Lecture notes: Algebra I, HKU, Fall 2024

About the lecture notes

This set of lecture notes is a modification of Professor JH Lu's "Lecture notes: Algebra I, HKU, Fall 2014". Lecture notes are used to facilitate lecturing, hence it is not sufficient to master the course contents without further reading and practices. Students should read the textbook/reference books **and** do enough exercises. Check out "Reading/Online materials" on moodle.

Contents

Chapter 1 — Preliminaries

1.1 Sets and Equivalence Relations

A collection of objects is called a **set**. A member x of a set S is also called an **element** of S and denoted as $x \in S$. If every element of a set T is also an element of a set S, we say that T is a **subset** of S and write $T \subset S$. If $T \subset S$ and $T \neq S$, we say that T is a **proper subset** of S. A set with no element is called the **empty set** and denoted by $\emptyset$, i.e. $\emptyset = \{\ \}$. Note: duplicate elements in a set are "ignored", thus as sets, $\{2,3,2\} = \{3,2\}$. Given a collection $\{S_\alpha : \alpha \in \mathcal{A}\}$ of subsets of a set S, one can form the intersection and union

$$\bigcap_{\alpha\in\mathcal{A}} S_\alpha = \{x \in S : x \in S_\alpha,\ \forall \alpha \in \mathcal{A}\}, \quad \bigcup_{\alpha\in\mathcal{A}} S_\alpha = \{x \in S : x \in S_\alpha \text{ for some } \alpha \in \mathcal{A}\}.$$

If $\{S_\alpha : \alpha \in \mathcal{A}\}$ is a collection of subsets of S such that $S_\alpha \cap S_\beta = \emptyset$ whenever $\alpha \neq \beta$ and if $S = \bigcup_{\alpha\in\mathcal{A}} S_\alpha$, we say that S is the **disjoint union** of the subsets $\{S_\alpha : \alpha \in \mathcal{A}\}$.

Here are some standard notation:

- $\mathbb{Z}$: the set of all integers;
- $\mathbb{N}$: the set of all positive integers also called natural numbers[†];
- $\mathbb{R}$: the set of all real numbers;
- $\mathbb{Q}$: the set of all rational numbers;
- $\mathbb{C}$: the set of all complex numbers;
- $\mathbb{Z}_{>0}$, $\mathbb{Q}_{>0}$, $\mathbb{R}_{>0}$: the sets of all positive numbers in $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$ respectively;
- $\mathbb{Z}^\times$, $\mathbb{Q}^\times$, $\mathbb{R}^\times$: the sets of all non-zero numbers in $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$ respectively[‡];
- $M(n, R)$: the set of all $n \times n$ matrices with entries in R, where $R = \mathbb{Z}, \mathbb{R}, \mathbb{Q}$, or $\mathbb{C}$;
- $GL(n, \mathbb{F})$: the set of all invertible $n \times n$ matrices with entries in $\mathbb{F}$, where $\mathbb{F} = \mathbb{R}, \mathbb{Q}$, or $\mathbb{C}$.

[†] In Lang's book "Undergraduate Algebra", the set of all positive integers is denoted by $\mathbb{Z}_{>0}$ and $\mathbb{N} = \{0\} \cup \mathbb{Z}_{>0}$, i.e. Prof Serge Lang viewed 0 as a natural number as well.

[‡] In Judson's book "Abstract Algebra", the notation $\mathbb{Z}^*, \mathbb{Q}^*, \mathbb{R}^*$ instead of $\mathbb{Z}^\times, \mathbb{Q}^\times, \mathbb{R}^\times$ is used.

Given two sets A and B. The **cartesian product** (or simply the **product**) $A \times B$ is the set

$$A \times B = \{(a,b) : \ a \in A, \, b \in B\}.$$

Let X be a set. By a **relation**, we mean a subset of $X \times X$. Suppose $R \subset X \times X$ and $x, y \in X$. If $(x,y) \in R$, then we say that x is **related** to y and write $x \sim y$.

By an **equivalence relation** in X we mean the relation satisfies the conditions:

1) **Reflexive property**: $(x,x) \in R$ for all $x \in X$;
2) **Symmetric property**: $(x,y) \in R$ implies $(y,x) \in R$;
3) **Transitive property**: $(x,y) \in R$ and $(y,z) \in R$ implies $(x,z) \in R$.

Remark. The three conditions can also be formulated as:
1) Reflexive: $x \sim x$ for all $x \in X$;
2) Symmetric: If $x \sim y$, then $y \sim x$;
3) Transitive: If $x \sim y$ and $y \sim z$, then $x \sim z$.

Let $\sim$ be an equivalence relation on X. For $a \in X$, the set $\{x \in X : \ x \sim a\}$ is called the **equivalence class of** a, denoted by $[a]$. We may also use $\bar{a}$ rather than $[a]$ for simplicity.

Every element in $[a]$ is called a **representative of the equivalence class**. Note:
(i) If $[a]$ and $[b]$ are two distinct equivalence classes, then $[a] \cap [b] = \emptyset$.
(ii) The set X is a disjoint union of $\{[a] : \ a \in X\}$. Thus the collection of distinct equivalence classes form a **partition of** X.

1.2 Integers

Here we review what we learned about integers in elementary school.

§ 1. Ordering on $\mathbb{Z}$ and the Well-ordering Principle

The ordering relation $\geq$ among integers is the usual ordering, e.g. $3 \geq 1$. A subset S of $\mathbb{Z}$ is said to be **bounded from below** if there exists an integer M such that $x \geq M$ for all $x \in S$. Similarly, S is said to be **bounded from above** if there exists an integer N such that $x \leq N$ for all $x \in S$.

Well-ordering Principle. *Every nonempty subset S of $\mathbb{Z}$ that is bounded from below has a unique smallest element, i.e., an element $a \in S$ such that $x \geq a$ for all $x \in S$.*

Remarks. (i) We take the well-ordering principle for granted (in elementary school as well as here), or more formally we take the well-ordering principle as an *axiom*.

(ii) By changing a set S to the set $-S \overset{\text{def}}{=} \{-x : x \in S\}$, one sees that *a subset S of $\mathbb{Z}$ that is bounded from above has a unique maximal element.*

§ 2. Congruences and the Division Algorithm

Notation. For two integers m, n with $n \neq 0$, we say that m *is divisible by* n or n *divides* m, and we write $n|m$, if $m = nq$ for some integer q. Otherwise we write $n \nmid m$.

For an integer $n > 1$, and two integers m_1 and m_2, we say that m_1 and m_2 are **congruent modulo** n, and we write $m_1 \equiv m_2 \pmod{n}$, if $n|(m_1 - m_2)$.

Theorem 1.2.1. (**Division Algorithm**) *Fix an integer $n > 0$. Every integer m is congruent modulo n to a unique integer r such that $0 \leq r < n$. Indeed, there exist unique integers q and r such that $m = qn + r$ with $0 \leq r < n$.*

Proof. Consider the set $Q = \{q \in \mathbb{Z} : qn \leq m\}$. Every $q \in Q$ satisfies $q \leq m/n$ because we are assuming that $n > 0$. Thus Q is bounded from above. By (Remark (ii) following) the well-ordering of $\mathbb{Z}$, Q has a maximal element q_0. Then $q_0 n \leq m$ but $(q_0 + 1)n > m$, so $0 \leq m - q_0 n < n$. Let $r = m - q_0 n$. Then $m \equiv r \pmod{n}$.

To show that such an r is unique, assume both $0 \leq r_1 < n$ and $0 \leq r_2 < n$ are such that $m = q_1 n + r_1$ and $m = q_2 n + r_2$. If $r_1 \neq r_2$, we may assume without loss of generality that $r_1 > r_2$. From $(q_2 - q_1)n = r_1 - r_2$, we deduce $q_1 < q_2$ so $(q_2 - q_1)n \geq n$ while $r_1 - r_2 < n$, which is a contradiction. Thus $r_1 = r_2$ and so $q_1 = q_2$. □

The integer r in Theorem 1.2.1 is called the **remainder of m when divided by** n.

It is easy to check that $\equiv \pmod{n}$ is an equivalence relation on the set $\mathbb{Z}$. The set of all equivalence classes of the congruence modulo n on $\mathbb{Z}$ is denoted by $\mathbb{Z}_n$. Theorem 1.2.1 thus says that for a given integer $n > 0$, there are precisely n equivalence classes of the congruence modulo n relation, and that each equivalence class has a unique representative which is an integer $0 \leq r < n$.

Example 1.2.2. For $n = 2$, there are two equivalence classes for the congruence modulo 2 relation: one formed by all even numbers and one formed by all the odd numbers. In other words $\mathbb{Z}_2 = \{[0], [1]\}$, the equivalence class $[0]$ is formed by all even numbers while $[1]$ consists of all odd numbers.

For $n = 3$, $\mathbb{Z}_3 = \{[0], [1], [2]\}$ where the three equivalence classes consisting, respectively, of all integers which have remainders 0, 1, or 2 when divided by 3. (Exercise: List out the elements in $\mathbb{Z}_{12}$.)

Lemma 1.2.3. *Let $n > 0$ be an integer and let $x, y, x', y' \in \mathbb{Z}$. If $x \equiv y \pmod{n}$ and $x' \equiv y' \pmod{n}$, then $x + x' \equiv y + y' \pmod{n}$, and $xx' \equiv yy' \pmod{n}$.*

Proof. Assume that $x \equiv y \pmod{n}$ and $x' \equiv y' \pmod{n}$. Then there exist $q, q' \in \mathbb{Z}$ such that $x - y = qn$ and $x' - y' = q'n$. Thus

$$x+x'-(y+y') = (q-q')n \quad \text{and} \quad xx'-yy' = (y+qn)(y'+q'n)-yy' = (yq'+y'q+qq'n)n.$$

Thus $x + x' \equiv y + y' \pmod{n}$, and $xx' \equiv yy' \pmod{n}$. □

Example 1.2.4. One has $89 \times 144 \equiv 5 \times 4 \pmod{7} \equiv 6 \pmod{7}$.

§ 3. Greatest Common Divisors

Definition 1.2.5. Given two non-zero integers m and n, the **greatest common divisor** of m and n is the largest positive integer d that divides both m and n, and is denoted by $\gcd(m, n)$ or simply (m, n). Also we may allow exactly one of m and n to be 0, in this case, say $n \neq 0$, we set $(0, n) = n$.

To give a conceptual treatment of greatest common divisors, and as a taster of what will come later in the course, we introduce the concept of *ideals* in $\mathbb{Z}$.

Definition 1.2.6. A subset J of $\mathbb{Z}$ is called an **ideal** if

1) $0 \in J$;
2) if $m, n \in J$, then $m + n \in J$;
3) if $n \in J$, then $nm \in J$ for all $m \in \mathbb{Z}$.

Example 1.2.7. The subset of $\mathbb{Z}$ consisting of only the zero element is an ideal called the *zero ideal.* The whole set $\mathbb{Z}$ is also an ideal.

Given any integer $m_1, m_2, \ldots, m_r$, it is easy to check that the set

$$J = \{x_1m_1 + x_2m_2 + \cdots + x_rm_r : \ x_1, x_2, \ldots, x_r \in \mathbb{Z}\}$$

is an ideal in $\mathbb{Z}$, which will be called **the ideal generated by** $m_1, m_2, \ldots, m_r$, and also denoted by $\langle m_1, m_2, \ldots, m_r\rangle$. In particular, $\langle a\rangle$ means $\{xa : x \in \mathbb{Z}\}$.

Example 1.2.8. Let $J = \langle 12, 16\rangle$. It is clear that every element in J is also divisible by 4, so $J \subset \langle 4\rangle$. On the other hand, as $4 = 16 - 12$, every integral multiple of 4 is an element in J. Thus $J = \langle 4\rangle$.

We now show that every ideal in $\mathbb{Z}$ is of the form $\langle d\rangle$ for a single integer $d \geq 0$. Clearly the zero ideal is generated by the element 0.

Theorem 1.2.9. *Let J be a non-zero ideal of $\mathbb{Z}$, and let d be the smallest positive integer in J. Then $J = \langle d\rangle$. We will call $d > 0$* **the generator for** *J.*

Proof. Note that if $n \in J$, then $-n \in J$, so J always contains a positive element. Let J_+ be the set of all positive elements in J. Then J_+ is a non-empty set and is bounded from below. By the well-ordered principle of the integers, J_+ has a smallest element d. Then $\langle d\rangle \subset J$. On the other hand, the Division Algorithm implies that every $n \in J$ is an integer multiple of d, i.e. $J \subset \langle d\rangle$. To see it, consider any $n \in J$. By Division Algorithm, $n = dq + r$ where $0 \leq r < d$. If $d \nmid n$, then $0 < r < d$ and $r = -qd + n \in J$. Hence $r \in J_+$ but $r < d$, contradicting to the minimality of d. In summary $J = \langle d\rangle$. □

Proposition 1.2.10. *Let $n, m \in \mathbb{Z}$, not both zero, and let $J = \langle m, n\rangle$. If $d > 0$ is the generator for J, then $d = \gcd(m, n)$.*

Proof. By definition, $J = \langle d\rangle$. Since $m, n \in J$, d divides both m and n, so d is a common divisor of m and n. We need to show that d is the largest common divisor of m and n. Suppose that e divides both m and n. Then $m = xe$ and $n = ye$ for some $x, y \in \mathbb{Z}$. On the other hand, since $d \in \langle m, n\rangle$, $d = am + bn$ for some $a, b \in \mathbb{Z}$. Then

$$d = aye + bxe = (ay + bx)e.$$

Thus $e|d$, so $e \leq d$. This shows that $d = \gcd(m, n)$. □

Theorem 1.2.9 and Proposition 1.2.10 give the following Corollary 1.2.11.

Corollary 1.2.11. *Let $m, n \in \mathbb{Z}$, not both zero, and let $d = \gcd(m, n)$. Then*

1) any linear combination $am + bn$ with $a, b \in \mathbb{Z}$ is a multiple of d;

2) there exist $a, b \in \mathbb{Z}$ such that $d = am + bn$;

3) any common divisor of m and n divide d (i.e. if $e \in \mathbb{Z}$ divides both m and n, then e divides d.)

Remarks. (i) Corollary 1.2.11 (1) tells when the equation $mx + ny = l$ (for given $m, n, l \in \mathbb{Z}$) can be solved with $x, y \in \mathbb{Z}$.

(ii) Corollary 1.2.11 (2) will imply that the solution set of $mx+ny = l$, *if solvable*, is $\{(x_0 + \frac{n}{d}t, y_0 - \frac{m}{d}t) : t \in \mathbb{Z}\}$ where (x_0, y_0) is a particular solution for $mx + ny = l$.

(iii) By definition, a common divisor of m and n must be less than or equal to the greatest common divisor $\gcd(m, n)$. Corollary 1.2.11 (3) tells an additional property.

(iv) *Euclidean Algorithm* can be applied to find one such pair of (a, b) in Corollary 1.2.11. (Google Euclidean Algorithm if you have forgotten.)

Definition 1.2.12. Two non-zero integers m and n are said to be **relatively prime** if $(m, n) = 1$.

Remark. By Corollary 1.2.11 we obtain: let $m, n \in \mathbb{Z}$,

$$(m, n) = 1 \text{ if and only if } 1 = am + bn \text{ for some } a, b \in \mathbb{Z}.$$

Theorem 1.2.13. (**The Chinese remainder theorem**) *Let m, n be positive integers such that $(m, n) = 1$. Then for any integers a, b, there exists $x \in \mathbb{Z}$ such that*

$$x \equiv a \pmod m \quad \textit{and} \quad x \equiv b \pmod n.$$

Proof. Since $(m, n) = 1$, the ideal generated by m and n is the whole $\mathbb{Z}$. Thus there exist $p, q \in \mathbb{Z}$ such that $a - b = pm + qn$. Then $x = a - pm = b + qn$ is as required. □

Example 1.2.14. Solve

$$x \equiv 2 \pmod{18} \quad \text{and} \quad x \equiv 5 \pmod{7}.$$

We use the method in Theorem 1.2.13. By inspection we have $1 = -35 + 36$, so

$$5 - 2 = -3 \times 35 + 3 \times 36,$$

so $x = 2 + 3 \times 36 = 5 + 3 \times 35 = 110$ is a solution. Every solution is of the form $x = 110 + (7 \times 18)m$ for an integer m (why?), and these are all the solutions.

§ 4. Unique Factorization

We define a **prime number** to be an integer $p \geq 2$ such that p is only divisible by ± 1 and $\pm p$. The first few prime numbers are $2,\ 3,\ 5,\ 7,\ 11,\ 13, \cdots$.

Lemma 1.2.15. *Let p be a prime number and let m, n be non-zero integers such that $p|mn$. Then $p|m$ or $p|n$.*

Proof. Suppose that p does not divide m. Then $(p, m) = 1$. Thus there exist integers a and b such that

$$1 = ap + bm$$

from which one gets $n = anp + bmn$. Since $p|mn$, we get $p|n$. □

Theorem 1.2.16. (**Fundamental Theorem of Arithmetic**) *For any non-zero integer n, there are unique $\varepsilon \in \{\pm 1\}$, (unique) prime numbers $p_1 < p_2 < \cdots < p_r$ and positive integers $m_1, m_2, \ldots, m_r$ such that*

$$n = \varepsilon p_1^{m_1} p_2^{m_2} \cdots p_r^{m_r}.$$

Proof. Skipped. □

It is easy to find $\gcd(m, n)$ using the unique factorization, for example:

$$\gcd(135, 320) = \gcd(3^2 \cdot 5,\ 2^6 \cdot 5) = 5.$$

Chapter 2 — Groups: Basic Concepts

2.1 Definition and Basic Examples

Definition 2.1.1. A **group** $(G, *)$ is a set G, together with a map

$$* : G \times G \longrightarrow G, \quad (x, y) \longmapsto x * y,$$

such that the following axioms are satisfied:

1) (**Associativity**): $(x * y) * z = x * (y * z)$;
2) There exists $e \in G$ such that $e * x = x * e = x$ for all $x \in G$;
3) For each $a \in G$, there exists $b \in G$ such that $a * b = b * a = e$.

For simplicity, we write xy for $x * y$. If $xy = yx$ for all $x, y \in G$, we say that the group G is **abelian**[†], and in this case we normally denote the group operation by $x + y$ instead of xy.

Lemma 2.1.2. *Let G be a group.*
1) The element e in Definition 2.1.1 is unique.
2) For any a, the element $b \in G$ such that $ab = ba = e$ is also unique.

Proof. If e, e' both have the required property, then $e = ee' = e'$. Let $a, b \in G$ be such that $ab = ba = e$. If $c \in G$ also satisfies $ac = ca = e$, then by the associativity, $c = ec = (ba)c = b(ac) = be = b$. □

Due to Lemma 2.1.2, the unique element e is called **the identity element** of the group, and the element b is called **the inverse of** a and denoted by a^{-1}. When the group is abelian, we denote e by 0 and a^{-1} by $-a$.

Definition 2.1.3. By a **finite group** we mean a group G that has finitely many elements, and we use $|G|$ to denote the number of elements in G and call it the **order of** G. If G is not a finite group, we say it is an **infinite group** and we write $|G| = \infty$.

For a finite group G, we may describe the map $*$ by a table, called the **Cayley**[‡] **table**. The table is read row by column.

[†] Abelian group is named after the Norwegian mathematician Abel (1802-1829).

[‡] Cayley (1821-1895) was a British mathematician.

Example 2.1.4. Here are some basic examples of groups:

1. **Trivial group**: The group with only one element: $G = \{e\}$.

2. **Groups of order 2**: Let $\mathrm{G} = \{\mathrm{e}, x\}$ be a group where $x \neq \mathrm{e}$, and its Cayley table is

$*$	e	x
e	e	x
x	x	e

   (from which we find, e.g., $\mathrm{e}x = x$), G is an abelian group of order 2.

   Remark. If G is a group with two elements a, e $(a \neq e)$ where e is the identity. Then the group operation *must be* given by

   $$ea = ae = a, \qquad a^2 = e.$$

   G has the *same group structure as* the above G.[§] Note: $a^{-1} = a$ and G is abelian.

3. The sets $\mathbb{R}, \mathbb{C}, \mathbb{Q}, \mathbb{Z}$ with addition as the group operation, also called the **additive group** $\mathbb{R}, \mathbb{C}, \mathbb{Q}$, or $\mathbb{Z}$. They are infinite abelian groups.

4. The sets $\mathbb{R}^\times, \mathbb{C}^\times, \mathbb{Q}^\times$ with multiplication as the group operation, also called the **multiplicative groups** $\mathbb{R}^\times, \mathbb{C}^\times$ or $\mathbb{Q}^\times$. They are again infinite abelian groups.

5. Let $\mathbb{F} = \mathbb{R}, \mathbb{C}$, or $\mathbb{Q}$, and let $G = GL(n, \mathbb{F})$, the set of all invertible $n \times n$ matrices with entries in $\mathbb{F}$. Then under matrix multiplication, G becomes a group. These are *infinite nonabelian* groups when $n \geq 2$, and they are called the **general linear groups over** $\mathbb{F}$.

6. A vector space V over $\mathbb{R}$ or $\mathbb{C}$ is an abelian group under the addition of vectors in V.

More examples can be found in reference books.

Remark 2.1.5. In the definition of a group, one must specify the set as well as the group operation. For example, the set $\mathbb{R}_{>0}$ of all positive real numbers is a group under multiplication but *not* under addition.

[§] *Same group structure* means the two groups G and G become the same after a suitable renaming of the group elements and the group operation. Such an identification (of group elements and operation) will be mathematically performed through a *group isomorphism* which will be discussed later.

2.2 Three Examples

§ 1. The group $\mathbb{Z}_n$

Consider the set $\mathbb{Z}_n = \{[0], [1], [2], \cdots, [n-1]\}$, and whenever no confusion arises, we shall omit the square bracket, i.e. use a to denote $[a]$, for simplicity. Thus we write $\mathbb{Z}_n = \{0, 1, 2, \cdots, n-1\}$. Next consider the operation defined by the map

$$\mathbb{Z}_n \times \mathbb{Z}_n \longrightarrow \mathbb{Z}_n, \quad (k,\, l) \longmapsto k + l \overset{\text{def}}{=} r \ \text{ if } \ 0 \leq r \leq n-1 \ \text{ and } \ k + l \equiv r \ (\text{mod}\ n).$$

Claim: The set $\mathbb{Z}_n$ with the operation $+$ is an abelian group of order n.

Proof. To check associativity, let $0 \leq k, l, h \leq n-1$. If $0 \leq s \leq n-1$ and $k + l \equiv s \ (\text{mod}\ n)$, then $[k] + [l] = [s]$ and thus

$$([k] + [l]) + [h] = [r] \quad \text{where } 0 \leq r \leq n-1 \text{ and } r \equiv s + h \equiv (k+l) + h \ (\text{mod}\ n).$$

Similarly, $[k] + ([l] + [h]) = [r]$ because $k + (l + h) \equiv r \ (\text{mod}\ n)$.

It is clear that for every $0 \leq k \leq n-1$,

$$[0] + [k] = [k] + [0] = [k] \quad \text{ and } \quad [k] + [n-k] = [n-k] + [k] = [0]. \, ^{\dagger}$$

Thus $\mathbb{Z}_n$ with the above operation is a group. It is abelian and $|\mathbb{Z}_n| = n$. □

§ 2. The group $\mathbb{Z}_n^\times$

Consider the set $\mathbb{Z}_n^\times = \{[a] \in \mathbb{Z}_n : \ (a, n) = 1\}$ and the operation

$$\mathbb{Z}_n^\times \times \mathbb{Z}_n^\times \longrightarrow \mathbb{Z}_n^\times, \quad (k,\, l) \longmapsto k \cdot l \overset{\text{def}}{=} r \ \text{ if } \ 0 \leq r \leq n-1 \ \text{ and } \ kl \equiv r \ (\text{mod}\ n).$$

Indeed, it is necessary to check that this operation is well-defined (because $(r, n) = 1$ is required in order that $[r] \in \mathbb{Z}_n^\times$). [‡]

Claim: The set $\mathbb{Z}_n^\times$ with the operation $\cdot$ is an abelian group.

The group $\mathbb{Z}_n^\times$ is called the **group of units of $\mathbb{Z}_n$**. [§]

Proof. We leave as an exercise to check associativity and to check 1 is the identity.

Let $a \in \mathbb{Z}_n^\times$. The existence of the inverse of a is justified with the following lemma whose proof is left as exercise. □

Lemma 2.2.1. *Let $n \in \mathbb{N}$ and $a \in \mathbb{Z}$. If $(a, n) = 1$, there exists $0 \leq b < n$ such that $ab \equiv 1 \ (\text{mod}\ n)$ and $(b, n) = 1$.*

[†] The class $[n]$ equals $[0]$.

[‡] Here is the justification: Suppose $d = (n, r) > 1$. From $kl \equiv r \ (\text{mod}\ n)$, we have $kl = nq + r$ and thus $d | kl$. However, by Proposition 1.2.10, $(k, n) = (l, n) = 1$ implies $(kl, n) = 1$ and so $1 = klx + ny$ for some integers x, y. This leads to $d | 1$ which is a contradiction.

[§] In Judson's book, the notation $U(n)$ instead of $\mathbb{Z}_n^\times$ is used.

§ 3. Product groups

Definition 2.2.2. Given two groups G_1 and G_2, the product set $G_1 \times G_2$ with the group operation

$$(g_1,\, g_2)(h_1,\, h_2) = (g_1 h_1,\, g_2 h_2), \qquad g_1, h_1 \in G_1,\; g_2, h_2 \in G_2,$$

is called the **product group** (or **direct product** [†]) of G_1 and G_2.

Remark. It is necessary to check $G_1 \times G_2$ in Definition 2.2.2 is indeed a group, which is an exercise.

Example 2.2.3. The product group $\mathbb{Z}_2 \times \mathbb{Z}_2$ is abelian and has 4 elements:

$$\mathbb{Z}_2 \times \mathbb{Z}_2 = \{(0,0),\ (0,1),\ (1,0),\ (1,1)\},$$

(recall 0 means [0], 1 means [1]) and the group operation is given by $(0,0) + a = a$ for all $a \in \mathbb{Z}_2 \times \mathbb{Z}_2$ and

$+$	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
$(0,0)$	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
$(0,1)$	$(0,1)$	$(0,0)$	$(1,1)$	$(1,0)$
$(1,0)$	$(1,0)$	$(1,1)$	$(0,0)$	$(0,1)$
$(1,1)$	$(1,1)$	$(1,0)$	$(0,1)$	$(0,0)$

Remarks. (i) Forming the products of groups is one way of obtaining more examples of groups, but it is not a very interesting one.

(ii) We have mentioned that groups of two elements have the same group structure. However groups of 4 elements may have different (group) structure, e.g. $\mathbb{Z}_2 \times \mathbb{Z}_2$ and $\mathbb{Z}_4$ have different structures, which is explained in the next section.

2.3 Isomorphisms

Definition 2.3.1. Let G_1 and G_2 be two groups.

1) A **homomorphism** from G_1 to G_2 is a map $\phi : G_1 \to G_2$ such that
$$\phi(gh) = \phi(g)\phi(h), \qquad \forall\ g, h \in G_1.$$

2) A bijective homomorphism from G_1 to G_2 is called an **isomorphism** from G_1 to G_2.

3) The two groups G_1 and G_2 are said to be **isomorphic** if there is a group isomorphism from G_1 to G_2, and in this case we write $G_1 \cong G_2$.

[†] In Judson's book, it is called the *external direct product* of G_1 and G_2.

Example 2.3.2. (a) Let G_1 and G_2 be groups of order 2. Write $G_1 = \{\mathrm{e}, x\}$ and $G_2 = \{e, a\}$ where $x^2 = \mathrm{e}$ and $a^2 = e$ by Example 2.1.4. Define

$$\phi : G_1 \to G_2,\ \phi(\mathrm{e}) = e,\ \phi(x) = a.$$

Clearly ϕ is a *bijective* function. A brute-force checking gives

(g,h)	(e,e)	(e,x)	(x,e)	(x,x)
$\phi(gh)$	$\phi(\mathrm{e}) = e$	$\phi(x) = a$	$\phi(x) = a$	$\phi(\mathrm{e}) = e$
$\phi(g)\phi(h)$	$ee = e$	$ea = a$	$ae = a$	$aa = e$

i.e. $\phi(gh) = \phi(g)\phi(h)\ \forall\ g, h \in G_1$, so ϕ is a *homomorphism.* Thus ϕ is a *bijective homomorphism*, i.e. an isomorphism. Thus $G_1 \cong G_2$.

(b) $\mathbb{Z}_4 \ncong \mathbb{Z}_2 \times \mathbb{Z}_2$ because we *never find* an isomorphism from $\mathbb{Z}_4$ to $\mathbb{Z}_2 \times \mathbb{Z}_2$.

This can be argued by contradiction. Suppose $\phi : \mathbb{Z}_4 \to \mathbb{Z}_2 \times \mathbb{Z}_2$ is an isomorphism. Then $\phi([1]) = (x, y)$ for some $(x, y) \in \mathbb{Z}_2 \times \mathbb{Z}_2$. In view of the table in Example 2.2.3, $(x, y) + (x, y) = (0, 0)$ *for all* $(x, y) \in \mathbb{Z}_2 \times \mathbb{Z}_2$. This implies

$$\phi([2]) = \phi([1] + [1]) = \phi([1]) + \phi([1]) = (x, y) + (x, y) = (0, 0).$$

As ϕ is injective and $\phi([0]) = (0, 0)$ (see Exercise 2.3.3 (1) below), we get $[2] = [0]$ in $\mathbb{Z}_4$, which is a contradiction.

Exercise 2.3.3. (a) Let G_1 and G_2 be groups, and $\phi : G_1 \to G_2$ be a homomorphism. Abusing the notation, e denotes the identity elements in *both* G_1 and G_2. Prove the following statements:

1) $\phi(e) = e$ and $\phi(g^{-1}) = \phi(g)^{-1}$ for all $g \in G_1$.
2) ϕ is injective if and only if ”$\phi(x) = e$ implies $x = e$”.
3) If $\phi : G_1 \to G_2$ is a group isomorphism, so is $\phi^{-1} : G_2 \to G_1$.
4) If two groups G_1 and G_2 are isomorphic, then $|G_1| = |G_2|$ (which may be infinity).

(b) Show that being isomorphic gives an equivalence relation on the set of groups.

2.4 Subgroups

Definition 2.4.1. Given a group $(G, *)$ and let $H \subset G$. If $(H, *)$ is a group, then H is called a **subgroup of** G.

Proposition 2.4.2. *A subset of H of a group G is a group if and only if H satisfies*

1. *$e \in H$ where e is the identity of G;*
2. *xy and x^{-1} are in H whenever $x, y \in H$.*

Proof. "If" part: Exercise.

"Only if" part: Let $(G, *)$ be a group and $H \subset G$. If $(H, *)$ is also a group, then H has its own identity element e_H. Being the identity in H, we have $e_He_H = e_H$. On the other hand, $e_H \in G$ and thus has the inverse $e_H^{-1} \in G$, satisfying $e_He_H^{-1} = e$. Also $ee_H = e_H$, hence $ee_H = e_He_H$. Multiplying both sides by e_H^{-1}, we get

$$e = e(e_He_H^{-1}) = (ee_H)e_H^{-1} = (e_He_H)e_H^{-1} = e_H(e_He_H^{-1}) = e_H.$$

This proves $e \in H$. As H is a group, by definition, $xy \in H$ whenever $x, y \in H$. Let $x \in H$. Write x' for the inverse of x in H, then $x'x = xx' = e_H = e$. This means $x' \in G$ and x' satisfies the condition to be the inverse of x in G. By the uniqueness, we conclude $x^{-1} = x' \in H$. □

Remarks. (i) Due to Proposition 2.4.2, one may define a subgroup of a group G to be a subset H of G satisfying the conditions (1) and (2). (ii) A subgroup of a group G is a group on its own. This is a very interesting way of obtaining new groups from known ones.

Example 2.4.3. 1) Every group G has the trivial subgroup $\{e\}$ and the whole group G as subgroups. The additive group $\mathbb{Q}$ is a subgroup of the additive group $\mathbb{R}$, which is in turn a subgroup of the additive group $\mathbb{C}$.

2) The product group $G_1 \times G_2$ has $G_1 \times \{e_2\}$ and $\{e_1\} \times G_2$ as subgroups.

3) The set $SL(n, \mathbb{R})$ consisting of all $n \times n$ real matrices of determinant 1 is a subgroup of $GL(n, \mathbb{R})$, and is called a **special linear group over** $\mathbb{R}$.

2.5 Cyclic subgroups

Let G be a group and $a \in G$. We define $a^0 = e$ (the identity of G), and for all $n \in \mathbb{N}$, write $a^n = aa \cdots a$ (n times) and $a^{-n} = a^{-1}a^{-1} \cdots a^{-1}$ (n times). Note that for any $m, n \in \mathbb{Z}$, (i) $a^m a^n = a^{m+n}$ and (ii) $(a^n)^{-1} = a^{-n}$. (Exercise: Prove (i) and (ii).)

Theorem 2.5.1. *Let G be a group and $a \in G$. Then (i) the set $\langle a \rangle \overset{\text{def}}{=} \{a^k : k \in \mathbb{Z}\}$ is a subgroup of G, and (ii) if $H \subset G$ is a subgroup and $a \in H$, then $\langle a \rangle \subset H$.*

Proof. For (i), it suffices to check Conditions (1) and (2) in Proposition 2.4.2 are satisfied, which is an easy exercise. (ii) Suppose H is a subgroup of G and $a \in H$. Observe that the definition of $\langle a \rangle$ does not depend on the group carrying a (but only on the group operation). Thus $\langle a \rangle$ is a subgroup of H by (i), so $\langle a \rangle \subset H$. □

Definition 2.5.2. For $a \in G$, we call $\langle a \rangle$ the **cyclic subgroup generated by** a. A group G is said to be **cyclic** if there exists $a \in G$ such that $G = \langle a \rangle$, and in this case we say that a is **a generator of** G.

Example 2.5.3. (1) The cyclic subgroup $\langle n\rangle$ of the additive group $\mathbb{Z}$ equals $\{nm : m \in \mathbb{Z}\}$, which is often written as $n\mathbb{Z}$.[†] Note that $n\mathbb{Z}$ is an infinite group.

(2) The cyclic subgroup $\langle 2\rangle$ of the multiplicative group $\mathbb{C}^\times$ is

$$\langle 2\rangle = \left\{\cdots \frac{1}{8},\ \frac{1}{4},\ \frac{1}{2},\ 1,\ 2,\ 4,\ 8, \cdots\right\}.$$

Note that $\langle 2\rangle$ is an infinite group. On the other hand, let $n \geq 1$ be an integer and let $\omega_n = e^{\frac{2\pi i}{n}} \in \mathbb{C}^\times$. Then the cyclic subgroup $\langle \omega_n\rangle$ of the multiplicative group $\mathbb{C}^\times$ consists of all the n-th roots of unity:

$$\langle \omega_n\rangle = \{1, \omega_n,\ \omega_n^2,\ \ldots,\ \omega_n^{n-1}\}$$

and is thus a group of order n.

(3) The element $a = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in GL(2, \mathbb{R})$ generates a subgroup of 4 elements:

$$\langle a\rangle = \left\{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right\} = \{e, a, a^2, a^3 = a^{-1}\}.$$

Example 2.5.4. For any integer $n \geq 1$, the group $\mathbb{Z}_n$ is a cyclic group of order n with the 1 ($= [1]$) as a generator. However $\mathbb{Z}_n^\times$ *may not* be cyclic, e.g., $\mathbb{Z}_3^\times$ is cyclic but $\mathbb{Z}_8^\times$ is *not* cyclic.

Question 2.5.5. Let G be a group and let $a \in G$. How to tell whether $\langle a\rangle$ is a finite group or not?

§ 1. The order of an element

Definition 2.5.6. Let G be a group and $a \in G$. If there exists an integer n such that $a^n = e$, we say that a is **of finite order**, and the smallest *positive* integer n such that $a^n = e$ is called the **order** (or **period**) **of** a and is denoted by $\operatorname{ord}(a)$.[‡] If there is no non-zero integer n such that $a^n = e$, we say that a is **of infinite order**.

Remark 2.5.7. The order of an element in G is not to be confused with the order of the group G, which is defined to be the number of elements in G.

Note that even if the group has infinite order, its elements may have finite orders.

Example 2.5.8. In the multiplicative group $\mathbb{C}^\times$, the element $i = \sqrt{-1}$ has order 4, while the element 2 has infinite order.

Proposition 2.5.9. *Let G be a group and let $a \in G$.*

1) If the order of a is infinite, then $a^i \neq a^j$ for any $i, j \in \mathbb{Z}$ with $i \neq j$, and the cyclic group $\langle a\rangle$ is an infinite abelian group;

2) If the order of a is an integer $n \geq 1$, then $\langle a\rangle = \{e, a, a^2, \cdots, a^{n-1}\}$ and the order of $\langle a\rangle$ is n. i.e. The order of $\langle a\rangle = \operatorname{ord}(a)$.

[†] We have used $\langle n\rangle$ to denote the ideal generated by n in $\mathbb{Z}$, which does not cause trouble because an ideal in $\mathbb{Z}$ is a subgroup of the additive group $\mathbb{Z}$.

[‡] In Judson's book, the notation $|a|$ instead of $\operatorname{ord}(a)$ is used for the order of a.

Proof. 1) Assume that a has infinite order and suppose that $a^i = a^j$ for some $i, j \in \mathbb{Z}$. Assume without loss of generality that $j \geq i$. Then $a^{j-i} = e$. Since a has infinite order, we must have $i = j$.

2) By writing every $k = r \pmod{n}$ for any integer k where $0 \leq r \leq n-1$, it is clear that $\langle a \rangle \subset \{e, a, a^2, \cdots, a^{n-1}\}$ and thus $\langle a \rangle = \{e, a, a^2, \cdots, a^{n-1}\}$. If $a^i = a^j$ for some $0 \leq i < j \leq n-1$, then $a^{j-i} = e$ with $0 < j - i < n$, which contradicts to $\mathrm{ord}(a) = n$. Thus the elements in $\{e, a, a^2, \cdots, a^{n-1}\}$ are pairwise distinct. □

§ 2. Subgroups of a cyclic group

Theorem 2.5.10. *Let $G = \langle a \rangle$ be a cyclic group of order n with generator a. Then any subgroup H of G is cyclic and can be expressed as $H = \langle a^m \rangle$ for some $m \in \{0\} \cup \mathbb{N}$. The order of H $(= \langle a^m \rangle)$ is $n/(m, n)$, hence the elements a^m with $(m, n) = 1$ are generators of G.*

Proof. If $H = \{e\}$, then it is cyclic and we take $m = 0$. Otherwise H contains a non-identity element in $\langle a \rangle$. Thus the set $Q = \{k \in \mathbb{N} : \ a^k \in H\}$ is nonempty, and by the well-ordering principle, has the smallest element, say m. We claim that a^m is a generator for H. Let $b \in H \subset G$. Then $b = a^k$ for some $k \in \mathbb{Z}$. By Division Algorithm, $k = mq + r$ where $0 \leq r < m$. Thus $a^k = (a^m)^q a^r$ and so $a^r \in H$ from $a^k, a^m \in H$. This implies $r = 0$ from the minimality of m in Q.

Exercise: Show the order of $\langle a^m \rangle$ is $n/(m, n)$. □

Remark. Does the argument in this proof apply to infinite cyclic group? I.e. *Is every subgroup of an infinite cyclic group cyclic?.*

§ 3. Classification of cyclic groups

Theorem 2.5.11. *Let $G = \langle a \rangle$ be a cyclic group with generator a.*

1) If $|G| = \infty$ (equivalently, the order of a is infinite), then $a^i \neq a^j$ for $i \neq j$, and the map

$$\phi : \quad G \longrightarrow \mathbb{Z}, \ \ a^k \longmapsto k, \qquad k \in \mathbb{Z}$$

is a group isomorphism. i.e. $G \cong \mathbb{Z}$.

2) If $|G| = n < \infty$ (equivalently, $\mathrm{ord}(a) = n$), then $G = \{e, a, a^2, \cdots, a^{n-1}\}$, and the map

$$\phi : \quad G \longrightarrow \mathbb{Z}_n, \ \ a^k \longmapsto [k], \qquad 0 \leq k \leq n-1,$$

is a group isomorphism. i.e. $G \cong \mathbb{Z}_n$.

Proof. It is straightforward to check that ϕ in both cases are homomorphisms (fill in the details as an exercise). It is also clear that ϕ is surjective in both cases. Injectivity of ϕ was proved in Proposition 2.5.9. Thus ϕ is an isomorphism. □

Corollary 2.5.12. *For any given $n \in \mathbb{N}$ or $n = \infty$, there is exactly one cyclic group of order n up to isomorphism.*

2.6 Lagrange's theorem

Theorem 2.5.10 implies that for a finite cyclic group G, the order of any its subgroup H divides the order of G, i.e. $|H|$ divides $|G|$. It turns out that this divisibility phenomenon holds for *all* finite groups.

Given a subgroup H of G and $a \in G$, the **left H-coset** of a is the subset

$$aH = \{ah : h \in H\}$$

of G.[‡] Define a relation $\sim_H$ on G by $a \sim_H b$ if $a \in bH$ for $a, b \in G$.

Lemma 2.6.1.[§] *Assume that H is subgroup of G. Then $\sim_H$ is an equivalence relation, and G is the disjoint union of all of its different left H-cosets. Moreover, for every $a \in G$, the map*

$$H \longrightarrow aH : \quad h \longmapsto ah, \quad h \in H, \tag{2.1}$$

is bijective. Consequently, if H is a finite subgroup, then all the left H-cosets have the same number of elements, namely $|H|$, the order of H.

Proof. It is clear that $a \sim_H a$ for all $a \in G$. If $a, b \in G$ are such that $a \in bH$, then $a = bh$ for some $h \in H$, so $b = ah^{-1}$. As $h^{-1} \in H$, $b \in aH$ so $b \sim_H a$. Finally, if $a = bh_1$ and $b = ch_2$, where $a, b, c \in G$ and $h_1, h_2 \in H$, then $a = ch_2h_1 \in cH$. Thus we have shown that the relation $\sim_H$ is reflexive, symmetric and transitive, i.e. $\sim_H$ is an equivalence relation. By definition, the equivalence classes in G of the equivalence relation $\sim_H$ are precisely the left H-cosets. Different left H-cosets are thus disjoint, and G is the disjoint union of all of its different left H-cosets.

If $a \in G$ and if $ah_1 = ah_2$ for $h_1, h_2 \in H$, then $h_1 = h_2$. Thus the map in (2.1) is injective. It is also surjective by the definition of aH, so it is bijective. Consequently, if H is a finite subgroup, then each left H-coset aH has exactly the same number of elements as there are in H. □

Theorem 2.6.2. *Let G be a finite group and H a subgroup. Then,*

$$|G| = k|H|$$

where k is the number of distinct left H-cosets in G.

Proof. Since G has finitely many elements, there are only finitely many left H-cosets. Let $C_1, \ldots, C_k$ be all the different left H-cosets. Then

$$G = C_1 \cup C_2 \cup \cdots \cup C_k$$

is a disjoint union and all the C_j's have the same number of elements, namely $|H|$. Thus $|G| = k|H|$. □

[‡] In Judson's book (see Chapter 6), the terminology for the left H-coset of a is the *left coset of H with representative a.*

[§] The proof here provides an alternative treatment to get the results ”Lemma 6.3 (1) ⇔ (4)”, ”Theorem 6.4” & ”Proposition 6.9” in Judson's book.

Corollary 2.6.3. (**Lagrange's**[†] **theorem**) *The order of a subgroup in a finite group G divides the order of G.*

Corollary 2.6.4. *The order of every element in a finite group G divides the order of G.*

Proof. For each $a \in G$, the integer $\mathrm{ord}(a)$, being the order of the subgroup $\langle a \rangle$ of G, divides the order of G. □

Theorem 2.6.5. *If p is a prime number, then every group of order p is cyclic and is isomorphic to $\mathbb{Z}_p$.*

Proof. Let G be a group of order p, where $p \geq 2$ is a prime number. Take any $a \in G$ and $a \neq e$. Then $\mathrm{ord}(a)$ divides p, so $\mathrm{ord}(a) = p$. As the cyclic subgroup $\langle a \rangle$ of G has the same number of elements as in G, we must have $\langle a \rangle = G$, i.e., G is cyclic. By Theorem 2.5.11, $G \cong \mathbb{Z}_p$. □

Definition 2.6.6. If G is a finite group and $H \subset G$ a subgroup, the integer $|G|/|H|$ is called the **index of H in G** and is also denoted by $[G : H]$.

Remark 2.6.7. For a group G and a subgroup H of G, define the **right H-coset** of $a \in G$ as the subset $Ha = \{ha : h \in H\}$ of G. All the arguments above apply to the right H-cosets. In particular, when G is finite, the index of H in G is also equal to the number of right H-cosets in G.

2.7 Classification problems

Corollary 2.5.12 classifies all cyclic groups up to isomorphisms. More generally one asks the following, which is a central problem in group theory:

Problem 2.7.1. *Classify, up to isomorphisms, all groups of a given order $n < \infty$.*

As this is a hard problem, it is a good idea to restrict ourselves to subclasses of groups: for example, we have classified all cyclic groups of a given order. Another special class of groups is the so-called *simple groups* which we will define later. You may search "*simple finite groups*" on the Internet or read books on finite groups to get some idea of what has been done in the classification of such groups.

In classifying groups up to isomorphisms, it is very useful to know some properties that are preserved under isomorphisms because they give necessary conditions for two groups to be isomorphic. For example, two groups that are isomorphic necessarily have the same order. We now give another such property.

Lemma 2.7.2. *If $\phi : G_1 \to G_2$ is an injective homomorphism, then for any $a \in G_1$, the order of a in G_1 is the same as the order of $\phi(a)$ in G_2 (both could be infinity).*

[†] Lagrange (1736-1813) was an Italian mathematician.

Proof. Let $\phi : G_1 \to G_2$ be an injective homomorphism and let $a \in G_1$. Let $n \in \mathbb{Z}$. If $a^n = e$, then $\phi(a)^n = \phi(a^n) = \phi(e) = e$. Conversely, if $\phi(a)^n = e$, then $\phi(a^n) = e$, which implies that $a^n = e$ because ϕ is injective. Thus $a^n = e$ if and only if $\phi(a)^n = e$. We now look at two possible cases:

Case 1: a has infinite order. In this case if $n \in \mathbb{Z}$ is such that $\phi(a)^n = e$, then $a^n = e$, so $n = 0$. Thus $\phi(a)$ also has infinite order.

Case 2: a has finite order $n \geq 1$. In this case $\phi(a)^n = e$, so $\phi(a)$ has finite order. Let $n' = \mathrm{ord}(\phi(a))$. Then $n' \leq n$ by the definition of $\mathrm{ord}(\phi(a))$. As $a^{n'} = e$ from the injectivity of ϕ, we have $n \leq n'$, so $n = n'$.

In both cases, we have proved that $\mathrm{ord}(a) = \mathrm{ord}(\phi(a))$. □

Corollary 2.7.3. *If G_1 and G_2 are two groups such that there exists $a \in G_1$ such that no element in G_2 has the same order as that of a, then G_1 and G_2 are not isomorphic.*

Proof. If G_1 were isomorphic to G_2, there would be an isomorphism $\phi : G_1 \to G_2$, and $\phi(a) \in G_2$ would have the same order of that of a, which is a contradiction. Thus G_1 cannot be isomorphic to G_2. □

Example 2.7.4. $\mathbb{Z}_2 \times \mathbb{Z}_2$ is not isomorphic to $\mathbb{Z}_4$ because the element $[1]$ in $\mathbb{Z}_4$ has order 4, while all the elements in $\mathbb{Z}_2 \times \mathbb{Z}_2$ have orders 1 or 2.

§ 1. Classification of groups of order up to 5

By Theorem 2.6.5, every group of order $n = 1, 2, 3$ or 5 is isomorphic to $\mathbb{Z}_n$. It remains to classify groups of order 4.

We have $\mathbb{Z}_4$ and $\mathbb{Z}_2 \times \mathbb{Z}_2$ which are groups of order 4, and we have seen in Example 2.7.4 that these two groups are not isomorphic. The question now is whether every group of order 4 is isomorphic to one of these two.

Lemma 2.7.5. *A group of order 4 is either isomorphic to $\mathbb{Z}_4$ or to $\mathbb{Z}_2 \times \mathbb{Z}_2$.*

Proof. Let G be a group of order 4 and let e be the identity element.

Case 1: G has an element a of order 4. In this case $G = \langle a \rangle$ is cyclic, so $G \cong \mathbb{Z}_4$ by Theorem 2.5.11.

Case 2: No element of G has order 4. By Lagrange's theorem, all elements in G that are not e have order 2. Let e, a, b, c be the 4 distinct elements of G, so $a^2 = b^2 = c^2 = e$. Consider $ab \in G$. We know that $ab \neq a, ab \neq b$. If $ab = e$, since $a^2 = e$, one gets $b = a$, a contradiction, so $ab \neq e$. Thus $ab = c$. Similarly, $ba = c, ac = ca = b, bc = cb = a$. Define $\phi : G \to \mathbb{Z}_2 \times \mathbb{Z}_2$ by

$$\phi(e) = (0,0), \quad \phi(a) = (0,1), \quad \phi(b) = (1,0), \quad \phi(c) = (1,1).$$

Then ϕ is an isomorphism. Hence $G \cong \mathbb{Z}_2 \times \mathbb{Z}_2$. □

§ 2. The isomorphic problem between the groups $\mathbb{Z}_m \times \mathbb{Z}_n$ and $\mathbb{Z}_{mn}$

After having classified all groups up to order 5, it is natural to (try to) classify all groups of order 6. In particular, we want to know whether $\mathbb{Z}_6$ and $\mathbb{Z}_2 \times \mathbb{Z}_3$ are isomorphic.

Lemma 2.7.6. *One has $\mathbb{Z}_2 \times \mathbb{Z}_3 \cong \mathbb{Z}_6$.*

Proof. The element $(1,1) \in \mathbb{Z}_2 \times \mathbb{Z}_3$ has order 6, so $\mathbb{Z}_2 \times \mathbb{Z}_3$ is cyclic. Thus $\mathbb{Z}_2 \times \mathbb{Z}_3 \cong \mathbb{Z}_6$. □

Now we have seen $\mathbb{Z}_2 \times \mathbb{Z}_2 \ncong \mathbb{Z}_4$ but $\mathbb{Z}_2 \times \mathbb{Z}_3 \cong \mathbb{Z}_6$, it is natural to wonder what happens to the two groups $\mathbb{Z}_m \times \mathbb{Z}_n$ and $\mathbb{Z}_{mn}$ for arbitrary $m, n \in \mathbb{N}$.

Lemma 2.7.7. *Let m and n be two positive integers. Then $\mathbb{Z}_m \times \mathbb{Z}_n \cong \mathbb{Z}_{mn}$ if and only if m and n are relatively prime.*

Proof. Consider the element $a = (1,1) \in \mathbb{Z}_m \times \mathbb{Z}_n$. For any integer $k > 0$, one has $ka = \underbrace{a + a + \cdots + a}_{(k \text{ times})} = (k,k)$. Now $(k,k) = (0,0)$ if and only if $n|k$ and $m|k$. If m and n are relatively prime, the smallest such k is $k = mn$. Thus the order of a is mn (which is the order of $\mathbb{Z}_m \times \mathbb{Z}_n$), so $\mathbb{Z}_m \times \mathbb{Z}_n \cong \mathbb{Z}_{mn}$.

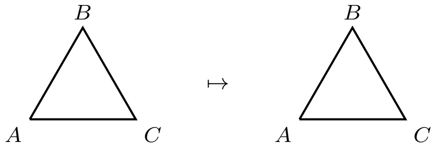
Conversely, suppose that m and n are not relatively prime. Then there exists $1 \leq r < mn$ such that $m|r, n|r$, e.g. $r = mn/(m,n)$. Then $ra = 0$ for every $a \in \mathbb{Z}_m \times \mathbb{Z}_n$, so the order of every $a \in \mathbb{Z}_m \times \mathbb{Z}_n$ is less than mn. Since no element in $\mathbb{Z}_m \times \mathbb{Z}_n$ has order mn while the element 1 in $\mathbb{Z}_{mn}$ has order mn, $\mathbb{Z}_m \times \mathbb{Z}_n$ is not isomorphic to $\mathbb{Z}_{mn}$. □

§ 3. A non-abelian Group of order 6

We have seen $\mathbb{Z}_2 \times \mathbb{Z}_3 \cong \mathbb{Z}_6$, tempting us to conclude only one group of order 6 up to isomorphisms. Unfortunately the world is not so simple, as illustrated below.

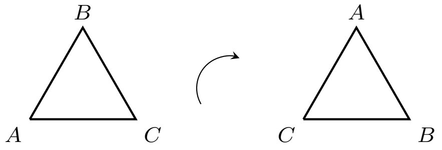
Example 2.7.8. Consider the 6 **symmetries** of an equilateral triangle ΔABC:

Identity id :



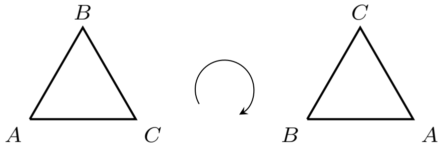
$\mathrm{id} = \begin{pmatrix} A & B & C \\ A & B & C \end{pmatrix}$

Rotation ρ_1 :



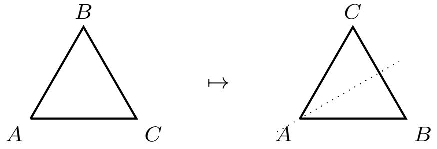
$\rho_1 = \begin{pmatrix} A & B & C \\ B & C & A \end{pmatrix}$

Rotation ρ_2 :



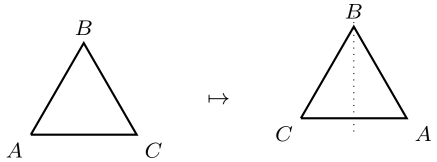
$\rho_2 = \begin{pmatrix} A & B & C \\ C & A & B \end{pmatrix}$

Reflection μ_1 :



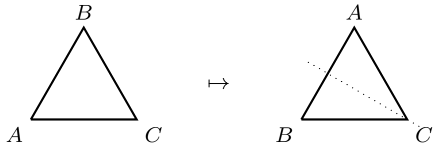
$$\mu_1 = \begin{pmatrix} A & B & C \\ A & C & B \end{pmatrix}$$

Reflection μ_2 :



$$\mu_2 = \begin{pmatrix} A & B & C \\ C & B & A \end{pmatrix}$$

Reflection μ_3 :



$$\mu_3 = \begin{pmatrix} A & B & C \\ B & A & C \end{pmatrix}$$

Observe that each symmetry can be regarded as a permutation of the vertices A, B, C and hence is a bijective function (from $\{A, B, C\}$ to $\{A, B, C\}$). Thus applying a sequence of the symmetries to the triangle may mathematically be formulated as composing a sequence of the functions $\mathrm{id}, \rho_1, \rho_2, \mu_1, \mu_2, \mu_3$.

The set $\mathrm{S}_3 = \{\mathrm{id}, \rho_1, \rho_2, \mu_1, \mu_2, \mu_3\}$ with function composition as the operation is a group, which can be easily checked. Moreover, one may check $\rho_1\mu_1 = \mu_3$ while $\mu_1\rho_1 = \mu_2$, so $\rho_1\mu_1 \neq \mu_1\rho_1$. Here $\rho_1\mu_1$ means $\rho_1 \circ \mu_1$, and thus

$$\rho_1 \circ \mu_1(A) = \rho_1(\mu_1(A)) = \rho_1(A) = B, \;\; \rho_1 \circ \mu_1(B) = A, \;\; \rho_1 \circ \mu_1(C) = C.$$

Consequently, the group S_3 is nonabelian and of order 6. Now $|\mathrm{S}_3| = 6 = |\mathbb{Z}_6|$ but the group S_3 is *not isomorphic* to $\mathbb{Z}_6$ simply because S_3 is nonabelian while $\mathbb{Z}_6$ is abelian.[‡]

Now we have seen 3 groups of order 6: $\mathbb{Z}_2 \times \mathbb{Z}_3$, $\mathbb{Z}_6$ and S_3. However $\mathbb{Z}_2 \times \mathbb{Z}_3 \cong \mathbb{Z}_6$, thus only two different groups (up to isomorphisms) are found. One would certainly ask:

Is there any group of order 6 that is not isomorphic to $\mathbb{Z}_6$ or S_3?

[‡] If the two groups G_1 and G_2 are isomorphic and G_1 is abelian, then G_2 is also abelian.

Chapter 3 — Permutation Groups

Highlights:

- Symmetric groups S_n, unless $n = 1, 2$, are *finite non-abelian* groups. (This furnishes examples of finite non-abelian groups!)
- Manipulate permutations with the **two-line** notation, **cycle** notation and **products of adjacent transpositions**.
- Uniqueness of parity in the decomposition of a permutation into products of adjacent transpositions.
- Dihedral groups, Automorphisms & Conjugations.

3.1 The permutation group of a set

Lemma 3.1.1. *For any set X, the set* $\mathrm{Perm}(X)$ *of all bijective maps from X to itself is a group with the group operation being composition of maps. We call* $\mathrm{Perm}(X)$, *with composition as the group operation, the* **permutation group of** X.

Proof. Exercise. □

When $X = \{1, 2, \ldots, n\}$, the group $\mathrm{Perm}(X)$ is also called the **symmetric group on** n **letters** and is denoted by S_n. Note that S_n has order $n!$.

For any finite set $X = \{x_1, x_2, \ldots, x_n\}$ with n elements, the map

$$I: \quad X \longrightarrow \{1, 2, \ldots, n\}, \quad x_j \longmapsto j, \qquad j = 1, 2, \ldots, n,$$

is a bijection. Thus the map $\phi_I : \mathrm{Perm}(X) \to S_n$, $f \mapsto I \circ f \circ I^{-1}$ is a group isomorphism (check it!), and hence $\mathrm{Perm}(X) \cong S_n$.

3.2 The symmetric group S_n

Now we study the group S_n in detail. First note that S_1 is the trivial group with only one element, and that $S_2 \cong \mathbb{Z}_2$.

One way of representing an element $\sigma \in S_n$ is by the following *two-line notation:*

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}.$$

The two-line notation is convenient for multiplying elements.

Example 3.2.1. The 6 elements in S_3 are[†]

$$e = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix},$$
$$\sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad \sigma_4 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \quad \sigma_5 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}.$$

Note that

$$\sigma_3 = \sigma_1\sigma_2, \quad \sigma_4 = \sigma_2\sigma_1, \quad \sigma_4 = \sigma_1\sigma_2\sigma_3, \quad \sigma_5 = \sigma_2\sigma_1\sigma_2 = \sigma_1\sigma_2\sigma_1.$$

As $\sigma_1\sigma_2 \neq \sigma_2\sigma_1$, the group S_3 is not abelian. A direct computation shows that the orders the above elements $e, \sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5$ are, respectively, $1, 2, 2, 3, 3, 2$.

Example 3.2.2. The group S_5 has order $5! = 120$. Using the two-line notation, we have

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 3 & 1 & 4 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 4 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 3 & 1 & 2 & 5 \end{pmatrix}.$$

Another way of denoting elements in S_n is by the so-called *cycle-notation.*

Notation. For distinct elements $i_1, i_2, \cdots, i_k \in \{1, 2, \cdots, n\}$, let

$$(i_1, i_2, \cdots, i_k)$$

denote the element $\tau \in S_n$ such that $\tau(i_j) = i_{j+1}$ for $1 \leq j \leq k-1$, $\tau(i_k) = i_1$, and $\tau(i) = i$ if $i \notin \{i_1, i_2, \ldots, i_k\}$.

Example 3.2.3. In S_6,

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 1 & 6 & 3 & 4 & 2 \end{pmatrix} = (1, 5, 4, 3, 6, 2) = (5, 4, 3, 6, 2, 1) = (3, 6, 2, 1, 5, 4).$$

As another example, we have

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 5 & 6 & 4 & 1 & 3 \end{pmatrix} = (1, 2, 5)(3, 6)(4) = (1, 2, 5)(3, 6).$$

Definition 3.2.4. Two cycles $(i_1, i_2, \cdots, i_k)$ and $(j_1, j_2, \cdots, j_l)$ are said to be **disjoint** if

$$\{i_1, \cdots, i_k\} \cap \{j_1, \cdots, j_l\} = \emptyset.$$

[†] The group S_3 is the "same as" S_3 in Example 3.2.1 when identifying $1, 2, 3$ with A, B, C respectively. Then $e = \mathrm{id}$, $\sigma_1 = \mu_3$, $\sigma_2 = \mu_1$, $\sigma_3 = \rho_1$, $\sigma_4 = \rho_2$, $\sigma_5 = \mu_2$. i.e. The symmetric group S_3 is the group of symmetries for an equilateral triangle.

Let $\sigma \in S_n$. First consider the sequence $(1, \sigma(1), \sigma^2(1), ...\sigma^k(1))$, where k is the smallest positive integer such that $\sigma^k(1) \neq 1$ and $\sigma^{k+1}(1) = 1$ and get the cycle $(1, \sigma^2(1), \ldots, \sigma^k(1))$. Proceed then by taking any element not in the sequence $(1, \sigma(1), \sigma^2(1), ...\sigma^k(1))$, if any, to build another cycle. Continue until no elements are left. This proves the following Lemma 3.2.5, and moreover, *provides an algorithm to decompose any $\sigma \in S_n$ into a product of disjoint cycles.*

Lemma 3.2.5. *Every element in S_n can be written as the product of disjoint cycles.*

Definition 3.2.6. For $1 \leq i < j \leq n$, the element $\tau_{ij} \in S_n$ that exchanges i and j and leaves every other element invariant is called the **transposition determined by i and j**. In the cycle notation,

$$\tau_{ij} = (i, j).$$

If $1 \leq i \leq n-1$, the transposition $\tau_{i,i+1} = (i, i+1)$ is called the ***i*'th adjacent transposition** and will also be denoted by τ_i. Note $\tau_i^{-1} = \tau_i$.

Lemma 3.2.7. *Every element in S_n ($n \geq 2$) is the product (not necessarily in a unique way) of adjacent transpositions.*

Proof. We shall prove by induction on n. The case $n = 2$ is obvious. Assume that $n \geq 2$, and let $\sigma \in S_n$. If $\sigma(n) = n$, then by regarding σ as a permutation on $\{1, \ldots, n-1\}$ and using induction assumption, we know that σ is the product of some adjacent transpositions. Assume thus that $\sigma(n) = k < n$. Let $\sigma' = \tau_{n-1}\tau_{n-2}\cdots\tau_k\sigma$. Then σ' fixes n and we have proved that σ' is the product of some adjacent transpositions, and thus so is $\sigma = \tau_k \cdots \tau_{n-2}\tau_{n-1}\sigma'$. □

A reformulation of Lemma 3.2.7 is the following

Corollary 3.2.8. *The group S_n is generated by the adjacent transpositions*

$$\tau_1 = (1, 2), \quad \tau_2 = (2, 3), \quad \ldots, \quad \tau_{n-1} = (n-1, n).$$

We have seen that every element in S_n is the product of adjacent transpositions, and thus also as products of transpositions. Indeed it is easy to decompose $\sigma \in S_n$ into a product of transpositions because of the following fact: *every cycle $(i_1, i_2, \ldots, i_k)$ is the product*

$$(i_1, i_2, \ldots, i_k) = (i_1, i_2)(i_2, i_3)\cdots(i_{k-1}, i_k), \tag{3.1}$$

in which, however, not all the transpositions are adjacent.[‡] Remember we know how to decompose σ into product of cycles, and hence into product of transpositions with the above fact. To decompose σ into a product of *adjacent* transpositions, we may apply the *argument in the proof of Lemma 3.2.7.* See Example 3.2.9 below.

[‡] Another possible decomposition is $(i_1, i_2, \ldots, i_k) = (i_1, i_k)(i_1, i_{k-1})\cdots(i_1, i_3)(i_1, i_2)$, see p.61 of Judson's book.

Example 3.2.9. In S_6,

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 5 & 2 & 3 & 4 & 6 \end{pmatrix} = (2,5,4,3)$$

Thus by (3.1), $\sigma = (2,5)(5,4)(4,3)$. To write σ into a product of adjacent transpositions, from $\sigma(5) = 4$ we consider

$$\tau_4\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 4 & 2 & 3 & 5 & 6 \end{pmatrix}.$$

Now $\tau_4\sigma(4) = 3$, we consider $\tau_3\tau_4\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 3 & 2 & 4 & 5 & 6 \end{pmatrix}$ and thus $\tau_2\tau_3\tau_4\sigma = e$. Hence $\sigma = \tau_4\tau_3\tau_2$.

Exercise 3.2.10. Try many examples of multiplying elements in S_n, decomposing them into products of disjoint cycles and products of adjacent transpositions.

As $e = (1,2)(1,2)$, a given $\sigma \in S_n$ may be written in more than one way as the product of transpositions. However, we now show that a $\sigma \in S_n$ cannot be written as the product of both an even number and an odd number of transpositions. We in fact prove something stronger.

Consider now the vector space $\mathbb{R}^n$ with the standard basis $e_1, \ldots, e_n$, where $e_j \in \mathbb{R}^n$ is the column vector with the entry 1 at the j'th place and 0 everywhere else. For $\sigma \in S_n$, define a linear map $T_\sigma : \mathbb{R}^n \to \mathbb{R}^n$ by

$$T_\sigma(x_1e_1 + x_2e_2 + \cdots + x_ne_n) = x_1e_{\sigma(1)} + x_2e_{\sigma(2)} + \cdots + x_ne_{\sigma(n)}, \quad x_1, \ldots, x_n \in \mathbb{R}.$$

Then for any $\sigma, \tau \in S_n$ and $x = x_1e_1 + x_2e_2 + \cdots + x_ne_n \in \mathbb{R}^n$, one has

$$\begin{aligned} T_{\sigma\tau}(x) &= x_1e_{\sigma(\tau(1))} + x_2e_{\sigma(\tau(2))} + \cdots + x_ne_{\sigma(\tau(n))} \\ (T_\sigma T_\tau)(x) &= T_\sigma(x_1e_{\tau(1)} + \cdots + x_ne_{\tau(n)}) = x_1e_{\sigma(\tau(1))} + x_2e_{\sigma(\tau(2))} + \cdots + x_ne_{\sigma(\tau(n))}. \end{aligned}$$

Thus $T_{\sigma\tau} = T_\sigma T_\tau$ for all $\sigma, \tau \in S_n$. In other words,

$$T : S_n \longrightarrow GL(n, \mathbb{R}), \quad \sigma \longmapsto T_\sigma$$

is a group homomorphism. It is also clear that T is injective. Thus S_n is isomorphic to a subgroup of $GL(n, \mathbb{R})$.[§]

In matrix notation, we have $T_\sigma = \begin{pmatrix} e_{\sigma(1)} & e_{\sigma(2)} & \cdots & e_{\sigma(n)} \end{pmatrix} \in GL(n, \mathbb{R})$. For example, we have

$$T_\sigma = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \qquad \text{for} \quad \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{pmatrix}.$$

What is the determinant of T_σ?

[§] The idea of reducing problems in abstract algebra to problems in linear algebra is the central spirit of *Representation Theory*.

Lemma 3.2.11. *Let $\sigma \in S_n$.*

1) If σ is the product of an even number of transpositions, then $\det(T_\sigma) = 1$.

2) If σ is the product of an odd number of transpositions, then $\det(T_\sigma) = -1$.

Proof. Since $\det(T_\sigma) = -1$ if σ is a transposition, both statements follow from the linear algebra fact that $\det(AB) = \det(A)\det(B)$ for all $n \times n$ matrices A and B. □

Corollary 3.2.12. *An element in S_n cannot be the products of both an even number and an odd number of transpositions.*

Proof. This is because $\det(T_\sigma)$ cannot be both 1 and -1 at the same time. □

Definition 3.2.13. For $\sigma \in S_n$, $\det(T_\sigma)$ is called the **sign of** σ and is also denoted by $\text{sign}(\sigma)$.

Exercise 3.2.14. The map $\text{sign} : S_n \to \{\pm 1\}$, $\sigma \mapsto \text{sign}(\sigma)$, is a group homomorphism, where $\{\pm 1\}$ has the subgroup structure of the multiplicative group $\mathbb{R}^\times$.

Notation. Denote by A_n the set of all elements in S_n that can be written as the product of even number of transpositions, i.e.,

$$A_n = \{\sigma \in S_n : \ \ \text{sign}(\sigma) = 1\} = \ker(\text{sign}).$$

Lemma 3.2.15. *For $n \geq 2$, A_n is a subgroup of S_n of index 2.*

Proof. As $e = (1,2)(2,1)$, $e \in A_n$. It is clear that A_n is closed under taking inverse and taking products, i.e. for $x, y \in A_n$, x^{-1} and $xy \in A_n$. Thus A_n is a subgroup of S_n. There are exactly two left A_n-cosets in S_n: A_n itself and $(1,2)A_n$, the set of all elements in S_n that are written as the product of odd number of transpositions. □

Definition 3.2.16. The subgroup A_n of S_n is called the **alternating group**.

3.3 The Dihedral group D_n

Definition 3.3.1. For any integer $n \geq 1$, the **dihedral group** is the subgroup D_n of $GL(2, \mathbb{R})$ generated by two elements a and b, where

$$a = \begin{pmatrix} \cos(2\pi/n) & \sin(2\pi/n) \\ -\sin(2\pi/n) & \cos(2\pi/n) \end{pmatrix}, \qquad b = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.^\dagger$$

[†] Recall "Let S be any subset of a group G. The subset $\langle S \rangle$ of G consisting of (i) the identity element e and (ii) all the products $x_1 x_2 \cdots x_n$ where for each $1 \leq j \leq n$, either $x_j \in S$ or $x_j^{-1} \in S$ is a subgroup of G". Now $D_n = \langle a, b \rangle$ is the subgroup of $GL(2, \mathbb{R})$ generated from $S = \{a, b\}$.

Remarks. (i) The matrices a and b represent rotation and reflection respectively. See Judson's §5.2 for D_n as the group of rigid motions of a regular n-gon.
(ii) Clearly $\operatorname{ord}(a) = n$ and $\operatorname{ord}(b) = 2$. Moreover, a direct calculation gives

$$ab = ba^{n-1} \qquad \text{(equivalently, } bab = a^{-1}\text{)}.$$

Thus a complete list of elements of D_n is

$$e,\ a,\ a^2, \cdots, a^{n-1},\ b,\ ab,\ a^2b, \cdots, a^{n-1}b.$$

In particular, D_n has $2n$ elements.
(iii) $S_3 \cong D_3$ because $\phi : \mathrm{id}, \rho_1, \rho_2, \mu_1, \mu_2, \mu_3 \mapsto e, a, a^2, b, ba, ba^2$ respectively is an isomorphism. However, $S_n \ncong D_n$ for all $n > 3$ (why?).

3.4 The permutation group of a group and Automorphisms of a group

Definition 3.4.1. For a group G and $a \in G$, define l_a and $r_a : G \to G$ by

$$l_a(g) = ag, \quad r_a(g) = ga, \quad g \in G,$$

and call them, respectively, the **left translation** and **the right translation** of G by a.

Remark. One may easily check that both maps $l_a, r_a \in \mathrm{Perm}(G)$.

Next for any group G, we consider the two maps

$$l : \quad G \longrightarrow \mathrm{Perm}(G), \quad a \longmapsto l_a \quad \forall\ a \in G, \tag{3.1}$$
$$r : \quad G \longrightarrow \mathrm{Perm}(G), \quad a \longmapsto r_a \quad \forall\ a \in G \tag{3.2}$$

which are well-defined by the above remark.

Lemma 3.4.2. *For any group G, the map $l : G \to \mathrm{Perm}(G)$ an embedding.*[†]

Proof. Exercise. □

Corollary 3.4.3. (**Cayley's Theorem**) *Every finite group of order n is isomorphic to a subgroup of the symmetric group S_n.*

Proof. By Lemma 3.4.2, G is isomorphic to the subgroup $l(G)$ of $\mathrm{Perm}(G) \cong S_n$. □

[†] Recall: The map ϕ is an *embedding of G_1 into G_2* if $\phi : G_1 \to G_2$ is an injective group homomorphism. For an embedding $\phi : G_1 \to G_2$, $\phi : G_1 \to \phi(G_1)$ is an isomorphism.

Corollary 3.4.3 implies that we may understand the group structure of finite groups through studying the subgroups of S_n.

It is a natural attempt to extend Lemma 3.4.2 to the map r, i.e., to show $r : G \to \mathrm{Perm}(G)$ is a group embedding. However, one will *fail* for many cases. Indeed from the definition, $r_{ab} = r_b r_a \in \mathrm{Perm}(G)$ for $a, b \in G$, so r is *in general not* a group homomorphism.

Definition 3.4.4. Let G_1 and G_2 be two groups. A map $\phi : G_1 \to G_2$ is called a group **anti-homomorphism** if $\phi(gh) = \phi(h)\phi(g)$ for all $g, h \in g_1$. A bijective anti-homomorphism from G_1 to G_2 is called an **anti-isomorphism** from G_1 to G_2.

Proposition 3.4.5. *Let G be any group. The map $r : G \to \mathrm{Perm}(G)$ in (3.2) is an injective group anti-homomorphism. Besides the map $\tau_G : G \to G, g \mapsto g^{-1}$ for $g \in G$ is an anti-isomorphism.*

Proof. Exercise. □

Definition 3.4.6. For a group G, an isomorphism from G to itself is called an **automorphism of** G, and the set of all automorphisms of G is denoted by $\mathrm{Aut}(G)$.

Let G be any group. As isomorphisms must be bijective, $\mathrm{Aut}(G)$ is a subset of $\mathrm{Perm}(G)$. As compositions of isomorphisms are again isomorphisms, $\mathrm{Aut}(G)$ *is a subgroup of* $\mathrm{Perm}(G)$.

Let G be any group and let $a \in G$. The left translation on G by $a \in G$ is, in general, *not* an automorphism of G, *nor* is the right translation r_a by a. On the other hand, for $a \in G$, define the map

$$c_a : \quad G \longrightarrow G, \quad g \longmapsto aga^{-1}, \qquad g \in G. \tag{3.3}$$

By definition, $c_a = l_a \circ r_{a^{-1}} = r_{a^{-1}} \circ l_a$. In particular, being the composition of two bijective maps, c_a is also bijective. It is also clear that for any $g, h \in H$, $c_a(gh) = c_a(g)c_a(h)$, so c_a *is an automorphism of* G.

Definition 3.4.7. Let G be a group. For $a \in G$, the map $c_a : G \to G$ in (3.3) is called the **conjugation on** G **by** a. An automorphism of G that is of the form c_a for some $a \in G$ is called an **inner automorphism** of G.

Automorphisms that are not inner are called **outer automorphisms** of G.

Example 3.4.8. 1) If G is an abelian group, only the identity map is an inner automorphism of G.

2) If G is abelian, the inverse map $\tau_G : G \to G, g \mapsto g^{-1}$ for $g \in G$ is an outer automorphism of G unless τ_G is the identity map.

3) For the additive group $(\mathbb{R}^n, +)$, every non-identity $M \in GL(n, \mathbb{R})$ defines an outer automorphism $\phi_M : \mathbb{R}^n \to \mathbb{R}^n, x \mapsto Mx$ for $x \in \mathbb{R}^n$. This example shows that an abelian group may have a large non-abelian automorphism group.

Chapter 4 — Normal subgroups and Factor groups

4.1 Normal subgroups

Definition 4.1.1. A subgroup H of a group G is said to be **normal** if $aH = Ha$ for every $a \in G$, or, equivalently, if $aHa^{-1} = H$ for all $a \in G$.

Example 4.1.2. 1) In the dihedral group D_n with generators a and b and relations $a^n = b^2 = e$ and $ab = ba^{n-1}$, the subgroup $\langle a \rangle$ generated by a is normal.

2) In the symmetric group S_n, the alternating subgroup A_n is a normal subgroup.

Example 4.1.3. If G is abelian, then any subgroup of G is normal.

Lemma 4.1.4. *If H is a subgroup of a group G such that $aHa^{-1} \subset H$ for every $a \in G$, then H is a normal subgroup.*

Proof. We only need to show that $H \subset aHa^{-1}$ for every $a \in G$. Let $a \in G$ be arbitrary. Then for every $h \in H$, one has $h = a(a^{-1}ha)a^{-1}$. Since $a^{-1}ha \in a^{-1}Ha \subset H$, we have $h \in aHa^{-1}$. Thus $H \subset aHa^{-1}$. □

Lemma 4.1.5. *For any group homomorphism $\phi : G_1 \to G_2$,*

$$\ker(\phi) \overset{\text{def}}{=} \{a \in G_1 : \phi(a) = e\}$$

is a normal subgroup of G_1.

Proof. We first assert that $\ker(\phi)$ is a subgroup of G_1, whose proof is an exercise.

Now let $g \in G_1$ be arbitrary and $a \in \ker(\phi)$. Then $\phi(gag^{-1}) = \phi(g)\phi(a)\phi(g^{-1}) = \phi(g)\phi(g)^{-1} = e$. Thus shows that $g \ker(\phi) g^{-1} \subset \ker(\phi)$ for every $g \in G_1$. By Lemma 4.1.4, $\ker(\phi)$ is a normal subgroup of G_1. □

4.2 Factor groups

Definition 4.2.1. If G is a group and $H \subset G$ any subgroup, we define G/H to be the set of all left H-cosets of G, i.e.

$$G/H = \{aH : a \in G\},$$

and G/H is called the **quotient space of G by H**. For $a \in G$, we will also use $\bar{a}$ to denote the point $aH \in G/H$. The map $G \to G/H, a \mapsto \bar{a}$, is called the **natural** (or **canonical**) **projection**.

Recall that being in the same left H-coset is an equivalence relation on G and that the left H-cosets are precisely all the equivalence classes. Thus a *point* in G/H is an *equivalence class* of the equivalence relation $\sim_H$ on G.

Example 4.2.2. For an integer $n \geq 2$, consider the alternating subgroup A_n of the symmetric group S_n. The set S_n/A_n has two elements. (Why?)

Theorem 4.2.3. *Let G be a group and let H be a normal subgroup of G. Then the set G/H becomes a group with the group operation given by*

$$G/H \times G/H \longrightarrow G/H : \quad (\bar{a}, \bar{b}) \longmapsto \bar{a}\bar{b} := \overline{ab}, \qquad a, b \in H \tag{4.1}$$

(or equivalently, $(aH)(bH) := abH$). The identity element is $\bar{e}$ and the inverse of $\bar{a}$ is $\overline{a^{-1}}$ (i.e. $\bar{a}^{-1} = \overline{a^{-1}}$).

Proof. We first need to show that the map in (4.1) is well-defined: if $a, b, a', b' \in H$ are such that $\bar{a} = \overline{a'}$ and $\bar{b} = \overline{b'}$, then $a' = ah_1$ and $b' = bh_2$ for some $h_1, h_2 \in H$. Thus $a'b' = ah_1bh_2 = ab(b^{-1}h_1bh_2)$. Since H is normal, $b^{-1}h_1b \in H$ and so $b^{-1}h_1bh_2 \in H$. Thus $\overline{a'b'} = \overline{ab}$. This shows that the map in (4.1) is well-defined. The associativity then follows from that of the group operation on G, and it also follows from the definition that $\bar{e}$ is really the identity element and that the inverse of $\bar{a}$ is $\overline{a^{-1}}$. □

Definition 4.2.4. When H is a *normal subgroup* of G, the set G/H with the group structure described in Theorem 4.2.3, i.e. $\bar{a}\bar{b} = \overline{ab}$ (or $(aH)(bH) = abH$), is called the **factor group** (or **quotient group**) **of G by H**.

It follows immediately from the definitions that, with the factor group structure on G/H, the natural projection $G \to G/H$ is a group homomorphism. Note that if $H = \{e\}$, then $G/H \cong G$, and if $H = G$, then $G/H \cong \{e\}$.

Theorem 4.2.5. (**First Isomorphism Theorem**)
Let $\phi : G_1 \to G_2$ be a group homomorphism. Then the map

$$\bar{\phi} : \quad G/\ker(\phi) \longrightarrow G_2, \quad \bar{\phi}(\bar{a}) := \phi(a), \qquad a \in G_1,$$

is a well-defined embedding. Consequently, $\bar{\phi}$ is also an isomorphism from the group $G/\ker(\phi)$ to the subgroup $\operatorname{Im}(\phi) := \{\phi(a) : a \in G_1\}$ of G_2.

Proof. Let $a, b \in G_1$. If $\bar{a} = \bar{b}$, then $b^{-1}a \in \ker(\phi)$, and thus $\phi(a) = \phi(b)$. This shows that the map $\bar{\phi}$ is well-defined. If $\bar{\phi}(\bar{a}) = \bar{\phi}(\bar{b})$, then $\phi(a) = \phi(b)$, so $\phi(b^{-1}a) = e$, so $\bar{a} = \bar{b}$. This shows that $\bar{\phi}$ is injective. Finally,

$$\bar{\phi}(\bar{a}\bar{b}) = \bar{\phi}(\overline{ab}) = \phi(ab) = \phi(a)\phi(b) = \bar{\phi}(\bar{a})\bar{\phi}(\bar{b}).$$

This shows that $\bar{\phi}$ is a group homomorphism. Thus $\bar{\phi}$ is an embedding. Regarded as a map from $G/\ker(\phi)$ to $\operatorname{Im}(\phi) \subset G_2$, $\bar{\phi}$ is bijective and thus an isomorphism. □

Example 4.2.6. Consider the additive group $(\mathbb{Z}, +)$. Let $n \geq 1$ be any integer. Then $n\mathbb{Z} := \{nx : \ x \in \mathbb{Z}\}$ is a normal subgroup of $\mathbb{Z}$. Consider the group homomorphism

$$\phi : \quad \mathbb{Z} \longrightarrow \mathbb{Z}_n, \quad \phi(k) = r \quad \text{where} \ \ 0 \leq r \leq n-1, \ \ k \equiv r \ (\mathrm{mod}\ n).$$

Then $\ker(\phi) = n\mathbb{Z}$ and $\bar{\phi} : \mathbb{Z}/n\mathbb{Z} \to Z_n$ is an isomorphism.

4.3 Simple and solvable groups

As the classification of all finite groups is a hard problem, it is natural to restrict to subclasses of groups. Simple and solvable groups are two such classes.

Definition 4.3.1. A group G is said to be **simple** if it contains no normal subgroups other than the trivial subgroup $\{e\}$ and G itself.

Example 4.3.2. We have seen that every group of order p, where p is a prime number, must be cyclic. By Lagrange's theorem, such a group has no subgroups, thus no normal subgroups, other than $\{e\}$ or the group itself, so it is simple.

Theorem 4.3.3. *For $n \geq 5$, the alternating group A_n is simple.*

Proof. For interested students only: see §10.2 in Chapter 10 of Judson's book. □

We state without proof the statements in Example 4.3.4 about simple groups.

Example 4.3.4. (i) The smallest non-abelian simple group is the alternating group A_5 of order 60, and (ii) every simple group of order 60 is isomorphic to A_5.

Definition 4.3.5. A group G is said to be **solvable** if there is a sequence of subgroups

$$G = H_0 \supset H_1 \supset H_2 \supset \cdots \supset H_r = \{e\}$$

such that H_{i+1} is normal in H_i for $i = 0, 1, \ldots, r-1$ [†] and H_i/H_{i+1} is abelian.[‡]

Example 4.3.6. Every abelian group is solvable.

Exercise 4.3.7. The symmetric groups S_1, S_2, S_3, S_4 are solvable.

Theorem 4.3.8. *The symmetric group S_5 is not solvable.*

Proof. Exercise. □

[†] A finite sequence of subgroups $G = H_0 \supset H_1 \supset H_2 \supset \cdots \supset H_r = \{e\}$ such that H_{i+1} is normal in H_i for $i = 0, 1, \ldots, r-1$ is called a **subnormal series of** G.

[‡] Interlude: As is well-known, the roots of a quadratic polynomial $q(x) = ax^2 + bx + c$ are given by the formula $(-b \pm \sqrt{b^2 - 4ac})/(2a)$. In this case we say that $q(x) = 0$ is *solvable in radicals.* Naturally one would ask, in general, when a polynomial equation is solvable in radicals. Shortly before the end of his life, the French mathematician É. Galois (1811-1832) proved

Theorem (Galois, 1832) For a polynomial $f(x)$ (over a field of characteristic zero), the equation $f(x) = 0$ is *solvable in radicals* if and only if the *Galois group* of f is *solvable.*

Chapter 5 — Finite Abelian Groups

5.1 Some properties of finite abelian groups

When G is an abelian group, every subgroup H of G is normal. When G is abelian and finite, for any non-trivial subgroup H of G, the factor group G/H is again abelian but of smaller size. Many properties of finite abelian groups can be proved by using induction on the order of G. Here is an example using this trick.

Theorem 5.1.1. (**Cauchy's theorem for abelian groups**) *If G is a finite abelian group and p is a prime number such that p divides the order of G, then there exists a subgroup of G with order p.*

Proof. In fact we are going to show: G has an element of order p (then $\langle a\rangle$ is a subgroup of order p). We first note that the statement holds if $|G| = p$. In general we use induction on the order of G. (As $p||G|$, we may write $|G| = pm$ and the induction is on m.)

If G has no non-trivial proper subgroup, then G is cyclic of order p and the statement holds. Assume that H is a non-trivial proper subgroup of G. If p divides the order of H, by induction, there exists an element of H with order p, which is also an element of G of order p, so we are done. If p does not divide the order of H, then p must divide the order of G/H (note $|G/H| = pm'$ with $m' < m$), so by induction, there exists $a \in G$ such that the order of $\bar{a} \in G/H$ is p. Let k be the order of a in G. Then $\bar{a}^k = \bar{e} \in G/H$, so p divides k. Let $k = pl$, where $1 \leq l < k$. Then $(a^l)^p = e$. Since $a^l \neq e$, a^l has order p and thus the subgroup $\langle a^l\rangle$ has order p. □

5.2 Fundamental theorem of finite abelian groups

We will state the Fundamental Theorem of Finite Abelian Groups and *apply it to find <u>all</u> abelian groups of a given order up to isomorphisms* but we will not prove it. The fundamental theorem essentially says that every finite abelian group G is isomorphic to a product of cyclic groups, where the orders of these cyclic groups depend on the order of G.

Let us explain the fundamental theorem with a special case before stating it.

Consider an abelian group G of order $p_1^3 p_2^4$ where $p_1 \neq p_2$ are primes. The fundamental theorem tells that G is isomorphic to a product group of two parts (because of two different prime factors) and each part is a product of cyclic groups; more precisely, if G is abelian and $|G| = p_1^3 p_2^4$, then there are positive integers $n_{1,1} \geq n_{1,2} \geq \cdots \geq n_{1,k_1}$ and $n_{2,1} \geq n_{2,2} \geq \cdots \geq n_{2,k_2}$ such that

$$G \cong \boxed{p_1\text{-part}} \times \boxed{p_2\text{-part}}$$

where

$$\begin{aligned}
\boxed{p_1\text{-part}} \quad &= \quad \mathbb{Z}_{p_1^{n_{1,1}}} \times \mathbb{Z}_{p_1^{n_{1,2}}} \times \cdots \times \mathbb{Z}_{p_1^{n_{1,k_1}}}\ , \quad n_{1,1} + n_{1,2} + \cdots + n_{1,k_1} = 3,\\
\boxed{p_2\text{-part}} \quad &= \quad \mathbb{Z}_{p_2^{n_{2,1}}} \times \mathbb{Z}_{p_2^{n_{2,2}}} \times \cdots \times \mathbb{Z}_{p_2^{n_{2,k_2}}}\ , \quad n_{2,1} + n_{2,2} + \cdots + n_{2,k_2} = 4.
\end{aligned}$$

Since $3 = 3 = 2 + 1 = 1 + 1 + 1$ (3 ways) and $4 = 4 = 3 + 1 = 2 + 2 = 2 + 1 + 1 = 1 + 1 + 1 + 1$ (5 ways), there are 15 non-isomorphic abelian groups of order $p_1^3 p_2^4$, which are

$$\mathbb{Z}_{p_1^3} \times \mathbb{Z}_{p_2^4}\ , \quad \mathbb{Z}_{p_1^3} \times \mathbb{Z}_{p_2^3} \times \mathbb{Z}_{p_2}\ , \quad \cdots .$$

Now we state the fundamental theorem.

Theorem 5.2.1. *(***Fundamental Theorem for Finite Abelian Groups***)* *Every finite abelian group is the product of cyclic groups whose orders are powers of prime numbers. More precisely, if $|G| = p_1^{m_1} p_2^{m_2} \cdots p_l^{m_l}$, where $p_1, p_2, \cdots, p_l$ are pairwise distinct prime numbers, then there exist unique positive integers*

$$n_{1,1} \geq n_{1,2} \geq \cdots \geq n_{1,k_1},\ n_{2,1} \geq n_{2,2} \geq \cdots \geq n_{2,k_2},\ \cdots\ ,\ n_{l,1} \geq n_{l,2} \geq \cdots \geq n_{l,k_l}$$

such that $n_{j,1} + n_{j,2} + \cdots + n_{j,k_j} = m_j$ for $j = 1, 2, \ldots, l$ and

$$G \cong \underbrace{\mathbb{Z}_{p_1^{n_{1,1}}} \times \cdots \times \mathbb{Z}_{p_1^{n_{1,k_1}}}}_{p_1-\text{part}} \times \cdots \times \underbrace{\mathbb{Z}_{p_l^{n_{l,1}}} \times \cdots \times \mathbb{Z}_{p_l^{n_{l,k_l}}}}_{p_l-\text{part}}\ .$$

Example 5.2.2. By the fundamental theorem on finite abelian groups, we have the following lists of all non-isomorphic abelian groups of

(i) order 8: $\mathbb{Z}_8, \quad \mathbb{Z}_4 \times \mathbb{Z}_2, \quad \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2;$

(ii) order 9: $\mathbb{Z}_9, \quad \mathbb{Z}_3 \times \mathbb{Z}_3;$

(iii) order 72: $\mathbb{Z}_8 \times \mathbb{Z}_9, \quad \mathbb{Z}_8 \times \mathbb{Z}_3 \times \mathbb{Z}_3,$
$\mathbb{Z}_2 \times \mathbb{Z}_4 \times \mathbb{Z}_9, \quad \mathbb{Z}_2 \times \mathbb{Z}_4 \times \mathbb{Z}_3 \times \mathbb{Z}_3,$
$\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_9, \quad \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_3 \times \mathbb{Z}_3.$

Exercise 5.2.3. List all non-isomorphic abelian groups of order 36.

Exercise 5.2.4. Suppose that $p_1, p_2, \ldots, p_k$ are k distinct prime numbers and that G is an abelian group of order $p_1 p_2 \cdots p_k$. Show that G is cyclic.

Chapter 6 — Group actions

Other than the concepts of groups, subgroups, group homomorphisms, the next important concept in group theory is that of group actions.

6.1 Group actions on a set

Definition 6.1.1. Let G be a group and X a set. A (left) **action of** G **on** X is a map

$$\sigma : \; G \times X \longrightarrow X : \;\; (g, x) \longmapsto g * x$$

such that

1) $e * x = x$ for all $x \in X$;
2) $g_1 * (g_2 * x) = (g_1 g_2) * x$

for all $g_1, g_2 \in G$ and $x \in X$. Most of the time we just write gx for $g * x$.

Note that if G acts on X, the action restricts to an action of any subgroup of G on X.

Definition 6.1.2. Let G be a group and assume that G acts on a set X. For $x \in X$, the set

$$Gx := \{gx \mid g \in G\}$$

is called the **orbit of** G **in** X **through** x. The subset given by

$$G_x = \{g \in G \mid gx = x\}$$

is called the **stabilizer subgroup of** x in G.

Remarks. (i) One checks easily that G_x *is a subgroup of* G *for any* $x \in X$.

(ii) For any group G and any set X, one always has the *trivial action* of G on X given by $gx = x$ for all $g \in G$ and $x \in X$. For this action, the G-orbit of any $x \in X$ is $Gx = \{x\}$ and the stabilizer subgroup of G at any $x \in X$ is $G_x = G$.

(iii) If there exists $x \in X$ such that $Gx = X$, we say that the action of G on X is **transitive**. In this case, there is only one orbit.

Lemma 6.1.3. *Let G act on a set X. Then two orbits of G in X are either equal or disjoint. Consequently, X is the disjoint union of the collection of distinct G-orbits in X.*

Proof. Exercise. Hint: Consider the relation: $x \sim y$ if $x \in Gy$. □

Example 6.1.4. $G = GL(n, \mathbb{F})$ acts on $X = \mathbb{F}^n$ for $\mathbb{F} = \mathbb{Q}, \mathbb{R}$ or $\mathbb{C}$ by $g * x = gx$, where $g \in G$, $x \in \mathbb{F}^n$ is regarded as an $n \times 1$ matrix with entries in $\mathbb{F}$, and gx means matrix multiplication. Then G acts on X and there are two orbits: $\{\underline{0}\}$ and $\mathbb{F}^n \setminus \{\underline{0}\}$. In fact, $Gx = \{\underline{0}\}$ if and only if $x = \underline{0}$, and for all $x \neq \underline{0}$, $Gx = \mathbb{F}^n \setminus \{\underline{0}\}$.

Example 6.1.5. Given a group G, we have the following two actions of G on itself:

1) G acts on itself by left translations: $G \times G \to G, (g, h) \mapsto gh$ for $g, h \in G$. This action is transitive. The orbit of any x equals G and the stabilizer is the trivial subgroup of G.

2) G acts on itself by conjugations:

$$G \times G \longrightarrow G, \quad (g, h) \longmapsto ghg^{-1}, \qquad g, h, \in G.$$

The orbits of the conjugation action are *conjugacy classes* and two elements in G are *conjugate to each other* if they lie in the same orbit (i.e. conjugacy class). If G is abelian, every conjugacy class has exactly one point.

Exercise 6.1.6. (**Conjugacy classes in** S_n)

(a) Let $\alpha = (1, 2, 3)$ and $\beta = (2, 4, 5)$ be two permutations in S_5. Find $\sigma \in S_5$ such that $\sigma\alpha\sigma^{-1} = \beta$. Could you find $\eta \in S_5$ such that $\eta(1, 2)\eta^{-1} = \beta$?

(b) Prove that two cycles τ and μ in S_n have the same length if and only if $\sigma\tau\sigma^{-1} = \mu$ for some $\sigma \in S_n$. (e.g. A cycle $(i_1, \cdots, i_k)$ is of length of k.)[†]

(c) If an element σ of the symmetric group S_n is written as the product of disjoint cycles of lengths $\lambda_1 \geq \cdots \geq \lambda_k$, we call $\lambda = (\lambda_1, \ldots, \lambda_n)$ the **cycle pattern** of σ. For example, the element

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 5 & 3 & 8 & 4 & 2 & 7 & 6 \end{pmatrix} = (2, 5, 4, 8, 6)(1)(3)(7)$$

in S_8 has cycle pattern $(5, 1, 1, 1)$, while the element

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 6 & 1 & 4 & 7 & 8 & 2 & 3 & 5 \end{pmatrix} = (1, 6, 2)(3, 4, 7)(5, 8)$$

in S_8 has cycle pattern $(3, 3, 2)$. Show that *two elements in the symmetric group S_n are conjugate to each other if and only if they have the same cycle pattern.* Hence *the conjugacy classes in S_n are the sets of elements of the same cycle pattern.*

(d) Use Part (c) to show that $V = \{e,\ (1, 2)(3, 4),\ (1, 3)(2, 4),\ (1, 4)(2, 3)\}$ is a normal subgroup of S_4.

[†] No idea how to prove?! See Theorem 6.16 of Judson's book.

Exercise 6.1.7. Suppose that a group G acts on a set X. Show that if $x_1, x_2 \in X$ are in the same G-orbit, then their stabilizer subgroups of G are conjugate to each other.

Example 6.1.8. Let G be a group and H a subgroup of G. Let $X = G/H$, the quotient space of G by H. Recall that elements in G/H are denoted by $\bar{a} = aH \in G/H$. Define

$$\sigma : \quad G \times G/H \longrightarrow G/H, \quad \sigma(g, \bar{a}) = \overline{ga}, \qquad g, a \in G.$$

Then σ is a well-defined left transitive action of G on G/H (i.e. $g * \bar{a} = \overline{ga}$). For $a \in G$, the stabilizer subgroup of G at $\bar{a} \in G/H$ is

$$G_{\bar{a}} = \{g \in G : \overline{ga} = \bar{a}\} = aHa^{-1}.$$

If H_1 is a subgroup of G, we may restrict the G action to an H_1 action (on G/H). The stabilizer of H_1 at $\bar{a}$ is $H_1 \cap (aHa^{-1})$.

Exercise 6.1.9. Let G be a group and let X be a set. Let $\sigma : G \times X \to X$ be a map. For $g \in G$, define $\rho(g) : X \to X$ by $\rho(g)(x) = \sigma(g, x)$ for $x \in X$. Recall that $\mathrm{Perm}(X)$ is the group of all bijections from X to itself with the group operation being composition of maps. Show that σ is a group action if and only if $\rho(g) : X \to X$ is bijective for every $g \in G$ and that $\rho : G \to \mathrm{Perm}(X)$ is a group homomorphism.

6.2 The class equation

Lemma 6.2.1. *Let G be a finite group acting on a set X. Then for each $x \in X$, the orbit Gx has finitely many elements and its size is equal to $[G : G_x] = |G|/|G_x|$. In particular, the size of each G-orbit in X divides the order of G.*

Proof. By definition, for $g_1, g_2 \in G$, $g_1x = g_2x$ if and only if $g_2^{-1}g_1x = x$, i.e., $g_2^{-1}g_1 \in G_x$, which is equivalent to $g_1G_x = g_2G_x$. Thus the map

$$G/G_x \longrightarrow Gx, \qquad gG_x \longmapsto gx, \qquad g \in G,$$

is a well-defined bijection, and hence the statement. □

Corollary 6.2.2. *Suppose that G is a finite group acting on a finite set X, and suppose $x_1, x_2, \cdots, x_k$ are elements in X such that $Gx_1, Gx_2, \cdots, Gx_k$ are exactly all the distinct G-orbits in X. Then one has the following orbit decomposition formula:*

$$|X| = \sum_{i=1}^{k} |Gx_i| = \sum_{i=1}^{k} |G|/|G_{x_i}|.$$

Definition 6.2.3. For an action $G \times X \to X$, an element $x \in X$ is called a **fixed point of** G is $gx = x$ for all $g \in G$, or, equivalently, if the G-orbit through x consists of only the point x. The set $X^G = \{x \in X \mid gx = x \ \forall g \in G\}$ is called the **fixed point set of G in X** or simply **G-fixed point set**.

Corollary 6.2.4. *Suppose that G is a finite group acting on a finite set X, and suppose $x_1, x_2, \cdots, x_k$ are elements in G such that $Gx_1, Gx_2, \cdots, Gx_k$ are exactly all the distinct G-orbits in X that have more than one element. Then*

$$|X| = |X^G| + \sum_{i=1}^{k} |Gx_i| = |X^G| + \sum_{i=1}^{k} |G|/|G_{x_i}|.$$

Remark. If the group G acts on itself by conjugation: $G \times G \ni (g, x) \mapsto gxg^{-1}$, then the *center* of G, $Z = \{x \in G : gx = xg, \ \forall g \in G\}$, is the set of points *fixed* by conjugation. i.e. $Z = G^G$. Also, the *conjugacy class* C containing x is the *orbit* of x under conjugation, thus $C = \{gxg^{-1} : \ g \in G\} = Gx$. Applying Corollary 6.2.2, we get the following *class equation.*[‡]

Theorem 6.2.5. (**Class Equation**) *Let G be a finite group, let Z be the center of G, and let $C_1, \cdots, C_k$ be the collection of all distinct non-trivial conjugacy classes in G. Then*

$$|G| = |Z| + \sum_{i=1}^{k} |C_k|.$$ [§]

Example 6.2.6. For the group S_3, and there are three conjugacy classes

$$Z = \{e\}, \quad C_1 = \{(1,2), \ (2,3), \ (1,3)\}, \quad C_2 = \{(1,2,3), \ (1,3,2)\},$$

the the class formula gives $6 = 1 + 3 + 2$.

Exercise 6.2.7. (a) What element(s) does the center of S_4 contain? How many conjugacy class does S_4 have? And how many elements does each conjugacy class of S_4 carry? Hence verify the class equation for S_4. [Hint: Use Exercise 6.1.6.]

(b) For an integer $n \geq 1$, a *partition of* n is by definition a sequence $(\lambda_1, \lambda_2, \ldots, \lambda_k)$ of positive integers such that

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k \quad \text{and} \quad \lambda_1 + \lambda_2 + \cdots + \lambda_k = n.$$

The number of partitions of n is denoted by $p(n)$. For example, $p(4) = 5$ and $p(5) = 7$, with

$$4 = 4 = 3 + 1 = 2 + 2 = 2 + 1 + 1 = 1 + 1 + 1 + 1,$$
$$5 = 5 = 4 + 1 = 3 + 2 = 3 + 1 + 1 = 2 + 2 + 1 = 2 + 1 + 1 + 1 = 1 + 1 + 1 + 1 + 1.$$

Show that the number of conjugacy classes in S_n is equal to $p(n)$.

[‡] Class equation is a vital tool to understand the group structure. You will see in a later chapter.

[§] In Judson's book, Section 14.2, the notation $C(x_i)$ denotes the *stabilizer* (subgroup) of x_i. Don't mix up $C(x_i)$ there and C_i here. In fact $|C_i| = |G|/|C(x_i)|$ in view of their definitions.

Chapter 7 — The Sylow Theorems

7.1 *p*-groups

Definition 7.1.1. Let p be a prime number. A ***p*-group** is a finite group whose order is a *power of* p. If G is a finite group, a subgroup H of G is called a ***p*-subgroup** if the order of H is a power of p.

Lemma 7.1.2. *Suppose that G is a p-group acting on a finite set X. Then p divides $|X| - |X^G|$.*[†] *Consequently, if p does not divide $|X|$, then $X^G \neq \emptyset$. If p divides $|X|$, then $p||X^G|$.*

Proof. Let $O_1, O_2, \cdots, O_k$ be the list of all G-orbits in X that have *more than one* point. Then by Corollary 6.2.4,

$$|X| = |X^G| + |O_1| + |O_2| + \cdots + |O_k|.$$

Now $|O_i| = |G/H_i|$ where H_i is a *proper* subgroup of G (using Lemma 6.2.1). Thus p divides $|O_i|$ for each $1 \leq i \leq k$. Hence p divides $|X| - |X^G|$. If p does not divide $|X|$, then p does not divide $|X^G|$, so $|X^G| \geq 1$. Otherwise $p||X|$ will imply $p||X^G|$. □

Lemma 7.1.3. *Let p be a prime number. Then any non-trivial p-group has a non-trivial center.*

Proof. This is is a direct application of Lemma 7.1.2 to the conjugation action of the p-group on itself. (Recall[‡] $Z = G^G$ when G acts on G by conjugation.) □

7.2 Sylow's three theorems and Cauchy's Theorem

The three Sylow's theorems[§] are fundamental in understanding subgroups of finite groups. They are all proved by applying the formula in Corollary 6.2.4 to various actions, such as the actions of the group on itself by conjugations, or on a coset space G/H, or on a subset of subgroups. We will prove Sylow's first theorem but will only state the other two theorems without proofs.

[†] Recall that for a set X with the action of a group G, we define $X^G = \{x \in X : gx = x \; \forall \; g \in G\}$ and call it the G-fixed point set in X.

[‡] Again the notation G^G for center is *seldom* used!

[§] L. Sylow (1832-1918) was a Norwegian mathematician.

Definition 7.2.1. A subgroup H of a finite group G is called a ***p*-Sylow subgroup** of G if H has order p^n, where p^n is the *highest* order of p dividing the order of G.

Theorem 7.2.2. (**Sylow's first theorem**) *Let G be a finite group and let p be a prime number dividing the order of G. Then there exists a p-Sylow subgroup of G.*

Proof. We use induction on the order of G. If the order of G is p, the statement is obvious. We assume that the statement holds for all finite groups whose orders are smaller than that of G.

Let p^n be the highest power of p dividing $|G|$, and consider the class formula

$$|G| = |Z| + \sum_{i=1}^{k} |C_i|,$$

where Z is the center of G and $C_1, C_2, \cdots, C_k$ is the list of all conjugacy classes of G consisting of at least two elements. Pick one element x_i from each C_i, and let G_{x_i} be the stabilizer[†] of x_i in G, so that $|C_i| = |G|/|G_{x_i}|$. Since each C_i contains at least two elements, G_{x_i} is a proper subgroup of G. If there exists an i such that p^n divides $|G_{x_i}|$, we can apply the induction hypothesis to G_{x_i} and any p-Sylow subgroup of G_{x_i} will be one for G and we are done. Assume then that p^n does not divide $|G_{x_i}|$ for any i. Then p divides $|C_i|$ for each i. Thus p divides $|Z|$. By Cauchy's theorem for abelian groups, there exists an element $a \in Z$ of order p. Consider $G/\langle a \rangle$ and the projection $\phi : G \to G/\langle a \rangle$. Now $G/\langle a \rangle$ is a quotient group[‡]. By induction hypothesis, there exists a subgroup H' of $G/\langle a \rangle$ of order p^{n-1}. Let $H = \phi^{-1}(H')$. Then $H/\langle a \rangle = H'$, so the order of H is p^n. □

We now prove Cauchy's theorem for any finite group G (not necessarily abelian).

Corollary 7.2.3. (**Cauchy's theorem for finite groups**) *Assume that G is a finite group and that p is a prime number dividing $|G|$. Then there exists a subgroup of G of order p.*

Proof. Let H be a p-Sylow subgroup of G. Since H is non-trivial, there exists $a \in H$ with $a \neq e$. The order of a must be a positive power of p, say p^r, where $r \geq 1$. It follows from $a^{p^r} = e$ that $(a^{p^{r-1}})^p = e$. Since $a^{p^{r-1}} \neq e$, the order of $a^{p^{r-1}}$ is p, thus the subgroup $\langle a^{p^{r-1}} \rangle$ has order p. □

One in fact has the following generalization of Cauchy's theorem.

[†] Under the action of conjugation, the stabilizer of x_i is $G_{x_i} = \{g \in G : gx_i = x_i g\}$.

[‡] Why do we know $\langle a \rangle$ is normal in G?

Theorem 7.2.4. *Let G be a finite group and let p be a prime number. Assume that p^n, where $n \geq 1$ is an integer, is the highest power of p that divides the order of G. Then for any $1 \leq r \leq n$, there exists a subgroup of G with order p^r.*

Proof. By Sylow's first theorem, there exists a subgroup H of G such that $|H| = p^n$. We will only need to show that H has a subgroup of order p^r for every $1 \leq r \leq n$. We use induction on n. If $n = 1$, the statement is obvious. Assume the statement holds for $n-1$. The center Z of this (non-trivial) p-group H is non-trivial, by Lemma 7.1.3. Since $p||Z|$, there exists an element a in Z of order p. Consider $H/\langle a\rangle$ (observing $\langle a\rangle$ is normal in H). By induction hypothesis, $H/\langle a\rangle$ has a subgroup H' of order p^{r-1}. Its preimage in H under the projection from H to $H/\langle a\rangle$ is a subgroup of H of order p^r. □

Example 7.2.5. By Theorem 7.2.4, every group of order $5^4 7^5 13^3$ has subgroups of order $5, 25, 125$ and 625, and similarly 7-subgroups and 13-subgroups.

Lemma 7.2.6.[†] *Every group of order p^2, where p is a prime number, is abelian and is thus either isomorphic to $\mathbb{Z}_{p^2}$ or $\mathbb{Z}_p \times \mathbb{Z}_p$.*

Proof. By the fundamental theorem on finite abelian groups, it is enough to prove that G must be abelian. Let Z be the center of G. By Lemma 7.1.3, Z is non-trivial. Hence $|Z| = p$ or $|Z| = p^2$. If $|Z| = p$, there must exist some $a \in G$ such that $a \notin Z$. Let Z_a be the centralizer of a in G.[‡] Then since $a \in Z_a$ and $Z \subset Z_a$, we have $|Z_a| > p$. Thus $|Z_a| = p^2$ and so $Z_a = G$, i.e. $a \in Z$ which is a contradiction. Hence $|Z| = p^2$ and $G = Z$ is abelian. □

Theorem 7.2.7. (**Sylow's second theorem.**) *Let G be a finite group and let p be a prime number dividing the order of G. Let P be a p-Sylow subgroup. Then any p-subgroup of G is contained in a conjugate of P. In particular, all p-Sylow subgroups of G are conjugate to each other.*

Theorem 7.2.8. (**Sylow's third theorem.**) *Let G be a finite group and let p be a prime number such that $|G| = p^n m$ for some $n \geq 1$ and p not dividing m. If r is the number of different p-Sylow subgroups of G, then*

$$r|m \quad \textit{and} \quad p|(r-1).$$

[†] This lemma classifies the structure of groups of order p^2. A prior we only know the particular case of $p = 2$: A group of order 4 is isomorphic to $\mathbb{Z}_4$ or $\mathbb{Z}_2 \times \mathbb{Z}_2$.

[‡] Let G be a group and $a \in G$. The centralizer of a is $\{g \in G : ga = ag\}$, i.e. the stabilizer of a w.r.t. the conjugation action of G on itself. Write Z for the center of G and Z_a for the centralizer of a in G. Then $a \in Z$ if and only if $Z_a = G$. (Verify it!)

Chapter 8 — Rings, Integral domains and Fields

8.1 Rings

Definition 8.1.1. A **ring with unity** is a set R with two operations

$$\begin{aligned} +: \quad & R \times R \longrightarrow R, \quad (x,y) \longmapsto x+y, \\ \cdot: \quad & R \times R \longrightarrow R, \quad (x,y) \longmapsto xy, \end{aligned}$$

respectively called **addition** and **multiplication**, such that

(1) $(R,+)$ is an abelian group; its identity element will be denoted as 0.

(2) For all $x, y, z \in R$, one has

(i) $(xy)z = x(yz)$,

(ii) $x(y+z) = xy + xz$ and $(y+z)x = yx + zx$.

(3) There exists an element $1 \in R$ such that $1x = x1 = x$ for all $x \in R$. [†]

The ring R is said to be **commutative** if $xy = yx$ for all $x, y \in R$.

Unless specified otherwise, we shall use "ring" to mean "ring with unity" from now on.

Remarks. (i) As with the case of groups, to specify a ring one needs to give the set R as well as the *two* operations.

(ii) The conditions in the definition of a ring $(R, +, \cdot)$ can be rephrased as:

- $(R,+)$ is an abelian group,
- $(R,\cdot)$ satisfies the first two group axioms,
- *distributive* properties in (2) (ii).

According to the commutativity of the ring multiplication, rings are divided into two categories: *commutative rings* and *non-commutative rings*. For example, $\mathbb{Z}$ under the addition and multiplication of integers is a commutative ring while $M_2(\mathbb{Z})$ with the addition and multiplication of matrices is a non-commutative ring.

[†] In many books including Judson's, $(R, +, \cdot)$ satisfying conditions (1) & (2) is called a **ring**. A ring that satisfies (3) is called a **ring with unity**. But in Lang's book, a ring is defined to fulfill conditions (1)-(3). i.e. We follow Lang.

(iii) Immediate consequences from the definition:

1) The element $1 \in R$ satisfying (3) in the definition is unique: If $1' \in R$ also satisfies (3), then $1 = 11' = 1'$. We call the element 1 the **unity of** R.[†]
2) $0x = 0$ for all $x \in R$: indeed, $0x + x = 0x + 1x = (0+1)x = 1x = x$, so $0x = 0$.
3) $(-1)x = -x = x(-1)$ for all $x \in R$: indeed, $(-1)x + x = (-1)x + 1x = (-1+1)x = 0x = 0$, so $(-1)x = -x$. (Here -1 is the *additive inverse of the unity*, not the integer -1. e.g. Think about $R = M_2(\mathbb{Z})$, what is -1 in R?)
4) For any $x, y \in R$, $(-x)y = ((-1)x)y = (-1)(xy) = -xy$, and similarly $x(-y) = -xy$, and $(-x)(-y) = xy$.

(iv) The smallest ring is the **zero ring**, i.e., $R = \{0\}$. Recall from the definition that a ring must contain the element 0 and the element 1. A ring R is the zero ring, i.e., $R = \{0\}$, if and only if $1 = 0$ in R. Indeed, if $R = \{0\}$ then $1 = 0$, and conversely if $1 = 0$, then for any $x \in R$, $x = 1x = 0x = 0$, so $R = \{0\}$. It is the *only* ring that has one element.

8.2 Fields

In the definition of rings, only the addition is required to satisfy the abelian group structure. One would naturally consider the case where multiplication also gives abelian group structure. However unless for the zero ring, we shall never be able to give a group structure to $(R, \cdot\,)$.[‡] If we remove the element 0, then it is possible and this algebraic structure is called a *field*.

In a ring R, we denote by $R\backslash\{0\}$ the set of all non-zero elements in R.

Definition 8.2.1. A *commutative* ring $(R, +, \cdot\,)$ such that $(R\backslash\{0\}, \cdot)$ is a group is called a **field**.[§]

Definition 8.2.2. Let R be a ring. A subset R' of R is called a **subring of** R if $(R', +)$ is a subgroup of $(R, +)$, $1 \in R'$ and $xy \in R'$ $\forall\ x, y \in R'$, or equivalently $0 \in R', 1 \in R'$, and $-x, x+y, xy \in R'$ for any $x, y \in R'$.

A subset F' of a field F is called a **subfield of** F if $(F', +)$ is a subgroup of $(F, +)$ and $(F' \setminus \{0\}, \cdot)$ is a subgroup of $(F \setminus \{0\}, \cdot)$.

Example 8.2.3. (i) $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ are rings under the addition and multiplication of numbers. Indeed, $\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$ are subrings of $\mathbb{C}$. The last three $\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$ are fields but $\mathbb{Z}$ is not. $\mathbb{Z}_n$ is also a ring under the addition and multiplication in §2.2.

(ii) If R is any ring, the set $M_n(R)$ of all $n \times n$ matrices with entries in R forms a ring under the addition and multiplication of matrices. We thus have the rings $M_n(\mathbb{Z})$, $M_n(\mathbb{Q})$, $M_n(\mathbb{R})$, and $M_n(\mathbb{C})$ which are *non-commutative* unless $n = 1$.

[†] In Lang's book the element 1 is called the *unit element of* R instead of unity.

[‡] Why not? Think about what is the possible multiplicative inverse for the element 0.

[§] For a commutative ring, multiplication is by definition commutative so if $(R \setminus \{0\}, \cdot\,)$ is a group, it will be an abelian group.

(iii) Let $\mathbb{Z}[t] = \{a_0 + a_1 t + \cdots + a_n t^n : a_0, \cdots, a_n \in \mathbb{Z}, n \geq 0\}$ be the set of all polynomials with integral coefficients and constant term. Under the usual polynomial addition and multiplication, $(\mathbb{Z}[t], +, \cdot)$ is a commutative ring with unity 1 (the constant polynomial 1). Similarly, $\mathbb{Q}[t]$, $\mathbb{R}[t]$ and $\mathbb{C}[t]$ are commutative rings, but none of them is a field.

8.3 Integral domains

Definition 8.3.1. For commutative ring R, $x \in R\backslash\{0\}$ is called a **zero-divisor of** R if there exists $y \in R\backslash\{0\}$ such that $xy = 0$ *or* $yx = 0$. A non-zero *commutative* ring without zero divisors is called **integral domain** (also called **integral ring**).

In other words, a non-zero commutative ring R is an integral domain if and only if for any $x, y \in R$, $xy = 0$ implies $x = 0$ *or* $y = 0$.

Example 8.3.2. The matrix ring $M_n(\mathbb{R})$ has many zero divisors: for example,

$$\text{in } M_n(\mathbb{R}),\ \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Example 8.3.3. The ring $(\mathbb{Z}, +, \cdot)$ is an integral domain. Any non-zero subring of an integral domain is again an integral domain.

Lemma 8.3.4. *A field is an integral domain, and a finite integral domain is a field.*

Proof. Let F be a field. We already know from the definition of a field that F is a non-zero commutative ring. If $x, y \in F$ are such that $xy = 0$ but $x \neq 0$, then $x^{-1}(xy) = y = 0$. Thus F has no zero divisor, so F is an integral domain.

Assume now that R is a finite integral domain. Let $R = \{x_1, \cdots, x_n\}$ have n (distinct) elements. Suppose $x \in R\backslash\{0\}$ and consider $S = \{xx_1, \cdots, xx_n\} \subset R$. If $1 \notin S$, then S contains less than n elements. i.e. $xx_i = xx_j$ for some $i \neq j$. Consequently $x(x_i - x_j) = 0$. Since x is not a zero divisor, $x_i = x_j$ which is a contradiction, i.e. $xy = 1$ for some $y \in R$. Thus R is a field. □

Lemma 8.3.5. *The ring $\mathbb{Z}_n$, where $n \geq 1$ in an integer, is an integral domain if and only if n is prime. Consequently, $\mathbb{Z}_n$ is a field if n is prime.*

Proof. Note that for $2 \leq k \leq n-1$, $[k] \in \mathbb{Z}_n$ is a zero divisor in $\mathbb{Z}_n$ if and only if $(n, k) > 1$. Thus $\mathbb{Z}_n$ is an integral domain if and only if n is prime. By Lemma 8.3.4, $\mathbb{Z}_n$ is a field if n is prime. Indeed, when p is prime, for $1 \leq k \leq p-1$, as $(k, p) = 1$, there exist $a, b \in \mathbb{Z}$ such that $ak + bp = 1$. Then $[a][k] = [1]$. □

Exercise 8.3.6. Consider the ring of real numbers $(\mathbb{R}, +, \cdot)$ and let $\mathbb{Z}[\sqrt{2}] = \{a + b\sqrt{2} : a, b \in \mathbb{Z}\}$ Show that $\mathbb{Z}[\sqrt{2}]$ is a subring of $(\mathbb{R}, +, \cdot)$ and is an integral domain.[¶] Is it a field? Repeat the same study for $\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$.

[¶] Although $\mathbb{Z}[t]$ and $\mathbb{Z}[\sqrt{2}]$ are defined differently, indeed no ambiguity will arise. (Why?)

Chapter 9 — Ideals, Homomorphisms, Quotient Rings

The idea of ideals in the ring $\mathbb{Z}$ is introduced in Chapter 1. Let us start with this notion in general context.

9.1 Left Ideals, Right Ideals and Two-sided Ideals

Definition 9.1.1. Let R be a ring, not necessarily commutative.

1) A **left ideal** of $(R, +, \cdot)$ is a subset J of R such that J is a subgroup of the additive group $(R, +)$ and that for every $a \in R$ and $x \in J$, one has $ax \in J$.

2) A **right ideal** of $(R, +, \cdot)$ is a subset J of R such that J is a subgroup of the additive group $(R, +)$ and that for every $a \in R$ and $x \in J$, one has $xa \in J$.

3) A **two-sided ideal** of R is a subset J of R that is both a left and a right ideal.

Note that for any ring R, the subset $\{0\}$ and the whole set R are both left and right ideals of R, which will be respectively called the **zero ideal** and the **unit ideal**.

Exercise 9.1.2. A (left or right) ideal J of R is equal to R if and only if $1 \in J$.

Notation. Given a ring $(R, +, \cdot)$ and $x_1, y_1, \ldots, x_n, y_n \in R$. The element $x_1y_1 + \cdots + x_ny_n$ in R is obtained by adding the products $x_1y_1, \ldots, x_ny_n$ in R in any order.

Definition 9.1.3. Let R be a ring and let $a_1, \ldots, a_n \in R$. The subset

$$Ra_1 + Ra_2 + \cdots + Ra_n = \{x_1a_1 + x_2a_2 + \cdots + x_na_n : x_1, x_2, \ldots, x_n \in R\}$$

is a left ideal of R called the **left ideal of R generated by** $a_1, a_2, \ldots, a_n$. Similarly,

$$a_1R + a_2R + \cdots + a_nR = \{a_1x_1 + a_2x_2 + \cdots + a_nx_n : x_1, x_2, \ldots, x_n \in R\}$$

is a right ideal of R called the **right ideal of R generated by** $a_1, a_2, \ldots, a_n$. For a single element $a \in R$, Ra and aR are respectively called the **principal left ideal** and the **principal right ideal** of R generated by a.

If R is commutative, a left ideal will be a right ideal and hence a two-sided ideal. We simply call it an ideal. An ideal of the form aR is called a **principal ideal**.

9.2 Units and ideals of commutative rings

Definition 9.2.1. An element a in a *commutative ring* is said to be a **unit** [†] if there exists $b \in R$ such that $ab = 1$.

Example 9.2.2. The commutative ring $\mathbb{Z}$ has two units 1 and -1.

Exercise 9.2.3. Let R be a commutative ring and I be its ideal. Show that

(1) I is the unit ideal if and only if I contains a unit.

(2) All elements in $R\backslash\{0\}$ are units if and only if R is a field.

Exercise 9.2.4. What are the units in $\mathbb{Z}[t]$? in $\mathbb{Q}[t]$? in $\mathbb{R}[t]$? in $\mathbb{C}[t]$?

Definition 9.2.5. An integral domain is called a **Principal ideal domain** or a **PID** in short if every ideal of R is principal.

Theorem 9.2.6. *Every ideal in $(\mathbb{Z}, +, \cdot)$ is of the form $n\mathbb{Z} = \{nx : x \in \mathbb{Z}\}$ for a unique $n \in \mathbb{Z}$ and $n \geq 0$. In particular, $(\mathbb{Z}, +, \cdot)$ is a PID.*

Proof. Proved in Chapter 1, §3. □

Now we impose the operations of *sum* and *product* on ideals in *commutative rings.*

Lemma 9.2.7. *Let R be a commutative ring. For two ideals $I, J \subset R$, define*

$$I + J = \{x + y : x \in I, y \in J\}$$
$$IJ = \{x_1y_1 + \cdots + x_ny_n : \ n \in \mathbb{N},\ x_j \in I, y_j \in J\}.$$

Then both $I+J$ and IJ are ideals of R, and so is $I \cap J$. Moreover, $IJ \subset I \cap J \subset I+J$.

Proof. The statements are proved by following directly the definition. For example, to prove that IJ is an ideal, first note that $0 \in IJ$ and that if $a = x_1y_1 + \cdots + x_ny_n$ and $a' = x'_1y'_1 + \cdots + x'_my'_m \in IJ$, then $-a = (-x_1)y_1 + \cdots + (-x_n)y_n \in IJ$ and $a + a' = x_1y_1 + \cdots + x_ny_n + x'_1y'_1 + \cdots + x'_my'_m \in IJ$, so IJ is a subgroup of the additive group $(R, +)$. Moreover, for any $x \in R$ and $a = x_1y_1 + \cdots + x_ny_n$, one has $xa = (xx_1)y_1 + \cdots + (xx_n)y_n \in IJ$. This shows that IJ is an ideal. □

Example 9.2.8. Let $R = (\mathbb{Z}, +, \cdot)$. We already know that every ideal is of the form $m\mathbb{Z} = \{mx : x \in \mathbb{Z}\}$ for some $m \in \mathbb{Z}$. We may always choose $m \geq 0$, and $m\mathbb{Z}$ is the zero ideal if and only if $m = 0$, and $m\mathbb{Z}$ is the unit ideal if and only if $m = 1$. Let $m, n \in \mathbb{Z}$ and $m, n > 0$. It follows from the definitions that

1) $m\mathbb{Z} \subset n\mathbb{Z}$ if and only if $n|m$;

2) $(m\mathbb{Z}) + (n\mathbb{Z}) = r\mathbb{Z}$, where $r = \gcd(m, n)$;

3) $(m\mathbb{Z}) \cap (n\mathbb{Z}) = q\mathbb{Z}$, where $q = \operatorname{lcm}(m, n)$;

4) $(m\mathbb{Z})(n\mathbb{Z}) = mn\mathbb{Z}$.

[†] In Lang's book, 1 is called the *unit element*, so don't mix up *unit* and *unit element.*

9.3 Homomorphisms and isomorphisms

Definition 9.3.1. Let R and R' be two rings with unities $1 \in R$ and $1' \in R'$.

1) By a **ring homomorphism** from R to R' we mean a map $f : R \to R'$ such that $f(1) = 1'$, $f(x+y) = f(x) + f(y)$ and $f(xy) = f(x)f(y)$, $\forall\, x, y \in R$.
2) A ring homomorphism $f : R \to R'$ is said to be a **ring isomorphism** if f is also bijective.
3) Two rings R and R' are said to be **isomorphic**, and written as $R \cong R'$, if there is a ring isomorphism from R to R'.
4) A ring isomorphism from R to itself is called an **automorphism** of R.

The following is easy to prove from the definitions:

Lemma 9.3.2. *1) For any ring homomorphism $f : R \to R'$, the kernel of f, defined as $\ker(f) = \{x \in R : f(x) = 0\}$, is a two sided ideal of R;*

2) If $f : R \to R'$ and $g : R' \to R''$ are ring homomorphisms, so is their composition $g \circ f : R \to R'' : x \mapsto g(f(x))$ for $x \in R$.

3) If $f : R \to R'$ is a ring isomorphism, so is its inverse $f^{-1} : R' \to R$.

Proof. Exercise. □

Example 9.3.3. For any integer $n \geq 1$, the map $\mathbb{Z} \to \mathbb{Z}_n, x \mapsto [x]$ for $x \in \mathbb{Z}$, is a ring homomorphism whose kernel is the ideal $n\mathbb{Z}$.

Example 9.3.4. The map $f : \mathbb{Q}[t] \to \mathbb{Q}$, $f(p) = p(2)\ \forall\, p \in \mathbb{Q}[t]$, is a ring homomorphism whose kernel is the ideal $\langle t - 2 \rangle$.[†]

9.4 Quotient (or Factor) Rings

Let R be a ring and let I be a two-sided ideal. Define a relation on R using I, by setting $x \equiv y \pmod I$ if $x - y \in I$. When it is clear which ideal I we are talking about, we simply write $x \equiv y$ in place of $x \equiv y \pmod I$.

The proof of the statements in the following Lemma 9.4.1 are all straightforward and we omit the details.

Lemma 9.4.1. *For any two-sided ideal I of R, we have*

1) $\equiv$ is an equivalence relation on R;

2) for any $x, y, z \in R$, if $x \equiv y$, then $zx \equiv zy$ and $xz \equiv yz$;

3) if $x \equiv x'$ and $y \equiv y'$, then $xy \equiv x'y'$ and $x + y \equiv x' + y'$.

[†] To see it, you may use Lemma 11.2.12.

Lemma 9.4.2. *Let R be a ring and $I \subset R$ a two-sided ideal. Let R/I be the set of equivalence classes with respect to the equivalence relation "$\equiv \pmod{I}$", and for $x \in R$, denote by $\bar{x}$ the equivalence class of x. Define two operations $+$ and $\cdot$ on R/I by*

$$\bar{x} + \bar{y} = \overline{x+y}, \quad \bar{x}\bar{y} = \overline{xy}, \quad x, y \in R.$$

Then $(R/I, +, \cdot)$ is a ring. Moreover, the map

$$p: \quad R \longrightarrow R/I, \quad p(x) = \bar{x}, \quad x \in R,$$

is a surjective ring homomorphism, and $\ker p = I$.

Remark. The equivalence class $\bar{x}$ equals $x + I = \{x + a : a \in I\}$, cf. Section 4.2.

Proof. We need to prove that the two operations "+" and "$\cdot$" are well-defined and that they satisfy all the axioms in the definition of a ring. Note that $(R/I, +)$ is nothing but the (abelian) quotient group of $(R, +)$ by I regarded as a normal subgroup of $(R, +)$. To show, for example, that "$\cdot$" is well-defined, suppose that $\bar{x} = \bar{x}'$ and $\bar{y} = \bar{y}'$, then $x \equiv x'$ and $y \equiv y'$. By 3) of Lemma 9.4.1, $xx' \equiv yy'$, so $\overline{xx'} = \overline{yy'}$. This shows that "$\cdot$" is well-defined. We leave the detail of rest of the proof as an exercise. □

Definition 9.4.3. Let R be a ring and $I \subset R$ a two-sided ideal. The ring $(R/I, +, \cdot)$ is called the **quotient** (or **factor**) **ring of R by I**. We call the ring homomorphism $p : R \to R/I$, $x \mapsto \overline{x}$ the **natural projection**.

We have the following result analogous to Theorem 4.2.5.

Lemma 9.4.4. *Let $f : R \to R'$ be a ring homomorphism. Then $f(R)$ is a subring of R', $\ker(f)$ is a two-sided ideal of R, and the map*

$$\bar{f}: \quad R/\ker(f) \longrightarrow f(R), \quad \bar{x} \longmapsto f(x), \quad x \in R,$$

is a well-defined ring isomorphism.

Proof. One first uses the definition to check that $f(R)$ is a subring of R'. If $x, x' \in R$ are such that $\bar{x} = \bar{x}'$, then $x - x' \in \ker(f)$, i.e., $f(x - x') = 0$, so $f(x) = f(x')$. This shows that the map $\bar{f}$ is well-defined. Clearly $\bar{f}$ is surjective. It follows from f being a ring homomorphism that $\bar{f}$ is a ring homomorphism. Finally, if $\bar{f}(\bar{x}) = \bar{f}(\bar{x}')$ for $x, x' \in R$, then $f(x) = f(x')$, so $x - x' \in \ker(f)$, or $\bar{x} = \bar{x}'$. This shows that $\bar{f}$ is injective. We have thus proved that $\bar{f}$ is an isomorphism. □

Example 9.4.5. For any integer $n \geq 1$, $\mathbb{Z}_n$ is the quotient ring of $(\mathbb{Z}, +, \cdot)$ by the ideal $n\mathbb{Z} \subset \mathbb{Z}$, because every $\bar{x} \in \mathbb{Z}/n\mathbb{Z}$ is the equivalence class $[x]$ with respect to the relation $\equiv$ mod n in Chapter 1.

Example 9.4.6. Consider the ring $\mathbb{Z}_{10} = \{\bar{0}, \bar{1}, \dots, \bar{9}\}$ with the ideal $I = \{\bar{0}, \bar{5}\}$. For $x \in \mathbb{Z}_{10}$, let $x + I = \{x + a : a \in I\}$ be the equivalence class of x for the relation "$\equiv \pmod{I}$" as a subset of $\mathbb{Z}_{10}$, and let $[x]$ be the equivalence class of x as a point in R/I. Then we have

$$\begin{aligned}
&\bar{0} + I = \bar{5} + I = \{\bar{0}, \bar{5}\}, \quad \bar{1} + I = \bar{6} + I = \{\bar{1}, \bar{6}\}, \quad \bar{2} + I = \bar{7} + I = \{\bar{2}, \bar{7}\},\\
&\bar{3} + I = \bar{8} + I = \{\bar{3}, \bar{8}\}, \quad \bar{4} + I = \bar{9} + I = \{\bar{4}, \bar{9}\}.
\end{aligned}$$

Thus the ring $R/I = \{[\bar{0}] = [\bar{5}], [\bar{1}] = [\bar{6}], [\bar{2}] = [\bar{7}], [\bar{3}] = [\bar{8}], [\bar{4}] = [\bar{9}]\}$ has five elements. It is straightforward to check that the map

$$R/I \longrightarrow \mathbb{Z}_5 : \quad [\bar{i}] \longmapsto \bar{i}, \qquad i = 0, 1, 2, 3, 4,$$

is a ring isomorphism.

Example 9.4.7. Let X be any subset of $\mathbb{R}^n$, and let R be the set of all real-valued functions on X with the ring operations $(f+g)(x) = f(x) + g(x)$ and $(fg)(x) = f(x)g(x)$ for $f, g \in R$ and $x \in X$. Let $x_0 \in X$ and let $I = \{f \in R : f(x_0) = 0\}$. We know that I is an ideal. The map

$$\mathrm{ev}_{x_0} : \quad R \longrightarrow \mathbb{R}, \quad \phi(f) = f(x_0), \qquad f \in R,$$

is a surjective ring homomorphism, and $\ker(\mathrm{ev}_{x_0}) = I$. Thus $R/I \cong (\mathbb{R}, +, \cdot)$.

Example 9.4.8. Consider the ring $\mathbb{Q}[t]$ with the ideal $I = \langle t \rangle$. For $p \in \mathbb{Q}[t]$,

$$\bar{p} = p + I \overset{\text{check it!}}{=\!=\!=\!=} p_0 + I$$

where p_0 is the constant term of the polynomial p. Moreover, one may check that $\mathbb{Q}[t]/I \longrightarrow \mathbb{Q} : \ \bar{p} \mapsto p_0$ is a ring isomorphism.

Exercise 9.4.9. Consider the map $f : \mathbb{R}[t] \longrightarrow \mathbb{C}$, $f(p) = p(i)$. Show that (i) f is a surjective ring homomorphism, (ii) $\ker f = \langle t^2 + 1 \rangle$, (iii) $\mathbb{R}[t]/\langle t^2 + 1 \rangle \cong \mathbb{C}$.

9.5 Characteristic of a Ring

Let R be any ring. For $a \in R$ and any integer $m \geq 1$, define

$$ma = \overbrace{a + a + \cdots + a}^{m} \qquad \text{and} \qquad (-m)a = -(ma),$$

and define $0a = 0$. This way gives a map $\phi_a : \mathbb{Z} \to R$ given by $\phi_a(m) = ma$ for any $m \in \mathbb{Z}$. We always have $\phi_a(m + n) = \phi_a(m) + \phi_a(n)$, but $\phi_a(mn) = \phi_a(m)\phi_a(n)$ for any $m, n \in \mathbb{Z}$ and $\phi_a(1) = 1$ only when $a = 1$. Thus $\phi_1 : \mathbb{Z} \to R$ is a ring homomorphism. As $\ker(\phi_1)$ is an ideal of $\mathbb{Z}$, we must have $\ker(\phi_1) = n\mathbb{Z}$ for a unique $n \geq 0$, and

$$\bar{\phi}_1 : \quad \mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z} \longrightarrow \mathrm{Im}(\phi_1) = \{n1 \in R : n \in \mathbb{Z}\}$$

is a ring isomorphism.

Definition 9.5.1. The unique integer $n \geq 0$ such that $\ker(\phi_1) = n\mathbb{Z}$ is called the **characteristic of** R and we write $\operatorname{char}(R) = n$.

REMARK. There are two cases:

Case 1. R having characteristic 0 means that $\ker(\phi_1) = \{0\}$, i.e., $m1 \neq 0$ for any $m \in \mathbb{Z}$, $m \neq 0$. In this case, the map $\phi_1 : \mathbb{Z} \to R$ is injective, so $\operatorname{Im}(\phi_1) = \{m1 : m \in \mathbb{Z}\}$ is a "copy of $\mathbb{Z}$" sitting inside R;

Case 2. R having characteristic $n > 0$ means that $k1 \neq 0$ for all $1 \leq k \leq n-1$ but $n1 = 0$. In this case, $na = n(1a) = (n1)a = 0$ for all $a \in R$. Note also that R has characteristic 1 if and only if R is the zero ring.

(This gives an alternative way to define the characteristics of a ring.)

Example 9.5.2. For any integer $n \geq 1$, the ring $\mathbb{Z}_n$ has characteristic n.

Lemma 9.5.3. *If R is a non-zero ring with no zero divisor, then* $\operatorname{char}(R)$ *is either* 0 *or a prime number.*

Proof. Let R be a non-zero ring with no zero divisor and let $n = \operatorname{char}(R)$. Assume that $n \neq 0$. We need to show that n must be a prime number. By the definition of $\operatorname{char}(R)$, $k1 \neq 0$ for all $1 \leq k \leq n-1$ and $n1 = 0$. If n is not prime, then $n = kl$ for some $1 \leq k, l \leq n-1$. It follows from $n1 = (k1)(l1)$ that $(k1)(l1) = 0$, so $k1 \neq 0$ and $l1 \neq 0$ are zero divisors, contradicting the assumption that R has no zero divisor. Thus n must be prime. □

Corollary 9.5.4. *The characteristic of any integral domain, and in particular any field, is either* 0 *or a prime number.*

Example 9.5.5.

1) The ring $(\mathbb{Z}, +, \cdot)$ has characteristic zero, and so does the fields $\mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$.

2) For a ring R, any subring R' of R has the same characteristic as R.

3) If R is a ring, then the ring $M_n(R)$ of all $n \times n$ matrices with coefficients in R has the same characteristic as R.

Chapter 10 — Prime ideals and Maximal ideals

10.1 Prime ideals and Maximal ideals

The definition of *prime ideals* imitates the notion of prime numbers.

Definition 10.1.1. Let R be a commutative ring.

1) An ideal I of R is called a **prime ideal** if $I \neq R$ and if whenever $x, y \in R$ are such that $xy \in I$ one has $x \in I$ or $y \in I$.

2) An ideal I of R is called a **maximal ideal** if $I \neq R$ and if there is no ideal J of R such that $I \subset J \subset R$ and $J \neq I$ and $J \neq R$.

Example 10.1.2. Consider the ring $(\mathbb{Z}, +, \cdot)$ and we know that every ideal I of $\mathbb{Z}$ is of the form $I = n\mathbb{Z}$ for a unique $n \in \mathbb{Z}$ and $n \geq 0$. Let's consider first the zero ideal $\{0\} = 0\mathbb{Z}$. If $x, y \in \mathbb{Z}$ are such that $xy = 0$ then $x = 0$ or $y = 0$, so $\{0\}$ is a prime ideal. As $\{0\} \subset 2\mathbb{Z} \subset \mathbb{Z}$, $\{0\}$ is not a maximal ideal. Assume now that $n \neq 0$. If $n = 1$, we have $1\mathbb{Z} = \mathbb{Z}$ which is not prime nor maximal by definition. Assume thus $n > 1$. Then $n\mathbb{Z}$ is a prime ideal if and only if it has the property that for any integers x, y, $n|(xy)$ if and only if $n|x$ or $n|y$. This happens if and only if n is prime. We thus conclude that for $n > 1$, the ideal $n\mathbb{Z}$ is a prime ideal if and only if n is a prime number.

Exercise 10.1.3. Let R be a non-zero commutative ring. Prove that R is an integral domain if and only if the zero ideal $\{0\}$ of R is a prime ideal.

10.2 Properties of prime ideals and maximal ideals

Lemma 10.2.1. *Let R be a commutative ring and $I \subset R$ an ideal such that $I \neq R$. Then I is a maximal ideal if and only if $I + xR = R$ for every $x \in R$ and $x \notin I$.*

Proof. Assume first that I is a maximal ideal. Then for any $x \in R$ and $x \notin I$, $I + xR$ is an ideal and $I \subset I + xR \subset R$. As $I + xR \neq I$, we must have $I + xR = R$. Conversely, assume that $I + xR = R$ for every $x \in R$ and $x \notin I$. Let J be any ideal of R such that $I \subset J \subset R$ and assume that $J \neq I$. Then there exists $x \in J$ and $x \notin I$. Then $R = I + xR \subset I + J \subset J$, so $J = R$. This shows that I is maximal. □

Proposition 10.2.2. *Let R be a commutative ring and I any ideal of R such that $I \neq R$. We have*

1) I is a prime ideal if and only if R/I is an integral domain;

2) I is a maximal ideal of and only if R/I is a field.

Proof. 1) As $I \neq R$, R/I is a non-zero commutative ring. Assume first that I is a prime ideal. To prove that R/I is an integral domain, suppose that $x, y \in R$ are such that $\bar{x}\bar{y} = \bar{0} \in R/I$. Then $\overline{xy} = \bar{0} \in R/I$, so $xy \in I$. By the definition of I being a prime ideal, either $x \in I$ or $y \in I$, so either $\bar{x} = \bar{0}$ or $\bar{y} = \bar{0}$. Thus R/I is an integral domain.

Conversely, assume that R/I is an integral domain. If $x, yı‘nR$ are such that $xy \in I$, then $\bar{x}\bar{y} = \bar{0} \in R/I$. Since R/I is an integral domain, either $\bar{x} = \bar{0}$ or $\bar{y} = \bar{0}$, i.e., either $x \in I$ or $y \in I$. Thus I is a prime ideal.

2) Assume first that I is a maximal ideal. As $I \neq R$, we know that R/I is a non-zero commutative ring. We need to show that every non-zero element in R/I has an inverse. Let thus $x \in R$ be such that $\bar{x} \neq \bar{0}$, i.e., $x \notin I$. By Lemma 10.2.1, $I + xR = R$, so $1 \in I + xR$, i.e., there exist $y \in R$ and $z \in I$ such that $1 = z + xy$. Then $xy - 1 \in I$, so $\bar{x}\bar{y} = \bar{1}$. This shows that $\bar{y}$ is the inverse of $\bar{x}$ in R/I. We have thus proved that every non-zero element in R/I has an inverse, so R/I is a field.

Conversely, assume that R/I is a field. We need to show that I is a maximal ideal. Assume that J is an ideal of R such that $I \subset J$ and $J \neq I$. We need to show that $J = R$. It is enough to show that $1 \in J$. Take any $x \in J$ and $x \notin I$. Then $\bar{x}$ is a non-zero element in R/I. As R/I is a field, the non-zero element $\bar{x}$ must have an inverse, i.e., there exists $y \in R$ such that $\bar{x}\bar{y} = \bar{1}$, which means that $xy - 1 \in I$. Since J is an ideal and $x \in J$ and $I \subset J$, we have $1 = xy - (xy - 1) \in J$. This shows that $J = R$. By definition, I is maximal. □

Corollary 10.2.3. *A maximal ideal is necessarily a prime ideal.*[†]

Example 10.2.4. We continue with Example 10.1.2. We have shown that the prime ideals of $\mathbb{Z}$ are precisely all ideals of the form $n\mathbb{Z}$, where $n = 0$ or n is a prime number. Among these, $\{0\}$ is not maximal, but if n is a prime number, since $\mathbb{Z}/n\mathbb{Z}$ is a field, $n\mathbb{Z}$ is a maximal ideal. Since every maximal ideal must be prime, we conclude that the maximal ideals of $\mathbb{Z}$ are precisely all ideals of the form $n\mathbb{Z}$, where n is a prime number.

Example 10.2.5. If R is a field, then R has only two ideals, the zero ideal and R itself. By definition, R, as the unit ideal, is neither prime nor maximal. The zero ideal, on the other hand, is maximal and thus also prime.

[†] Naturally one wonders under what circumstances all prime ideals are maximal. If R is an integral domain, then it is easy to answer when all its prime ideals are maximal. (Hint: Use Proposition 10.2.2 and the fact "the zero ideal is a prime ideal".) Thus one considers when all *nonzero* prime ideals are maximal. Later we shall give a sufficient (but not necessary) condition.

10.3 Prime and maximal ideals of the ring $\mathbb{Z}_n$

For an integer $n \geq 2$, we now describe all prime and maximal ideals of the ring $\mathbb{Z}_n$. We first need to determine all different ideals of $\mathbb{Z}_n$. Let $D_n = \{1 \leq m \leq n : m|n\}$, and for $m \in D_n$, let $\bar{m}\mathbb{Z}_n = \{x\bar{m} : x \in \mathbb{Z}_n\}$ be the principal ideal of $\mathbb{Z}_n$ generated by $\bar{m} \in \mathbb{Z}_n$. Note that $\bar{1}\mathbb{Z}_n = \mathbb{Z}_n$ is the unit ideal and $\bar{n}\mathbb{Z}_n = \{\bar{0}\}$ is the zero ideal.

Lemma 10.3.1.

1) For every $m \in D_n$, $\overline{m}\mathbb{Z}_n = \{\bar{0},\ \overline{m},\ \overline{2m},\ \ldots, \overline{(k-1)m}\}$, where $k = n/m$.

2) Every ideal I of $\mathbb{Z}_n$ is equal to $\overline{m}\mathbb{Z}_n$ for some $m \in D_n$, and $\overline{m}\mathbb{Z}_n \neq \overline{m'}\mathbb{Z}_n$ if $m, m' \in D_n$ and $m \neq m'$; consequently, the number of different ideals of $\mathbb{Z}_n$ is equal to the number of positive divisors of n.

Proof. To prove 1), let $x = \bar{l} \in \mathbb{Z}_n$ be arbitrary, where $0 \leq l \leq n-1$. Let $0 \leq c \leq k-1$ be the remainder of l when divided by k, i.e., $l = bk+c$ for some $b \in \mathbb{Z}$. Since $k\overline{m} = \bar{0}$, we have $x\overline{m} = c\overline{m} = \overline{cm}$. This shows that $\overline{m}\mathbb{Z}_n \subset \{\bar{0},\ \overline{m},\ \overline{2m},\ \ldots, \overline{(k-1)m}\}$. As $\overline{m}\mathbb{Z}_n$ is an ideal, one has $\{\bar{0},\ \overline{m},\ \overline{2m},\ \ldots, \overline{(k-1)m}\} \subset \overline{m}\mathbb{Z}_n$. This proves 1).

To prove 2), let I be any non-zero ideal of $\mathbb{Z}_n$ and let m be the smallest integer between 1 and n such that $\overline{m} \in I$. Then $\overline{m}\mathbb{Z}_n \subset I$. If a is any integer such that $\overline{a} \in I$, by writing $a = qm + r$ for some $q \in \mathbb{Z}$ and $0 \leq r \leq m-1$, we see that $\overline{r} \in I$, so $r = 0$ by the definition of m, and thus $m|a$ and $\overline{a} = \overline{qm} \in \overline{m}\mathbb{Z}_n$. Thus shows that $I = \overline{m}\mathbb{Z}_n$. By taking $a = n$ we see that $m|n$, so $m \in D_n$. As $|\overline{m}\mathbb{Z}_n| = n/m$ for every $m \in D_n$, if $m' \in D_n$ is different from m, then $n/m \neq n/m'$, so $\overline{m}\mathbb{Z}_n \neq \overline{m'}\mathbb{Z}_n$. □

For the prime and maximal ideals of the ring $\mathbb{Z}_n$, we have

Lemma 10.3.2. *For an integer $n \geq 2$, an ideal I of $\mathbb{Z}_n$ is a prime ideal if and only if $I = \overline{m}\mathbb{Z}_n$, where m is a prime number that divides n; Every prime ideal of $\mathbb{Z}_n$ is also a maximal ideal. (Recall a prime number must be at least 2 by definition).*[‡]

Proof. We know from Lemma 10.3.1 that the assignment $m \to \overline{m}\mathbb{Z}_n$ is a one to one correspondence between the set of all positive divisors of n and the set of all different ideals of $\mathbb{Z}_n$. Let $m \geq 1$ be a divisor of n. It remains to determine when $\overline{m}\mathbb{Z}_n$ is prime and when $\overline{m}\mathbb{Z}_n$ is maximal. As the unit ideal $\mathbb{Z}_n$ is not prime nor maximal by definition, we assume that $m > 1$. Since $m|n$, one readily checks that the map

$$\phi : \quad \mathbb{Z}_n \longrightarrow \mathbb{Z}_m, \quad \mathbb{Z}_n \ni \bar{x} \longmapsto \bar{x} \in \mathbb{Z}_m, \quad x \in \mathbb{Z},$$

is a well-defined ring homomorphism, and that $\ker(\phi) = \overline{m}\mathbb{Z}_n$. Thus $\mathbb{Z}_n/(\overline{m}\mathbb{Z}_n) \cong \mathbb{Z}_m$ as rings. By Proposition 10.2.2, $\overline{m}\mathbb{Z}_n$ is a prime (resp. maximal) ideal of $\mathbb{Z}_n$ if and only if $\mathbb{Z}_m$ is an integral domain (resp. a field). Since we are assuming that $m \geq 2$, we know that $\mathbb{Z}_m$ is an integral domain if and only if m is prime number, and in such a case $\mathbb{Z}_m$ is also a field. This finishes the proof of Lemma 10.3.2. □

[‡] Example: By Lemmas 10.3.1-2, there are precisely 6 different ideals of $\mathbb{Z}_{12}$, namely, $\bar{1}\mathbb{Z}_{12} = \mathbb{Z}_{12}$, $\bar{2}\mathbb{Z}_{12} = \{\bar{0},\ \bar{2},\ \bar{4},\ \bar{6},\ \bar{8},\ \overline{10}\}$, $\bar{3}\mathbb{Z}_{12} = \{\bar{0},\ \bar{3},\ \bar{6},\ \bar{9}\}$, $\bar{4}\mathbb{Z}_{12} = \{\bar{0},\ \bar{4},\ \bar{8}\}$, $\bar{6}\mathbb{Z}_{12} = \{\bar{0},\ \bar{6}\}$, $\overline{12}\mathbb{Z}_{12} = \{\bar{0}\}$, and among which, $\bar{2}\mathbb{Z}_{12}$ and $\bar{3}\mathbb{Z}_{12}$ are prime ideals, and these two are also all the maximal ideals.

Chapter 11 — Polynomials

11.1 Polynomial rings over a commutative ring

§ 1. The polynomial ring $R[t]$

Let R be a commutative ring and t denote an indeterminate. Consider the set

$$R[t] = \{f = a_0 + a_1t + a_2t^2 + \cdots + a_nt^n :\ n \in \mathbb{N}_0,\ a_0, a_1, \ldots, a_n \in R\}.$$

For $f = a_0 + a_1t + a_2t^2 + \cdots + a_nt^n \in R[t]$, the elements $a_0, a_1, \ldots, a_n$ are called the **coefficients of** f, a_0 is also called the **constant term**, and if $a_n \neq 0$, we say that f has **degree** n and write $n = \deg(f)$, and we also call a_n the **leading coefficient** of f. We define that the degree of the zero polynomial is $-\infty$. Two elements in $R[t]$ are said to be equal if they have the same degrees and if all their coefficients are equal. Besides, let $f = a_0 + a_1t + a_2t^2 + \cdots + a_nt^n$ and $g = b_0 + b_1t + b_2t^2 + \cdots + b_mt^m$, define

$$\begin{aligned} f + g &= (a_0 + b_0) + (a_1 + b_1)t + \cdots \qquad \text{(a finite sum)}, \\ fg &= c_0 + c_1t + \cdots + c_{m+n}t^{m+n} \quad \text{where} \quad c_k = a_0b_k + a_1b_{k-1} + \cdots + a_kb_0 \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \text{for } 0 \leq k \leq m + n. \end{aligned}$$

Then $(R[t], +, \cdot)$ is a commutative ring whose unity is 1 (the unity of R) and whose units are the constant polynomials u where u is a unit in R. (Check it!) The ring $R[t]$ is called the **polynomial ring over** R **with one variable**, and elements in $R[t]$ are called **polynomials in one variable with coefficients in** R.

Remark. If R is an integral domain, then $R[t]$ is an integral domain. (Exercise.)

§ 2. Polynomials vs Polynomial Functions

Each $f = a_0 + a_1t + a_2t^2 + \cdots + a_nt^n \in R[t]$ defines an R-valued **polynomial function** on R, namely the map

$$f_R : R \longrightarrow R, \qquad \alpha \longmapsto a_0 + a_1\alpha + a_2\alpha^2 + \cdots + a_n\alpha^n,$$

which maps $\alpha \in R$ to the element $f_R(\alpha) = a_0 + a_1\alpha + a_2\alpha^2 + \cdots + a_n\alpha^n$ in R.

Remark. It is important to distinguish between the polynomial $f \in R[t]$ and the polynomial function f_R.

Example 11.1.1. Consider, for example, $R = \mathbb{Z}_p$ where p is a prime number. We know that R is a field, so $R \backslash \{0\}$ is a group of order $p-1$ under multiplication. Thus every non-zero element $\alpha \in R$ satisfies $\alpha^{p-1} = 1$, and thus also $\alpha^p = \alpha$. Clearly the zero element also satisfies $\alpha^p = \alpha$. Thus the two *different polynomials* $f = t^p$ and $g = t$ give rise to the *same polynomial function* on R.

Later on we will show (in Example 11.2.14) that *if R is an infinite field, then two different polynomials $f, g \in R[t]$ do give rise to two different polynomial functions.*

11.2 The polynomial ring $K[t]$ over a field

Exercise 11.2.1.

(a) Let K be any field. Prove that for any $f, g \in K[t]$, $\deg(fg) = \deg(f) + \deg(g)$.

(b) Consider the ring $\mathbb{Z}_4[t]$. Evaluate $\deg(fg)$ where $f = 2t^2+1$, $g = 2t^2 - 1 \in \mathbb{Z}_4[t]$. What do you observe?

§ 1. Division algorithm for $K[t]$

Theorem 11.2.2. (Division algorithm)[†] *Let f, g be polynomials over a field K and assume that* $\deg(g) \geq 0$. *Then there exist unique polynomials q, r in $K[t]$ such that*

$$f = qg + r$$

and $\deg(r) < \deg(g)$.

Proof. Write $f = a_n t^n + \cdots + a_0$ and $g = b_m t^m + \cdots + b_0$, where $a_n \neq 0$ and $b_m \neq 0$, so that $\deg(f) = n$ and $\deg(g) = m$. If $n < m$, let $q = 0$ and $r = f$, and we are done. Assume that $n \geq m$. Consider

$$f_1 = f - a_n b_m^{-1} t^{n-m} g.$$

Then $\deg(f_1) < \deg(f)$. Continue this way to find $q_1, r \in K[t]$ such that $f_1 = q_1 g + r$ with $\deg(r) < \deg(g)$. Then

$$f = (a_n b_m^{-1} t^{n-m} + q_1)g + r$$

so g and r are as required. To prove uniqueness, suppose that $f = q_1 g + r_1 = q_2 g + r_2$ with $\deg(r_1) < \deg(g)$ and $\deg(r_2) < \deg(g)$. Then $(q_1 - q_2)g = r_2 - r_1$. By looking at the degree on each side and using the fact that $K[t]$ is an integral domain, we see that $q_1 = q_2$ and $r_1 = r_2$.[‡] □

[†] Often, e.g. in Lang's book, it is called the **Euclidean algorithm**.

[‡] If $q_1 \neq q_2$, then $q_1 - q_2$ is a nonzero polynomial, thus $\deg(g(q_1 - q_2)) = \deg(g) + \deg(q_1 - q_2) \geq \deg(g)$. However, $\deg(r_2 - r_1) \leq \max(\deg(r_1), \deg(r_2)) < \deg(g)$. Contradiction arises and so $g_1 = g_2$. Consequently $r_1 = r_2$.

Example 11.2.3. Evaluate the quotient and remainder in the division of $4t^2+3$ by $3t+4$ in the polynomial ring $\mathbb{Z}_5[t]$. Here we abbreviate 4 for $\overline{4}$ (or [4]) in $\mathbb{Z}_5$, and etc., for other coefficients.

From $3\cdot 3 = 4$ in $\mathbb{Z}_5$, $4t^2+3 = 3t(3t+4)+(3t+3) = (3t+1)(3t+4)+4$. The quotient is $3t+1$ and the remainder is 4.

Proposition 11.2.4. *Let K be a field. Then $K[t]$ is a principal ideal domain.*

See Corollary 13.1.3 for a proof of Proposition 11.2.4.

Let $f(x), g(x) \in K[x]$ where K is a field. A *monic* polynomial $d(x) \in K[x]$ is called a **greatest common divisor** of $f(x)$ and $g(x)$ if

(i) $d(x)|f(x)$ and $d(x)|g(x)$,

(ii) if $h(x) \in K[x]$ divides both $f(x)$ and $g(x)$, then $h(x)|d(x)$.

We denote $d(x)$ by $\gcd(f(x), g(x))$ or simply $(f(x), g(x))$. The polynomials $f(x)$ and $g(x)$ are **relatively prime** if $(f(x), g(x)) = 1$.

Proposition 11.2.5. *Let K be a field, $f(x), g(x) \in K[x]$. There exist polynomials $\alpha(x), \beta(x) \in K[x]$ such that $(f(x), g(x)) = \alpha(x)f(x) + \beta(x)g(x)$.*

We leave the proof of Propositions 11.2.4 and 11.2.5 as exercises.[§]

Example 11.2.6. $\mathbb{Z}[t]$ is *not* a PID (though $\mathbb{Z}[t]$ is an integral domain, $\mathbb{Z}$ is a PID).[††]

§ 2. Roots of polynomials in $K[t]$

Definition 11.2.7. Let K be a field. Let $f \in K[t]$ and let $f_K : K \to K$ be the corresponding polynomial function on K. An element $\alpha \in K$ is called a **root** of f if $f_K(\alpha) = 0$.

Example 11.2.8. The polynomial $t^2+1 \in \mathbb{R}[t]$ has no roots, while the polynomial $t^2+1 \in \mathbb{C}[t]$ has two roots, namely i and $-i$.

Example 11.2.9. Let $K = \mathbb{Z}_p$ be the field of p elements, where p is a prime number. Then $\alpha^p = \alpha$ for any $\alpha \in \mathbb{Z}_p$. Consider $f = t^p - 1 \in K[t]$. Then for every $\alpha \in K$, $f_K(\alpha) = \alpha^p - 1 = 0$, so α is a root of f if and only if $\alpha = 1$. In other words, the polynomial $f = t^p - 1$ has only one root, namely $\alpha = 1$. On the other hand, the polynomial $g(t) = t^p - t$ has p roots, namely, every element in $\mathbb{Z}_p$ is a root of g.

Definition 11.2.10. If a field K has the property that every non-constant polynomial has a root in K, we say that K is **algebraically closed**.

[§] Hint: For any non-zero ideal I of $K[t]$, any non-zero element in I of the smallest degree is a generator for I. cf. Theorem 1.2.9.

[††] Indeed, consider $I = \langle 2, t\rangle = 2R + tR$ (here $R = \mathbb{Z}[t]$) which consists of all polynomials over $\mathbb{Z}$ whose constant terms are even integers. If $I = \langle f\rangle$, then $I \ni 2 = f(t)g(t)$ for some $g \in \mathbb{Z}[t]$, so $f(t) = \pm 2$ or $f(t) = \pm 1$, and in either case $I \neq fR$. (Why? Note: $t \notin 2\mathbb{Z}[t]$ and $I \neq \mathbb{Z}[t]$.)

Example 11.2.11.

1. The field $\mathbb{R}$ of real numbers is *not* algebraically closed, while by the *fundamental theorem of algebra*, the field $\mathbb{C}$ of complex numbers is algebraically closed.
2. The polynomial $t^2 - 2 \in \mathbb{Q}[t]$ has no root in $\mathbb{Q}$ so $\mathbb{Q}$ is not algebraically closed.
3. Consider the field $K = \mathbb{Z}_p = \{\bar{0}, \bar{1}, \ldots, \overline{p-1}\}$. The polynomial
$$f = \bar{1} + (t - \bar{0})(t - \bar{1}) \cdots (t - \overline{p-1}) \in \mathbb{Z}_p[t]$$
has no roots in $\mathbb{Z}_p$, so $\mathbb{Z}_p$ is not algebraically closed.

Lemma 11.2.12. *Let K be a field and let $f \in K[t]$ be a non-zero polynomial. If $\alpha \in K$ is a root of f, then there exists $q \in K[t]$ such that*
$$f = (t - \alpha)q.$$

Proof. By the Euclidean algorithm, there exist $q \in K[t]$ and a constant $r \in K$ such that $f = (t - \alpha)q + r$. It follows from $f_K(\alpha) = 0$ that $r = 0$. □

Corollary 11.2.13. *Let K be a field and let $f \in K[t]$ be a non-zero polynomial. Then f has at most n roots where $n = \deg f$.*

Example 11.2.14. If K is an infinite field, then for any $f, g \in K[t]$ such that $f \neq g$, the polynomial $f - g$ is non-zero and thus $f - g$ has at most n roots where $n = \deg(f - g)$. As K is infinite, we can find $x \in K$ that is not a root of $f - g$. Hence $f_K(x) - g_K(x) \neq 0$ where f_K and g_K are the corresponding polynomial functions of f_K and g_K. Thus f_K and g_K are *not* the same as maps from K to K.

Remark. Once we understand the difference between a polynomial f and its associated polynomial function f_K, we may write f for f_K for simplicity.

§ 3. Irreducible polynomials in $K[t]$

Let K be a field. A non-constant polynomial $f \in K[t]$ is **irreducible** if whenever $f = gh$ where $g, h \in K[t]$, either g or h is a non-zero constant polynomial. For example, $t^2 + 1 \in \mathbb{R}[t]$ is irreducible but $t^2 - 1 \in \mathbb{R}[t]$ is not irreducible. (cf. §17.3.)

Example 11.2.15. Let K be a field and $f \in K[t]$ be a quadratic (i.e. having degree 2) polynomial. If f has no root in K, then f is irreducible.

Proof. Suppose that $f = gh$ for some $g, h \in K[t]$ and neither g nor h is a non-zero constant polynomial. Then both g and h would have degree 1 so g (and h) would have a root in K. This implies f has a root in K, which contradicts the assumption on f, and so we conclude that f is irreducible. □

Theorem 11.2.16. *If K is a field and if $f \in K[t]$ is irreducible, then the ideal $\langle f \rangle$ is maximal and thus the quotient ring $K[t]/\langle f \rangle$ is a field.*

Theorem 11.2.16 follows from Theorem 13.2.4 (i) $\Leftrightarrow$ (iii). Example 11.2.15 gives a sufficient condition for irreducible *quadratic* polynomials. It is not easy to check the irreducibility of a high degree polynomial. The theorem below (whose proof is omitted) gives a test but bear in mind: *the criterion does not conclude anything if the polynomial does not fulfill the condition.*

Theorem 11.2.17. (Eisenstein's Criterion) *Let p be a prime and suppose*

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0 \in \mathbb{Z}[x] \subset \mathbb{Q}[x].$$

If (i) $p | a_0, a_1, \cdots, a_{n-1}$ but $p \nmid a_n$ and (ii) $p^2 \nmid a_0$, then $f(x)$ is irreducible over $\mathbb{Q}$ (i.e. $f(x)$ is an irreducible polynomial in $\mathbb{Q}[x]$).

For example, $x^3 + 5x^2 - 15$ and $25x^5 - 9x^4 + 3x^2 - 12$ are irreducible in $\mathbb{Q}[x]$.

11.3 Use irreducible polynomials to construct fields

By Theorem 11.2.16, *irreducible polynomials are very useful in constructing fields.*

Example 11.3.1. Let $f = t^2 + 1 \in \mathbb{R}[t]$. Then f is irreducible, and

$$\mathbb{R}[t]/\langle f \rangle \cong \mathbb{C} \qquad \text{as fields.}$$

Proof. The map $\varphi : \mathbb{R}[t] \to \mathbb{C}$, $\varphi(p) = p(i)$, is a ring homomorphism (check it), $\ker \varphi = \langle f \rangle$ and φ is onto (as $a + bt \in \mathbb{R}[t]$, $\forall\, a, b \in \mathbb{R}$ and $\varphi(a + bt) = a + bi$). By Lemma 9.4.4, $\mathbb{R}[t]/\langle f \rangle \cong \mathbb{C}$. □

Example 11.3.2. Let $K = \mathbb{F}_2 = \{0, 1\}$ be the field with 2 elements (i.e. $\mathbb{F}_2 \cong \mathbb{Z}_2$ as fields) and let $f = t^2 + t + 1$. Since neither 0 nor 1 is a root for f, f is irreducible by Example 11.2.15. Thus $\mathbb{F}_2[t]/\langle f \rangle$ is a field. How many elements in $\mathbb{F}_2[t]/\langle f \rangle$?

Answer. Every element in $\mathbb{F}_2[t]/\langle f \rangle$ is uniquely of the form $a_0 + a_1\bar{t}$ with $a_0, a_1 \in \mathbb{F}_2$ and the product is determined by the relation $\bar{t}^2 = -\bar{t} - 1 = \bar{t} + 1$ (e.g. $\bar{t}^3 = 1$), so

$$\mathbb{F}_2[t]/\langle t \rangle = \{a_1 + a_1\bar{t} : a_0, a_1 \in \mathbb{F}_2\}$$

is a field with 4 elements; the addition and multiplication are given by

$$\begin{aligned}(a_0 + a_1\bar{t}) + (b_0 + b_1\bar{t}) &= (a_0 + b_0) + (a_1 + b_1)\bar{t},\\ (a_0 + a_1\bar{t})(b_0 + b_1\bar{t}) &= a_0b_0 + (a_0b_1 + a_1b_0)\bar{t} + a_1b_1(\bar{t} + 1)\\ &= a_0b_0 + a_1b_1 + (a_0b_1 + a_1b_0 + a_1b_1)\bar{t}.\end{aligned}$$

Alternatively we have the following addition and multiplication tables:

$+$	0	1	$\bar{t}$	$1+\bar{t}$
0	0	1	$\bar{t}$	$1+\bar{t}$
1	1	0	$1+\bar{t}$	$\bar{t}$
$\bar{t}$	$\bar{t}$	$1+\bar{t}$	0	1
$1+\bar{t}$	$1+\bar{t}$	$\bar{t}$	1	0

$\cdot$	1	$\bar{t}$	$1+\bar{t}$
1	1	$\bar{t}$	$1+\bar{t}$
$\bar{t}$	$\bar{t}$	$1+\bar{t}$	1
$1+\bar{t}$	$1+\bar{t}$	1	$\bar{t}$

Chapter 12 — Fields

12.1 Quotient fields of integral domains

In this section, we will mimic the construction of the field of rational numbers

$$\mathbb{Q} = \left\{ \frac{m}{n} : n, m \in \mathbb{Z}, m \neq 0 \right\}$$

from the ring of integers $(\mathbb{Z}, +, \cdot)$.

Let R be an integral domain. Let $P(R) = \{(a, b) \in R \times R : b \neq 0\}$ and define a relation $\sim$ on $P(R)$ by $(a, b) \sim (c, d)$ if $ad = bc$.

Lemma 12.1.1. *The relation $\sim$ on $P(R)$ is an equivalence relation.*

Proof. It is clear that $(a, b) \sim (a, b)$ for all $(a, b) \in P(R)$. It is also clear that if $(a, b) \sim (c, d)$, then $(c, d) \sim (a, b)$. Assume that $(a, b), (c, d)$ and (e, f) in $P(R)$ are such that $(a, b) \sim (b, d)$ and $(b, d) \sim (e, f)$. We need to show that $(a, b) \sim (e, f)$. By definition, $ad = bc$ and $cf = de$. Thus

$$d(af - be) = adf - bcf + bcf - dbe = f(ad - bc) + b(cf - de) = 0.$$

Since $d \neq 0$ and since R is an integral domain, we must have $af = be$. Thus shows that $(a, b) \sim (e, f)$. □

Lemma 12.1.2. *Let $Q(R)$ denote the set of all equivalence classes in $P(R)$, and for $(a, b) \in P(R)$, let $\frac{a}{b} \in Q(R)$ denote the equivalence class of (a, b). Define two operations on $Q(R)$, called addition and multiplication and denoted respectively as $+$ and $\cdot$, by*

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}, \qquad \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}, \qquad (a, b),\ (c, d) \in P(R).$$

Then $(Q(R), +, \cdot)$ is a field with the zero element given by $\frac{0}{1}$ and 1 given by $\frac{1}{1}$.

Proof. One first needs to prove that both operations are well-defined, and then check that they make $Q(R)$ into a non-zero commutative ring. This is done carefully in Lang's book (p. 100 - 102). Note that for $(a, b) \in P(R)$, $\frac{a}{b} = \frac{0}{1}$ if and only if $(a, b) \sim (0, 1)$ which means $a = 0$. Thus the non-zero elements in $Q(R)$ are those $\frac{a}{b}$ such that $a \neq 0$. For such an element, it follows from the definition that $\frac{a}{b}\frac{b}{a} = \frac{1}{1}$. This shows that $Q(R)$ is a field. □

Definition 12.1.3. Let R be an integral domain. The field $Q(R)$ constructed in Lemma 12.1.2 is called the **quotient field** (or **Field of Fractions**) **of** R.

Remark. Formally the quotient field $Q(R)$ is the set of equivalence classes, e.g. for $R = \mathbb{Z}$, the element $\frac{a}{b}$ in $Q(\mathbb{Z})$ is the equivalence class of (a, b), so $\frac{a}{b}$ is *not* our familiar fraction in $\mathbb{Q}$. However we may *identify* them via the field isomorphism[†] $Q(\mathbb{Z}) \cong \mathbb{Q}$. That's the proper meaning of saying "the quotient field of $\mathbb{Z}$ *is* $\mathbb{Q}$".

We state without proof a key feature of quotient fields in Lemma 12.1.4, asserting that the quotient field $Q(R)$ is the *smallest field containing* the integral domain R.

Lemma 12.1.4. *Every field F containing an integral domain R contains the quotient field $Q(R)$. i.e If the field $F \supset R$, then $F \supset Q(R)$.*

Example 12.1.5. Recall that $\mathbb{Q}[t], \mathbb{R}[t], \mathbb{C}[t]$ are integral domains.

1. The quotient field of the ring $\mathbb{Q}[t]$ consists of *rational functions* $\frac{p(t)}{q(t)}$ where $p, q \in \mathbb{Q}[t]$ and $q \neq 0$. We denote this quotient field by $\mathbb{Q}(t)$, i.e.

$$\mathbb{Q}(t) = \left\{ \frac{p(t)}{q(t)} : \ p, q \in \mathbb{Q}[t],\, q \neq 0 \right\}.$$

Caveat: Don't mix up $\mathbb{Q}[t]$ and $\mathbb{Q}(t)$, in fact, $\mathbb{Q}[t] \subsetneqq \mathbb{Q}(t)$.

2. The quotient fields $\mathbb{R}(t)$ and $\mathbb{C}(t)$ of the rings $\mathbb{R}[t]$ and $\mathbb{C}[t]$ are resp. given by

$$\mathbb{R}(t) = \left\{ \frac{p(t)}{q(t)} : \ p, q \in \mathbb{R}[t],\, q \neq 0 \right\} \text{ and } \mathbb{C}(t) = \left\{ \frac{p(t)}{q(t)} : \ p, q \in \mathbb{C}[t],\, q \neq 0 \right\}.$$

Exercise 12.1.6. Show that the quotient field of the integral domain $\mathbb{Z}[t]$ is $\mathbb{Q}(t)$.

Let $\alpha \in \mathbb{C}$. We may view $\mathbb{Q}[\alpha]$ and $\mathbb{Q}(\alpha)$ obtained by putting $t = \alpha$ into $\mathbb{Q}[t]$ and $\mathbb{Q}(t)$[‡] respectively. Then $\mathbb{Q}[\alpha]$ is an integral domain and $\mathbb{Q}(\alpha)$ is the quotient field of the $\mathbb{Q}[\alpha]$ (and $\mathbb{Q}(\alpha) \supset \mathbb{Q}[\alpha]$).

Example 12.1.7. The polynomial $f = t^2 + 1$ is irreducible in $\mathbb{Q}[t]$, and thus $\mathbb{Q}[t]/\langle f \rangle$ is a field; moreover $\mathbb{Q}[t]/\langle f \rangle \cong \mathbb{Q}[i]$. So $\mathbb{Q}[i]$ is a field (and in fact, $\mathbb{Q}[i] = \mathbb{Q}(i)$) and the field $\mathbb{Q}[t]/\langle f \rangle$ is different from the field $\mathbb{R}[t]/\langle f \rangle$. (Note $\mathbb{R}[t]/\langle f \rangle \cong \mathbb{C}$ and $\mathbb{Q}(i) \subsetneqq \mathbb{C}$.)

Proof. By Corollary 11.2.13, f is irreducible in $\mathbb{Q}[t]$, and by Theorem 11.2.16, $\mathbb{Q}[t]/\langle f \rangle$ is a field. The argument in the proof of Example 11.3.1 shows $\mathbb{Q}[t]/\langle f \rangle \cong \mathbb{Q}[i]$, so $\mathbb{Q}[i]$ is a field (containing the integral domain $\mathbb{Q}[i]$). By Lemma 12.1.4, $\mathbb{Q}[i] \supset \mathbb{Q}(i)$ and thus $\mathbb{Q}(i) = \mathbb{Q}[i]$ $(= \{a + bi : a, b \in \mathbb{Q}\})$. □

Example 12.1.8. The polynomial $t^2 - 2$ is irreducible in $\mathbb{Q}[t]$, thus $\mathbb{Q}[\sqrt{2}] \cong \mathbb{Q}[t]/\langle t^2 - 2 \rangle$ is a field and $\mathbb{Q}(\sqrt{2}) = \mathbb{Q}[\sqrt{2}]$.

[†] i.e. $\varphi : F \to \mathbb{F}$ is a field isomorphism, meaning that φ is a bijective map, $\varphi(x+y) = \varphi(x) + \varphi(y)$, $\varphi(xy) = \varphi(x)\varphi(y)$ and $\varphi(1) = 1$. So viewing a field as a ring, field isomorphisms are just ring isomorphisms.

[‡] Here $\frac{p(\alpha)}{q(\alpha)}$ will be ignored if $q(\alpha) = 0$ and this is NOT the conventional meaning of $\mathbb{Q}(\alpha)$.

12.2 Field extensions

Lemma 12.2.1. *If K and L are two fields and if $\phi : K \to L$ is a ring homomorphism, then ϕ is injective.*

Proof. The kernel $\ker(\phi)$ of ϕ, being an ideal of K, must be either the zero ideal or all of K, but since $\phi(1) = 1$, $\ker(\phi) \neq K$, so $\ker(\phi) = 0$. □

Definition 12.2.2. If K and L are fields and if there is a ring homomorphism $\phi : K \to L$, we say that L is a **field extension of** K, or say that K is a **sub-field of** L.

The notion of vector spaces over any field can be formulated and in particular we can talk about bases of vector spaces over any fields and about finite or infinite dimensional vector spaces over any field.

Definition 12.2.3. A field extension of a field K can be regarded as a vector space over K. A field extension L of K is called a **finite extension** if L is a finite dimensional vector space over K. Finite extensions of the field $\mathbb{Q}$ of rational numbers are called **number fields** and are the central topics of *algebraic number theory.*

Example 12.2.4. The field $\mathbb{C}$ is a finite extension of $\mathbb{R}$, as $\mathbb{C}$ is a vector space over $\mathbb{R}$. (A basis for the vector space $\mathbb{C}$ over $\mathbb{R}$ is $\{1, i\}$.) However, $\mathbb{C}$ is *not* a finite extension of $\mathbb{Q}$, in fact $\mathbb{R}$ is also *not* a finite extension of $\mathbb{Q}$.[†] i.e. Both $\mathbb{C}$ and $\mathbb{R}$ are *not* number fields.

Clearly a field extension of an infinite field K (i.e. K has infinite many elements) is infinite. We turn to another "extreme case" in the next section.

12.3 Finite fields

Lemma 12.3.1. *The characteristic of any finite field is a prime number.*

Proof. By Corollary 9.5.4, the characteristic of a field must be either 0 or a prime number. If a field F has characteristic 0, the map $\mathbb{Z} \to F, n \mapsto n1$ is injective, so F must have infinitely many elements. Thus the characteristic of any finite field is a prime number. □

For a prime number p, the finite field $\mathbb{Z}_p$ is also commonly denoted as $\mathbb{F}_p$.

[†] Here is an explanation for your interest only: $\mathbb{R}$ (and hence $\mathbb{C}$) contains some interesting elements, called *transcendental numbers*, e.g. $\pi = 3.1416\cdots$ is a transcendental number. By definition, a transcendental number is a number which is *not* a root of any polynomial in $\mathbb{Q}[t]$. Thus the subset $\{\pi, \pi^2, \cdots, \pi^n, \cdots\}$ of $\mathbb{R}$ is a linearly independent set over $\mathbb{Q}$. So $\dim_{\mathbb{Q}} \mathbb{R} = \infty$.

Lemma 12.3.2. *For any prime number p, every finite field of characteristic p is a field extension of $\mathbb{F}_p$. Consequently, every field with p elements is isomorphic to $\mathbb{F}_p$.*

Proof. Let F be a finite field with characteristic p and let 1 be the unity of F. Then the map $\bar{r} \mapsto 1 + 1 + \cdots + 1$ (r-times) for $0 \le r \le p-1$ is a ring homomorphism $\mathbb{F}_p$ to F. Clearly $\dim_{\mathbb{F}_p} F < \infty$ (as F is finite), so F is a finite extension of $\mathbb{F}_p$.

The above ring homomorphism from $\mathbb{F}_p$ to F is injective by Lemma 12.2.1. If $|F| = p$ $(= |\mathbb{F}_p|)$, then the homomorphism is bijective, i.e., a ring isomorphism. □

Theorem 12.3.3. *The cardinality of a finite field F with characteristic p must be p^n for some positive integer n.*

Proof. Let n be the dimension of F as a finite dimensional vector space over $\mathbb{F}_p$. Then $|F| = p^n$. □

Theorem 12.3.4. *(***Existence and uniqueness of fields with p^n elements***) For any prime number p and any integer $n \ge 1$, there is a field with p^n elements, and any two such fields are isomorphic.*

The field with p^n elements will be denoted as $\mathbb{F}_{p^n}$. The proof of Theorem 12.3.4 has to wait till Algebra II. Here we give a way to construct a field of p^n elements using the technique in §11.3.

Lemma 12.3.5. *Let p be a prime number. If f is a degree n irreducible polynomial in $\mathbb{F}_p[t]$, then $K = \mathbb{F}_p[t]/\langle f \rangle$ is a field with p^n elements.*

Proof. Without loss of generality, assume that f is monic, i.e.,

$$f = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0,$$

where $a_j \in \mathbb{F}_p$ for $j = 0, 1, \ldots, n-1$. Let $\bar{t}$ be the image[‡] of t in K. Then

$$\bar{t}^n = -a_{n-1}\bar{t}^{n-1} - \cdots - a_1\bar{t} - a_0 \in K, \qquad (*)$$

and elements in K are precisely of the form[§]

$$b_{n-1}\bar{t}^{n-1} + \cdots + b_1\bar{t} + b_0,$$

where $b_0, b_1, \ldots, b_{n-1} \in \mathbb{F}_p$. Thus $|K| = p^n$. □

Remark. From the proof, we see that the field $K = \mathbb{F}_p[t]/\langle f \rangle$ is a vector space over $\mathbb{F}_p$ of dimenision n. Indeed, $K = \mathrm{Span}_{\mathbb{F}}(1, \bar{t}, \cdots, \bar{t}^{n-1})$ where $\{1, \bar{t}, \cdots, \bar{t}^{n-1}\}$ is linearly independent over $\mathbb{F}_p$, i.e. $\{1, \bar{t}, \cdots, \bar{t}^{n-1}\}$ is a basis for K.

[‡] i.e. $\bar{t} = \pi(t)$ where $\pi : \mathbb{F}_p[t] \to \mathbb{F}_p[t]/\langle f \rangle$ is the natural projection from the ring $\mathbb{F}_p[t]$ to the quotient ring $\mathbb{F}_p[t]/\langle f \rangle$. Explicitly, $\bar{t} = t + \langle f \rangle$ in coset notation.

[§] Every element in K can be expressed as $\overline{g}$ for some $g \in \mathbb{F}_p[t]$. If $g = c_m t^m + \cdots + c_1 t + c_0$, then $\overline{g} = c_m \bar{t}^m + \cdots + c_1\bar{t} + c_0$. By $(*)$, we can write $\overline{g}$ into a polynomial of degree $\le n-1$ in $\bar{t}$. See Example 11.3.2 as well.

Example 12.3.6. Let's construct $\mathbb{F}_9$. So we need a quadratic irreducible polynomial over $\mathbb{F}_3$. There are altogether 9 monic quadratic polynomial over $\mathbb{F}_3$

$$\begin{gathered} t^2, \quad t^2+1, \quad t^2+2, \quad t^2+t, \quad t^2+t+1, \\ t^2+t+2, \quad t^2+2t, \quad t^2+2t+1, \quad t^2+2t+2, \end{gathered}$$

of which t^2+1, t^2+t+2, t^2+2t+2 are irreducible. Take $f = t^2+t+2$ and let

$$\mathbb{F}_9 = \mathbb{F}_3[t]/\langle t^2+t+2\rangle = \{a + b\bar{t} : \ a, b \in \mathbb{F}_3\},$$

and multiplication in $\mathbb{F}_9$ is governed by the role that $\bar{t}^{\,2} = -\bar{t} - 2 = 2\bar{t} + 1$. For instance, the product of $2 + \bar{t}$ and $1 + 2\bar{t}$ in $\mathbb{F}_9$ is

$$(2+\bar{t})(1+2\bar{t}) = 2 + (2^2+1)\bar{t} + 2\bar{t}^{\,2} = 2 + 2\bar{t} + 2(2\bar{t}+1) = 1.$$

As a vector space over $\mathbb{F}_3$, $\{1, \bar{t}\}$ is a basis for $\mathbb{F}_9$.

Theorem 12.3.7. *If F is a finite field with q elements (we know that q must be a power of a prime number), then the multiplicative group $(F\backslash\{0\}, \cdot)$ is cyclic.*

Proof. Let $G = F\backslash\{0\}$. We know that G is abelian and of order $q-1$. Let $p_1, \ldots, p_l$ be all the distinct prime factors of $q-1$, so $q-1 = p_1^{r_1} \cdots p_l^{r_l}$ for some $r_1, \cdots, r_l \in \mathbb{N}$. By the fundamental theorem of finite abelian groups, G is a product of p_j-groups[¶] for $j = 1, \ldots, l$. If there is some $1 \leq j \leq l$ such that there are more than one p_j-groups appearing in the factorization of G, say one G_1 of order p_j^r and another G_2 of order p_j^s with $r \geq s$, then all the elements in $G_1 \times G_2$ satisfy $t^{p_j^r} = 1$, and there are p_j^{r+s} elements in $G_1 \times G_2$. This contradicts the fact that, over a field, every polynomial of degree n can have at most n roots. Thus there is exactly one p_j-group in the factorization of G into the product of cyclic groups. Consequently,

$$G \cong \mathbb{Z}_{p_1^{r_1}} \times \cdots \times \mathbb{Z}_{p_l^{r_l}} \cong \mathbb{Z}_{p_1^{r_1} \cdots p_l^{r_l}}$$

by Lemma 2.7.7. So $G \cong \mathbb{Z}_{q-1}$ is cyclic. □

Example 12.3.8. The element $\bar{t} \in \mathbb{F}_9$ (see Example 12.3.6) is a generator for the cyclic group $\mathbb{F}_9\backslash\{0\}$ whose order is 8:

$$\begin{gathered} \bar{t}^{\,1} = \bar{t}, \quad \bar{t}^{\,2} = 2\bar{t}+1, \quad \bar{t}^{\,3} = 2\bar{t}+2, \quad \bar{t}^{\,4} = 2, \\ \bar{t}^{\,5} = 2\bar{t}, \quad \bar{t}^{\,6} = \bar{t}+2, \quad \bar{t}^{\,7} = \bar{t}+1, \quad \bar{t}^{\,8} = 1. \end{gathered}$$

REMARK. $\mathbb{F}_p \cong \mathbb{Z}_p$ but $\mathbb{F}_{p^n} \not\cong \mathbb{Z}_{p^n}$ for $n \geq 2$. There are two simple ways to see:
(i) $\mathbb{F}_{p^n}$ is a field but $\mathbb{Z}_{p^n}$ is not an integral domain (for $n \geq 2$);
(ii) $\mathrm{char}(\mathbb{F}_{p^n}) = p$ but $\mathrm{char}(\mathbb{Z}_{p^n}) = p^n$ (see Section 9.5).

[¶] See Definition 7.1.1. Thus $G = \underbrace{G_{1,1} \times \cdots \times G_{1,k_1}}_{\text{product of } p_1-\text{groups}} \times \underbrace{G_{2,1} \times \cdots \times G_{2,k_2}}_{\text{product of } p_2-\text{groups}} \times \cdots \times G_{l,1} \times \cdots \times G_{l,k_l}$.
And we say G is factorized into a product of k_1 p_1-groups, k_2 p_2-groups, etc.

Chapter 13 — Irreducibles, Primes and UFDs

Poetry is an art of giving different names to the same thing.
"Mathematics is the art of giving the same name
to different things", said Henri Poincaré.

13.1 Euclidean domains

Definition 13.1.1. By an **Euclidean domain** we mean a pair (D, v), where D is an integral domain, and $v : D\backslash\{0\} \to \mathbb{N} \cup \{0\}$ a map such that

1) $v(a) \leq v(ab)$ for all $a, b \in D\backslash\{0\}$;
2) for all $a \in D\backslash\{0\}$ and $b \in D$, there exist $q, r \in D$ such that $b = aq + r$, where either $r = 0$ or $r \neq 0$ and $v(r) < v(a)$.

Thus far we have seen two examples of Euclidean domains: (1) $D = \mathbb{Z}$ with $v(a) = |a|$ for $a \in \mathbb{Z}\backslash\{0\}$, and (2) $D = K[t]$ for any field K and $v(f) = \deg(f)$ for $f \in K[t]\backslash\{0\}$.

Theorem 13.1.2. *Every Euclidean domain (D, v) is a PID.*

Proof. Assume that (D, v) is an Euclidean domain. Let I be any ideal of D. If $I = \{0\}$, then $I = \langle 0 \rangle = 0D$ is principal. Assume that I is not zero. Choose $a \in I$, $a \neq 0$, such that $v(a)$ is the smallest among all $b \in I \setminus \{0\}$. Since I is an ideal, $aD \subset I$. Assume that $b \in I$ is arbitrary. As (D, v) is an Euclidean domain, there exist $q, r \in D$ such that $b = aq + r$, and either $r = 0$ or $r \neq 0$ but $v(r) < v(a)$. By the choice of a, we must have $r = 0$. Thus $b = aq \in I$. This shows that $I = aD$ is principal. □

Remark. There are PIDs which are *not* Euclidean domains. ■ For your interest: the ring $\mathbb{Z}[\theta]$ where $\theta = (1 + \sqrt{-19})/2$ is a PID but not an Euclidean domain. ■

Corollary 13.1.3. *Let K be a field. Then $K[t]$ is a principal ideal domain. In fact, for any non-zero ideal I of $K[t]$, any non-zero element in I of the smallest degree is a generator for I.*

13.2 Irreducible elements and primes

Definition 13.2.1. Let R be an integral domain. A non-zero and non-unit $p \in R$ is said to be **irreducible** if $p = ab$ for $a, b \in R$ implies that either a or b is a unit. A non-zero and non-unit $p \in R$ is called a **prime** if $p|ab$ implies $p|a$ or $p|b$. Here $p|a$ means $a = px$ for some $x \in R$.

Lemma 13.2.2. *Let R be an integral domain and $a \in R$ be a non-zero non-unit.*

1. *If a is a prime, then a is irreducible.*
2. *The element a is a prime if and only if the principal ideal aR is a prime ideal.*

Proof. 1. Suppose $a = uv$ where u, v are non-units. As $a|uv$ and a is a prime, $a|u$ or $a|v$. WLOG, let $a|u$, i.e. $u = ax$ for some $x \in R$ and thus $a = axv$. As $a \neq 0$ and R has no zero divisor, $xv = 1$ and so v is a unit. Hence we conclude a is irreducible.

2. Exercise. □

Example 13.2.3. ■ (For your interest only) Consider the subring $R = \mathbb{Z}[\sqrt{-5}] = \{m + n\sqrt{-5} : m, n \in \mathbb{Z}\}$ of $\mathbb{C}$. R is an integral domain whose units are 1 or -1 (by the facts $(m + n\sqrt{-5})\overline{(m + n\sqrt{-5})} = m^2 + 5n^2 \in \mathbb{Z}$ and $\alpha\overline{\alpha} = 1$ if $\alpha \in \mathbb{Z}[\sqrt{-5}]$ is a unit). Then $x = 1 + \sqrt{-5}$ is an *irreducible element* in R but x is *not a prime.*[†] ■

In PID, Irreducible elements = Primes. (See Corollary 13.2.5 below.)

Theorem 13.2.4. *Let R be a principal ideal domain (PID), and let I be a non-zero ideal of R. The following statements are equivalent:*
1) I is a maximal ideal;
2) I is a prime ideal;
3) I is generated by an irreducible element.

Proof. We have seen that 1) implies 2).

Assume 2), and let p be a generator for I. We want to show that p is irreducible. Assume that $p = ab$, where $a, b \in R$. Since $I = pR$ is a prime ideal and $ab \in I$, we get $a \in pR$ or $b \in pR$. Thus, $a = px$ for some $x \in R$, or $b = py$ for some $y \in R$. Suppose that $a = px$ for some $x \in R$. Then $p = ab = pxb$, so $p(1 - xb) = 0$. Since R has no zero divisor, we have $xb = 1$, so b is a unit. Similarly, if $b = py$ for some $y \in R$, then a is a unit. This shows shows that p is irreducible. Thus 2) implies 3).

[†] 1° x is irreducible: Suppose $x = uv$ where $u, v \in R$ are not units. Then $6 = |x|^2 = |u|^2|v|^2$ where both $|u|^2$ and $|v|^2$ are (positive) integers. So $|u|^2 = 1, 2, 3, 6$. As the argument works for $|v|^2$, we may assume $|u|^2 = 1$ or 2. If $|u|^2 = 1$, then u is a unit. On the other hand, there is *no* element $y \in R$ with $|y|^2 = 2$. Thus x is an irreducible element.

2° x is not a prime: Note $x(1 - \sqrt{-5}) = 6 = 2 \cdot 3$, i.e. $x|2 \cdot 3$. However, $x \nmid 2$ and $x \nmid 3$. It is because, for example, if $x|3$, then $xy = 3$ for some $y \in R$. Multiplying with their complex conjugates, we get $|x|^2|y|^2 = 9$, i.e. $6|y|^2 = 9$ where $|y|^2 \in \mathbb{Z}$. Absurdity!

Assume 3), i.e., $I = pR$ and p is irreducible. Let J be any ideal of R such that $J \supset I$. Since R is a PID, J has a generator q. Now $p \in J = qR$ implies that $p = qx$ for some $x \in R$. Since p is irreducible, x is a unit or q is a unit. If x is a unit, $I = J$. If q is a unit, $J = R$. Thus I is maximal. □

Corollary 13.2.5. *In a PID, a non-unit element is a prime if and only if it is irreducible.*

13.3 Unique factorization domain

We have seen in PID, an irreducible element is a prime. But in fact the same result holds for those integral domains that satisfy the unique factorization property of $\mathbb{Z}$.

Definition 13.3.1. Let R be an integral domain in which every non-zero and non-unit element can be written as a finite product of irreducible elements, and furthermore, whenever a non-zero and non-unit element $x \in R$ is written as

$$x = p_1 p_2 \cdots p_n = q_1 q_2 \cdots q_m$$

where $p_1, p_2, \ldots, p_n$ and $q_1, q_2, \ldots, q_m$ are irreducible, then $m = n$ and after a suitable re-ordering, $p_i = q_i a_i$ for some unit elements a_i for every $i = 1, 2, \ldots, n$. Then we say that R is a **unique factorization domain** or a UFD.

We skip both the proofs of Theorem 13.3.2 and Theorem 13.3.3.

Theorem 13.3.2. *Every PID is a UFD*

Theorem 13.3.3. *If R is a UFD, so is $R[t]$.*

REMARK. (1) The ring $\mathbb{Z}$ is a UFD, as a particular case of Theorem 13.3.2. Also $K[t]$ (K a field) is a PID and hence a UFD. (2) Theorem 13.3.3 implies that $\mathbb{Z}[t]$ is an example of UFD but *not* PID by Example 11.2.6..

Theorem 13.3.4. *In UFD, an irreducible element is a prime.*

Proof. Exercise. □

REMARK.

1. Regarding the various notions we have introduced so far, we have

$$\{\text{Fields}\} \subsetneqq \left\{\begin{matrix}\text{Euclidean}\\ \text{domains}\end{matrix}\right\} \subsetneqq \{\text{PIDs}\} \subsetneqq \{\text{UFDs}\} \subsetneqq \left\{\begin{matrix}\text{Integral}\\ \text{domains}\end{matrix}\right\} \subsetneqq \left\{\begin{matrix}\text{Commutative}\\ \text{rings}\end{matrix}\right\}.$$

2. Let I be a nonzero ideal in PID. Then
   (i) I is maximal iff I is prime.
   (ii) I is a prime ideal iff $I = \langle p \rangle$ for some prime element p;
   (iii) I is a maximal ideal iff $I = \langle q \rangle$ for some irreducible element q.
   (iv) In PID, a is an irreducible element iff a is a prime.

End