{k=1}^{(n-1)/2} \left(\zeta^k x - \zeta^{-k} y\right)\left(\zeta^{-k} x - \zeta^{k} y\right).$$

[Hint. If $k > n/2$, then $j = n-k < n/2$ and $\zeta^k x - \zeta^{-k} y = \zeta^{-j} x - \zeta^{j} y$.]

(c) Prove that $f(z+1) = f(z)$ and $f(-z) = -f(z)$, and that the only real zeros of $f(z)$ are the numbers $n/2$, where n is an integer.

(d) Prove that if $n \in \mathbb{N}$ is odd, then

$$\frac{f(nz)}{f(z)} = \prod_{k=1}^{(n-1)/2} f\left(z + \frac{k}{n}\right) f\left(z - \frac{k}{n}\right).$$

[Hint. Use the identity from part (b) with $x = e^{2\pi \mathrm{i} z}$ and $y = e^{-2\pi \mathrm{i} z}$.]

(e) Prove that if $(a, p) = 1$, then

$$\prod_{j=1}^{(p-1)/2} f\left(\frac{ja}{p}\right) = \left(\frac{a}{p}\right) \prod_{j=1}^{(p-1)/2} f\left(\frac{j}{p}\right).$$

[Hint. For $j = 1, 2, \ldots, \frac{1}{2}(p-1)$, define $r_j \in \{1, 2, \ldots, \frac{1}{2}(p-1)\}$ and $\varepsilon_j = \pm 1$ as in Exercise 2.8(a). Prove that $f(ja/p) = f(\varepsilon_j r_j/p) = \varepsilon_j f(r_j/p)$. Take the product of these identities, then use Exercise 2.8(a, b) and Gauss' lemma.]

(f) Prove that

$$\left(\frac{q}{p}\right) = \prod_{k=1}^{(q-1)/2} \prod_{j=1}^{(p-1)/2} f\left(\frac{j}{p} + \frac{k}{q}\right) f\left(\frac{j}{p} - \frac{k}{q}\right).$$

[Hint. By part (e) with $a = q$, $\left(\frac{q}{p}\right) = \prod_{j=1}^{(p-1)/2} f(jq/p)/f(j/p)$. Apply the identity in part (d) with $n = q$ and $z = j/p$ to the jth factor of the last product.]

(g) Use part (f) and the analogous expression for $\left(\frac{p}{q}\right)$ to prove the law of quadratic reciprocity.

7.4. An algebraic proof: Gauss sums are algebraic integers

Recall that the Gauss sum is defined for $m \in \mathbb{N}$ and $a \in \mathbb{Z}$ by

$$G(m, a) = \sum_{n=1}^{m} e^{2\pi \mathrm{i} a n^2/m}. \tag{7.1}$$

In this section, we present a proof that uses some algebraic properties of $G(m, a)$. In particular, we use that $G(m, a)$ is an algebraic integer.

[2] Application de l'algèbre à l'arithmétique transcendante, *J. Reine Angew. Math.* **29** (1845), 177-184.

Exercise 7.5 (Fourth proof of the law of quadratic reciprocity)**.** Let p and q be distinct odd primes and let $G(m, a)$ be the sum defined in (7.1). Also, set $p^* = (-1)^{(p-1)/2}p$.

(a) Use Exercise 1.14 to show that the numbers

$$e^{2\pi i n^2/pq} \qquad (1 \le n \le pq)$$

are the same as

$$e^{2\pi i (qx+py)^2/pq} \qquad (1 \le x \le p,\ 1 \le y \le q).$$

(b) Use part (a) to prove that $G(pq, 1) = G(p, q)G(q, p)$.

(c) Let $(a, p) = 1$. Prove that

$$G(p, a) = \sum_{x=1}^{p-1} \left(\frac{x}{p}\right) e^{2\pi i ax/p}.$$

[Hint. Show that $G(p, a) = 1 + 2\sum_r e^{2\pi i ar/p}$, where r runs through the quadratic residues modulo p. Then show that $\left(\frac{x}{p}\right) + 1$ equals 2 when x is a quadratic residue and 0 when x is a quadratic nonresidue. Deduce that $G(p, a) = 1 + \sum_{x=1}^{p-1} \left(\left(\frac{x}{p}\right) + 1\right) e^{2\pi i ax/p}$. Use Lemma 3.25 to complete the proof.]

(d) Let $(a, p) = 1$. Use part (c) to prove that $G(p, a) = \left(\frac{a}{p}\right) G(p, 1)$.
[Hint. By Theorem 1.31, as x runs through $1, 2, \ldots, p-1$, the products ax also run through a reduced residue system modulo p.]

(e) Show that $\overline{G(p, 1)} = G(p, -1)$ (the bar denotes complex conjugation). Use this identity and part (d) to prove that $|G(p, 1)|^2 = \left(\frac{-1}{p}\right) G(p, 1)^2$.

(f) Use part (e), Corollary 2.13 and Theorem 3.24 to prove that $G(p, 1)^2 = p^*$.

(g) Use part (f) and Euler's criterion to prove that $G(p, 1)^{q-1} \equiv \left(\frac{p^*}{q}\right) \pmod{q}$.

(h) Prove that $G(p, 1)$ is an algebraic integer.

(i) Prove that $G(p, 1)^q \equiv G(p, q) \pmod{q}$. Use this congruence and part (d) to deduce that $G(p, 1)^q \equiv \left(\frac{q}{p}\right) G(p, 1) \pmod{q}$.

(j) Use parts (f), (g) and (i) to prove that $G(p, 1)^{q+1} \equiv \left(\frac{p^*}{q}\right) p^* \equiv \left(\frac{q}{p}\right) p^* \pmod{q}$.

(k) Observe that $(p^*, q) = 1$ and deduce from part (j) that $\left(\frac{p^*}{q}\right) \equiv \left(\frac{q}{p}\right) \pmod{q}$.

(l) Use part (k) to prove that $\left(\frac{p^*}{q}\right) = \left(\frac{q}{p}\right)$.

(m) Deduce the law of quadratic reciprocity from part (l).

7.5. A proof using the formula for the Gauss sum

In this section, we demonstrate that the law of quadratic reciprocity follows from a closed formula for the Gauss sum (7.1). The idea goes back to Gauss. Note that there is some overlap between this proof and the previous one, but this proof requires the exact formula in Theorem 2.21. In the following two sections, we present two different proofs of that formula.

Exercise 7.6 (Fifth proof of the law of quadratic reciprocity)**.** Let p and q be distinct odd primes. For $m \in \mathbb{N}$ and $a \in \mathbb{Z}$, define

$$G(m, a) = \sum_{n=1}^{m} e^{2\pi i a n^2/m}.$$

(a) Use Exercise 1.14 to show that the numbers

$$e^{2\pi i n^2/pq} \qquad (1 \le n \le pq)$$

are the same as

$$e^{2\pi \mathrm{i}(qx+py)^2/pq} \qquad (1 \le x \le p,\ 1 \le y \le q).$$

(b) Use part (a) to prove that $G(pq,1) = G(p,q)G(q,p)$.

(c) Let $(a,p) = 1$. Prove that

$$G(p,a) = \sum_{x=1}^{p-1} \left(\frac{x}{p}\right) e^{2\pi \mathrm{i} ax/p}.$$

[Hint. Show that $G(p,a) = 1 + 2\sum_r e^{2\pi \mathrm{i} ar/p}$, where r runs through the quadratic residues modulo p. Then show that $\left(\frac{x}{p}\right) + 1$ equals 2 when x is a quadratic residue and 0 when x is a quadratic nonresidue. Deduce that $G(p,a) = 1 + \sum_{x=1}^{p-1} \left(\left(\frac{x}{p}\right) + 1\right) e^{2\pi \mathrm{i} ax/p}$. Use Lemma 3.25 to complete the proof.]

(d) Let $(a,p) = 1$. Use part (c) to prove that $G(p,a) = \left(\frac{a}{p}\right) G(p,1)$.
[Hint. By Theorem 1.31, as x runs through $1, 2, \ldots, p-1$, the products ax also run through a reduced residue system modulo p.]

(e) Use parts (b) and (d) to prove that $G(pq,1) = \left(\frac{p}{q}\right)\left(\frac{q}{p}\right) G(p,1)G(q,1)$.

(f) Deduce the law of quadratic reciprocity from part (e) and the explicit formula for $G(m,1)$ in Theorem 2.21.

7.6. An evaluation of the Gauss sum using Fourier series

In this section, we deduce Theorem 2.21 from Poisson's summation formula (Theorem 5.12).

Exercise 7.7 (First evaluation of the Gauss sum)**.** Suppose that m is a positive integer.

(a) Define the function $f : \mathbb{R} \to \mathbb{C}$ by

$$f(x) = \frac{f_0(x-0) + f_0(x+0)}{2}, \qquad f_0(x) = \begin{cases} e^{2\pi \mathrm{i} x^2/m} & \text{if } 0 \le x \le m, \\ 0 & \text{otherwise.} \end{cases}$$

Apply Poisson's summation formula to f to prove that

$$G(m) = m \sum_{n\in\mathbb{Z}} e^{-\pi \mathrm{i} mn^2/2} \int_{-n/2}^{-n/2+1} e^{2\pi \mathrm{i} my^2}\, dy. \qquad (*)$$

(b) Show that the improper integral $\gamma = \int_{-\infty}^{\infty} e^{2\pi \mathrm{i} t^2}\, dt$ is convergent if and only if the improper integrals

$$\int_1^\infty \frac{\sin(2\pi u)}{\sqrt{u}}\, du \quad \text{and} \quad \int_1^\infty \frac{\cos(2\pi u)}{\sqrt{u}}\, du$$

are both convergent.

(c) The following test for convergence of improper integrals is due to Dirichlet:

Let f and g be functions continuous on the interval $[a,\infty)$ and such that:

i) *there is a number $K > 0$ such that $\left|\int_a^b f(x)\,dx\right| \le K$ for all $b > a$;*

ii) *g is a decreasing function and $\lim_{x\to\infty} g(x) = 0$.*

Then the improper integral $\int_a^\infty f(x)g(x)\,dx$ is convergent.

Use Dirichlet's test and part (b) to show that the improper integral γ is convergent.

(d) Use part (a) to prove that $G(m) = \gamma(1 + \mathrm{i}^{-m})\sqrt{m}$. [Hint. The contribution to the sum in $(*)$ from the even values of n is $\sum_k \int_{-k}^{-k+1} e^{2\pi \mathrm{i} m y^2}\,dy$, and the contribution from the odd values of n is $\mathrm{i}^{-m}\sum_k \int_{-k-1/2}^{-k+1/2} e^{2\pi \mathrm{i} m y^2}\,dy$.]

(e) Deduce Theorem 2.21 from part (d).

7.7. An evaluation of the Gauss sum using matrices

Note that our fifth proof of the quadratic reciprocity law requires Theorem 2.21 only in the special cases $m = p$, $m = q$ and $m = pq$. In this section, we give another proof of Theorem 2.21 in those special cases. This proof uses linear algebra over the complex numbers.

Exercise 7.8 (Second evaluation of the Gauss sum)**.** Let m be an odd integer, and A be the $m \times m$ matrix with entries $e^{2\pi \mathrm{i} jk/m}$, $0 \le j, k \le m-1$: that is, the entry in the $(j+1, k+1)$ position of A is $e^{2\pi \mathrm{i} jk/m}$.

(a) Show that $G(m, 1) = \sum_{j=1}^{m} \lambda_j$, where $\lambda_1, \lambda_2, \dots, \lambda_m$ are the eigenvalues of A.

(b) Use Lemma 3.25 to prove that

$$A^2 = \begin{bmatrix} m & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{bmatrix}, \quad \text{where } B = \begin{bmatrix} 0 & \cdots & 0 & m \\ 0 & \cdots & m & 0 \\ \vdots & & \vdots & \vdots \\ m & \cdots & 0 & 0 \end{bmatrix}.$$

(c) Prove that the characteristic polynomial of A^2 is $-(x-m)^{(m+1)/2}(x+m)^{(m-1)/2}$. [Hint. See Exercise 4.18.]

(d) Show that among the eigenvalues $\lambda_1^2, \lambda_2^2, \dots, \lambda_m^2$ of A^2 there are $(m+1)/2$ equal to m and $(m-1)/2$ equal to $-m$. Deduce that all the eigenvalues of A are of the forms $\pm\sqrt{m}$ or $\pm\mathrm{i}\sqrt{m}$ and use part (a) to show that

$$G(m, 1) = ((a-b) + (c-d)\mathrm{i})\sqrt{m}, \qquad (*)$$

where a, b, c and d are the numbers of eigenvalues of A equal to $\sqrt{m}, -\sqrt{m}, \mathrm{i}\sqrt{m}$ and $-\mathrm{i}\sqrt{m}$, respectively.

(e) Show that $a + b = (m+1)/2$ and $c + d = (m-1)/2$ and that $\det A = \lambda_1\lambda_2\cdots\lambda_m = \mathrm{i}^{2b+c-d} m^{m/2}$.

(f) Use part (c) to show that $\det A^2 = (-1)^{m(m-1)/2} m^m$. Deduce that

$$\det A = \pm \mathrm{i}^{m(m-1)/2}\sqrt{m^m}.$$

(g) Use Theorem 4.22 to evaluate $\det A$. Then use the identity

$$e^{2\pi \mathrm{i} k/m} - e^{2\pi \mathrm{i} j/m} = (-1)^{(k+j)/m} 2\mathrm{i}\sin\big(\pi(k-j)/m\big)$$

to obtain that

$$\det A = \mathrm{i}^{m(m-1)/2} \prod_{0 \le j < k \le m-1} \big(2\sin\big(\pi(k-j)/m\big).$$

(h) Combine parts (f) and (g) to show that $\det A = \mathrm{i}^{m(m-1)/2}\sqrt{m^m}$.

(i) Let p be an odd prime. Show that $\overline{G(p,1)} = G(p,-1)$ (the bar denotes complex conjugation). Use this identity and Exercise 7.6(d) to prove that $|G(p,1)|^2 = \left(\frac{-1}{p}\right) G(p,1)^2$.

(j) Use part (i), Corollary 2.13 and Theorem 3.24 to show that $G(p,1)^2 = (-1)^{(p-1)/2}p$. Deduce that

$$G(p,1) = \begin{cases} \pm\sqrt{p} & \text{if } p \equiv 1 \pmod 4, \\ \pm\mathrm{i}\sqrt{p} & \text{if } p \equiv 3 \pmod 4. \end{cases}$$

(k) For $m = p$, use parts (d) and (j) to prove that:

$$\begin{aligned} a &= b \pm 1, & c &= d & &\text{when } p \equiv 1 \pmod 4; \\ a &= b, & c &= d \pm 1 & &\text{when } p \equiv 3 \pmod 4. \end{aligned}$$

(l) Use parts (e), (h) and (k) to deduce from $(*)$ that

$$G(p,1) = \mathrm{i}^{(p-1)^2/4}\sqrt{p} = \begin{cases} \sqrt{p} & \text{if } p \equiv 1 \pmod 4, \\ \mathrm{i}\sqrt{p} & \text{if } p \equiv 3 \pmod 4. \end{cases}$$

(m) Let p and q be two distinct odd primes. Use part (j) and Exercise 7.6(e) to prove that $G(pq,1)^2 = (-1)^{(p-1)/2+(q-1)/2}pq$. Deduce that

$$G(pq,1) = \begin{cases} \pm\sqrt{pq} & \text{if } p \equiv q \pmod 4, \\ \pm\mathrm{i}\sqrt{pq} & \text{if } p \not\equiv q \pmod 4. \end{cases}$$

(n) For $m = pq$, use parts (d) and (m) to prove that

$$\begin{aligned} a &= b \pm 1, & c &= d & &\text{when } p \equiv q \pmod 4, \\ a &= b, & c &= d \pm 1 & &\text{when } p \not\equiv q \pmod 4. \end{aligned}$$

(o) Use parts (e), (h) and (n) to deduce from $(*)$ that

$$G(pq,1) = \mathrm{i}^{(pq-1)^2/4}\sqrt{pq} = \begin{cases} \sqrt{pq} & \text{if } p \equiv q \pmod 4, \\ \mathrm{i}\sqrt{pq} & \text{if } p \not\equiv q \pmod 4. \end{cases}$$

7.8. A proof using polynomials and resultants

Exercise 7.9 (Sixth proof of the quadratic reciprocity law)**.** Given variables x, y and an integer $k \geq 1$, define the polynomials

$$\sigma_1 = x + y, \quad \sigma_2 = xy, \quad S_k = x^k + y^k.$$

Throughout this exercise, m and n denote odd positive integers and p an odd prime.

(a) Show that $S_n = \sigma_1 T_n$, where

$$T_n = x^{n-1} - x^{n-2}y + \cdots - xy^{n-2} + y^{n-1}.$$

Use the fundamental theorem for symmetric polynomials to deduce that

$$S_n = \sigma_1 f_n(\sigma_1^2, \sigma_2),$$

where $f_n(u,v)$ is a homogeneous polynomial with integer coefficients of degree $(n-1)/2$.

(b) Let $f_n(u,v)$ be the homogeneous polynomial from part (a), and define $f_n(z)$ by $f_n(z) = f_n(z,1)$. Show that $f_n(z)$ is a monic polynomial with integer coefficients of degree $(n-1)/2$ and $f_n(0) = (-1)^{(n-1)/2}n$. [Hint. Use $S_n(x,0)$ and $S_n(x,-x)$.]

(c) Use the binomial theorem to show that $S_p = \sigma_1^p + pG(x,y)$, where $G(x,y)$ is a homogeneous polynomial with integer coefficients of degree p which is also symmetric in x and y.

(d) Let $f_p(z)$ be the polynomial from part (b). Write

$$f_p(z) = z^{(p-1)/2} + \sum_{k=0}^{(p-3)/2} c_{p,k} z^k \qquad (c_{p,k} \in \mathbb{Z}).$$

Use part (c) and the fundamental theorem for symmetric polynomials to show that p divides all the coefficient $c_{p,k}$, $0 \leq k < (p-1)/2$. Deduce that

$$\operatorname{res}(f_n, f_p) \equiv \operatorname{res}\left(f_n, z^{(p-1)/2}\right) \pmod{p}.$$

(e) Use part (b), Exercise 4.23 and the properties of resultants to show that

$$\operatorname{res}\left(f_n, z^{(p-1)/2}\right) = n^{(p-1)/2}.$$

(f) Use parts (d) and (e) and Euler's criterion to show that if $(n, p) = 1$, then

$$\operatorname{res}(f_n, f_p) \equiv \left(\frac{n}{p}\right) \pmod{p}.$$

(g) Show that if $(m, n) = 1$, then 1 is the only common root of the polynomials $z^m - 1$ and $z^n - 1$. Deduce that $\gcd(z^m - 1, z^n - 1) = z - 1$.

(h) Let $(m, n) = 1$. Use part (g) and the analog of Theorem 1.10 for polynomials to show that there exist polynomials $P(z)$ and $Q(z)$ with integer coefficients such that $z - 1 = P(z)(z^m - 1) + Q(z)(z^n - 1)$.

(i) Let $(m, n) = 1$. Apply the result of part (h) with $z = -x/y$ to show that there exist an integer k and homogeneous polynomials with integer coefficients $H(x, y)$ and $K(x, y)$ such that $\sigma_1 y^k = H(x, y)S_m + K(x, y)S_n$.

(j) Let $(m, n) = 1$. Use the result of part (i) to derive the identities

$$\sigma_1 x^{k+1} y^k = x^{k+1}\left(H(x, y)S_m + K(x, y)S_n\right),$$
$$\sigma_1 x^k y^{k+1} = y^{k+1}\left(H(y, x)S_m + K(y, x)S_n\right).$$

Use these identities to show that there exist homogeneous symmetric polynomials with integer coefficients $M(x, y)$ and $N(x, y)$ such that

$$\sigma_1^2 \sigma_2^k = M(x, y)S_m + N(x, y)S_n.$$

(k) Use the fundamental theorem for symmetric polynomials to show that the polynomials $M(x, y)$ and $N(x, y)$ in part (j) can be rewritten as $\sigma_1 U(\sigma_1^2, \sigma_2)$ and $\sigma_1 V(\sigma_1^2, \sigma_2)$, respectively, where $U(u, v)$ and $V(u, v)$ are also homogeneous polynomials with integer coefficients. Deduce that

$$\sigma_2^k = U(\sigma_1^2, \sigma_2) f_m(\sigma_1^2, \sigma_2) + V(\sigma_1^2, \sigma_2) f_n(\sigma_1^2, \sigma_2).$$

(l) Use part (k) to show that if $(m, n) = 1$, then there exist polynomials $U(z)$ and $V(z)$ with integer coefficients such that $U(z)f_m(z) + V(z)f_n(z) = 1$.

(m) Use part (l) and Proposition 4.25 to show that if $(m, n) = 1$, then

$$\operatorname{res}(f_m, f_n) \operatorname{res}(f_m, V) = 1.$$

Since both resultants are integers, deduce that $\operatorname{res}(f_m, f_n) = \pm 1$. [Hint. Start from $\operatorname{res}(f_m, 1) = 1$.]

(n) Combine parts (f) and (m) to show that if $(n, p) = 1$, then $\left(\frac{n}{p}\right) = \operatorname{res}(f_n, f_p)$.

(o) Deduce the law of quadratic reciprocity from part (n) and Proposition 4.25.