# MTL 105

**Course Instructor: Dr. Ekata Saha**
**email : ekata at maths.iitd.ac.in**

**Department of Mathematics**
**Indian Institute of Technology Delhi**
**New Delhi**


**Session: Semester 1, 2022**

# Contents


**1 Preliminaries** **1**

1.1 Relations and partitions on a set . . . . . 1

1.1.1 Equivalence relation . . . . . 2

1.1.2 Partial order relation . . . . . 4

1.2 Zorn's lemma . . . . . 6

1.3 Binary operations . . . . . 6

**2 Group Theory** **9**

2.1 Introduction . . . . . 9

2.2 Subgroups . . . . . 13

2.3 Cyclic groups . . . . . 17

2.4 Permutation groups . . . . . 20

2.5 Cosets and Lagrange's theorem . . . . . 28

2.5.1 An application of Lagrange's theorem . . . . . 31

2.6 Normal subgroups . . . . . 33

2.7 Quotient groups . . . . . 36

2.8 Homomorphism of groups . . . . . 37

2.9 Direct product of groups . . . . . 49

2.10 Conjugacy relation and Class equation . . . . . 54

2.10.1 Applications of class equation . . . . . 55

2.11 Partial converses of the Lagrange's theorem . . . . . 56

2.11.1 Applications of Cauchy's theorem . . . . . 57

2.12 Group actions . . . . . 59

2.12.1 Orbit and stabilizer . . . . . 60

2.12.2 Kernel of a group action . . . . . 62

2.13 Automorphism group . . . 62
2.14 Characteristic subgroups . . . 64
2.15 Sylow theorems and applications . . . 64
2.15.1 Applications of Sylow's theorems . . . 69
2.16 Structure theorem of finite abelian groups . . . 72
2.16.1 Applications of the structure theorem of finite abelian groups . . . 74
2.16.2 Proof of the structure theorem of finite abelian groups . . . 74

**3 Ring theory** **77**
3.1 Definition and examples . . . 77
3.2 Zero divisors . . . 80
3.3 Idempotent elements in a ring . . . 81
3.4 Nilpotent elements in a ring . . . 82
3.5 Subrings . . . 84
3.6 Boolean ring . . . 86
3.7 Polynomial ring . . . 87
3.8 Power series ring . . . 88
3.9 Ring of Gaussian integers . . . 89
3.10 Integral domain . . . 90
3.10.1 Fields . . . 92
3.11 Characteristic of a non-zero commutative ring with unity . . . 93
3.12 Ideals . . . 94
3.12.1 Sum and product of ideals . . . 95
3.12.2 Principal ideals . . . 96
3.12.3 Maximal ideals . . . 98
3.12.4 Prime ideals . . . 102
3.13 Quotient rings . . . 104
3.14 Characterisation of prime and maximal ideals with quotient rings . . . 105
3.15 Homomorphism of rings . . . 107
3.15.1 Isomorphism theorem of rings . . . 109
3.15.2 Correspondence theorem . . . 112
3.16 Quotient field/Field of fractions . . . 113
3.16.1 Construction of the field of fraction of an integral domain . . . 114
3.17 Division in a commutative ring . . . 119
3.17.1 Associates . . . 120
3.17.2 Greatest common divisor and least common multiplier . . . 121

3.18 Irreducible and prime elements in a commutative ring with unity . . . . . 123
3.19 Principal ideal domains . . . . . 127
3.20 Factorization domains . . . . . 129
3.21 Unique factorization domains . . . . . 131
3.21.1 Primitive polynomials over a UFD . . . . . 134
3.21.2 Gauss's lemma . . . . . 136
3.21.3 Gauss's theorem on UFD . . . . . 138
3.21.4 Eisenstein's criterion . . . . . 142
3.22 Euclidean domains . . . . . 144
3.23 Some counter examples . . . . . 148

**Bibliography** **150**

CHAPTER 1

# Preliminaries

## 1.1 Relations and partitions on a set

We know that a **Set** is a **well defined** collection of **distinct** objects. We call a set $A$ is **finite** if the number of elements in the set $A$ is finite, otherwise we call it an **infinite** set. A very first example of an infinite set which one encounters is the set of natural numbers which is denoted by $\mathbb{N}$. Below is the list of notations which we will use throughout.

**Notations**

- For a set $A$, by $|A|$ we mean the cardinality of $A$.
- By $A \subseteq B$ we mean that the set $A$ is a subset of the set $B$ and the set $B$ is a superset of the set $A$ i.e. all the elements of $A$ belong to $B$.
- By $A \subset B$ we mean that the set $A$ is a proper subset of $B$ i.e. all the elements of $A$ belong to $B$ and there exists at least one element in $B$ which does not belong to $A$.
- By $A \cup B$ we mean the set $\{c : c \in A \text{ or } c \in B\}$. We call this set as the union of the sets $A$ and $B$.
- By $A \cap B$ we mean the set $\{c : c \in A \text{ and } c \in B\}$. We call this set as the intersection of the sets $A$ and $B$.
- We denote the empty set by $\varnothing$.
- We say that the sets $A$ and $B$ are disjoint if $A \cap B = \varnothing$.

- By the Cartesian product of the sets $A$ and $B$ we mean the set $\{(a, b) : a \in A, b \in B\}$. We denote this set by $A \times B$.

**Definition 1.1.1** Let $A \neq \varnothing$. A **relation** $R$ on $A$ is a subset of $A \times A$. Let $a, a' \in A$. We say that $a$ is related to $a'$ with respect to the relation $R$ if $(a, a') \in R$. If $a$ is related to $a'$ with respect to $R$, we denote it by $aRa'$. If the set $A = \varnothing$, then on $A$ we have only one relation in particular the empty relation.

**Remark 1.1.2** In general for any relation $R$ on a non-empty set $A$, $aRa'$ does not imply that $a'Ra$ for all $a, a' \in A$.

**Definition 1.1.3** Let $A$ be a set and $R$ be a relation on $A$. We say that

- $R$ is **reflexive** on $A$ if for all $a \in A$, we have $(a, a) \in R$.
- $R$ is **symmetric** on $A$ if for any $a, a' \in A$ such that $(a, a') \in R$, we have $(a', a) \in R$.
- $R$ is **anti-symmetric** on $A$ if for any $a, a' \in A$ if $(a, a'), (a', a) \in R$, then we have $a$ and $a'$ are same in $A$. We denote $a = a'$.
- $R$ is **transitive** on $A$ if for any $a, a', a'' \in A$ such that $(a, a'), (a', a'') \in R$, we have $(a, a'') \in R$.

**Example 1.1.4** Let $A = \{1, 2, 3, 4\}$ and $R = \{(1, 2), (2, 1), (1, 1)\}$. We see that $R$ is not reflexive as $(2, 2), (3, 3)$ and $(4, 4) \notin R$, $R$ is symmetric, and $R$ is not transitive as $(2, 1), (1, 2) \in R$ but $(2, 2) \notin R$. Also $R$ is not anti-symmetric as $(2, 1), (1, 2) \in R$, but $1 \neq 2$.

### 1.1.1 Equivalence relation

**Definition 1.1.5** Let $A$ be a set and $R$ be a relation on $A$. We say that $R$ is an **equivalence relation** on $A$ if $R$ is reflexive, symmetric and transitive.

**Remark 1.1.6** If the set $A$ is empty then the empty relation is vacuously reflexive, symmetric and transitive, and hence it is an equivalence relation. The empty relation is not refliexive on any non-empty set $A$, hence on any non-empty set $A$, the empty relation is not an equivalence relation.

**Exercise 1.1.7** Let $\mathbb{R}$ denote the set of all real numbers. Check which of the following relations on $\mathbb{R}$ are equivalence relations:

1) $\{(x, y) \in \mathbb{R} \times \mathbb{R} : y = 2x\}$.

2) $\{(x, y) \in \mathbb{R} \times \mathbb{R} : x < y\}$.

3) $\{(x, y) \in \mathbb{R} \times \mathbb{R} : xy > 0\}$.

4) $\{(x, y) \in \mathbb{R} \times \mathbb{R} : y \neq x\}$.

5) $\{(x, y) \in \mathbb{R} \times \mathbb{R} : y \neq 2 + x\}$.

6) $\{(x, y) \in \mathbb{R} \times \mathbb{R} : x \leq y\}$.

7) $\{(x, y) \in \mathbb{R} \times \mathbb{R} : xy \geq 0\}$.

8) $\{(x, y) \in \mathbb{R} \times \mathbb{R} : y = x\}$.

**Definition 1.1.8** Let $R$ be an equivalence relation on a non-empty set $A$. Now for $a \in A$, let us define the set

$$[a] := \{a' \in A : a'Ra\}.$$

Then the set $[a]$ is called the **equivalence class of the element** $a \in A$ with respect to the equivalence relation $R$.

**Exercise 1.1.9** Let $R$ be an equivalence relation on a non-empty set $A$. Then show the following:

1) For all $a \in A$, $[a] \neq \varnothing$.

2) If $a' \in [a]$, then $[a'] = [a]$.

3) For any $a, a' \in A$, either $[a] = [a']$ or $[a] \cap [a'] = \varnothing$.

4) Let $\{[a_\alpha] : \alpha \in I\}$ be the collection of all disjoint equivalence classes of $A$ with respect to $R$, where $I$ is an indexing set. Then

$$A = \bigcup_{\alpha \in I} [a_\alpha].$$

Now we define what do we mean by a partition on a non-empty set.

**Definition 1.1.10** Let $A \neq \varnothing$ and $P$ be a collection of some non-empty subsets of $A$. Let us write $P = \{A_i\}_{i \in I}$ where $A_i \neq \varnothing$ and $A_i \subseteq A$. We say that $P$ is a **partition** on $A$ if and only if $A = \bigcup_{i \in I} A_i$ and $A_i \cap A_j = \varnothing$ for all $i \neq j$.

**Remark 1.1.11** Let $A \neq \varnothing$. We can see immediately from the definition of a partition that every equivalence relation on $A$ gives rise to a partition on $A$.

**Question 1.1.12** We can now ask the converse of the above remark i.e. for given any partition on $A \neq \varnothing$, do we get an equivalence relation on $A \neq \varnothing$?

The answer is yes. Let $P = \{A_i\}_{i \in I}$ be a partition on $A \neq \varnothing$. We can define a relation $R$ on $A$ as follows:
For any two elements $a, a' \in A$, we say that $aRa'$ if and only if both $a, a' \in A_i$ for some $i \in I$. One can check that $R$ is an equivalence relation on $A$.

Hence we now know that the set of all partitions on a non-empty set $A$ is in one to one correspondence with the set of all equivalence relations on $A$. We end this section by giving an interesting example of an equivalence relation on the set of integers $\mathbb{Z}$.

**Example 1.1.13 (Congruence relation on $\mathbb{Z}$)** Let $n \in \mathbb{N}$. Now we define a relation '$\equiv \pmod{n}$' on $\mathbb{Z}$ as follows:
For any $a, b \in \mathbb{Z}$, we say that $a \equiv b \pmod{n}$ if and only if $n|(a - b)$. Check that '$\equiv \pmod{n}$' is an equivalence relation on $\mathbb{Z}$. This equivalence relation is called the **congruence relation modulo** $n$ on $\mathbb{Z}$ and the equivalence classes are called the **congruence classes modulo** $n$. Check that there are $n$ disjoint congruence classes modulo $n$.

### 1.1.2 Partial order relation

**Definition 1.1.14** Let $A$ be a set and $R$ be a relation on $A$. We say that $R$ is a **partial order relation** on $A$ if $R$ is reflexive, antisymmetric and transitive. We usually denote a partial order relation by the notation '$\leq$' instead of $R$.

**Remark 1.1.15** If the set $A$ is empty then the empty relation is vacuously reflexive, antisymmetric and transitive, and hence it is a partial order relation. The empty relation is not reflexive on any non-empty set $A$, hence on any non-empty set $A$, the empty relation is not a partial order relation.

**Definition 1.1.16** Let $A$ be a set and $\leq$ be a partial order relation on $A$. Then the pair $(A, \leq)$ is called a **partial ordered set** or a **poset**.

**Examples 1.1.17**

1. Consider the set of integers $\mathbb{Z}$ with the standard $\leq$(less than or equal to) ordering. Then $(\mathbb{Z}, \leq)$ is a poset as '$\leq$' is reflexive, anti-symmetric and transitive on $\mathbb{Z}$.

2. Consider the division of integers in $\mathbb{Z}$. Note that $a \mid a$ for all $a \in \mathbb{Z}$, hence '$\mid$' is reflexive on $\mathbb{Z}$. Also if $a \mid b$ and $b \mid c$ for $a, b, c \in \mathbb{Z}$, then $a \mid c$, therefore '$\mid$' is transitive on $\mathbb{Z}$. But note that $2 \mid -2$ and $-2 \mid 2$, but $2 \neq -2$, therefore '$\mid$' is not anti-symmetric on $\mathbb{Z}$, and hence division of integers is not a partial order relation.

3. Consider the set of natural numbers $\mathbb{N}$ with the division of natural numbers. Again '$\mid$' is reflexive and transitive on $\mathbb{N}$. Further $a \mid b$ and $b \mid a$ in $\mathbb{N}$, implies that $a = b$, therefore '$\mid$' is also anti-symmetric on $\mathbb{N}$, and hence a partial order relation on $\mathbb{N}$.

**Exercise 1.1.18** Let $X$ be a set and $\mathcal{P}(X)$ be the power set of $X$. Let $Y$ be a subset of $\mathcal{P}(X)$. Consider the relation '$\leq$' on $Y$ as follows: for $A, B \in Y$, we define $A \leq B$ if and only if $A \subseteq B$. Show that $(Y, \leq)$ is a poset.

**Definition 1.1.19** Let $(X, \leq)$ be a poset. Then a subset $Y$ of $X$ is said to be a **chain** or a **totally ordered set** if any two elements of $Y$ are comparable i.e. for any $a, b \in Y$ we have $a \leq b$ or $b \leq a$.

**Example 1.1.20** Consider the set of natural numbers $\mathbb{N}$ with the relation division '$\mid$' of natural numbers. Then $(\mathbb{N}, \mid)$ is a poset. Consider the subset $A := \{1, 2, 4, 8\}$ of $\mathbb{N}$. The set $A$ is a chain of $(\mathbb{N}, \mid)$. Let $B := \{1, 2, 3\} \subset \mathbb{N}$. The set $B$ is not a chain of the poset $(\mathbb{N}, \mid)$ as $2, 3$ are not comparable.

**Remark 1.1.21** Let $(X, \leq)$ be a poset and $Y$ a chain of $(X, \leq)$. Then for any $a_1, \ldots, a_n \in Y$, there exists a permutation $\sigma$ of $n$ symbols that

$$a_{\sigma(1)} \leq \cdots \leq a_{\sigma(n)}.$$

We get this as any two elements in $Y$ are comparable and the relation $\leq$ is transitive.

**Remark 1.1.22** Let $X \neq \varnothing$ and '$\leq$' be a partial order relation on $X$. Then there is always a chain in $X$. If no two elements of $X$ are comparable with respect to $\leq$, then consider $Y := \{x\}$ where $x \in X$. Note that $Y$ is a chain in $(X, \leq)$.

**Definition 1.1.23** Let $(X, \leq)$ be a poset and $Y$ a subset of $(X, \leq)$. Then the subset $Y$ is called **bounded above** if there exists $\alpha \in X$ such that $y \leq \alpha$ for all $y \in Y$.

**Example 1.1.24** Consider $\mathbb{N}$ and $\leq$ (standard less than or equal to relation) on $\mathbb{N}$. Then we know that $(\mathbb{N}, \leq)$ is a poset. Consider the subset $A := \{1, 2, 3, 4\}$ of $\mathbb{N}$. The set $A$ is bounded above. Let $B := 2\mathbb{N}$. The subset $B$ is not bounded above.

## 1.2 Zorn's lemma

**Definition 1.2.1** Let $(X, \leq)$ be a poset. An element $\alpha \in X$ is said to be a **maximal element** of $X$ if $\alpha \leq \beta$ for some $\beta \in X$ implies that $\alpha = \beta$.

**Definition 1.2.2** Let $(X, \leq)$ be a totally ordered set i.e. any two elements of $X$ are comparable with respect to the partial order $\leq$ on $X$. An element $\alpha \in X$ is said to be a **least element** if $\alpha \leq \beta$ for all $\beta \in X$.

**Definition 1.2.3** Let $(X, \leq)$ be a totally ordered set. We say $X$ is **well-ordered** if every non-empty subset $Y$ of $X$ has a least element in $Y$.

The following three statements are considered to be the three pillars of modern set theory. They are equivalent and taken as standard axioms of modern set theory.

1. **Axiom of choice:** Let $\{X_\lambda\}_{\lambda \in \Lambda}$ be a collection of non-empty sets. Then we can construct a set $X$ such that $X \cap X_\lambda$ is a singleton set for all $\lambda \in \Lambda$. In other words, $\prod_{\lambda \in \Lambda} X_\lambda \neq \varnothing$.
2. **Zorn's lemma:** A non-empty poset in which every non-empty chain is bounded above, has a maximal element.
3. **Well-ordering theorem:** Every set can be given a total ordering so that it becomes well-ordered.

**Remark 1.2.4** The set of integers $\mathbb{Z}$ with respect to the standard less than or equal to relation $\leq$ is not well-ordered as the subset $2\mathbb{Z}$ does not have any least element.

## 1.3 Binary operations

In this section we will discuss about the binary operations on a non-empty set.

**Definition 1.3.1** Let $A \neq \varnothing$. By a **binary operation** $*$ on $A$ we mean a map from $A \times A$ to $A$.

**Example 1.3.2** The usual addition '+' is a binary operation on the set of natural numbers $\mathbb{N}$.

**Definition 1.3.3** Let $A \neq \varnothing$ and $*$ be a binary operation on $A$.

- We say that $*$ is **associative** if for all $a, b, c \in A$, we have $(a * b) * c = a * (b * c)$.
- We say that $*$ is **commutative** if for all $a, b \in A$, we have $a * b = b * a$.

**Examples 1.3.4** The following are few examples of various binary operations.

1) Define $*$ on $\mathbb{Z}$ by $a * b = a + b$ for all $a, b \in \mathbb{Z}$. This binary relation is commutative as well as associative.
2) Define $*$ on $\mathbb{Z}$ by $a * b = a - b$ for all $a, b \in \mathbb{Z}$. This binary relation is neither commutative nor associative.
3) Define $*$ on $\mathbb{Z}$ by $a * b = |a + b|$ for all $a, b \in \mathbb{Z}$. This binary relation is commutative but not associative.
4) Define $*$ on $\mathbb{Z}$ by $a * b = a$ for all $a, b \in \mathbb{Z}$. This binary relation is not commutative but associative.

**Question 1.3.5** Let $A$ be a set of $n$ elements, $n \in \mathbb{N}$. How many binary operations one can get on $A$?

A binary operation $*$ is a map from $A \times A$ to $A$. So $*$ is a map from a set of $n^2$ elements to a set of $n$ elements. For each element in $A \times A$, we have exactly $n$ choices in $A$ to map to. Since there are $n^2$ elements in $A \times A$, one can have $n^{n^2}$ many binary operations on $A$.

**Question 1.3.6** Let $A$ be a set of $n$ elements, $n \in \mathbb{N}$. How many commutative binary operations one can get on $A$?

We already know that there are total $n^{n^2}$ many binary operations on $A$. A binary operation $*$ is commutative means $a * b = b * a$ for all $a, b \in A$. So for two elements $a, b \in A$, it is enough to to know where does $(a, b)$ map to. Now there are exactly $n$ many elements of the form $(a, a)$ in $A \times A$. So there are $n^2 - n$ many elements of the form $(a, b)$ with $a \neq b$. Hence we need to know where do these $(n^2 - n)/2 + n = (n^2 + n)/2$ many elements map to. Again each element has $n$ choices to get mapped to. So one can get total $n^{(n^2+n)/2}$ many commutative binary operations on $A$.

**Definition 1.3.7** Let $A \neq \varnothing$ and $*$ be a binary operation on $A$.

- We say, $*$ satisfies **left cancellation law** if for any $a, b, c \in A$, $a * b = a * c$ implies that $b = c$.

- We say, $*$ satisfies **right cancellation law** if for any $a, b, c \in A$, $a*c = b*c$ implies that $a = b$.

**Example 1.3.8** Define $*$ on $\mathbb{Z} \setminus \{0\}$ by $a * b := |a|b$ for all non-zero integers $a, b$. This binary operation $*$ on $\mathbb{Z} \setminus \{0\}$ satisfies left cancellation law but does not satisfy right cancellation law.

**Definition 1.3.9** Let $A \neq \varnothing$ and $*$ be a binary operation on $A$.

- An element $e' \in A$ is called a **left identity** of $(A, *)$ if for all $a \in A$, $e' * a = a$.
- An element $e'' \in A$ is called a **right identity** of $(A, *)$ if for all $a \in A$, $a * e'' = a$.
- An element $e \in A$ is called an **identity** of $(A, *)$ if for all $a \in A$, $a * e = a = e * a$.

**Examples 1.3.10** The following are few examples of various binary operations with different types of identity elements.

1) Define $*$ on $\mathbb{Z}$ by $a * b = a + b$ for all $a, b \in \mathbb{Z}$. This binary relation has a left identity, a right identity as well as an identity.

2) Define $*$ on $\mathbb{Z}$ by $a * b = a - b$ for all $a, b \in \mathbb{Z}$. This binary relation has a right identity but no left identity and no identity.

3) Define $*$ on $\mathbb{Z}$ by $a * b = |a + b|$ for all $a, b \in \mathbb{Z}$. This binary relation has no left identity, no right identity and no identity.

4) Define $*$ on $\mathbb{Z}$ by $a * b = b$ for all $a, b \in \mathbb{Z}$. This binary relation has a left identity, but no right identity and no identity.

**Theorem 1.3.11** Let $A \neq \varnothing$ and $*$ be a binary operation on $A$. If there exists an identity of $A$, then it is unique.

*Proof.* Let us assume that $e_1, e_2 \in A$ be two identities of $A$. So we have for all $a \in A$,

$$e_1 * a = a \quad \text{and} \quad a * e_2 = a.$$

Hence we get,

$$e_1 = e_1 * e_2 = e_2.$$

This completes the proof.

CHAPTER 2

# Group Theory

## 2.1 Introduction

**Definition 2.1.1** A **group** is an ordered pair $(G, *)$ where $G$ is a non-empty set and $*$ is a binary operation on $G$ such that the following hold:

- For all $a, b, c \in G$, $a * (b * c) = (a * b) * c$.
- There exists an element $e \in G$ such that $a * e = a = e * a$ for all $a \in G$.
- For each $a \in G$, there exists $b \in G$ such that $a * b = e = b * a$.

So $(G, *)$ is a group if $*$ is associative, the identity exists in $G$, and for each element in $G$, we have an inverse of that element in $G$.

**Theorem 2.1.2** Let $(G, *)$ be a group. Then for each $a \in G$, there exists unique $b \in G$ such that $a * b = e = b * a$ where $e$ is the identity of $(G, *)$.

*Proof.* Let $a \in G$. Suppose there are two elements $b, c \in G$ such that

$$a * b = e = b * a \text{ and } \quad a * c = e = c * a.$$

We shall show that $b = c$. Note that

$$b = b * e = b * (a * c) = (b * a) * c = e * c = c.$$

Hence the proof.

**Remark 2.1.3** For each $a \in G$, the unique element $b \in G$ such that $a * b = e = b * a$, is called the inverse of $a$. Generally $b$ is denoted by $a^{-1}$.

**Definition 2.1.4** Let $(G, *)$ be a group.

- $(G, *)$ is called **abelian** or **commutative** if $a * b = b * a$ for all $a, b \in G$. Otherwise $(G, *)$ is called **non-abelian** or **non-commutative**.
- $(G, *)$ is called **finite group** if the number of elements of $G$ is finite. Otherwise it is called an **infinite group**.

**Definition 2.1.5** By the **order of a group** $(G, *)$ we mean the number of elements of $G$. If the number of elements of $G$ is finite, then $G$ is called a group of finite order or a finite group, otherwise $G$ is called a group of infinite order or infinite group.

**Examples 2.1.6** The following are few examples of groups.

1) $(\mathbb{Z}, +)$ is an infinite abelian group.

2) For $n \in \mathbb{N}$, we denote the set of all the congruence classes modulo $n$ by $\mathbb{Z}/n\mathbb{Z}$. We define the binary operation '+' on $\mathbb{Z}/n\mathbb{Z}$ as follows:

   For $[a], [b] \in \mathbb{Z}/n\mathbb{Z}$, $[a] + [b] := [a + b]$. First we show that the operation '+' is well-defined. Let $[a] = [a']$ and $[b] = [b']$. We need to show that $[a + b] = [a' + b']$. Since $[a] = [a']$, we have $n \mid a - a'$. Similarly since $[b] = [b']$, we have $n \mid b - b'$. Therefore $n \mid (a + b) - (a' + b')$. This implies that $[a + b] = [a' + b']$.

   Note that $[0], [1], \ldots, [n - 1]$ are distinct elements of $\mathbb{Z}/n\mathbb{Z}$. Also for any integer $a \in \mathbb{Z}$, by division algorithm we can check that $a \equiv r \pmod{n}$ for some $0 \leq r < n$. Therefore the cardinality of $\mathbb{Z}/n\mathbb{Z}$ is exactly $n$. Under the operation '+', $\mathbb{Z}/n\mathbb{Z}$ forms a group. So $(\mathbb{Z}/n\mathbb{Z}, +)$ is a finite abelian group.

3) Consider $GL_2(\mathbb{R})$, the set of all $2 \times 2$ invertible matrices (with respect to the usual matrix multiplication) with the entries from $\mathbb{R}$. This set under the usual matrix multiplication is a group. This is an infinite non-abelian group.

4) Consider $GL_2(\mathbb{Z}/2\mathbb{Z})$, the set of all $2 \times 2$ invertible matrices (with respect to usual matrix multiplication) with the entries from $\mathbb{Z}/2\mathbb{Z}$. Under the usual matrix multiplication, $GL_2(\mathbb{Z}/2\mathbb{Z})$ is a group. This is a finite non-abelian group.

**Exercise 2.1.7** Let $(G, *)$ be a group. Show that

1) For all $a \in G$, $(a^{-1})^{-1} = a$.

2) For all $a, b \in G$, $(a * b)^{-1} = b^{-1} * a^{-1}$.

From now on, for simplicity we will write $\underbrace{a * \cdots * a}_{n \text{ times}} = a^n$.

**Exercise 2.1.8** Let $(G, *)$ be an abelian group. Show that for all $a, b \in G$ and $n \in \mathbb{Z}$,

$$(a * b)^n = a^n * b^n.$$

**Exercise 2.1.9** Let $(G, *)$ be a group such that for all $a, b \in G$,

$$(a * b)^{-1} = a^{-1} * b^{-1}.$$

Show that $(G, *)$ is an abelian group.

**Definition 2.1.10** Let $(G, *)$ be a group and $a \in G$. Suppose there exists a positive integer $m$ such that $a^m = e$, where $e$ is the identity of $(G, *)$. Then the least positive integer $n$ such that $a^n = e$, is called the **order** of $a$. If no such integer exists, then we call that $a$ is of **infinite order**. By the notation $o(a)$, we refer to the order of $a$.

**Remark 2.1.11** Note that in a group $(G, *)$, for an element $a \in G$ if $o(a) = n$, then $o(a^{-1}) = n$.

**Theorem 2.1.12** Let $(G, *)$ be a group and $a \in G$ be such that $o(a) = n$. Let $m \in \mathbb{N}$ be such that $a^m = e$, then $n|m$.

*Proof.* By division algorithm we know there exist integers $q, r$ such that

$$m = nq + r, \text{ where } 0 \leq r < n.$$

Therefore,

$$a^r = a^{m-qn} = a^m * a^{-qn} = e * (a^{-1})^{qn} = e, \text{ where } e \text{ is the identity of } G.$$

Since $r < n$ and $o(a) = n$, we get that $r = 0$. This proves that $m|n$.

**Theorem 2.1.13** Let $(G, *)$ be a finite group. Then every element of $G$ is of finite order.

*Proof.* Let us assume that $|G| = n$. We have the order of the identity $e$ is 1. Let $a \in G$ be such that $a \neq e$. Consider the collection $\{e, a, a^2, \ldots, a^n\} \subseteq G$. By convention, $a^0 := e$. Since the order of $G$ is $n$, there exist $0 \leq m < r \leq n$ such that $a^m = a^r$. By convention, $a^0 := e$. Therefore, $a^{r-m} = e$, which implies that $o(a) \leq r - m \leq n$. This proves that order of any element in $(G, *)$ is finite.

**Question 2.1.14** Does there exist an infinite group whose every element is of finite order?

Let us first consider the finite group $(\mathbb{Z}/2\mathbb{Z}, +)$. Every non-zero element of this group is of order 2. Now consider the infinite Cartesian product of $\mathbb{Z}/2\mathbb{Z}$. Since infinite Cartesian product of sets having more than one element is infinite, $G := \prod_{\infty} \mathbb{Z}/2\mathbb{Z}$ is infinite. We consider point-wise addition as the binary operation on $G$. Every non-zero element of $(G, +)$ is of order 2.

**Theorem 2.1.15** Let $(G, *)$ be a finite group with $|G| = n$ and $a \in G$. Then $a^n = e$.

*Proof.* If $a = e$, then clearly $a^n = e$. So let $a \neq e$. We know that $o(a) \leq n$. Let $o(a) = m$. Consider $\{e, a, \ldots, a^{m-1}\} \subseteq G$. If $\{e, a, \ldots, a^{m-1}\} = G$, then $m = n$. Therefore $a^n = e$. So let $\{e, a, \ldots, a^{m-1}\} \neq G$, we consider $x \in G \setminus \{e, a, \ldots, a^{m-1}\}$ and the set $\{x, xa, \ldots, xa^{m-1}\}$. We claim that $\{e, a, \ldots, a^{m-1}\} \cap \{x, xa, \ldots, xa^{m-1}\} = \varnothing$. If $xa^k = a^l$, for some $0 \leq k, l < m$, then $x = a^{l-k}$. If $l \geq k$, then $0 \leq l - k < m$ and so $x \in \{e, a, \ldots, a^{m-1}\}$. If $l < k$, then $0 < k - l < m$, and so $x = (a^{-1})^{k-l} \in \{e, a, \ldots, a^{m-1}\}$ as $a^{-1} = a^{m-1}$. In both the cases we get $x \in \{e, a, \ldots, a^{m-1}\}$, a contradiction. So the sets $\{e, a, \ldots, a^{m-1}\}$ and $\{x, xa, \ldots, xa^{m-1}\}$ are disjoint.

Now if $\{e, a, \ldots, a^{m-1}\} \cup \{x, xa, \ldots, xa^{m-1}\} = G$, then $n = 2m$ and hence $a^n = e$. If $\{e, a, \ldots, a^{m-1}\} \cup \{x, xa, \ldots, xa^{m-1}\} \neq G$, then we choose $y \in G$ so that $y \notin \{e, a, \ldots, a^{m-1}\}$ and $y \notin \{x, xa, \ldots, xa^{m-1}\}$. Following the above argument, we similarly get

$$\{e, a, \ldots, a^{m-1}\} \cap \{y, ya, \ldots, ya^{m-1}\} = \varnothing \text{ and } \{y, ya, \ldots, ya^{m-1}\} \cap \{x, xa, \ldots, xa^{m-1}\} = \varnothing.$$

If $\{e, a, \ldots, a^{m-1}\} \cup \{x, xa, \ldots, xa^{m-1}\} \cup \{y, ya, \ldots, ya^{m-1}\} = G$, we get $n = 3m$ and hence $a^n = e$, otherwise we continue the above process which will stop after finitely many steps as $G$ is finite. Hence $G$ can be written as a disjoint union of finitely many sets each having $m$ elements i.e. $n = rm$ for some positive integer $r$ and hence $a^n = e$.

**Theorem 2.1.16** Let $(G, *)$ be a group and $a \in G$ be such that $o(a) = n$. Then for every $m \in \mathbb{N}$,

$$o(a^m) = \frac{n}{\gcd(m, n)}.$$

*Proof.* Let us assume that $o(a^m) = k$. So $a^{mk} = e$, where $e$ is the identity of $G$. Now since $o(a) = n$, we get that $n \mid mk$. Therefore, $mk = nr$ for some $r \in \mathbb{N}$.

Also let $\gcd(m, n) = d$. So there exist $u, v \in \mathbb{N}$ such that $m = du$ and $n = dv$ with $\gcd(u, v) = 1$. So substituting the values of $m, n$ in $mk = nr$ we get that, $duk = dvr$ i.e. $uk = vr$. This implies that $v \mid uk$. Since $\gcd(u, v) = 1$, we get that $v|k$ i.e. $(n/d) \mid k$. Again,

$$(a^m)^{\frac{n}{d}} = a^{\frac{mn}{d}} = a^{\frac{dun}{d}} = a^{un} = e.$$

As $o(a^m) = k$, we also get that $k \mid (n/d)$. This proves that $k = n/d$ i.e.

$$o(a^m) = \frac{n}{\gcd(m, n)}.$$

## 2.2 Subgroups

**Definition 2.2.1** Let $(G, *)$ be a group and $H$ be a non-empty subset of $G$. We say that $H$ is a **subgroup** of $(G, *)$ if $(H, *)$ is also a group.

**Remark 2.2.2** In a group $(G, *)$, we always have two subgroups $H = \{e\}$, where $e$ is the identity of $(G, *)$ and $H = G$. They are called **trivial subgroups** of $(G, *)$. A subgroup $H$ of $G$ such that $H \neq \{e\}$ and $H \neq G$, is called a **non-trivial** subgroup of $(G, *)$.

**Example 2.2.3** The set of all even numbers i.e. $2\mathbb{Z}$ is a subgroup of $(\mathbb{Z}, +)$.

**Theorem 2.2.4** All subgroups of $(G, *)$ have same identity.

*Proof.* Let $H$ be a subgroup of $(G, *)$. We denote the identity of $(G, *)$ by $e$. We show that $e$ is also the identity of $(H, *)$. Let us denote the identity of $H$ by $e_H$. So we have,

$$e_H = e_H * e_H.$$

Also we have,

$$e_H = e_H * e, \text{ as } e \text{ is the identity of } (G, *).$$

Therefore, $e_H * e_H = e_H * e$ and hence using left cancellation law we get $e_H = e$.

**Remark 2.2.5** For $a \in H$, where $H$ is a subgroup of $(G, *)$, the inverse of $a$ in $G$ and $H$ are same.

**Theorem 2.2.6** Let $(G, *)$ be a group and $H$ be a non-empty subset of $G$. Then $H$ is a subgroup of $(G, *)$ if and only if for all $a, b \in H$, $a * b^{-1} \in H$.

*Proof.* First we assume that $H$ is a subgroup of $(G, *)$. We shall show that for all $a, b \in H$, $a * b^{-1} \in H$. Let $a, b \in H$. Since $H$ is a subgroup of $G$, we have $a, b^{-1} \in H$ and therefore $a * b^{-1} \in H$.

Next assume that for all $a, b \in H$, $a * b^{-1} \in H$. We show that $H$ is a subgroup of $(G, *)$. Since $H \neq \varnothing$, there exists an element $a \in H$. Now from the hypothesis we get $e = a * a^{-1} \in H$, where $e$ is the identity of $(G, *)$.
Now let $b \in H$, so for $e, b \in H$, we get from the hypothesis that $b^{-1} = e * b^{-1} \in H$. So we have proved that every element in $H$ has an inverse in $H$.
Let $a, b \in H$. So $a, b^{-1} \in H$ as we have already shown that each element of $H$ has an inverse in $H$. Now from the hypothesis we obtain that $a * b = a * (b^{-1})^{-1} \in H$. This proves that $*$ is a binary operation on $H$.
The associativity of the elements of $H$ under $*$ follows from the fact that $H$ is a subset of $G$. This completes the proof of the fact that $H$ is a subgroup of $(G, *)$.

**Theorem 2.2.7** Let $(G, *)$ be a group and $\{H_\alpha : \alpha \in I\}$ be a collection of subgroups of $G$. Then $H := \bigcap_{\alpha \in I} H_\alpha$ is a subgroup of $(G, *)$.

*Proof.* Since $H_\alpha$'s are subgroups of $(G, *)$, $e \in H_\alpha$ for all $\alpha \in I$ where $e$ is the identity of $(G, *)$. Therefore $H \neq \varnothing$. Now let $a, b \in H$. Then $a, b \in H_\alpha$ for all $\alpha \in I$. Since each $H_\alpha$ is a subgroup of $(G, *)$, we have $a, b^{-1} \in H_\alpha$ and therefore $a * b^{-1} \in H_\alpha$ for all $\alpha \in I$. This proves that $a * b^{-1} \in H$ and hence $H$ is a subgroup of $(G, *)$.

**Definition 2.2.8** Let $(G, *)$ be a group. By **center** of the group $(G, *)$ we mean the set $Z(G) := \{x \in G : x * g = g * x \text{ for all } g \in G\}$.

**Remark 2.2.9** Note that $Z(G) \neq \varnothing$ as the identity of $G$ is always in $Z(G)$. Also note that a group $(G, *)$ is abelian if and only if $G = Z(G)$.

**Theorem 2.2.10** The center $Z(G)$ is a subgroup of $(G, *)$.

*Proof.* We have already observed that $Z(G) \neq \varnothing$. So let $a, b \in Z(G)$. We show that $a * b^{-1} \in Z(G)$. Since $b \in Z(G)$, we have $b * g = g * b$ for all $g \in G$. Therefore $g * b^{-1} = b^{-1} * g$ for all $g \in G$. Hence we have for all $g \in G$,

$$(a * b^{-1}) * g = a * (b^{-1} * g) = a * (g * b^{-1}) = (a * g) * b^{-1} = (g * a) * b^{-1} = g * (a * b^{-1}).$$

This completes the proof.

**Definition 2.2.11** Let $A \neq \varnothing$ be a subset of a group $(G, *)$. By the **centralizer** of $A$ in $(G, *)$ we mean the following set,

$$C_G(A) := \{g \in G : g * a = a * g \text{ for all } a \in G\}.$$

**Exercise 2.2.12** Prove that $C_G(A)$ is a subgroup of $(G, *)$.

**Remark 2.2.13** Note that $Z(G) \subseteq C_G(A)$ and $C_G(G) = Z(G)$.

Let $H, K$ be two subgroups of a group $(G, *)$. Then we define,

$$HK := \{h * k : h \in H, k \in K\}.$$

In general $HK$ need not be a subgroup of $(G, *)$. The following theorem gives us a necessary and sufficient condition on $HK$ so that $HK$ becomes a subgroup of $(G, *)$.

**Theorem 2.2.14** The following are equivalent:

1) $HK$ is a subgroup of $(G, *)$.

2) $HK = KH$.

3) $KH$ is a subgroup of $(G, *)$.

*Proof.* First we assume that $HK$ is a subgroup of $(G, *)$. We will show that $HK = KH$.

Let $h * k \in HK$ where $h \in H$ and $k \in K$. So $(h * k)^{-1} = k^{-1} * h^{-1} \in HK$. We write $k^{-1} * h^{-1} = h_1 * k_1$ where $h_1 \in H$ and $k_1 \in K$. Therefore,

$$h * k = (k^{-1} * h^{-1})^{-1} = (h_1 * k_1)^{-1} = k_1^{-1} * h_1^{-1} \in KH.$$

This proves that $HK \subseteq KH$.

Now let $k * h \in KH$ where $k \in K$ and $h \in H$. Note that,

$$k = e * k \in HK \quad \text{and} \quad h = h * e \in HK,$$

where $e$ denotes the identity of $(G, *)$. Since $HK$ is a subgroup of $(G, *)$, we obtain that $k * h \in HK$. Hence $KH \subseteq HK$. This completes the proof of the fact that if $HK$ is a subgroup of $(G, *)$ then $HK = KH$.

Next we assume that $HK = KH$. We shall show that $HK$ is a subgroup of $(G, *)$. Let $h_1 * k_1, h_2 * k_2 \in HK$ where $h_1, h_2 \in H$ and $k_1, k_2 \in K$. Now,

$$(h_1 * k_1) * (h_2 * k_2)^{-1} = h_1 * (k_1 * k_2^{-1}) * h_2^{-1}.$$

We know $(k_1 * k_2^{-1}) * h_2^{-1} \in KH$. We have $KH = HK$. So we can write $(k_1 * k_2^{-1}) * h_2^{-1} = h * k$ for some $h \in H$ and $k \in K$. Hence $(h_1 * k_1) * (h_2 * k_2)^{-1} = (h_1 * h) * k \in HK$. This proves that $HK$ is a subgroup of $(G, *)$.

**Remarks 2.2.15**

- If $(G, *)$ is an abelian group, then for any two subgroups $H, K$ of $(G, *)$, we get $HK$ is a subgroup of $(G, *)$.

- In general for non-abelian groups $(G, *)$, for any two subgroups $H, K$ of $(G, *)$, $HK$ need not be a subgroup of $(G, *)$. For example, consider $G = GL_2(\mathbb{Z}/2\mathbb{Z})$ with usual matrix multiplication. We take the subgroups

$$H = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \right\},$$

$$K = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\}.$$

  Now

$$HK = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \right\}.$$

  Note that order of $\left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ is 3 and as $3 \nmid 4$, hence we conclude that $HK$ is not a subgroup of $G$.

## 2.3 Cyclic groups

**Definition 2.3.1** A group $(G, *)$ is called **cyclic** if there exists an element $a \in G$ such that $G = \{a^n : n \in \mathbb{Z}\}$. The set $\{a^n : n \in \mathbb{Z}\}$ is denoted by $\langle a \rangle$.

**Definition 2.3.2** Let $(G, *)$ be a cyclic group such that $G = \langle a \rangle$ for some $a \in G$. The element $a$ is called a **generator** of $G$.

**Remark 2.3.3** For a cyclic group $(G, *)$, if $a$ is a generator of $G$, then $a^{-1}$ is also a generator of $G$.

**Examples 2.3.4** The following are few examples of cyclic groups.

1) The group $(\mathbb{Z}, +)$ is an infinite cyclic group and $\mathbb{Z} = \langle 1 \rangle$.

2) For any $n \in \mathbb{N}$, the group $(\mathbb{Z}/n\mathbb{Z}, +)$ is a finite cyclic group and $\mathbb{Z}/n\mathbb{Z} = \langle [1] \rangle$.

**Theorem 2.3.5** Every cyclic group is abelian.

*Proof.* Let $G = \langle a \rangle$ be a cyclic group. Let $b, c \in G$. Then $b = a^m$ and $c = a^n$ for some $m, n \in \mathbb{Z}$. Therefore,

$$b * c = a^m * a^n = a^{m+n} = a^{n+m} = a^n * a^m = c * b.$$

This proves that $(G, *)$ is an abelian group.

The converse of the above theorem is not true i.e. there are abelian groups which are not cyclic. Here we give one example of an abelian group which is not cyclic. Let $G = \{e, a, b, c\}$ and define $*$ on $G$ by the following:

$$a * a = b * b = c * c = e,$$
$$a * b = b * a = c,\ b * c = c * b = a,\ a * c = c * a = b,$$
and $e$ is the identity element.

This group $(G, *)$ is known as Klein's 4 group. Clearly this group is abelian but not cyclic.

**Theorem 2.3.6** Let $(G, *)$ be a finite group of order $n$. Then $(G, *)$ is cyclic if and only if there exists an element $a \in G$ such that $o(a) = n$.

*Proof.* Throughout the proof we denote $e$ as the identity of $(G, *)$.

First we assume that we have an element $a \in G$ such that $o(a) = n$. We show that $G = \langle a \rangle$, and hence $G$ is cyclic. Since $o(a) = n$, we have that the number of elements of the set $\{e, a, \ldots, a^{n-1}\}$ is $n$. As $\{e, a, \ldots, a^{n-1}\} \subseteq G$ and both the sets have same number of elements, we obtain that $G = \{e, a, \ldots, a^{n-1}\} = \langle a \rangle$.

Next we assume that $(G, *)$ is a cyclic group of order $n$. Let $a$ be a generator of $(G, *)$. We shall show that $o(a) = n$. We know that in a finite group, order of each element is finite. So let $o(a) = m$. Then the number of elements in the set $S := \{e, a, \ldots, a^{m-1}\} \subseteq G$ is $m$. Now let $a^k \in G$. Using division algorithm we get two integers $m, r$ such that $k = mq + r$ where $0 \leq r < m$. Therefore,

$$a^k = a^{mq+r} = a^{mq} * a^r = e * a^r = a^r \in S, \text{ as } r < m.$$

This proves that $S = G$. Therefore $m = |S| = |G| = n$. Hence we have found an element in $G$ whose order is $n$.

**Corollary 2.3.7** Every group of prime order is cyclic.

*Proof.* Let $(G, *)$ be a group of prime order $p$. It is enough for us to find an element of order $p$ in $(G, *)$. As $|G| = p > 1$, there exists $a \neq e$ in $G$. We know that in a finite group order of every element divides order of the group. So $o(a) = 1$ or $p$. Since $a \neq e$, we get $o(a) = p$.

**Exercise 2.3.8** Let $G = \langle a \rangle$ and $o(a) = n$. Let $k \in \mathbb{N}$ be such that $1 \leq k \leq n$. Then show that $a^k$ is also a generator of $G$ if and only if $\gcd(k, n) = 1$.

**Theorem 2.3.9** All subgroups of a cyclic group are cyclic.

*Proof.* Let $G = \langle a \rangle$ and $H$ be a subgroup of $G$. If $H = \{e\}$, where $e$ is the identity of $(G, *)$, then clearly $H$ is cyclic.

So let $H \neq \{e\}$, then there exists an element $b \in H$ such that $b \neq e$. Since $b \in G$ also, there exists $m \in \mathbb{Z}$ such that $b = a^m$. Also $b^{-1} = a^{-m} \in H$. Among $m, -m$, one is positive. So we choose $n$ to be the least positive integer so that $a^n \in H$. We show that $H = \langle a^n \rangle$. Clearly $\langle a^n \rangle \subseteq H$ as $a^n \in H$. It only remains to show that $H \subseteq \langle a^n \rangle$. Let $h \in H$. Now $h = a^s$, for some $s \in \mathbb{Z}$. By division algorithm we can find integers $q, r$ such that $s = qn + r$ and $0 \leq r < n$. Therefore,

$$a^r = a^{s-qn} = a^s * (a^{-1})^{qn} \in H.$$

By the choice of $n$ we get $r = 0$ as $r < n$. Hence $s = qn$. This proves that $h \in \langle a^n \rangle$. This completes the proof that $H$ is cyclic.

**Theorem 2.3.10** Let $G$ be a finite cyclic group and $H$ is a subgroup of $G$. Then $|H|$ divides $|G|$.

*Proof.* If $H = \{e\}$, then $|H| = 1$, so $|H|$ divides $|G|$. So let $H \neq \{e\}$. Since $H$ is a subgroup of a cyclic group $G$, $H$ is also cyclic. So there exists $b \in H$ such that $o(b) = |H|$. Now $b \in G$ also. Let $G = \langle a \rangle$. Then $b = a^m$ for some integer $m$. Also $b^{-1} = a^{-m} \in H$ and $o(b) = o(b^{-1})$. Among $m, -m$, one is positive. Without loss of generality let $m > 0$. Then

$$o(a^m) = \frac{o(a)}{\gcd(o(a), m)}.$$

This proves that $o(a^m)$ divides $o(a)$ i.e. $|H|$ divides $|G|$.

**Theorem 2.3.11** Let $G$ be a finite cyclic group of order $n$. Let $m \in \mathbb{N}$ be such that $m \mid n$. Then there exists a unique subgroup of $G$ of order $m$.

*Proof.* Let $G = \langle a \rangle$. Then $o(a) = |G| = n$. First we show the existence of a subgroup of $G$ of order $m$. Given that $m \mid n$. Therefore, $n = mk$ for some $k \in \mathbb{N}$. Consider $H := \langle a^k \rangle$. Now we have,

$$o(a^k) = \frac{o(a)}{\gcd(o(a), k)} = \frac{n}{k} = m.$$

Now $|H| = o(a^k) = m$. This proves the existence of a subgroup of $G$ of order $m$.

Now we show the uniqueness of such a subgroup. Let $K$ be a subgroup of $G$ of order $m$. We shall show that $K = H$. We know that $K$ is also cyclic as $K$ is a subgroup of a cyclic group. Let $K := \langle a^t \rangle$ for some $t \in \mathbb{Z}$. Now $m = |K| = o(a^t)$. Therefore $a^{tm} = e$, where $e$ is the identity of $(G, *)$. Since $o(a) = n$, we have $n \mid tm$. Hence $tm = nr$ for some $r \in \mathbb{Z}$. This implies that $t = kr$. So $a^t = a^{kr} \in H$. We have $|H| = |K|$ and $K \subseteq H$. Therefore, $K = H$.

**Remark 2.3.12** As a consequence of the above theorem we can say that in a cyclic group of order $n \in \mathbb{N}$, there are exactly $d(n)$ many subgroups where $d(n)$ denotes the number of positive divisors of $n$.

## 2.4 Permutation groups

**Definition 2.4.1** Let $A$ be a non-empty set. By a **permutation** of $A$, we mean a bijective map from $A$ to itself.

**Definition 2.4.2** A group $(G, *)$ is called a **permutation group**, if $G$ consists of 'some' permutations of a non-empty set $A$ and the operation $*$ is the composition of maps.

As an example of permutation group, we look at the symmetric group with $n$ symbols in more detail. Let $I_n = \{1, 2, \ldots, n\}$. Denote, $S_n$ as the collection of all permutations of $I_n$. Under the operation composition of maps, $S_n$ forms a group which we refer to as the **symmetric group with $n$ symbols**. Note that $|S_n| = n!$. Let $\sigma \in S_n$ be such that for any $i \in I_n$, $\sigma(i) := a_i$, $a_i \in I_n$ and $a_i \neq a_j$ for all $i \neq j$. Then one can denote $\sigma$ by the following two row notation:

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ a_1 & a_2 & \cdots & a_n \end{pmatrix}.$$

**Example 2.4.3** We consider the group $(S_3, \circ)$. We know that $|S_3| = 6$. Below we list down all the elements of $S_3$ in two row notation.

$$\begin{pmatrix} 1\ 2\ 3 \\ 1\ 2\ 3 \end{pmatrix}, \begin{pmatrix} 1\ 2\ 3 \\ 1\ 3\ 2 \end{pmatrix}, \begin{pmatrix} 1\ 2\ 3 \\ 2\ 1\ 3 \end{pmatrix}, \begin{pmatrix} 1\ 2\ 3 \\ 2\ 3\ 1 \end{pmatrix}, \begin{pmatrix} 1\ 2\ 3 \\ 3\ 2\ 1 \end{pmatrix}, \begin{pmatrix} 1\ 2\ 3 \\ 3\ 1\ 2 \end{pmatrix}.$$

Note that,

$$\begin{pmatrix} 1\ 2\ 3 \\ 1\ 3\ 2 \end{pmatrix} \circ \begin{pmatrix} 1\ 2\ 3 \\ 2\ 1\ 3 \end{pmatrix} = \begin{pmatrix} 1\ 2\ 3 \\ 3\ 1\ 2 \end{pmatrix},$$

and

$$\begin{pmatrix} 1\ 2\ 3 \\ 2\ 1\ 3 \end{pmatrix} \circ \begin{pmatrix} 1\ 2\ 3 \\ 1\ 3\ 2 \end{pmatrix} = \begin{pmatrix} 1\ 2\ 3 \\ 2\ 3\ 1 \end{pmatrix}.$$

So we see that

$$\begin{pmatrix} 1\ 2\ 3 \\ 1\ 3\ 2 \end{pmatrix} \circ \begin{pmatrix} 1\ 2\ 3 \\ 2\ 1\ 3 \end{pmatrix} \neq \begin{pmatrix} 1\ 2\ 3 \\ 2\ 1\ 3 \end{pmatrix} \circ \begin{pmatrix} 1\ 2\ 3 \\ 1\ 3\ 2 \end{pmatrix}.$$

This shows that $(S_3, \circ)$ is non-abelian.

**Exercise 2.4.4** Show that $(S_n, \circ)$ is non-abelian for all $n \geq 3$.

Now we give examples of subgroups $H, K$ of $(S_3, \circ)$ such that $HK$ is not a subgroup of $(S_3, \circ)$.

**Definition 2.4.5** A permutation $\sigma$ of $I_n$ is called a $k$-**cycle** or a **cycle of length** $k$ if

there exist distinct $a_1, a_2, \ldots, a_k \in I_n$ such that $\sigma(a_i) = a_{i+1}$ for all $1 \leq i \leq (k-1)$, $\sigma(a_k) = a_1$, and $\sigma(a) = a$ for all $a \in I_n \setminus \{a_1, a_2, \ldots, a_k\}$. We denote $\sigma$ as $(a_1\ a_2\ \cdots\ a_k)$.

By convention we consider the identity permutation as 1-cycle.

**Remarks 2.4.6** Note that

1) In $(S_3, \circ)$, all the non-identity permutations of can be written as cycles of length less than or equal to 3. Below are the non-identity elements of $(S_3, \circ)$:

$$(2\ 3), (1\ 2), (1\ 2\ 3), (1\ 3), (1\ 3\ 2).$$

2) As we see in $(S_3, \circ)$, all non-identity elements are cycles, we may wonder whether for any $n$, all non-identity elements of $(S_n, \circ)$ are cycles?

   Let us consider,
   $$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix} \in S_4.$$
   Note that
   $$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix} = (1\ 2)(3\ 4) = (3\ 4)(1\ 2).$$
   So we see by this example that not all non-identity permutations are cycles for $n \geq 4$.

**Theorem 2.4.7** Let $\sigma \in S_n$ be a cycle. Then $\sigma$ is a $k$-cycle if and only if $o(\sigma) = k$.

*Proof.* First assume that $\sigma$ is a $k$-cycle. Let us write,

$$\sigma = (a_1\ \cdots\ a_k).$$

Note that for any $1 \leq m < k$,

$$\sigma^m(a_1) = a_{m+1} \neq a_1.$$

Also, $\sigma^k(a_i) = a_i$ for all $1 \leq i \leq k$. This proves that $o(\sigma) = k$.

Next assume that $\sigma$ is a cycle and $o(\sigma) = k$. We show that $\sigma$ is a $k$-cycle. Let length of $\sigma$ be $m$. So, $o(\sigma) = m$ as order of an $m$-cycle is $m$. Given that $o(\sigma) = k$. Therefore $m = k$. This proves that $\sigma$ is a $k$-cycle.

**Remark 2.4.8** So it justifies the convention of considering the identity permutation as 1-cycle.

**Definition 2.4.9** Two cycles $(a_1\ a_2\ \cdots\ a_k)$ and $(b_1\ b_2\ \cdots\ a_r)$ in $(S_n, \circ)$ are called **disjoint cycles** if $\{a_1, a_2, \ldots,\ a_k\} \bigcap \{b_1, b_2, \ldots,\ b_r\} = \varnothing$.

**Exercise 2.4.10** Let $\alpha, \beta$ be two disjoint cycles in $(S_n, \circ)$. Show that $\alpha \circ \beta = \beta \circ \alpha$.

**Theorem 2.4.11** Let $\sigma \in S_n$, $n \geq 2$ be such that $\sigma = \sigma_1 \circ \cdots \circ \sigma_r$, where $\sigma_i$'s are disjoint cycles for all $1 \leq i \leq r$. Let $o(\sigma_i) = n_i$ for all $1 \leq i \leq r$. Then $o(\sigma) = \text{lcm}(n_1, \ldots, n_r)$.

*Proof.* Let $o(\sigma) = t$ and $\text{lcm}(n_1, \ldots, n_r) = m$. We want to show that $t = m$. Note that,

$$\begin{aligned} \sigma^m = (\sigma_1 \circ \cdots \circ \sigma_r)^m &= \sigma_1^m \circ \cdots \circ \sigma_r^m, \text{ as } \sigma_i\text{'s are disjoint cycles.} \\ &= e. \end{aligned}$$

Therefore, $t \mid m$. Now to show $m \mid t$, it is enough to show that $n_i \mid t$ for all $1 \leq i \leq r$. For that we show, $\sigma_i^t = e$ for all $1 \leq i \leq r$. Let $k \in I_n$.

- If $\sigma_i(k) = k$, then $\sigma_i^t(k) = k$.
- If $\sigma_i(k) \neq k$, then $\sigma_j(k) = k$ for all $j \neq i$ as $\sigma_i$'s are disjoint cycles. Therefore, $\sigma(k) = \sigma_i(k)$. So $\sigma_i^t(k) = \sigma^t(k) = k$.

This proves that $n_i \mid t$, and hence $m \mid t$. This completes the proof.

We saw earlier that

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix} \in S_4,$$

is not a cycle but a product of two disjoint 2-cycles. So we may wonder that for any $n$, can we say that any non-identity permutation in $(S_n, \circ)$ is a product of disjoint cycles? To answer this, we have the following theorem.

**Theorem 2.4.12** Let $n \geq 2$ and $\sigma \in S_n$ be such that $\sigma \neq e$. Then $\sigma$ can be written as a product of disjoint cycles of length $\geq 2$.

*Proof.* We have that $\sigma \neq e$ and $n \geq 2$, so there exists $m \in I_n$ such that $\sigma(m) \neq m$. We know that $(S_n, \circ)$ is a finite group. So the order of $\sigma$ is finite. Let $o(\sigma) = t$ i.e. $\sigma^t = e$. This implies that $\sigma^t(m) = m$. We choose $1 < k \leq t$ to be the least positive integer so that $\sigma^k(m) = m$. Now we consider,

$$\alpha_1 := (m\ \sigma(m)\ \cdots\ \sigma^{k-1}(m)).$$

**Case: 1** All the elements of $I_n$, which are not fixed by $\sigma$ appear in $\alpha_1$.

Let $a \in I_n$ be such that $\sigma(a) \neq a$ and $a \in \alpha_1$, then the map $\alpha_1$ takes $a$ to $\sigma(a)$. Now let $b \in I_n$ be such that $\sigma(b) = b$, then $b \notin \alpha_1$ as $\sigma^l(b) = b$ for any positive integer $l$, so $\alpha_1(b) = b$. This proves that the two maps $\alpha_1$ and $\sigma$ and same. So $\sigma$ itself is a cycle.

**Case: 2** There are elements of $I_n$, which are not fixed by $\sigma$ and do not appear in $\alpha_1$.

Let $m' \in I_n$ be such that $\sigma(m') \neq m'$ and $m'$ does not appear in $I_n$. Again we have, $\sigma^t(m') = m'$. So we choose the least positive integer $1 < k' \leq t$ such that $\sigma^{k'}(m') = m'$. We write,

$$\alpha_2 := (m' \ \sigma(m') \ \cdots \ \sigma^{k'-1}(m')).$$

We show that $\alpha_1, \alpha_2$ are disjoint cycles. We know that $m \neq m'$. Now if $\alpha_1, \alpha_2$ are not disjoint, then there exist integers $p, q$, $1 \leq p \leq k$, $1 \leq q \leq k'$ such that $\sigma^p(m) = \sigma^q(m')$. Now,

$$\sigma^{p+1}(m) = \sigma(\sigma^p(m)) = \sigma(\sigma^q(m')) = \sigma^{q+1}(m').$$

Therefore,

$$\sigma^{p+i}(m) = \sigma^{q+i}(m'), \text{ for all } i \in \mathbb{N} \cup \{0\}.$$

As $q \leq k'$, we can choose an integer $j \in \mathbb{N} \cup \{0\}$ such that $q + j = k'$. So,

$$m' = \sigma^{k'}(m') = \sigma^{q+j}(m') = \sigma^{p+j}(m).$$

This shows that $m' \in \alpha_1$, a contradiction. So $\alpha_1, \alpha_2$ are disjoint cycles and hence

$$\alpha_1 \circ \alpha_2 = \alpha_2 \circ \alpha_1.$$

**Case: 2A** All the elements which are not fixed by $\sigma$ appear in either $\alpha_1$ or $\alpha_2$.

Let $\sigma(a) \neq a$. If $a \in \alpha_1$, then $a \notin \alpha_2$, as $\alpha_1, \alpha_2$ are disjoint cycles. Now $a \notin \alpha_2$ implies that $\alpha_2(a) = a$. So,

$$\alpha_1 \circ \alpha_2(a) = \alpha_1(a) = \sigma(a).$$

Similarly if $a \in \alpha_2$, then

$$\alpha_1 \circ \alpha_2(a) = \alpha_2(a) = \sigma(a).$$

If $\sigma(b) = b$, then following the justification given in Case: 1 we get that $b$ is fixed by both $\alpha_1, \alpha_2$. This proves that $\sigma = \alpha_1 \circ \alpha_2$.

**Case: 2B** There are elements of $I_n$, which are not fixed by $\sigma$ and do not appear in $\alpha_1, \alpha_2$.

Then again with such an element we construct a cycle $\alpha_3$. We repeat this process,

it terminates at a finite stage as $I_n$ is finite. In the process, the cycles we construct are disjoint. If we finally get $r$ many disjoint cycles $\alpha_1, \ldots, \alpha_r$, then following the justification given in Case 2A we get that

$$\sigma = \alpha_1 \circ \cdots \circ \alpha_r.$$

This completes the proof.

**Proposition 2.4.13** Let $\sigma$ be an $n$ - cycle and $k \in \mathbb{N}$. Then $\sigma^k$ is a product of $d := \gcd(k, n)$ many disjoint $n/d$ - cycles.

*Proof.* We write $\sigma = (a_1 \ \cdots \ a_n)$. We have, the order of $\sigma^k$ is $n/d$. By Theorem 2.4.11 we know that $\sigma^k$ can be written as a product of disjoint cycles. So it is enough to show that for each $i$, the cycle containing $a_i$ in this cycle decomposition is of length $n/d$. Then clearly $d$ many such disjoint cycles appear in this cycle decomposition of $\sigma^k$. Now,

$$\begin{aligned}
\sigma^k(a_i) &= a_{(k+i) \pmod n}, \\
\sigma^k(a_{(k+i) \pmod n}) &= a_{(2k+i) \pmod n}, \\
&\vdots \\
\sigma^k(a_{(tk+i) \pmod n}) &= a_{((t+1)k+i) \pmod n}.
\end{aligned}$$

So the length of the cycle containing $a_i$ in the cycle decomposition of $\sigma^k$ is the least positive integer $r$ such that $rk + i \equiv i \pmod n$. Therefore we get $n \mid rk$. Hence $r = n/d$. This completes the proof.

**Remark 2.4.14** Let $\tau = (1\ 2)(3\ 4\ 5)$. As a consequence of Proposition 2.4.13 we can immediately conclude tha there is no $n$ - cycle $\sigma$ such that $\tau = \sigma^k$.

**Definition 2.4.15** A 2-cycle in $S_n$ is called a **transposition**.

**Theorem 2.4.16** Every cycle of length greater than or equal to 2 can be written as a product of transpositions.

*Proof.* Let $\sigma = (a_1\ a_2\ \cdots\ a_k)$ be a $k$-cycle and $k \geq 2$. We claim that

$$\sigma = (a_1\ a_k) \circ \cdots \circ (a_1\ a_3) \circ (a_1\ a_2).$$

We denote $\alpha = (a_1\ a_k) \circ \cdots \circ (a_1\ a_3) \circ (a_1\ a_2)$. Note that $(a_1\ a_2)$ maps $a_1$ to $a_2$, and $a_2$ does not belong to the rest of the transpositions, and hence $a_2$ is fixed by the rest $(k-2)$ many transpositions. So $\alpha$ maps $a_1$ to $a_2$. Now $(a_1\ a_2)$ maps $a_2$ to $a_1$, $(a_1\ a_3)$ maps $a_1$ to $a_3$, and $a_3$ does not belong to the rest of the transpositions, and hence $a_3$ is fixed by

the rest $(k-3)$ many transpositions. So $\alpha$ maps $a_2$ to $a_3$. Now $(a_1\ a_2)$ maps $a_3$ to $a_3$, $(a_1\ a_3)$ maps $a_3$ to $a_1$, $(a_1\ a_4)$ maps $a_1$ to $a_4$ and $a_4$ does not belong to the rest of the transpositions, and hence $a_4$ is fixed by the rest $(k-4)$ many transpositions. So $\alpha$ maps $a_3$ to $a_4$. Proceeding this way we get that $\alpha$ maps $a_i$ to $a_{i+1}$ for all $1 \leq i \leq (k-1)$ and $\alpha$ maps $a_k$ to $a_1$ as $a_k$ belongs to only $(a_1\ a_k)$. So we get that $\alpha = \sigma$ and this completes the proof of the theorem.

So from the above two theorems we can conclude the following:

**Theorem 2.4.17** Every non-identity permutation in $S_n$ can be written as a product of transpositions.

At this point, we may wonder what about the identity permutation? Note that we can write $e = (1\ k) \circ (1\ k)$ for any $2 \leq k \leq n$. So $e$ can be written as a product of two transpositions. In fact, we can show that $e$ can only be a product of even number of transpositions.

**Theorem 2.4.18** The identity permutation $e$ in $S_n$ can not be written as a product of odd many permutations.

*Proof.* We consider a polynomial in $n$ variables as follows:

$$P(X_1, \ldots, X_n) = \prod_{1 \leq k < l \leq n} (X_k - X_l).$$

For any permutation $\sigma \in S_n$, we define

$$\sigma(P(X_1, \ldots, X_n)) := P(X_\sigma(1), \ldots, X_\sigma(n)) = \prod_{1 \leq k < l \leq n} (X_{\sigma(k)} - X_{\sigma(l)}).$$

So according to our definition,

$$e(P(X_1, \ldots, X_n)) = \prod_{1 \leq k < l \leq n} (X_k - X_l) = P(X_1, \ldots, X_n).$$

Suppose, $e = \sigma_1 \circ \cdots \circ \sigma_r$ where $r$ is an odd number and $\sigma_1, \ldots, \sigma_r$ are transpositions. Let $\sigma = (i\ j)$ and $i < j$ be a transposition. We show that

$$\sigma(P(X_1, \ldots, X_n)) = -P(X_1, \ldots, X_n).$$

Below we write down the factors of $P(X_1, \ldots, X_n)$:

$$
\begin{aligned}
&(X_1 - X_2)\cdots(X_1 - X_n)\\
&(X_2 - X_3)\cdots(X_2 - X_n)\\
&\vdots\\
&(X_{i-1} - X_i)\cdots(X_{i-1} - X_j)\cdots(X_{i-1} - X_n)\\
&(X_i - X_{i+1})\cdots(X_i - X_{j-1})(X_i - X_j)(X_i - X_{j+1})\cdots(X_i - X_n)\\
&(X_{i+1} - X_{i+2})\cdots(X_{i+1} - X_{j-1})(X_{i+1} - X_j)(X_{i+1} - X_{j+1})\cdots(X_{i+1} - X_n)\\
&\vdots\\
&(X_{j-1} - X_j)(X_{j-1} - X_{j+1})\cdots(X_{j-1} - X_n)\\
&(X_j - X_{j+1})\cdots(X_j - X_n)\\
&(X_{j+1} - X_{j+2})\cdots(X_{j+1} - X_n)\\
&\vdots\\
&(X_{n-1} - X_n)
\end{aligned}
$$

We see that under $\sigma_1$,

$$
\begin{aligned}
(X_i - X_{i+1})\cdots(X_i - X_{j-1}) \mapsto &(X_j - X_{i+1})\cdots(X_j - X_{j-1})\\
&= (-1)^{j-i-1}(X_{i+1} - X_j)\cdots(X_{j-1} - X_j).
\end{aligned}
$$

Also under $\sigma$, each factor of the form $(X_{i+k} - X_j) \mapsto (X_{i+k} - X_i) = -(X_i - X_{i+k})$ for every $1 \leq k \leq j - i - 1$. Hence

$$
\begin{aligned}
(X_{i+1} - X_j)\cdots(X_{j-1} - X_j) \mapsto &(X_{i+1} - X_i)\cdots(X_{j-1} - X_i)\\
&= (-1)^{j-i-1}(X_i - X_{i+1})\cdots(X_i - X_{j-1}).
\end{aligned}
$$

Hence under $\sigma$, the factor

$$(X_i - X_{i+1})\cdots(X_i - X_{j-1})(X_{i+1} - X_j)\cdots(X_{j-1} - X_j)$$

remains fixed. The factor $(X_i - X_j) \mapsto (X_j - X_i) = -(X_i - X_j)$.

Now for the remaining factors, each of them either gets fixed under $\sigma$ or gets mapped to another factor among them. Note that $\sigma$ is a bijection of $I_n$, therefore two different factors can never get mapped to a same factor. Hence the product of all the remain-

ing factors is fixed under $\sigma$. This proves that $\sigma(P(X_1, \ldots, X_n)) = -P(X_1, \ldots, X_n)$. Therefore,

$$\begin{aligned} e(P(X_1, \ldots, X_n)) &= (\sigma_1 \circ \cdots \circ \sigma_r)(P(X_1, \ldots, X_n)) \\ &= (\sigma_1 \circ \cdots \circ \sigma_{r-1})(P(X_{\sigma_r(1)}, \ldots, X_{\sigma_r(n)})) \\ &= -(\sigma_1 \circ \cdots \circ \sigma_{r-1})(P(X_1, \ldots, X_n)) \\ &= (-1)^r P(X_1, \ldots, X_n) \\ &= -P(X_1, \ldots, X_n), \text{ as } r \text{ is odd.} \end{aligned}$$

As $e(P(X_1, \ldots, X_n)) = P(X_1, \ldots, X_n)$, this gives a contradiction and hence completes the proof.

**Theorem 2.4.19** A permutation $\sigma \in S_n$ is a product of either even or odd number of transpositions, but never both.

*Proof.* Suppose $\sigma$ can be written as both. We write,

$$\sigma = \sigma_1 \circ \cdots \circ \sigma_m, \text{ where } m \text{ is even and } \sigma_i\text{'s are transpositions for all } 1 \leq i \leq m,$$

and

$$\sigma = \tilde{\sigma}_1 \circ \cdots \circ \tilde{\sigma}_r, \text{ where } r \text{ is odd and } \tilde{\sigma}_i\text{'s are transpositions for all } 1 \leq i \leq r.$$

So, $\sigma^{-1} = \tilde{\sigma}_r^{-1} \circ \cdots \tilde{\sigma}_1^{-1}$. Note that the inverse of a transposition is always a transposition, as $\alpha \circ \alpha = e$ for any transposition $\alpha$. Now we write,

$$e = \sigma \circ \sigma^{-1} = \sigma_1 \circ \cdots \sigma_m \circ \tilde{\sigma}_r^{-1} \circ \cdots \tilde{\sigma}_1^{-1}.$$

This shows that $e$ is a product of $r + m$, an odd many transpositions. A contradiction to the fact that $e$ can never be a product of odd many transpositions. Therefore any permutation $\sigma \in S_n$ can be product of either even or odd many transpositions, but never both.

**Definition 2.4.20** If a permutation in $S_n$ is a product of even number of transpositions, then it is called an **even** permutation, otherwise it is called an **odd** permutation.

Define $A_n$ to be the set of all even permutations of $S_n$. Clearly $A_n \neq \varnothing$ as $e \in A_n$.

**Exercise 2.4.21** Show that $A_n$ is a subgroup of $(S_n, \circ)$.

**Definition 2.4.22** The set $A_n$ of all even permutations of $S_n$ form a group under the composition of maps, which is called as the **alternating group** of $n$ symbols.

**Theorem 2.4.23** For $n \geq 2$, $|A_n| = n!/2$.

*Proof.* Let $\tau = (1\ 2)$. Clearly, if $\sigma \in A_n$, then $\sigma \circ \tau \in S_n \setminus A_n$. So define $\phi : A_n \to S_n \setminus A_n$ by $\phi(\sigma) := \sigma \circ \tau$ for all $\sigma \in A_n$. This map is injective as $\sigma_1 \circ \tau = \sigma_2 \circ \tau$ gives $\sigma_1 = \sigma_2$ due to cancellation law. Now $\phi$ is surjective, as for $\alpha \in S_n \setminus A_n$, we see that $\phi(\alpha \circ \tau) = \alpha \circ \tau \circ \tau = \alpha$. Hence $|A_n| = |S_n \setminus A_n|$. As $S_n$ is disjoint union of $A_n$ and $S_n \setminus A_n$, $|A_n| = \frac{|S_n|}{2} = \frac{n!}{2}$.

## 2.5 Cosets and Lagrange's theorem

Let $(G, *)$ be a group and $H$ be a subgroup of $(G, *)$. Let $a \in G$. Then the set

$$aH := \{a * h : h \in H\},$$

is called a **left coset of $H$ in $G$**. The set

$$Ha := \{h * a : h \in H\},$$

is called a **right coset of $H$ in $G$**.

**Example 2.5.1** Let us consider the Klein's 4 group i.e. the set $G = \{e, a, b, c\}$ with the operation $*$ as follows: $a * a = b * b = c * c = e, a * b = b * a = c, a * c = c * a = b, b * c = c * b = a$. Let us consider the subgroup $H := \{e, a\}$ of $(G, *)$. We look at all the left and right cosets of $H$ in $(G, *)$.
The below are the left cosets of $H$ in $(G, *)$:

$$\begin{aligned} eH &= \{e, a\} = H, \\ aH &= \{a, a * a\} = \{e, a\} = H, \\ bH &= \{b, ba\} = \{b, c\}, \\ cH &= \{c, ca\} = \{c, b\}. \end{aligned}$$

The below are the right cosets of $H$ in $(G, *)$:

$$\begin{aligned} He &= \{e, a\} = H, \\ Ha &= \{a, a * a\} = \{e, a\} = H, \end{aligned}$$

$$Hb = \{b, a * b\} = \{b, c\},$$
$$Hc = \{c, a * c\} = \{c, b\}.$$

We also note that here the number of distinct left and right cosets of $H$ in $(G, *)$ are same.

**Exercise 2.5.2** Consider the subgroup $H := \{e, (1\ 2)\}$ of the symmetric group with 3 symbols $S_3$. Find out all the distinct left and right cosets of $H$ in $S_3$.

**Exercise 2.5.3** Consider the subgroup $H := 3\mathbb{Z}$ of the group $(\mathbb{Z}, +)$. Find out all the distinct left and right cosets of $H$ in $(\mathbb{Z}, +)$.

**Remark 2.5.4** Note that a subgroup $H$ is both left as well as right coset of $H$ in a group $(G, *)$ as $eH = H = He$. So the collection of all left cosets of $H$ in $(G, *)$ and the collection of all right cosets of $H$ in $(G, *)$ are non-empty.

**Theorem 2.5.5** Let $(G, *)$ be a group and $H$ be a subgroup of $(G, *)$. Then the following statements hold:

1) $aH = H$ if and only if $a \in H$.

2) $Ha = H$ if and only if $a \in H$.

3) $aH = bH$ if and only if $a^{-1} * b \in H$.

4) $Ha = Hb$ if and only if $a * b^{-1} \in H$.

5) Either $aH = bH$ or $aH \cap bH = \varnothing$.

6) Either $Ha = Hb$ or $Ha \cap Hb = \varnothing$.

*Proof.* This can be proved as follows:

1) First let $aH = H$. So $a = a * e \in aH$, as $e \in H$. This implies that $a \in H$. Next let $a \in H$. So $aH \subseteq H$. Now for any $h \in H$, we can write $h = a * (a^{-1} * h)$. This proves that $h \in aH$ as $a^{-1} * h \in H$, and hence $H \subseteq aH$. Therefore finally we get $aH = H$.

2) First let $Ha = H$. So $a = e * a \in Ha$, as $e \in H$. This implies that $a \in H$. Next let $a \in H$. So $Ha \subseteq H$. Now for any $h \in H$, we can write $h = (h * a^{-1}) * a$. This proves that $h \in Ha$ as $h * a^{-1} \in H$, and hence $H \subseteq Ha$. Therefore finally we get $Ha = H$.

3) First let $aH = bH$. So $H = a^{-1}bH$, and now using 1), we get, $a^{-1}b \in H$. Next let $a^{-1}b \in H$. So $a^{-1}bH = H$. This implies, $aH = bH$.

4) First let $Ha = Hb$. So $H = H(a * b^{-1})$, and now using 2), we get, $a * b^{-1} \in H$. Next let $a * b^{-1} \in H$. So $H(a * b^{-1}) = H$. This implies, $Ha = Hb$.

5) Let $aH \cap bH \neq \varnothing$. We show that $aH = bH$. Let $x \in aH \cap bH$, so $x = a * h_1$ and also $x = b * h_2$ for some $h_1, h_2 \in H$. Therefore $a * h_1 = b * h_2$ and hence $a^{-1} * b = h_1 * h_2^{-1} \in H$. This proves that $aH = bH$.

6) Let $Ha \cap Hb \neq \varnothing$. We show that $Ha = Hb$. Let $x \in Ha \cap Hb$, so $x = h_1 * a$ and also $x = h_2 * b$ for some $h_1, h_2 \in H$. Therefore $h_1 * a = h_2 * b$ and hence $a * b^{-1} = h_1^{-1} * h_2 \in H$. This proves that $Ha = Hb$.

**Remark 2.5.6** Let $H$ be a subgroup of $(G, *)$. From the above theorem we can conclude that collections of all distinct left (resp. right) cosets of $H$ forms a partition of $G$.

**Exercise 2.5.7** Let $H$ be a subgroup of $(G, *)$. For any two elements $a, b \in G$, we define $aRb$ if and only if $a^{-1} * b \in H$. Show that $R$ is an equivalence relation on $G$ and the equivalence classes of $G$ with respect to $R$ is the collections of all distinct left cosets of $H$ in $G$

**Theorem 2.5.8** Let $H$ be a subgroup of $(G, *)$ and $a \in G$. Then

$$|aH| = |H| = |Ha|.$$

*Proof.* To show that $|aH| = |H|$, we need to show that there is a bijection from $H$ to $aH$. We define, $f : H \to aH$, by $f(h) = a * h$ for all $h \in H$. First we show that $f$ is injective. Let $f(h_1) = f(h_2)$ i.e. $a * h_1 = a * h_2$. This implies that $h_1 = a^{-1} * (a * h_1) = a^{-1} * (a * h_2) = h_2$. Now we show that $f$ is surjective. Let $b \in aH$. Then we consider $h \in H$ such that $b = a * h$. Clearly $f(h) = a * h$. This proves that $|aH| = |H|$.

To show $|H| = |Ha|$, define $g : H \to Ha$, by $g(h) = h * a$ for all $h \in H$ and then similarly we get $g$ is a bijection. Thus $|H| = |Ha|$.

Therefore $|aH| = |H| = |Ha|$.

**Theorem 2.5.9** Let $H$ be a subgroup of $(G, *)$ and $\mathcal{L}$ (resp. $\mathcal{R}$) be the set of all left (resp. right) cosets of the subgroup $H$ in $G$. Then $|\mathcal{L}| = |\mathcal{R}|$.

*Proof.* We need to show that there is a bijection from $\mathcal{L}$ to $\mathcal{R}$. We define, $\phi : \mathcal{L} \to \mathcal{R}$ as follows: for $aH \in \mathcal{L}$ set $\phi(aH) := Ha^{-1}$. First we need to show that $\phi$ is well-defined.

Let $aH = bH$ for some $a, b \in G$. So $a^{-1} * b \in H$. This implies that $H(a^{-1} * b) = H$ i.e. $Ha^{-1} = Hb^{-1}$ and hence $\phi$ is well-defined. Now we show that $\phi$ is injective. Let $\phi(aH) = \phi(bH)$. So $Ha^{-1} = Hb^{-1}$ and therefore $a^{-1} * b \in H$. This implies that $aH = bH$. Thus $\phi$ is injective. Now we show that $\phi$ is surjective. Let $Ha \in \mathcal{R}$. Consider $a^{-1}H \in \mathcal{L}$. Note that $\phi(a^{-1}H) = Ha$. Thus $\phi$ is surjective. This completes the proof.

**Definition 2.5.10** Let $H$ be a subgroup of a group $(G, *)$. Then the **index of $H$ in $G$** is defined as the number of distinct left or right cosets of $H$ in $(G, *)$. We denote the index of $H$ in $G$ by $[G : H]$.

**Theorem 2.5.11 (Lagrange's theorem)** Let $H$ be a subgroup of a finite group $(G, *)$. Then $|H|$ divides $|G|$. In fact,

$$[G : H] = \frac{|G|}{|H|}.$$

*Proof.* Since $(G, *)$ is a finite group we have $[G : H]$ is finite. Let $x_1H, \ldots, x_rH$ be all the distinct left cosets of $H$ in $(G, *)$. Now as they form a partition of $(G, *)$, we have

$$G = \bigcup_{1 \leq i \leq r} x_iH.$$

Therefore,

$$|G| = \sum_{1 \leq i \leq r} |x_iH|.$$

Also we know that $|x_iH| = |H|$ for all $1 \leq i \leq r$. This implies that

$$|G| = r|H|.$$

This proves that $|H|$ divides $|G|$. Also note that

$$[G : H] = r = \frac{|G|}{|H|}.$$

### 2.5.1 An application of Lagrange's theorem

**Theorem 2.5.12** Let $H, K$ be two finite subgroups of a group $(G, *)$. Then

$$|HK| = \frac{|H||K|}{|H \bigcap K|}.$$

*Proof.* Write $A = H \bigcap K$. Note that $A$ is a subgroup of $(H, *)$. We have, $H$ is a finite group. Let $x_1A, \ldots, x_rA$ be all the distinct left cosets of $A$ in $H$. Therefore, $|H| = r|A|$. Now $AK \subseteq K$ and $K \subseteq AK$ as $e \in A$. So $K = AK$. Then,

$$HK = \left( \bigcup_{1 \leq i \leq r} x_iA \right) K = \bigcup_{1 \leq i \leq r} x_iK.$$

We show that $x_iK \bigcap x_jK = \varnothing$ for all $i \neq j$. If $x_iK \bigcap x_jK \neq \varnothing$, then $x_iK = x_jK$ i.e. $x_i^{-1} * x_j \in K$. Also $x_i^{-1} * x_j \in H$ as $x_i, x_j \in H$. This implies that $x_i^{-1} * x_j \in A$ and hence $x_jA = x_jA$, a contradiction. Therefore,

$$|HK| = \sum_{1 \leq i \leq r} |x_iK| = r|K|.$$

So finally we have

$$|HK| = \frac{|H||K|}{|A|} = \frac{|H||K|}{|H \bigcap K|}.$$

**Remark 2.5.13** From the above theorem we get that for any finite subgroups $H, K$ of $(G, *)$, $|HK| = |KH|$.

**Question 2.5.14** Does the converse of Lagrange's theorem hold i.e. if $(G, *)$ is a finite group and if $m \mid |G|$, then does there exists a subgroup of $(G, *)$ of order $m$?

We know that if $(G, *)$ is cyclic then for every divisor $m$ of $|G|$, there exists an unique subgroup of $(G, *)$ of order $m$. But it is not true for all finite groups. For example consider the alternating group of 4 symbols $A_4$. We have $|S_4| = 24$. So $|A_4| = 12$. Consider all the 3-cycles of $S_4$:

$$(1\ 2\ 3), (1\ 3\ 2), (1\ 2\ 4), (1\ 4\ 2), (2\ 3\ 4), (2\ 4\ 3), (1\ 4\ 3), (1\ 3\ 4).$$

They are all even, so all the 3-cycles belong to $A_4$. We see that the number of 3-cycles in $S_4$ is 8. Now $6 \mid |A_4|$. We show that $A_4$ has no subgroup of order 6. If there is a subgroup $H$ of $A_4$ of order 6, then there exists a 3-cycle $\sigma = (a\ b\ c) \in S_4$ such that $\sigma \notin H$. We know that $o(\sigma) = 3$, so $K := \{e, \sigma, \sigma^2\}$ is a subgroup of $A_4$. Note that $\sigma^2 = \sigma^{-1} \notin H$. Thus we get,

$$|HK| = \frac{|H||K|}{|H \bigcap K|} = 18.$$

But $HK \subseteq A_4$ and $|A_4| = 12$. This leads us to a contradiction. So there is no subgroup of $A_4$ of order 6.

## 2.6 Normal subgroups

**Definition 2.6.1** Let $H$ be a subgroup of a group $(G, *)$. We say that $H$ is a **normal subgroup** of $G$ if $gH = Hg$ for all $g \in G$ i.e. if every left coset is a right coset and vice-versa.

**Remarks 2.6.2** Note that

1) The trivial subgroups i.e. $\{e\}$ and $G$ are always normal in $(G, *)$.

2) In general, $gH = Hg$ for some $g \in G$ does not imply that $gh = hg$ for all $h \in H$. Let $G = S_3$ and $H = \{e, (1\ 2\ 3), (1\ 3\ 2)\}$. We can check that for all $\sigma \in S_3$, $\sigma H = H\sigma$. But
$$(2\ 3)(1\ 2\ 3) \neq (1\ 2\ 3)(2\ 3).$$

**Theorem 2.6.3** Let $H$ be a subgroup of a group $(G, *)$. Then the following are equivalent:

1) $H$ is a normal subgroup of $(G, *)$.

2) $gHg^{-1} \subseteq H$ for all $g \in G$.

3) $gHg^{-1} = H$ for all $g \in G$.

*Proof.* First we assume that $H$ is a normal subgroup of $(G, *)$. Since $H$ is a normal subgroup of $(G, *)$, by definition we have $gH = Hg$ for all $g \in G$. Therefore $gHg^{-1} \subseteq H$ for all $g \in G$.

Next we assume that $gHg^{-1} \subseteq H$ for all $g \in G$. We show that $H \subseteq gHg^{-1}$ for all $g \in G$. Let $h \in H$. Now $h = g * (g^{-1} * h * g) * g^{-1} \in gHg^{-1}$ as $g^{-1} * h * g \in H$. This shows that $H \subseteq gHg^{-1}$ for all $g \in G$ and hence $gHg^{-1} = H$ for all $g \in G$.

Now assume that $gHg^{-1} = H$ for all $g \in G$. Then $gH = Hg$ for all. $g \in G$. So $H$ is a normal subgroup of $(G, *)$.

**Theorem 2.6.4** The center $Z(G)$ of a group $(G, *)$ is a normal subgroup.

*Proof.* We have seen that $Z(G)$ is a subgroup of $(G, *)$. Now we show that $gZ(G)g^{-1} \subseteq Z(G)$ for all $g \in G$. Let $g \in G$ and $a \in Z(G)$. Now $a \in \mathbb{Z}(G)$ implies that $a * g = g * a$,

so $g*a*g^{-1} = a*g*g^{-1} = a \in Z(G)$. This is true for all $a \in Z(G)$, hence $gZ(G)g^{-1} \subseteq Z(G)$.

**Theorem 2.6.5** Let $H$ be a subgroup of $(G,*)$ such that $[G : H] = 2$. Then $H$ is normal in $(G,*)$.

*Proof.* We are given that $H$ has two distinct left cosets and two distinct right cosets in $(G,*)$. We want to show that $gH = Hg$ for all $g \in G$. Let $g \in H$, then we have $gH = H = Hg$. So let $g \notin H$. Then $gH \cap H = \varnothing$ and $Hg \cap H = \varnothing$. We can then write for any $g \notin H$,

$$G = H \bigcup gH \quad \text{and} \quad G = H \bigcup Hg.$$

Therefore,

$$gH = G \setminus H = Hg \quad \text{for all } g \notin G.$$

This prove that $H$ is a normal subgroup of $(G,*)$.

**Example 2.6.6** The alternating group with $n$ symbols $A_n$ is a normal subgroup of the symmetric group with $n$ symbols $S_n$. We know $|A_n| = \frac{n!}{2} = \frac{|S_n|}{2}$. So $[S_n : A_n] = 2$ and hence $A_n$ is a normal subgroup of $S_n$.

**Theorem 2.6.7** Let $H, K$ be two subgroups of a group $(G,*)$. Then,

1) If $H$ is normal then $HK = KH$, and hence $HK$ is a subgroup of $(G,*)$.
2) If $H, K$ both are normal then $HK(= KH)$ is a normal subgroup of $(G,*)$.
3) If $H, K$ both are normal then $H \bigcap K$ is a normal subgroup of $(G,*)$.

*Proof.* This can be proved as follows:

1) Since $H$ is normal for any $k \in K$, we have $kH = Hk$, so

$$HK = \bigcup_{k \in K} Hk = \bigcup_{k \in K} kH = KH.$$

Therefore $HK$ is a subgroup of $(G,*)$.

2) As $H, K$ both are normal we have, $gHg^{-1} = H$ and $gKg^{-1} = K$ for all $g \in G$. For any $g \in G$,

$$gHKg^{-1} = gH(g^{-1} * g)Kg^{-1} = (gHg^{-1})(gKg^{-1}) = HK.$$

This proves that $HK(= KH)$ is a normal subgroup of $(G,*)$.

3) Let $a \in H \cap K$ and $g \in G$. Now $g * a * g^{-1} \in H$ and also $g * a * g^{-1} \in K$ as $H, K$ both are normal subgroups of $(G, *)$. Therefore $g * a * g^{-1} \in H \cap K$, and hence $g(H \cap K)g^{-1} \subseteq HK$. This is true for all $g \in G$, therefore the proof.

**Remarks 2.6.8** Note that

1) If $HK = KH$ is a subgroup of $(G, *)$, that does not imply that at least one of $H, K$ is a normal subgroup. Consider the following:

$$G := S_4, \ H := \{e, (1\ 2)\}, \text{ and } K = \{e, (3\ 4)\}.$$

Now we note that $H, K$ are not normal in $S_4$, as

$$(1\ 3)(1\ 2)(1\ 3) = (2\ 3) \notin H, \text{ and } (1\ 3)(3\ 4)(1\ 3) = (1\ 4) \notin K.$$

But $HK = \{e, (1\ 2), (3\ 4), (1\ 2)(3\ 4)\} = KH$ is a subgroup of $S_4$.

2) If $HK = KH$ is a normal subgroup of $(G, *)$, that does not imply that both of $H, K$ are normal subgroups. Consider the following:

$$G := S_5, \ H := \langle (1\ 2\ 3\ 4\ 5) \rangle, \text{ and } K = A_4.$$

So we have two subgroups $H, K$ of $S_5$. We show that $A_5 = \langle (1\ 2\ 3\ 4\ 5) \rangle A_4$. Note that $H \cap K = \{e\}$. Therefore,

$$|HK| = \frac{|H||K|}{|H \cap K|} = \frac{5 \cdot 12}{1} = 60 = |A_5|.$$

As $HK \subseteq A_5$ and $|HK| = |A_5| < \infty$, we conclude that $A_5 = \langle (1\ 2\ 3\ 4\ 5) \rangle A_4$. Now $A_5$ is normal in $S_5$, so $\langle (1\ 2\ 3\ 4\ 5) \rangle A_4$ is normal in $S_5$. We show that $H$ is not normal in $S_5$. We have,

$$H = \{e, (1\ 2\ 3\ 4\ 5), (1\ 3\ 5\ 2\ 4), (1\ 4\ 2\ 5\ 3), (1\ 5\ 4\ 3\ 2)\},$$

and

$$(1\ 2)(1\ 2\ 3\ 4\ 5)(1\ 2) = (1\ 3\ 4\ 5\ 2) \notin H.$$

This shows that $H$ is not normal in $S_5$.

3) If $H \cap K$ is a normal subgroup of $(G, *)$, that does not imply that both of $H, K$

are normal subgroups. Consider the following:

$$G := S_3,\ H := \{e, (1\ 2)\},\ \text{and}\ K = \{e, (1\ 3)\}.$$

We have already seen that $HK \neq KH$, and hence $H, K$ both are not normal subgroups of $S_3$. But $H \cap K = \{e\}$ is a normal subgroup of $S_3$.

## 2.7 Quotient groups

Let $H$ be a normal subgroup of a group $(G, *)$. So each left coset of $H$ is also a right coset of $H$ in $(G, *)$ and vice-versa. Therefore in this case we refer to them as just cosets. Let $G/H$ denote the set of all distinct cosets of $H$ in $(G, *)$. We can induce an operation on $G/H$ from the operation on $G$. Define $*$ on $G/H$ as follows: for any $aH, bH \in G/H$,

$$aH * bH = (a * b)H.$$

First we show that this operation is well-defined. Let $a_1H = a_2H$ and $b_1H = b_2H$ i.e. $a_1^{-1} * a_2 \in H$ and $b_1^{-1} * b_2 \in H$. We shall show that $(a_1 * b_1)H = (a_2 * b_2)H$. Now,

$$\begin{aligned}(a_1 * b_1)^{-1} * (a_2 * b_2) &= (b_1^{-1} * a_1^{-1}) * (a_2 * b_2)\\ &= b_1^{-1} * (a_1^{-1} * a_2) * b_2\\ &= (b_1^{-1} * (a_1^{-1} * a_2) * b_1) * (b_1^{-1} * b_2).\end{aligned}$$

This shows that $(a_1 * b_1)^{-1} * (a_2 * b_2) \in H$ as $H$ is normal in $(G, *)$. Hence,

$$(a_1 * b_1)H = (a_2 * b_2)H.$$

We note that the operation $*$ is associative in $G/H$. Let $aH, bH, cH \in G/H$. We have,

$$(aH * bH) * cH = (a * b)H * cH = (a * b * c)H = aH * (b * c)H = aH * (bH * cH),$$

as $*$ is associative in $G$. Also note that $H$ is the identity of $G/H$ as $aH * H = aH = H * Ha$ for any $a \in G$. We now show that for $aH \in G/H$, we have an inverse element in $G/H$. As $a \in G$, we have $a^{-1} \in G$. Consider $a^{-1}H \in G/H$. Now,

$$aH * a^{-1}H = (a * a^{-1})H = H = (a^{-1} * a)H = a^{-1}H * aH.$$

So we see that $(G/H, *)$ forms a group. This group $(G/H, *)$ is called the **quotient group** of $G$ by $H$.

**Remarks 2.7.1** Note that

1) If $(G, *)$ is abelian, then $(G/H, *)$ is abelian as for any $aH, bH \in G/H$, $aH * bH = (a * b)H = (b * a)H = bH * aH$.

2) If $(G, *)$ is cyclic, then $(G/H, *)$ is cyclic as for any $aH \in G/H$,

$$aH = (g^m)H = (gH)^m, \text{ where } G := \langle g \rangle \text{ and } a = g^m \text{ for some } m \in \mathbb{Z}.$$

So $G/H = \langle gH \rangle$.

3) If $(G/H, *)$ is abelian, then $(G, *)$ need not be abelian. Note that $S_n/A_n$ is a group of order 2, and hence abelian but $S_n$ is non-abelian.

4) If $(G/H, *)$ is cyclic, then $(G, *)$ need not be cyclic. Note that $S_n/A_n$ is a group of order 2, and hence cyclic but $S_n$ is not cyclic.

**Example 2.7.2** Let $(G, *) = (\mathbb{Z}, +)$. Consider the subgroup $3\mathbb{Z}$ in $(\mathbb{Z}, +)$. As $(\mathbb{Z}, +)$ is abelian, $3\mathbb{Z}$ is a normal subgroup. Note that $\mathbb{Z}/3\mathbb{Z} = \{3\mathbb{Z}, 1 + 3\mathbb{Z}, 2 + 3\mathbb{Z}\}$. It is interesting to note that this is the group of distinct congruence classes modulo 3.

## 2.8 Homomorphism of groups

Let $(G_1, *_1)$ and $(G_2, *_2)$ be two groups. We would like to see the interplay between them. A map $f : G_1 \to G_2$ is called a **homomorphism** if $f(a *_1 b) = f(a) *_2 f(b)$ for all $a, b \in G_1$.

**Remark 2.8.1** Between $(G_1, *_1)$ and $(G_2, *_2)$, there is always at least one homomorphism as the map $f : G_1 \to G_2$, defined by $f(a) = e_2$ for all $a \in G_1$, is a homomorphism $(G_1, *_1)$ into $(G_2, *_2)$. Here $e_2$ denotes the identity of $(G_2, *_2)$. This homomorphism is called the **trivial homomorphism** of $(G_1, *_1)$ into $(G_2, *_2)$.

**Examples 2.8.2** The following are examples of group homomorphisms.

1) Consider $(G_1, *_1)$ as the group of real numbers $\mathbb{R}$ with usual addition and $(G_2, *_2)$ as the group of non-zero real numbers $\mathbb{R} \setminus \{0\}$ with usual multiplication. Define $f : \mathbb{R} \to \mathbb{R} \setminus \{0\}$ by $f(a) := \exp(a)$. The notation exp denotes the exponential map.

2) Consider $(G_1, *_1)$ as the group of invertible $2 \times 2$ matrices with real entries, with the operation usual matrix multiplication. This group is denoted by $GL_2(\mathbb{R})$, and is called the **general linear group of order 2 with real entries**. The name linear comes from the fact that $GL_2(\mathbb{R})$ is essentially the set of all bijective linear transformations from the vector space $\mathbb{R}^2$ (over $\mathbb{R}$) to itself. Consider $(G_2, *_2)$ as the group of non-zero real numbers $\mathbb{R} \setminus \{0\}$ with usual multiplication. Define $f : GL_2(\mathbb{R}) \to \mathbb{R} \setminus \{0\}$ by $f(A) := \det(A)$. The notation $\det(A)$ denotes the determinant of the matrix $A$.

3) Consider $(G_1, *_1)$ as $S_n$, the symmetric group with $n$ symbols. Consider $(G_2, *_2)$ as the group $\{1, -1\}$ with usual multiplication. Define $f : S_n \to \{1, -1\}$ by

$$f(\sigma) := \begin{cases} 1 & \text{if } \sigma \text{ is even,} \\ -1 & \text{otherwise.} \end{cases}$$

**Theorem 2.8.3** Let $(G_1, *_1)$ and $(G_2, *_2)$ be two groups and $f : G_1 \to G_2$ be a group homomorphism. Then,

1) $f(e_1) = e_2$, where $e_1, e_2$ denote the identities of $(G_1, *_1)$ and $(G_2, *_2)$ respectively.

2) $f(a^{-1}) = f(a)^{-1}$ for all $a \in G_1$.

3) $f(a^n) = f(a)^n$ for all $n \in \mathbb{Z}$.

*Proof.* This can be proved as follows:

1) Note that $f(e_1) = f(e_1 *_1 e_1) = f(e_1) *_2 f(e_1)$. Therefore by cancellation we obtain, $f(e_1) = e_2$.

2) Note that $e_2 = f(e_1) = f(a *_1 a^{-1}) = f(a) *_2 f(a^{-1})$ for all $a \in G_1$. Therefore, $f(a^{-1}) = f(a)^{-1}$ for all $a \in G_1$.

3) Let $a \in G_1$. For $n = 0$, we have already seen that $f(e_1) = e_2$. Let $n \in \mathbb{N}$. Assume the induction hypothesis for $n = k$, i.e. $f(a^k) = f(a)^k$. Now,

$$f(a^{k+1}) = f(a^k *_1 a) = f(a^k) *_2 f(a) = f(a)^{k+1}.$$

Now let $n < 0$. Let $n = -m$ where $m > 0$. Therefore,

$$f(a^n) = f(a^{-m}) = f(a^{-1})^m = f(a)^{-m} = f(a)^n.$$

This completes the proof.

As an application we see that $f : \mathbb{Z} \to \mathbb{Z}$ defined by $f(n) = n + 1$ is not a group homomorphism as $f(0) \neq 0$.

**Exercise 2.8.4** Show that all the group homomorphisms from $(\mathbb{Z}, +)$ to itself are of the form $f_m(n) := mn$ for all $n \in \mathbb{Z}$, for some $m \in \mathbb{Z}$.

**Remarks 2.8.5** Let $(G_1, *_1)$ and $(G_2, *_2)$ be two groups and $f : G_1 \to G_2$ be a group homomorphism. Then,

1) For a subgroup $H$ of $(G_1, *_1)$, $f(H)$ is a subgroup of $(G_2, *_2)$. Note that $f(H) \neq \varnothing$ as $e_2 = f(e_1) \in f(H)$. Let $a, b \in f(H)$. So $a = f(h_1)$ and $b = f(h_2)$ for some $h_1, h_2 \in H$. Therefore,

$$a *_2 b^{-1} = f(h_1) *_2 f(h_2)^{-1} = f(h_1 *_1 h_2^{-1}) \in f(H).$$

This proves that $f(H)$ is a subgroup of $(G_2, *_2)$.

2) For a subgroup $H$ of $(G_2, *_2)$, the inverse image of $H$ under $f$ i.e. the set $f^{-1}(H) := \{g \in G_1 : f(g) \in H\}$ is a subgroup of $(G_1, *_1)$. Note that $f^{-1}(H) \neq \varnothing$ as $e_1 \in f^{-1}(H)$. Let $a, b \in f^{-1}(H)$. So $f(a), f(b) \in H$. As $H$ is a subgroup of $(G_2, *_2)$, we have $f(a) *_2 f(b)^{-1} \in H$. So $f(a *_1 b^{-1}) \in H$. Therefore $a *_1 b \in f^{-1}(H)$. This proves that $f^{-1}(H)$ is a subgroup of $(G_1, *_1)$.

3) For a normal subgroup $H$ of $(G_1, *_1)$, $f(H)$ need not be normal in $(G_2, *_2)$. Consider the homomorphism $f : S_3 \to S_3$ defined by

$$f(\sigma) := \begin{cases} (1\ 2) & \text{if } \sigma \text{ is a transposition,} \\ e & \text{otherwise.} \end{cases}$$

We know $S_3$ is normal subgroup of $S_3$, but $\{e, (1\ 2)\}$ is not normal in $S_3$.

4) If $f$ is surjective, for a normal subgroup $H$ of $(G_1, *_1)$, $f(H)$ is normal in $(G_2, *_2)$. Let $a \in f(H)$ and $g_2 \in G_2$. We have $a = f(h)$ for some $h \in H$. Since $f$ is surjective there exists $g_1 \in G_1$ such that $f(g_1) = g_2$. Therefore,

$$g_2 *_2 a *_2 g_2^{-1} = f(g_1 *_1 h *_1 g_1^{-1}) \in f(H) \quad \text{as} \quad g_1 *_1 h *_1 g_1^{-1} \in H.$$

This proves that $f(H)$ is normal in $(G_2, *_2)$.

5) For a normal subgroup $H$ of $(G_2, *_2)$, $f^{-1}(H) := \{g \in G_1 : f(g) \in H\}$ is a normal subgroup of $(G_1, *_1)$. We have already seen that $f^{-1}(H)$ is a subgroup of $(G_1, *_1)$. Now let $g \in G_1$ and $a \in f^{-1}(H)$. To show $g *_1 a *_1 g^{-1} \in f^{-1}(H)$, we need to show that $f(g *_1 a *_1 g^{-1}) \in H$. Note that

$$f(g *_1 a *_1 g^{-1}) = f(g) *_2 f(a) *_2 f(g)^{-1} \in H \quad \text{as} \quad f(g) \in G_2 \quad \text{and} \quad f(a) \in H.$$

This proves that $f^{-1}(H)$ is a normal subgroup of $(G_1, *_1)$.

6) Let $a \in G_1$ be such that $o(a) = n$. Then $o(f(a)) \mid n$. Note that $f(a)^n = f(a^n) = f(e_1) = e_2$. So $o(f(a)) \mid n$.

In general $o(f(a))$ need not be equal to $n$. For example, consider the group homomorphism $f : S_n \to \{1, -1\}$ by

$$f(\sigma) := \begin{cases} 1 & \text{if } \sigma \text{ is even,} \\ -1 & \text{otherwise.} \end{cases}$$

If $\sigma$ is a 3-cycle then $o(\sigma) = 3$ and $o(f(\sigma)) = 1$.

7) Let $a \in G_1$ be such that $o(a) = n$ and $f$ be injective. Then $o(f(a)) = n$. Note that if $o(f(a)) = m$ then $f(a^m) = f(a)^m = e_2 = f(e_1)$. So $a^m = e_1$, so $n \mid m$ and hence $m = n$.

**Definition 2.8.6** A group $(G_2, *_2)$ is called a **homomorphic image** of $(G_1, *_1)$ if there exists a surjective homomorphism $f : G_1 \to G_2$.

**Theorem 2.8.7** Let $f$ be a surjective homomorphism from $(G_1, *_1)$ onto $(G_2, *_2)$. Then

1) If $(G_1, *_1)$ is abelian then so is $(G_2, *_2)$.

2) If $(G_1, *_1)$ is cyclic then so is $(G_2, *_2)$.

*Proof.* This can be proved as follows:

1) Let $a, b \in G_2$. So there exist $a', b' \in G_1$ such that $f(a') = a$ and $f(b') = b$. Now,

$$a *_2 b = f(a') *_2 f(b') = f(a' *_1 b') = f(b' *_1 a') = f(b') *_2 f(a') = b *_2 a.$$

This proves that $(G_2, *_2)$ is abelian.

2) Let $G_1 := \langle g \rangle$ and $b \in G_2$. So there exists $a \in G_1$ be such that $f(a) = b$. Now $a = g^m$ for some $m \in \mathbb{Z}$. Therefore,

$$b = f(g^m) = f(g)^m \in \langle f(g) \rangle.$$

Thus $G_2 \subseteq \langle f(g) \rangle$. Clearly $\langle f(g) \rangle \subseteq G_2$ as $f(g) \in G_2$. This proves that $G_2 = \langle f(g) \rangle$, and hence cyclic.

**Remark 2.8.8** If $(G_2, *_2)$ is abelian, $(G_1, *_1)$ need not be abelian. Also if $(G_2, *_2)$ is cyclic, $(G_1, *_1)$ need not be cyclic. For example, consider the homomorphism $f : S_n \to \{1, -1\}$ defined as

$$f(\sigma) := \begin{cases} 1 & \text{if } \sigma \text{ is even,} \\ -1 & \text{otherwise.} \end{cases}$$

Here $\{1, -1\}$ is abelian and cyclic with usual multiplication, but $S_n$ is neither abelian nor cyclic.

**Definition 2.8.9** Let $(G_1, *_1)$ and $(G_2, *_2)$ be two groups and $f : G_1 \to G_2$ be a group homomorphism. Then the set

$$\{g \in G_1 : f(g) = e_2\},$$

where $e_2$ is the identity of $G_2$, is called the **kernel** of $f$ and is denoted by $\mathrm{Ker} f$.

**Definition 2.8.10** Let $(G_1, *_1)$ and $(G_2, *_2)$ be two groups and $f : G_1 \to G_2$ be a group homomorphism. Then the set

$$\{f(g) : g \in G_1\}$$

is called the **image** of $f$ and is denoted by $\mathrm{Im} f$.

**Remark 2.8.11** Both the sets $\mathrm{Ker} f$ and $\mathrm{Im} f$ are non-empty as $e_1 \in \mathrm{Ker} f$ and $e_2 \in \mathrm{Im} f$.

**Exercise 2.8.12** Show that $\mathrm{Ker} f$ is a normal subgroup of $G_1$.

**Exercise 2.8.13** Show that $\mathrm{Im} f$ is a subgroup of $G_2$.

**Remark 2.8.14** The subgroup $\mathrm{Im} f$ of $(G_2, *_2)$ need not be normal. Consider the homomorphism $f : S_3 \to S_3$ defined by

$$f(\sigma) := \begin{cases} (1\ 2) & \text{if } \sigma \text{ is a transposition,} \\ e & \text{otherwise.} \end{cases}$$

Here $\text{Im}f = \{e, (1\ 2)\}$ is not normal in $S_3$.

**Theorem 2.8.15** Let $(G_1, *_1)$ and $(G_2, *_2)$ be two groups and $f : G_1 \to G_2$ be a group homomorphism. Then $f$ is injective if and only if $\text{Ker}f = \{e_1\}$.

*Proof.* First assume that $\text{Ker}f = \{e_1\}$. We show that $f$ is injective. Let $f(a) = f(b)$ for some $a, b \in G_1$. Then we have $f(a) *_2 f(b)^{-1} = e_2$ i.e. $f(a *_1 b^{-1}) = e_2$ as $f$ is a group homomorphism. This shows that $a *_1 b^{-1} \in \text{Ker}f$, and hence $a *_1 b^{-1} = e_1$ i.e. $a = b$.

Next we assume that $f$ is injective. We show that $\text{Ker}f = \{e_1\}$. Let $a \in \text{Ker}f$. So $f(a) = e_2 = f(e_1)$. As $f$ is injective we get $a = e_1$. This proves that $\text{Ker}f = \{e_1\}$.

**Definition 2.8.16** Let $(G_1, *_1)$ and $(G_2, *_2)$ be two groups and $f : G_1 \to G_2$ be a group homomorphism. Then $f$ is called an **isomorphism** of $(G_1, *_1)$ and $(G_2, *_2)$ if the map $f$ is bijective.

**Theorem 2.8.17** Let $f : G_1 \to G_2$ be an isomorphism of $(G_1, *_1)$ and $(G_2, *_2)$. Then we have, $f^{-1} : G_2 \to G_1$ is an isomorphism of $(G_2, *_2)$ and $(G_1, *_1)$.

*Proof.* Since $f$ is bijective, the inverse map of $f$, denoted by $f^{-1}$ exists. We write the map $f^{-1} : G_2 \to G_1$ below:

$$f^{-1}(a) := a' \text{ for all } a \in G_2, \text{ where } f(a') = a.$$

Let $a, b \in G_2$. So there exist unique $a', b' \in G_1$ such that $f(a') = a$ and $f(b') = b$. Now $a *_2 b = f(a') *_2 f(b') = f(a' *_1 b')$ as $f$ is a homomorphism. So we have,

$$f^{-1}(a *_2 b) = a' *_1 b' = f^{-1}(a) *_1 f^{-1}(b).$$

This shows that $f^{-1}$ is a homomorphism from $(G_2, *_2)$ to $(G_1, *_1)$. Now $f^{-1}$ is bijective as $f$ is bijective. This completes the proof.

**Remarks 2.8.18** If there is an isomorphism between $(G_1, *_1)$ and $(G_2, *_2)$, then we say that the groups $(G_1, *_1)$ and $(G_2, *_2)$ are isomorphic. In this case we denote $G_1 \simeq G_2$.

1) If the groups $(G_1, *_1)$ and $(G_2, *_2)$ are isomorphic, then $(G_1, *_1)$ is abelian if and only if $(G_2, *_2)$ is abelian. We have already seen that if the group $(G_2, *_2)$ is a homomorphic image of the group $(G_1, *_1)$ with $(G_1, *_1)$ abelian, then $(G_2, *_2)$ is abelian. In this case the groups $(G_1, *_1)$ and $(G_2, *_2)$ are homomorphic images of each other, and hence the remark follows.

2) If the groups $(G_1, *_1)$ and $(G_2, *_2)$ are isomorphic, then $(G_1, *_1)$ is cyclic if and only if $(G_2, *_2)$ is cyclic. We have already seen that if the group $(G_2, *_2)$ is a homomorphic image of the group $(G_1, *_1)$ with $(G_1, *_1)$ cyclic, then $(G_2, *_2)$ is cyclic. In this case the groups $(G_1, *_1)$ and $(G_2, *_2)$ are homomorphic images of each other, and hence the remark follows.

3) If $f$ is an isomorphism of $(G_1, *_1)$ and $(G_2, *_2)$, then we have $o(a) = o(f(a))$ for all $a \in G_1$. We have already seen that $o(a) = o(f(a))$ for all $a \in G_1$ if $f$ is injective. Hence it is a particular case.

**Exercise 2.8.19** Let $(G_1, *_1)$, $(G_2, *_2)$ and $(G_3, *_3)$ be three groups. Then show that,

1) $G_i \simeq G_i$ for all $1 \leq i \leq 3$.

2) $G_1 \simeq G_2$ implies that $G_2 \simeq G_1$.

3) If $G_1 \simeq G_2$ and $G_2 \simeq G_3$ then $G_1 \simeq G_3$.

**Remark 2.8.20** The above exercise shows us that $\simeq$ is an equivalence relation on the collection of all groups. For a group $(G, *)$, by the **isomorphic class** of $(G, *)$ we mean the collection of all groups which are isomorphic to $(G, *)$.

**Example 2.8.21** The groups $(\mathbb{R}, +)$ and $(\mathbb{R} \setminus \{0\}, \cdot)$ are not in the same isomorphic class. There is no element in $(\mathbb{R}, +)$ of finite order but $-1$ in $(\mathbb{R} \setminus \{0\}, \cdot)$ is of order 2.

**Remarks 2.8.22** We observe the following:

- A finite group can never be isomorphic to an infinite group.
- A cyclic group can never be isomorphic to a non-cyclic group.
- An abelian group can never be isomorphic to a non-abelian group.
- Two groups can be of same order, but they may not be isomorphic. For example, we know $\mathbb{Z}/6\mathbb{Z}$ is abelian group of order 6 and $S_3$ is non-abelian group of order 6.

**Theorem 2.8.23** Any infinite cyclic group is isomorphic to the group $(\mathbb{Z}, +)$.

*Proof.* Let $(G, *)$ be an infinite cyclic group. Let $a$ be a generator of $(G, *)$. Clearly $a \neq e$. Let us define, $f : \mathbb{Z} \to G$ by $f(n) := a^n$. Now $f(n+m) = a^{n+m} = a^n * a^m = f(n) * f(m)$. So $f$ is a homomorphism from $(\mathbb{Z}, +)$ to $(G, *)$. Now as $G$ is infinite, $a$ can not be of finite order. So by convention, $a^n = e$ if and only if $n = 0$. Therefore $\mathrm{Ker} f = \{0\}$

and hence $f$ is injective. Now let $b \in G$. So $b = a^r$ for some $r \in \mathbb{Z}$. We can see that $f(r) = a^r = b$. Thus $f$ is surjective. This proves that $G \simeq \mathbb{Z}$.

**Theorem 2.8.24** Any finite cyclic group of order $n \in \mathbb{N}$ is isomorphic to the group $(\mathbb{Z}/n\mathbb{Z}, +)$.

*Proof.* Let $(G, *)$ be a finite cyclic group of order $n \in \mathbb{N}$. Let $a$ be a generator of $(G, *)$. So $a^n = e$. Let us define, $f : G \to \mathbb{Z}/n\mathbb{Z}$ by $f(a^m) := [m]$. First we show that $f$ is well-defined. Let $a^s = a^t$. So $a^{s-t} = e$. So $n \mid (s-t)$. Therefore $[s] = [t]$ i.e. $f(a^s) = f(a^t)$. Now $f(a^m * a^r) = f(a^{m+r}) = [m+r] = [m] + [r] = f(a^m) + f(a^r)$. So $f$ is a homomorphism. Note that $\mathrm{Ker}f = \{g \in G : f(g) = [0]\} = \{a^m \in G : [m] = 0\} = \{a^m \in G : n \mid m\} = \{e\}$, as $a^m = e$ for all $m$ such that $n \mid m$. So $f$ is injective. For $[r] \in \mathbb{Z}/n\mathbb{Z}$, we can consider $a^r \in G$. We see that $f(a^r) = [r]$ and hence $f$ is surjective. This proves that $G \simeq \mathbb{Z}/n\mathbb{Z}$.

**Remark 2.8.25** We conclude that up to isomorphism there is only one cyclic group of order $n \in \mathbb{N}$.

**Theorem 2.8.26 (First isomorphism theorem)** Let $f : G_1 \to G_2$ be a homomorphism of the groups $(G_1, *_1)$ and $(G_2, *_2)$. Then,

$$G_1/\mathrm{Ker}f \simeq \mathrm{Im}f.$$

*Proof.* We know that $\mathrm{Ker}f$ is a normal subgroup of the group $(G_1, *_1)$. So we can consider the quotient group $(G_1/\mathrm{Ker}f, *_1)$ where the operation $*_1$ on $(G_1/\mathrm{Ker}f, *_1)$ is induced from the operation $*_1$ of $(G_1, *_1)$.

We define $\tilde{f} : G_1/\mathrm{Ker}f \to \mathrm{Im}f$ by $\tilde{f}(a\mathrm{Ker}f) := f(a)$ for all $a\mathrm{Ker}f \in G_1/\mathrm{Ker}f$. We need to show that the map $\tilde{f}$ is well-defined. Note that,

$$\begin{aligned} & a\mathrm{Ker}f = b\mathrm{Ker}f \\ \iff & a^{-1}b \in \mathrm{Ker}f \\ \iff & f(a^{-1}b) = e_2 \\ \iff & f(a) = f(b) \\ \iff & \tilde{f}(a\mathrm{Ker}f) = \tilde{f}(b\mathrm{Ker}f). \end{aligned}$$

This proves that the map $\tilde{f}$ is well-defined and also injective. Let $g \in \mathrm{Im}f$. So there exists $a \in G_1$ such that $f(a) = g$. Consider $a\mathrm{Ker}f \in G_1/\mathrm{Ker}f$. We see that $\tilde{f}(a\mathrm{Ker}f) = f(a) = g$. This proves that $\tilde{f}$ is surjective.

Now we show that $\tilde{f}$ is a group homomorphism of $(G_1/\mathrm{Ker}f, *_1)$ onto $(\mathrm{Im}f, *_2)$. Let $a\mathrm{Ker}f, b\mathrm{Ker}f \in G_1/\mathrm{Ker}f$. Now,

$$\begin{aligned}\tilde{f}(a\mathrm{Ker}f *_1 b\mathrm{Ker}f) &= \tilde{f}((a *_1 b)\mathrm{Ker}f)\\ &= f(a *_1 b) = f(a) *_2 f(b), \text{ as } f \text{ is a homomorphism.}\\ &= \tilde{f}(a\mathrm{Ker}f) *_2 \tilde{f}(b\mathrm{Ker}f).\end{aligned}$$

Therefore we conclude that $\tilde{f}$ is an isomorphism of the groups $(G_1/\mathrm{Ker}f, *_1)$ and $(\mathrm{Im}f, *_2)$. Thus $G_1/\mathrm{Ker}f \simeq \mathrm{Im}f$.

**Theorem 2.8.27 (Cayley)** Every group $(G, *)$ is isomorphic to some subgroup of the permutation group of the set $G$.

*Proof.* Let us denote the permutation group of the set $G$ by $A(G)$. We shall define a homomorphism from the group $(G, *)$ to the group $(A(G), \circ)$.

For that let $a \in G$, define $\phi_a : G \to G$ by $\phi_a(g) := a * g$ for all $g \in G$. First we show that $\phi_a \in A(G)$. Let $b \in G$. Consider $a^{-1} * b \in G$, note that $\phi_a(a^{-1} * b) = a * (a^{-1} * b) = b$. So the map $\phi_a$ is onto. Now let $\phi_a(g_1) = \phi_a(g_2)$, so $a * g_1 = a * g_2$. Therefore $g_1 = a^{-1} * (a * g_1) = a^{-1} * (a * g_2) = g_2$. This proves that $\phi_a$ is injective and hence $\phi_a \in A(G)$. Now we define $\psi : G \to A(G)$ by $\psi(g) := \phi_g$ for all $g \in G$. Let $g_1, g_2 \in G$. We show that $\psi(g_1 * g_2) = \psi(g_1) \circ \psi(g_2)$. Note that for $a \in G$,

$$\phi_{g_1 * g_2}(a) = (g_1 * g_2) * a = g_1 * (g_2 * a) = g_1 * \phi_{g_2}(a) = \phi_{g_1}(\phi_{g_2}(a)) = \phi_{g_1} \circ \phi_{g_2}(a).$$

This proves that the functions $\phi_{g_1 * g_2}$ and $\phi_{g_1} \circ \phi_{g_2}$ are same, and hence $\psi$ is a group homomorphism. Now,

$$\begin{aligned}\mathrm{Ker}\psi &:= \{g \in G : \psi(g) = e_{A(G)}\}\\ &= \{g \in G : \phi_g = e_{A(G)}\}\\ &= \{g \in G : g * a = a \text{ for all } a \in G\}\\ &= \{e_G\}.\end{aligned}$$

Here $e_G, e_{A(G)}$ denote the identities of the group $(G, *)$ and $(A(G), \circ)$, respectively. Therefore by first isomorphism theorem,

$$G \simeq G/\mathrm{Ker}\psi \simeq \mathrm{Im}\psi.$$

We know that Im$\psi$ is a subgroup of $(A(G), \circ)$, hence the proof.

**Corollary 2.8.28** If $(G, *)$ is a finite group of order $n \in \mathbb{N}$, then the group $(G, *)$ is isomorphic to some subgroup of $S_n$.

**Theorem 2.8.29 (Second isomorphism theorem)** Let $H, K$ be two subgroups of a group $(G, *)$ with $K$ normal in $(G, *)$. Then,

$$H/\ (H \cap K) \simeq HK/K.$$

*Proof.* We define, $\psi : H \to HK/K$ by $\psi(h) := hK$ for all $h \in H$. First we show that $\psi$ is a homomorphism. Let $h_1, h_2 \in H$. Note that,

$$\psi(h_1 * h_2) = (h_1 * h_2)K = h_1K * h_2K = \psi(h_1) * \psi(h_2).$$

So $\psi$ is a homomorphism. Let $xK \in HK/K$. Now there exist $h \in H, k \in K$ such that $x = hk$. So $xK = hkK = hK$. Consider $h \in H$, so $\psi(h) = hK = xK$. So $\psi$ is onto i.e. Im$\psi = HK/K$. Now,

$$\text{Ker}\psi = \{h \in H : \psi(h) = K\} = \{h \in H : hK = K\} = \{h \in H : h \in K\} = H \cap K.$$

Therefore by first isomorphism theorem,

$$H/(H \cap K) \simeq HK/K.$$

**Remark 2.8.30** We know that even if one of $H, K$ is not normal in $(G, *)$, then $H \cap K$ need not be normal in $(G, *)$. For example, take $K = G$ and $H$ a subgroup which is not normal in $(G, *)$. Then $H \cap K = H$ is not normal in $(G, *)$. But here we see that $H \cap K$ is normal in $H$ whenever $K$ is normal in $(G, *)$.

**Theorem 2.8.31 (Third isomorphism theorem)** Let $H, K$ be two normal subgroups of a group $(G, *)$ with $H \subseteq K$. Then,

$$(G/H)\Big/(K/H) \simeq G/K.$$

*Proof.* We define, $\psi : G/H \to G/K$ by $\psi(gH) := gK$ for all $g \in G$. First we show that $\psi$ is well-defined. Let $g_1H = g_2H$. Then $g_1^{-1}g_2 \in H \subseteq K$. So $g_1K = g_2K$, this proves that $\psi$ is well-defined. Now we show that $\psi$ is a homomorphism. Let $g_1H, g_2H \in G/H$.

Now,

$$\psi(g_1H * g_2H) = \psi(g_1 * g_2H) = g_1 * g_2K = g_1K * g_2K = \psi(g_1H) * \psi(g_2H).$$

Therefore $\psi$ is a homomorphism. For $gK \in G/K$, consider $gH \in G/H$, so $\psi(gH) = gK$, and hence $\psi$ is onto. Note that,

$$\text{Ker}\psi = \{gH \in G/H : \psi(gH) = K\} = \{gH \in G/H : gK = K\} = \{gH \in G/H : g \in K\} = K/H.$$

Thus by first isomorphism theorem,

$$(G/H)\Big/(K/H) \simeq G/K.$$

Now we try to understand the connection between the subgroups of $(G, *)$ and the subgroups of $(G/H, *)$ where $H$ is normal in $(G, *)$.

**Theorem 2.8.32** Let $H$ be a normal subgroup of $(G, *)$. If $K$ is a subgroup of $(G, *)$ with $H \subseteq K$, then $K/H := \{kH : k \in K\}$ is a subgroup of $(G/H, *)$.

*Proof.* Define $\psi : G \mapsto G/H$ by $\psi(g) := gH$ for all $g \in G$. Note that for $g_1, g_2 \in G$, $\psi(g_1 * g_2) = (g_1 * g_2)H = g_1H * g_2H = \psi(g_1) * \psi(g_2)$, so $\psi$ is a homomorphism. We know that under a homomorphism, image of a subgroup of the domain is again a subgroup of the range. So for a subgroup $K$ of $(G, *)$, the image of $K$ under $\psi$, i.e. $\psi(K) = \{kH : k \in K\}$ is a subgroup of $(G/H, *)$. As $H \subseteq K$, we can write $\psi(K) = \{kH : k \in K\} = K/H$, and hence the proof.

**Theorem 2.8.33** Let $H$ be a normal subgroup of $(G, *)$. Suppose $T$ is a subgroup of $(G/H, *)$, then there exists a subgroup $K$ of $(G, *)$ with $H \subseteq K$ such that $T = K/H$.

*Proof.* Define the set $K := \{g \in G : gH \in T\} \subseteq G$. We show that $K$ is a subgroup of $(G, *)$. Note that $e \in K$ as $H \in T$. So $K \neq \varnothing$. Let $a, b \in K$. Therefore $aH, bH \in T$. As $T$ is a subgroup of $(G/H, *)$, we have $(bH)^{-1} = b^{-1}H \in T$. Thus $a * b^{-1}H \in T$. Therefore $a * b^{-1} \in K$. This proves that $K$ is a subgroup of $(G, *)$. Clearly $H \subseteq K$ as for all $h \in H$, $hH = H \in T$. Now from the definition of $K$ we can note that $T = K/H$. This completes the proof.

**Theorem 2.8.34** Let $H$ be a normal subgroup of $(G, *)$. The subgroups of $(G, *)$ containing $H$ and the subgroups of $(G/H, *)$ are in bijective correspondence.

*Proof.* Let $A$ denote the set of all subgroups of $(G, *)$ that contain $H$ and $B$ denote the set of all subgroups of $(G/H, *)$. We show that there is a bijection from the set $A$ to the set $B$. We have seen that for a subgroup $K$ of $(G, *)$ containing $H$, $K/H$ is a subgroup of $(G/H, *)$. So we can define $f : A \to B$ by $f(K) = K/H$ for all $K \in A$. Also we have seen that for any subgroup $T$ of $(G/H, *)$, there exists a subgroup $K$ of $(G, *)$ such that $H \subseteq K$ and $T = K/H$. So the map $f$ is onto. It only remains to show that $f$ is injective. Let $f(K_1) = f(K_2)$ i.e. $K_1/H = K_2/H$. We show that $K_1 = K_2$. Let $k_1 \in K_1$. So $k_1H \in K_1/H = K_2/H$. This implies that $k_1H = k_2H$ for some $k_2 \in K_2$. Thus we have, $k_1^{-1} * k_2 \in H \subseteq K_2$. Hence $k_1^{-1} = (k_1^{-1} * k_2) * k_2^{-1} \in K_2$. Therefore $k_1 \in K_2$ as $K_2$ is a subgroup of $(G, *)$. This proves that $K_1 \subseteq K_2$. Similarly we can show that $K_2 \subseteq K_1$ and hence $K_1 = K_2$. This proves that the sets $A$ and $B$ are in bijective correspondence.

**Example 2.8.35** As an application we can find out all the subgroups of $(\mathbb{Z}/12\mathbb{Z}, +)$. We know that all the subgroups of $(\mathbb{Z}/12\mathbb{Z}, +)$ are of the form $H/12\mathbb{Z}$ where $12\mathbb{Z} \subseteq H$ and $H$ is a subgroup of $(\mathbb{Z}, +)$. We know that $(\mathbb{Z}, +)$ is cyclic, so all the subgroups of $(\mathbb{Z}, +)$ are of the form $(m\mathbb{Z}, +)$ for $m \in \mathbb{Z}$. Note that $-m\mathbb{Z} = m\mathbb{Z}$. So we can restrict ourselves to $m \in \mathbb{N} \cup \{0\}$. Now $12\mathbb{Z} \subseteq m\mathbb{Z}$ if and only if $m \mid 12$. So the possibilities for $m$ are $1, 2, 3, 4, 6, 12$. So there are total 6 subgroups of $(\mathbb{Z}/12\mathbb{Z}, +)$ and they are $(\mathbb{Z}/12\mathbb{Z}, +)$, $(2\mathbb{Z}/12\mathbb{Z}, +)$, $(3\mathbb{Z}/12\mathbb{Z}, +)$, $(4\mathbb{Z}/12\mathbb{Z}, +)$, $(6\mathbb{Z}/12\mathbb{Z}, +)$ and $(12\mathbb{Z}/12\mathbb{Z}, +)$. Note that $(\mathbb{Z}/12\mathbb{Z}, +)$ and $(12\mathbb{Z}/12\mathbb{Z}, +)$ are the trivial subgroups of $(\mathbb{Z}/12\mathbb{Z}, +)$.

**Theorem 2.8.36** Let $H$ be a normal subgroup of $(G, *)$. Let $K_1, K_2$ be two subgroups of $(G, *)$ containing $H$. Then $K_1 \subseteq K_2$ if and only if $K_1/H \subseteq K_2/H$.

*Proof.* Let $H \subseteq K_1 \subseteq K_2$. Then clearly $K_1/H \subseteq K_2/H$. So we assume that $K_1/H \subseteq K_2/H$. We show that $K_1 \subseteq K_2$. Let $k_1 \in K_1$. Therefore $k_1H \in K_1/H$. As $K_1/H \subseteq K_2/H$, we have $k_1H = k_2H$ for some $k_2 \in K_2$. This implies that $k_1^{-1} * k_2 \in H \subseteq K_2$. So,

$$k_1^{-1} = (k_1^{-1} * k_2) * k_2^{-1} \in K_2.$$

As $K_2$ is a subgroup of $(G, *)$, we have $k_1 \in K_2$. Thus $K_1 \subseteq K_2$.

**Theorem 2.8.37** Let $H$ be a normal subgroup of $(G, *)$ and $K$ be a subgroup of $(G, *)$ containing $H$. Then $K$ is normal in $(G, *)$ if and only if $K/H$ is normal in $(G/H, *)$.

*Proof.* First suppose that $K$ is normal in $(G, *)$. We have already seen that $\psi : G \to G/H$ defined by $\psi(g) := gH$ for all $g \in G$ is a surjective homomorphism. We know that under

a surjective homomorphism, image of a normal subgroup is normal. We have $K$ is normal in $(G, *)$, so the image $\psi(K) = K/H$ is normal in $(G/H, *)$.

Next suppose that $K/H$ is normal in $(G/H, *)$. We show that $K$ is normal in $(G, *)$. Let $g \in G$ and $k \in K$. So $gH \in G/H$ and $kH \in K/H$. Now as $K/H$ is normal in $(G/H, *)$,

$$g * k * g^{-1}H = gH * kH * g^{-1}H = gH * kH * (gH)^{-1} \in K/H.$$

Thus $g * k * g^{-1} = k'$ for some $k' \in K$. This proves that $g * k * g^{-1} \in K$ and hence $K$ is normal in $(G, *)$.

## 2.9 Direct product of groups

Let $(G_1, *_1)$ and $(G_2, *_2)$ be two groups. We consider the Cartesian product of the sets $G_1$ and $G_2$, i.e.

$$G_1 \times G_2 := \{(g_1, g_2) : g_1 \in G_1 \text{ and } g_2 \in G_2\}.$$

We define an operation $*$ on $G_1 \times G_2$ as follows: for $(a, b), (c, d) \in G_1 \times G_2$,

$$(a, b) * (c, d) = (a *_1 c, b *_2 d).$$

As $a *_1 c \in G_1$ and $b *_2 d \in G_2$, we see that $*$ is a binary operation on $G_1 \times G_2$. Now $G_1 \times G_2 \neq \varnothing$ as both $G_1, G_2$ are non-empty. The set $G_1 \times G_2$ with the operation $*$ forms a group, which we call as the **direct product** of the groups $(G_1, *_1)$ and $(G_2, *_2)$. The element $(e_1, e_2)$ is the identity of $(G_1 \times G_2, *)$ where $e_1, e_2$ denote the identities of $(G_1, *_1)$ and $(G_2, *_2)$ respectively. For each $(a, b) \in G_1 \times G_2$, consider $a^{-1} \in G_1$ and $b^{-1} \in G_2$, then the element $(a^{-1}, b^{-1})$ is the inverse of $(a, b)$ in $(G_1 \times G_2, *)$. Now one can check that the associativity of $*$ on $G_1 \times G_2$ follows from the associativity properties of $*_1$ and $*_2$. With these we are done showing that $(G_1 \times G_2, *)$ is a group. We can take $n$ many groups $(G_i, *_i)$ for $1 \leq i \leq n$ and consider their direct product $G_1 \times \cdots \times G_n$ defining $*$ on $G_1 \times \cdots \times G_n$, in a similar manner.

**Exercise 2.9.1** The group $(G_1 \times G_2, *)$ is abelian if and only if both $(G_1, *_1)$ and $(G_2, *_2)$ are abelian.

**Example 2.9.2** The direct product $S_3 \times \mathbb{Z}$ is an infinite non-abelian group.

**Exercise 2.9.3** If the group $(G_1 \times G_2, *)$ is cyclic then both $(G_1, *_1)$ and $(G_2, *_2)$ are cyclic.

**Remark 2.9.4** The converse of the above exercise need not be true.

- The group $(\mathbb{Z} \times \mathbb{Z}, +)$ is not cyclic though $(\mathbb{Z}, +)$ is cyclic. Suppose $(\mathbb{Z} \times \mathbb{Z}, +)$ is cyclic and let $(a, b)$ be a generator of $(\mathbb{Z} \times \mathbb{Z}, +)$. Now $(1, 0), (0, 1) \in \mathbb{Z} \times \mathbb{Z}$. So there exist $m, n \in \mathbb{Z}$ such that

  $$(1, 0) = m(a, b) \quad \text{and} \quad (0, 1) = n(a, b).$$

  Now $ma = 1$ and $mb = 0$ gives that $b = 0$. Also we have $nb = 1$. This is a contradiction. Thus $(\mathbb{Z} \times \mathbb{Z}, +)$ is not cyclic.

- The group $(\mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/6\mathbb{Z}, +)$ is not cyclic though $(\mathbb{Z}/4\mathbb{Z}, +)$, $(\mathbb{Z}/6\mathbb{Z}, +)$ both are cyclic. If $(\mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/6\mathbb{Z}, +)$ is cyclic, then we must have an element of order 24 in it. But we can see that order of no element in $(\mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/6\mathbb{Z}, +)$ exceeds $\text{lcm}(4, 6) = 12$.

At this point we have the following theorem which is equivalent to the well-known Chinese remainder theorem.

**Theorem 2.9.5** The group $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ is cyclic if and only if $\gcd(m, n) = 1$. In other words, $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \simeq \mathbb{Z}/mn\mathbb{Z}$ if and only if $\gcd(m, n) = 1$.

*Proof.* First suppose that $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ is a cyclic group. We show that $\gcd(m, n) = 1$. Let $\gcd(m, n) = d > 1$. Since $\mathbb{Z}/m\mathbb{Z}, \mathbb{Z}/n\mathbb{Z}$ are finite groups of order $m, n$ respectively, we have $ma = [0]$ and $nb = [0]$ for all $a \in \mathbb{Z}/m\mathbb{Z}$ and $b \in \mathbb{Z}/n\mathbb{Z}$. Note that,

$$\frac{mn}{d}(a, b) = \left(\frac{n}{d}ma, \frac{m}{d}nb\right) = ([0], [0]).$$

Now $mn/d \in \mathbb{N}$ and $mn/d < mn$ as $d > 1$. Therefore order of every element in $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ is strictly less than $mn$, a contradiction to the fact that $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ is a cyclic group. Therefore $d = 1$.

Next suppose that $\gcd(m, n) = 1$. We show that $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ is a cyclic group. Since $\mathbb{Z}/m\mathbb{Z}, \mathbb{Z}/n\mathbb{Z}$ are cyclic groups of order $m, n$ respectively, there exist $a \in \mathbb{Z}/m\mathbb{Z}$ and $b \in \mathbb{Z}/n\mathbb{Z}$ such that $o(a) = m$ and $o(b) = n$. Now,

$$mn(a, b) = (n(ma), m(nb)) = ([0], [0]).$$

So $o(a,b) \leq mn$. Now if $d(a,b) = ([0],[0])$, then $da = [0]$ and $db = [0]$. Therefore $m \mid d$ and $n \mid d$. We have $\gcd(m,n) = 1$, so $mn \mid d$. So $mn$ is the least positive integer so that $mn(a,b) = ([0],[0])$ and hence $o(a,b) = mn$. This proves that $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ is cyclic.

Note that $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ is cyclic of order $mn$ if and only if $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \simeq \mathbb{Z}/mn\mathbb{Z}$.

**Corollary 2.9.6** If $(G_1, *_1)$ and $(G_2, *_2)$ are two finite cyclic groups of co-prime orders, then $(G_1 \times G_2, *)$ is a cyclic group.

*Proof.* Suppose order of $(G_1, *_1)$ and $(G_2, *_2)$ are $m, n$ respectively. Then,

$$G_1 \simeq \mathbb{Z}/m\mathbb{Z} \quad \text{and} \quad G_2 \simeq \mathbb{Z}/n\mathbb{Z}.$$

So, $G_1 \times G_2 \simeq \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ and as $\gcd(m,n) = 1$, we have $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ is cyclic. Thus $G_1 \times G_2$ is cyclic.

**Exercise 2.9.7** Let $(G_1, *_1)$, $(G_2, *_2)$ be two groups such that $G_1 \simeq G_2$. Then for any group $(G, *)$ show that $G \times G_1 \simeq G \times G_2$.

**Remark 2.9.8** The converse of the above exercise need not be true i.e. $G \times G_1 \simeq G \times G_2$ need not imply $G_1 \simeq G_2$. For example let $(G, *)$ be the group $(\prod_{\mathbb{N}} \mathbb{Z}, +)$. Consider $(G_1, *_1)$ and $(G_2, *_2)$ as $(\mathbb{Z}, +)$ and $\{0\}$ respectively. Then although $G \times G_1 \simeq G \times G_2$, we see that $G_1 \not\simeq G_2$.

When for any group $(G_1, *_1)$ and $(G_2, *_2)$, we have $G \times G_1 \simeq G \times G_2$ implies $G_1 \simeq G_2$, we say that the group $(G, *)$ is **cancellable group**. One can show that any finite group is a cancellable.

**Definition 2.9.9** Let $H, K$ be two subgroups of a group $(G, *)$. Then $(G, *)$ is said to be an **internal direct product** of $H$ and $K$ if

1) $G = HK$,

2) $H \cap K = \{e\}$,

3) $h * k = k * h$ for all $h \in H$ and $k \in K$.

**Examples 2.9.10**

- Consider the Klein's 4 group. Consider the subgroups $H = \{e, a\}$ and $K = \{1, b\}$ of it. Note that Klein's 4 group is an internal direct product of $H$ and $K$.

- Consider $A_4$, the alternating group of 4 symbols. Consider the subgroups

$$H = \{e, (1\ 2)(3\ 4), (1\ 3)(2\ 4), (1\ 4)(2\ 3)\} \text{ and } K = \langle (1\ 2\ 3) \rangle.$$

  Note that $|HK| = 12 = |A_4|$ and $HK \subseteq A_4$. Thus $A_4 = HK$. But $A_4$ is not an internal direct product of $H$ and $K$.

**Lemma 2.9.11** Let $H, K$ be two normal subgroups of the group $(G, *)$ and $H \cap K = \{e\}$. Then $h * k = k * h$ for all $h \in H$ and $k \in K$.

*Proof.* Let $h \in H$ and $k \in K$. Now $(h * k) * (h * k)^{-1} = h * k * h^{-1} * k^{-1}$. We see that as $H, K$ are normal, $h * k * h^{-1} * k^{-1} \in H \cap K$. Therefore, $(h * k) * (h * k)^{-1} = e$. Hence $h * k = k * h$.

The next theorem gives us an equivalent definition of the internal direct product of groups.

**Theorem 2.9.12** Let $H, K$ be two subgroups of the group $(G, *)$. Then $(G, *)$ is an internal direct product of $H$ and $K$ if and only if

1) $G = HK$,

2) $H \cap K = \{e\}$,

3) $H, K$ are normal subgroups.

*Proof.* We see that the if part of the theorem follows from the previous lemma. So we assume that $(G, *)$ is an internal direct product of the subgroups $H$ and $K$. We show that $H, K$ are normal subgroups. Let $g \in G$ and $h \in H$. As $G = HK$, we can have $h_1 \in H$ and $k_1 \in K$ such that $g = h_1 * k_1$. Therefore,

$$\begin{aligned} g * h * g^{-1} &= h_1 * k_1 * h * k_1^{-1} * h_1^{-1} \\ &= h_1 * h * k_1 * k_1^{-1} * h_1^{-1}, \text{ as the elements of } H, K \text{ commute with each other.} \\ &= h_1 * h * h_1^{-1} \in H. \end{aligned}$$

This proves that $H$ is a normal subgroup. Similarly one can show that $K$ is a normal subgroup.

**Remark 2.9.13** Let $n \in \mathbb{N}$ and $H_1, \cdots, H_k$ be normal subgroups of $(G, *)$ such that $G = H_1 \cdots H_k$ and $H_i \cap H_1 \cdots H_{i-1} H_{i+1} \cdots H_k = \{e\}$ for all $1 \leq i \leq k$. Then $(G, *)$ is called the **internal direct product** of $H_1, \ldots, H_k$. Note that $H_i \cap H_1 \cdots H_{i-1} H_{i+1} \cdots H_k = \{e\}$

implies that $H_j \cap H_j = \{e\}$ for all $i \neq j$. But the other way is not true. For example consider the symmetric group of three symbols $S_3$. Let us consider,

$$H_1 := \{e, (1\ 2)\},\ H_2 := \{e, (1\ 3)\} \text{ and } H_3 := \{e, (1\ 2\ 3), (1\ 3\ 2)\}.$$

Here $H_i \cap H_j = \{e\}$ for all $i \neq j$ but $H_1 \cap H_2H_3 = H_1 \cap S_3 = H_1$.

**Theorem 2.9.14** Let $(G, *)$ be an internal direct product of the subgroups $H$ and $K$. Then,

1) $G \simeq H \times K$.

2) $G/H \simeq K$ and $G/K \simeq H$.

*Proof.*

1) Define $f : H \times K \to G$ by $f(h, k) = h * k$ for all $h \in H, k \in K$. It is easy to note that $f$ is a group homomorphism. We show that $f$ is bijective.

   Let $g \in G$. Since $(G, *)$ is an internal direct product of $H, K$, there exist $h \in H$ and $k \in K$ such that $g = h * k$. We consider $(h, k) \in H \times K$ so that $f(h, k) = h * k = g$.

   Let $f(h_1, k_1) = f(h_2, k_2)$ i.e. $h_1 * k_1 = h_2 * k_2$ for $h_1, h_2 \in H$ and $k_1, k_2 \in K$. Therefore, $h_1^{-1} * h_2 = k_1 * k_2^{-1}$. Since $H \cap K = \{e\}$ we get that $h_1 = h_2$ and $k_1 = k_2$.

2) Define $\phi : K \to G/H$ by $\phi(k) = kH$ for all $k \in K$. Note that $\phi$ is a group homomorphism. Clearly $\phi$ is onto. Now,

$$\text{Ker}\phi = \{k \in K : \phi(k) = H\} = \{k \in K : kH = H\} = H \cap K = \{e\}.$$

   Therefore by first isomorphism theorem we obtain, $G/H \simeq K$. Similarly by defining $\psi : H \to G/K$ by $\psi(h) = hK$ for all $h \in H$, we can show that $G/K \simeq H$.

**Remark 2.9.15** Similarly one can show that if $(G, *)$ is an internal direct product of $H_1, \ldots, H_k$, then $G \simeq H_1 \times \cdots \times H_k$ and $G/H_i \simeq H_1 \times H_{i-1} \times H_{i+1} \times H_k$ for all $1 \leq i \leq k$.

**Remark 2.9.16** It is noteworthy that if $(G, *)$ is direct product of two of its subgroups then we can view $(G, *)$ as an internal direct product of its subgroups too. So we see that the notions of direct product and internal direct product are not different.

## 2.10 Conjugacy relation and Class equation

Let $(G, *)$ be a group. We define a relation $\sim$ on $G$ as follows: for $a, b \in G$, we say that $a \sim b$ if and only if there exist a $g \in G$ such that $b = g * a * g^{-1}$.

**Exercise 2.10.1** Show that $\sim$ is an equivalence relation on $(G, *)$.

This equivalence relation is called the **conjugacy relation** and if $a, b \in G$ are related by conjugacy relation, then they are called the **conjugates** of each other. We know that every equivalence relation gives rise to a partition on the set. Here the equivalence class of $a \in G$ is called the **conjugacy class** of $a \in G$ and is denoted by $cl(a)$. Note that $cl(a) \neq \varnothing$ as $a \in cl(a)$. So if we have $(G, *)$ is a group of finite order then we can write,

$$G = \bigcup_{1 \leq i \leq n} cl(a_i),$$

where $a_i$'s for all $1 \leq i \leq n$ are all the distinct representatives of all the conjugacy classes of $G$.

**Exercise 2.10.2** Show that $a \in Z(G)$ if and only if $cl(a) = \{a\}$.

Therefore for distinct representatives $x_i \notin Z(G)$, $1 \leq i \leq t$, we can write,

$$G = \bigcup_{1 \leq i \leq t} cl(x_i) \bigcup Z(G).$$

Thus,

$$|G| = |Z(G)| + \sum_{1 \leq i \leq t} |cl(x_i)|.$$

The above equation is known as the **class equation** of the group $(G, *)$.

Recall that in a group $(G, *)$, for $a \in G$, by the **centraliser** of $a \in G$, we mean the set $C_G(a) := \{g \in G : g * a = a * g\}$. We know that $Z(G) \subseteq C_G(a)$ for every $a \in Z(G)$.

**Theorem 2.10.3** Let $(G, *)$ be a group. For any $a \in G$, we have $[G : C_G(a)] = |cl(a)|$.

*Proof.* Let $\mathcal{L}$ denote the set of all left cosets of $C_G(a)$ in $(G, *)$. We shall show that there is a bijection from $\mathcal{L}$ to $cl(a)$. We define, $f : \mathcal{L} \to cl(a)$ by $f(gC_G(a)) := g * a * g^{-1}$. We need to show first that $f$ is well-defined. Note that,

$$\begin{aligned} & xC_G(a) = yC_G(a) \\ \iff & x^{-1} * y \in C_G(a) \end{aligned}$$

$$\iff (x^{-1} * y) * a = a * (x^{-1} * y)$$
$$\iff x * a * x^{-1} = y * a * y^{-1}.$$

From the above we get that $f$ is well-defined and injective. It only remains to show that $f$ is onto. Let $b \in cl(a)$. Then there exists $g \in G$ such that $b = g * a * g^{-1}$. We consider $gC_G(a) \in \mathcal{L}$. Then $f(gC_G(a)) = g * a * g^{-1} = b$. Thus $f$ is a bijective map from $\mathcal{L}$ to $cl(a)$. Hence

$$|\mathcal{L}| = |cl(a)|.$$

So we can rewrite the **class equation** as

$$|G| = |Z(G)| + \sum_{1 \le i \le t} \frac{|G|}{|C_G(x_i)|}.$$

### 2.10.1 Applications of class equation

1) Let $(G, *)$ be a group of order $p^n$ where $p$ is a prime and $n \in \mathbb{N}$. Then $Z(G) \neq \{e\}$. If $Z(G) = G$, then the statement is true. So let $Z(G) \neq G$. Consider $a \notin Z(G)$. Therefore, the centraliser of $a$, $C_G(a) \neq G$. Also $C_G(a) \neq \{e\}$ as $a \in C_G(a)$. So $C_G(a)$ is a proper subgroup of $(G, *)$. Thus we have,

$$\frac{|G|}{|C_G(a)|} = p^{n-r} \text{ for some } 0 < r < n.$$

We have,

$$|Z(G)| = |G| - \sum_{1 \le i \le t} \frac{|G|}{|C_G(x_i)|}.$$

Note that

$$p \mid \left( |G| - \sum_{1 \le i \le t} \frac{|G|}{|C_G(x_i)|} \right),$$

so we get $|Z(G)|$ is divisible by $p$. Also $|Z(G)| \geq 1$ as $e \in Z(G)$. Therefore $|Z(G)| \geq p$ and hence $Z(G) \neq \{e\}$.

2) Every group of order $p^2$ is abelian where $p$ is prime. We have already seen that $Z(G) \neq \{e\}$. So $|Z(G)|$ is either $p$ or $p^2$. Suppose $|Z(G)| = p$. Let $a \notin Z(G)$ and consider $C_G(a)$. We have,

$$Z(G) \subseteq C_G(a) \subseteq G.$$

So $|C_G(a)|$ is either $p$ or $p^2$. If $|C_G(a)| = p$, then $Z(G) = C_G(a)$, a contradiction as $a \notin Z(G)$. If $|C_G(a)| = p^2$, then $C_G(a) = G$. This implies that $a \in Z(G)$, a contradiction. Therefore $|Z(G)| = p^2$ and hence $Z(G) = G$. This completes the proof.

## 2.11 Partial converses of the Lagrange's theorem

We know that in general for any group of order $n$, if $m \mid n$, there need not be a subgroup of order $m$. For example, we have seen that $A_4$ does not have a subgroup of order 6 though $6 \mid |A_4|$. In this section we discuss the partial converses of the Lagrange's theorem. Recall that we have seen that for a finite cyclic group of order $n$ if $m \in \mathbb{N}$ be such that $m \mid n$, then there exists a **unique** subgroup of order $m$. So for finite cyclic groups the converse of Lagrange's theorem holds. The theorem below by Cauchy tells us that in general for any finite group, the converse of Lagrange's theorem holds for prime divisors of the order of the group.

**Theorem 2.11.1 (Cauchy)** Let $(G, *)$ be a finite group and $p$ be a prime such that $p \mid |G|$. Then $(G, *)$ has an element of order $p$.

*Proof.* First we prove the statement for all abelian groups of finite order. Let $(G, *)$ be a finite abelian group and $p$ be a prime such that $p \mid |G|$. We use induction on $|G|$ to prove the statement. Suppose $|G| = 1$, then there is no prime $p$ such that $p \mid |G|$, so the statement is true vacuously. Suppose the statement is true for all abelian groups of order strictly less than $|G|$. Now we consider the abelian group $(G, *)$. Since $|G| > 1$, there exists an element $a \in G$ such that $a \neq e$.
**Case 1:** Suppose, $p \mid o(a)$. Then $o(a) = pm$ for some $m \in \mathbb{N}$. We consider $a^m \in G$. Note that,

$$o(a^m) = \frac{pm}{\gcd(pm, m)} = p.$$

**Case 2:** Suppose, $p \nmid o(a)$. Let $o(a) = t$. Then $\gcd(t, p) = 1$. We consider the subgroup $H := \langle a \rangle$. Since $(G, *)$ is abelian, we have $H$ is normal in $(G, *)$. Therefore we can consider the quotient group $(G/H, *)$. Now,

$$|G/H| = \frac{|G|}{|H|} < |G|, \text{ as } t > 1.$$

Now $(G/H, *)$ is also abelian, so by induction hypothesis $(G/H, *)$ has an element say $y$ of order $p$. We know that $f : G \to G/H$ defined by $f(g) = gH$ for all $g \in G$ is an

onto homomorphism. So for $y \in G/H$, there exists an $x \in G$ such that $f(x) = y$. We also know that $o(\phi(g)) \mid o(g)$ for any group homomorphism $\phi$. So here we get $p \mid o(x)$ as $o(y) \mid o(x)$. This gives us an element $x \in G$ such that $p \mid o(x)$. Hence following the argument in Case 1, we have an element of order $p$.

Therefore we have proved the statement for all finite abelian groups. Now we shall prove the statement for all finite groups. Let $(G, *)$ be a finite group and $p$ be a prime such that $p \mid |G|$. We use induction on $|G|$ to prove it. Suppose $|G| = 1$, then there is no prime $p$ such that $p \mid |G|$, so the statement is true vacuously. Suppose the statement is true for all groups of order strictly less than $|G|$. Now we consider the group $(G, *)$.
**Case 1:** There exist a proper subgroup $H$ of $(G, *)$ such that $p \mid |H|$. Now as $|H| < |G|$, there exists an element of order $p$ in $(H, *)$ and hence in $(G, *)$.
**Case 2:** There is no proper subgroup of $(G, *)$, whose order is divisible by the prime $p$. We write down the class equation of $(G, *)$:

$$|G| = |Z(G)| + \sum_{1 \leq i \leq k} \frac{|G|}{|C_G(a_i)|},$$

where for all $1 \leq i \leq k$, $a_i \notin Z(G)$ and they are representatives of all distinct conjugacy classes of the group $(G, *)$. For all $1 \leq i \leq k$, we have $a_i \notin Z(G)$, so $C_G(a_i) \subset G$. Therefore in this case we have $p \nmid C_G(a_i)$ for all $1 \leq i \leq k$. So for all $1 \leq i \leq k$, we have $p \mid (|G|/|C_G(a_i)|)$. This implies that $p \mid |Z(G)|$. So $Z(G)$ can not be a proper subgroup of $(G, *)$. Hence $G = Z(G)$, i.e. $(G, *)$ is abelian and for an abelian group we have already proved the statement.

### 2.11.1 Applications of Cauchy's theorem

1) Up to isomorphism there are only two groups of order 6.
Let $(G, *)$ be a group of order 6. By Cauchy's theorem there exist elements of order 2 and 3 in $(G, *)$. Ley $a \in G$ be such that $o(a) = 2$ and $b \in G$ such that $o(b) = 3$.
**Case I :** Suppose $ab = ba$. Therefore $(a * b)^n = a^n * b^n$ for any $n \in \mathbb{N}$. Thus $o(ab) = \text{lcm}(o(a), o(b)) = o(a)o(b) = 6$. Therefore $(G, *)$ is cyclic group of order 6 and hence $G \simeq \mathbb{Z}/6\mathbb{Z}$.
**Case II :** Suppose $ab \neq ba$. Consider the following subgroups:

$$H := \{e, a\} \quad \text{and} \quad K := \{e, b, b^2\}.$$

Note that $K$ is a normal subgroup of $(G, *)$ as $[G : K] = 2$. Therefore $HK$ is a subgroup of $(G, *)$. Now orders of $H, K$ are co-prime, so $|H \cap K| = 1$. Therefore,

$$|HK| = \frac{|H||K|}{|H \cap K|} = 6 = |G|.$$

This proves that $G = HK := \{e, a, b, b^2, a * b, a * b^2\}$. It is a good point to note that $(G, *)$ is not an internal direct product of $H, K$ as $ab \neq ba$. We define a map $f : G \to S_3$ as follows:

$$f(e) := e,\ f(a) := (1\ 2),\ f(b) := (1\ 2\ 3),$$
$$f(b^2) := f(b)^2,\ f(a * b) := f(a) \circ f(b),\ f(a * b^2) := f(a) \circ f(b)^2.$$

So from the definition of $f$ we see that $f$ is a group homomorphism. We show that $f$ is onto. As $\{e, (1\ 2), (1\ 2\ 3), (1\ 3\ 2)\} \subseteq \mathrm{Im}f$, we have $|\mathrm{Im}f| \geq 4$. Also we know that $\mathrm{Im}f$ is a subgroup of $S_3$, therefore by Lagrange's theorem $|\mathrm{Im}f| \mid 6$. This proves that $|\mathrm{Im}f| = 6$ and hence $\mathrm{Im}f = S_3$. Since $f$ is an onto map from a set of 6 elements to a set of 6 elements, we get that $f$ is injective. By first isomorphism theorem we obtain $G \simeq S_3$.
Thus we get $(G, *)$ is either isomorphic to $\mathbb{Z}/6\mathbb{Z}$ or $S_3$ and hence the proof.

2) Let $(G, *)$ be a group of order $p^n$ where $p$ is a prime and $n \in \mathbb{N}$. Then $(G, *)$ contains a normal subgroup of order $p$.
Since $|G| = p^n$, we have $|Z(G)| \geq 1$. So $|Z(G)| = p^r$ for some $0 < r \leq n$. Now since $p \mid |Z(G)|$, by Cauchy's theorem there exists an element say $x$ of order $p$. Consider $H := \langle x \rangle \subseteq Z(G)$. So $H$ is normal and $|H| = p$. This proves the existence of a normal subgroup of order $p$ in $(G, *)$.

3) Let $(G, *)$ be a group of order $pn$ where $p$ is a prime and $n \in \mathbb{N}$ with $p > n$. Then $(G, *)$ contains a normal subgroup of order $p$.
Since $p \mid |G|$, by Cauchy's theorem there exists an element say $x$ of order $p$. We consider $H := \langle x \rangle$. This proves the existence of one subgroup of order $p$. To show that there exists one normal subgroup of order $p$, we show that $(G, *)$ has only one subgroup of order $p$. Because for every subgroup $H$ of order $p$, $gHg^{-1}$ is also a subgroup of order $p$ for each $g \in G$. As there is only one subgroup of order $p$, it follows that $H = gHg^{-1}$ for all $g \in G$ i.e. $H$ is normal in $(G, *)$. So we show now that $H$ is the only subgroup of $(G, *)$ of order $p$. Suppose $K$ is another subgroup of order $p$, then $H \cap K = \{e\}$ as $H \neq K$ and $H \cap K$ is subgroup of both $H$ and

$K$. Therefore,

$$|HK| = \frac{|H||K|}{|H \cap K|} = \frac{p^2}{1} > pn.$$

A contradiction as $|G| = pn$. So there is only one subgroup of order $p$.

**Exercise 2.11.2** Let $(G, *)$ be a finite abelian group of order $n$. Let $m \in \mathbb{N}$ be such that $m \mid n$. Then $(G, *)$ has a subgroup of order $m$.

The above exercise also tells us that converse of Lagrange's theorem holds for all finite abelian groups.

## 2.12 Group actions

Let $(G, *)$ be a group and $A$ be a non-empty set. By an **action of the group** $(G, *)$ **on** $A$ we mean a map from $G \times A \to A$ such that

1) $(g_1 * g_2) \cdot a = g_1 \cdot (g_2 \cdot a)$ for all $g_1, g_2 \in G$ and $a \in A$,

2) $e_G \cdot a = a$ for all $a \in A$,

where $g \cdot a$ denotes the image of $(g, a)$ and $e_G$ denotes the identity of $(G, *)$.

**Examples 2.12.1**

1. Let $V$ be a $\mathbb{R}$ vector space. Note that the scalar multiplication is an action of the group of non-zero real numbers with multiplication denoted by $(\mathbb{R} \setminus \{0\}, \cdot)$ on $V$.

2. Let $(G, *)$ be a group and $A \neq \varnothing$ be a set. Define $G \times A \to A$ as $g \cdot a = a$ for all $a \in A, g \in G$. This action is called the trivial action of $(G, *)$ on $A$.

3. Let $(G, *)$ be a group. Then $g_1 \cdot g_2 = g_1 * g_2$ for all $g_1, g_2 \in G$ is an action of the group $(G, *)$ on the set $G$.

4. Let $H$ be a normal subgroup of the group $(G, *)$. Then $g \cdot h = g * h * g^{-1}$ for all $g \in G, h \in H$ is an action of $(G, *)$ on $H$.

5. Let $n \in \mathbb{N}$ and $A = \{1, \ldots, n\}$. Note that the symmetric group with $n$ symbols $S_n$ acts on $A$ as follows: $\sigma \cdot m = \sigma(m)$.

**Remark 2.12.2** We can note that if $(G, *)$ is a non-abelian group then the map $g_1 \cdot g_2 = g_2 * g_1$ for all $g_1, g_2 \in G$ is not an action of $(G, *)$ on $G$.

**Proposition 2.12.3** Let $(G, *)$ be a group and $A \neq \varnothing$ a set. Then every action of $(G, *)$ on $A$ gives rise to a group homomorphism from $(G, *)$ to the permutation group on $A$, denoted by $S_A$.

*Proof.* Suppose $(G, *)$ acts on the set $A$. Now for $g \in G$, we have the map $\sigma_g : A \to A$ defined by $\sigma_g(a) = g \cdot a$ for all $a \in A$. Note that,

$$\begin{aligned} g \cdot a_1 = g \cdot a_2 &\implies g^{-1} \cdot (g \cdot a_1) = g^{-1} \cdot (g \cdot a_2) \\ &\implies (g^{-1} * g) \cdot a_1 = (g^{-1} * g) \cdot a_2 \\ &\implies e \cdot a_1 = e \cdot a_2 \\ &\implies a_1 = a_2. \end{aligned}$$

For $a \in A$, consider $g^{-1} \cdot a \in A$. Then $\sigma_g(g^{-1} \cdot a) = g \cdot (g^{-1} \cdot a) = a$. Therefore $\sigma_g \in S_A$. We define a map $f : G \to S_A$ defined by $f(g) = \sigma_g$ for all $g \in G$. We show that $f$ is a group homomorphism. Let $g_1, g_2 \in G$. We need to show that

$$\sigma_{g_1 * g_2} = \sigma_{g_1} \circ \sigma_{g_2}.$$

Note that for $a \in A$,

$$\sigma_{g_1 * g_2}(a) = (g_1 * g_2) \cdot a = g_1 \cdot (g_2 \cdot a) = g_1 \cdot (\sigma_{g_2}(a)) = \sigma_{g_1}(\sigma_{g_2}(a)) = (\sigma_{g_1} \circ \sigma_{g_2})(a).$$

**Remark 2.12.4** On the other hand, for a group $(G, *)$ and $A \neq \varnothing$ set, every group homomorphism $f : G \to S_A$ gives rise to an action of $(G, *)$ on $A$. Define $G \times A \to A$ by $g \cdot a = f(g)(a)$. It is easy to check that this is a group action. Also the associated group homomorphism $G \to S_A$ that we get from this action is nothing but $f$.

### 2.12.1 Orbit and stabilizer

Suppose a group $(G, *)$ is acting on a non-empty set $A$. We can define $\sim$ on $A$ as follows:

$$a \sim b \text{ if and only if there exists } g \in G \text{ such that } b = g \cdot a.$$

**Exercise 2.12.5** Prove that $\sim$ is an equivalence relation on $A$.

**Definition 2.12.6** For $a \in A$, by the **orbit of** $a$, we mean the equivalence class of $a$. We usually denote the orbit of $a$ by $\mathrm{Orb}(a)$. So,

$$\mathrm{Orb}(a) = \{g \cdot a : g \in G\}.$$

As we know that every equivalence relation gives a partition on the set, we can write $A$ as disjoint union of orbits. Below we revisit the important theorem by Lagrange in group theory.

**Theorem 2.12.7 (Lagrange)** Let $(G, *)$ be a finite group and $H$ be a subgroup of it. Then $|H|$ divides $|G|$.

*Proof.* We consider the action of $H$ on $G$ as follows: $h \cdot g := h * g$ for all $h \in H, g \in G$. Since $G$ is finite we can write, $G$ as finite disjoint union of orbits of its elements as $G$. For $g \in G$, we can note that $\mathrm{Orb}(g) = \{h \cdot g : h \in H\}$ is in bijective correspondence with $H$. So $|\mathrm{Orb}(g)| = |H|$ for each $g \in G$. Hence clearly $|H|$ divides $|G|$.

**Definition 2.12.8** For $a \in A$, by the **stabilizer of** $a$ **in** $(G, *)$, we refer to the set $\{g \in G : g \cdot a = a\}$, denoted by $G_a$.

**Exercise 2.12.9** Prove that $G_a$ is a subgroup of $(G, *)$.

**Examples 2.12.10**

1) Consider the action of the symmetric group with $n$-symbols on the set $\{1, \ldots, n\}$ as follows: $\sigma \cdot m := \sigma(m)$ for all $\sigma \in S_n$ and $m \in \{1, \ldots, n\}$. Note that the stabilizer of $m$ in $S_n$ is isomorphic to $S_{n-1}$.

2) Let $(G, *)$ be a group and $\mathcal{P}$ be the power set of $G$. Consider the action of $(G, *)$ on $\mathcal{P}$ as follows: $g \cdot A := gAg^{-1}$ for all $g \in G$ and $A \in \mathcal{P}$. Note that the stabilizer of $A$ in $G$ is $N_G(A)$.

**Proposition 2.12.11** Suppose a group $(G, *)$ acts on a set $A \neq \varnothing$. Let $a \in A$. Then $\mathrm{Orb}(a)$ contains exactly $[G : G_a]$ many elements.

*Proof.* Let $\mathcal{L}$ denote the set of left cosets of $G_a$ in $(G, *)$. We show that $\mathcal{L}$ and $\mathrm{Orb}(a)$ are in one-to-one correspondence. Define $f : \mathrm{Orb}(a) \to \mathcal{L}$ by $f(g \cdot a) = gG_a$. Note that,

$$\begin{aligned} & g_1 \cdot a = g_2 \cdot a \\ \iff & g_1^{-1} \cdot (g_2 \cdot a) = (g_1^{-1} * g_2) \cdot a = e \cdot a = a \end{aligned}$$

$$\iff g_1^{-1} * g_2 \in G_a$$
$$\iff g_1 G_a = g_2 G_a.$$

This shows $f$ is well-defined and injective. Clearly $f$ is surjective and hence the proof. As an application of this we prove the following theorem:

**Theorem 2.12.12** Every $\sigma \in S_n$ can be written as a composition of disjoint cycles.

*Proof.* Consider the cyclic subgroup of $S_n$ generated by $\sigma$. We call it $G = \langle \sigma \rangle$. We know that $(G, \circ)$ acts on $\{1, \ldots, n\}$ by $\sigma^i \cdot m = \sigma^i(m)$. So we can write $\{1, \ldots, n\}$ as disjoint union of orbits of elements from the set $\{1, \ldots, n\}$.

Let $\mathcal{O}$ be an orbit and $m \in \mathcal{O}$. We know that $|\mathcal{O}| = [G : G_m]$ and $\sigma^i \cdot m \to \sigma^i G_m$ gives us a bijective correspondence between $\mathcal{O}$ and the set of all left cosets of the stabilizer $G_m$ in $(G, \circ)$. Since $(G, \circ)$ is cyclic, $G/G_m$ is cyclic. Let $|G/G_m| = d$. So the distinct left cosets of $G_m$ in $(G, \circ)$ are $G_m, \sigma G_m, \ldots, \sigma^{d-1} G_m$. Hence $\mathcal{O} = \{m, \sigma \cdot m, \ldots, \sigma^{d-1} \cdot m\}$. This completes the proof.

### 2.12.2 Kernel of a group action

Suppose a group $(G, *)$ acts on a set $A \neq \varnothing$. Then by the **kernel of the group action** we refer to the following set:

$$\text{Kernel} := \{g \in G : g \cdot a = a \text{ for all } a \in A\}.$$

Clearly Kernel $\subseteq G_a$ for all $a \in A$.

**Exercise 2.12.13** Prove that Kernel is a subgroup of $(G, *)$.

**Example 2.12.14** Let $A$ be a subset of a group $(G, *)$. We consider the normalizer of $A$ in $G$ denoted by $N_G(A)$ and the action of $N_G(A)$ on $A$ as follows: $g \cdot a = g * a * g^{-1}$ for all $g \in N_G(A)$ and $a \in A$. Note that the Kernel of this action is $C_G(A)$, the centralizer of $A$ in $(G, *)$. In particular taking $A = G$ we get that the Kernel is $Z(G)$.

## 2.13 Automorphism group

An isomorphism from the group $(G, *)$ to $(G, *)$ is called an **automorphism of** $(G, *)$. We denote the set of all automorphisms of $(G, *)$ by $\text{Aut}(G)$.

**Exercise 2.13.1** Show that with the composition of functions $\mathrm{Aut}(G)$ is a group.

**Example 2.13.2** A primary set of examples of automorphisms of $(G, *)$ come from conjugation. Fix $g \in G$. Define $f_g : G \to G$ by $f_g(a) = g * a * g^{-1}$ for all $a \in G$. It can be checked that $f_g$ is an automorphism of $(G, *)$. This particular type of automorphisms are called **inner automorphisms** of $(G, *)$. We denote the set of all inner automorphisms of $(G, *)$ by $\mathrm{Inn}(G)$.

**Exercise 2.13.3** Show that $\mathrm{Inn}(G)$ is a subgroup of $\mathrm{Aut}(G)$ under composition of functions.

A priori, we may wonder about a possible bijective correspondence between $(G, *)$ and $(\mathrm{Inn}(G), \circ)$. We shall note that it happens only if $Z(G) = \{e\}$. Because in general

$$\mathrm{Inn}(G) \simeq G/Z(G).$$

And therefore $G \simeq \mathrm{Inn}(G)$ if $Z(G) = \{e\}$.

**Proposition 2.13.4** Let $H$ be a normal subgroup of $(G, *)$. Then the action of $(G, *)$ on $H$ by $g \cdot h = g * h * g^{-1}$ induces a homomorphism from $(G, *)$ to $(\mathrm{Aut}(H), \circ)$ with kernel $C_G(H)$.

*Proof.* Define $F : G \to \mathrm{Aut}(H)$ by $F(g) := f_g$, where $f_g : H \to H$ is defined by $f_g(h) = g * h * g^{-1}$. One can check that $F$ is a group homomorphism. Note that,

$$\begin{aligned}\mathrm{Ker}(F) &= \{g \in G : f_g = I\}, \text{ where } I(h) = h \text{ for all } h \in H \\ &= \{g \in G : g * h * g^{-1} = h \text{ for all } h \in H\} \\ &= C_G(H).\end{aligned}$$

**Remark 2.13.5** By the first isomorphism theorem we get from the above proposition that

$$G/C_G(H) \simeq \mathrm{Im}(F).$$

Now if we take $H = G$, we get $C_G(H) = Z(G)$ and $\mathrm{Im}(F) = \mathrm{Inn}(G)$. Hence

$$\mathrm{Inn}(G) \simeq G/Z(G).$$

**Remark 2.13.6** If we take $(G, *)$ an abelian group then from the above remark we can observe that $\text{Inn}(G)$ contains only the identity. Let us consider the abelian group $G = \mathbb{Z}/3\mathbb{Z}$. We can note that $\text{Aut}(\mathbb{Z}/3\mathbb{Z})$ contains exactly two maps;

$$0 \to 0,\ 1 \to 1,\ 2 \to 2 \text{ and } 0 \to 0,\ 1 \to 2,\ 2 \to 1.$$

So $\text{Aut}(\mathbb{Z}/3\mathbb{Z}) \simeq \mathbb{Z}/2\mathbb{Z}$. In this case we see that $\text{Aut}(\mathbb{Z}/3\mathbb{Z}) \neq \text{Inn}(\mathbb{Z}/3\mathbb{Z})$.

**Exercise 2.13.7** Prove that the automorphism group of the Klein's 4 - group is isomorphic to $S_3$.

## 2.14 Characteristic subgroups

Let $(G, *)$ be a group. A subgroup $H$ of $(G, *)$ is called a **characteristic subgroup** of $(G, *)$ if every automorphism of $(G, *)$ maps $H$ onto itself i.e. if $\sigma(H) = H$ for all $\sigma \in \text{Aut}(G)$.

Note that every characteristic subgroup of $(G, *)$ is normal as $\text{Inn}(G) \subseteq \text{Aut}(G)$. But the converse is not true i.e. every normal subgroup need not be a characteristic subgroup. For example in Klein's 4 - group the subgroup $\{e, a\}$ is normal but not characteristic.

**Exercise 2.14.1** Prove that if $H$ is the unique subgroup of $(G, *)$ of a given order, then $H$ is characteristic subgroup of $(G, *)$.

## 2.15 Sylow theorems and applications

In this section we present three theorems which are named after the Norwegian mathematician Peter Ludwig Sylow. The first theorem of Sylow is again a partial converse of Lagrange's theorem.

**Theorem 2.15.1 (Sylow's first theorem)** Let $(G, *)$ be a group and $p$ be a prime number such that $p^n \mid |G|$, where $n \in \mathbb{N}$. Then there exist subgroups of order $p^n$.

*Proof.* We shall prove it by induction on $|G|$.

Let $|G| = 1$. Then there is no prime $p$ and $n \in \mathbb{N}$ such that $p^n \mid |G|$. So the statement is vacuously true.

Now we assume the induction hypothesis i.e. for all groups of order strictly less than $|G|$, if any prime power $p^i$ where $i \in \mathbb{N}$ divides their order then they have subgroups of order $p^i$.

Now we consider the group $(G, *)$ and let $p^n \mid |G|$. We need to show that $(G, *)$ has a subgroup of order $p^n$. We consider two cases.
**Case 1:** Suppose $(G, *)$ has a proper subgroup $H$ such that $p^n \mid |H|$. Then by induction hypothesis $(H, *)$ has a subgroup of order $p^n$ and hence $(G, *)$ has a subgroup of order $p^n$.
**Case 2:** Suppose $(G, *)$ has no proper subgroup $H$ such that $p^n \mid |H|$. We write down the class equation of $(G, *)$:

$$|G| = |Z(G)| + \sum_{1 \leq j \leq k} \frac{|G|}{|C_G(a_j)|},$$

where for all $1 \leq j \leq k$, $a_j \notin Z(G)$ and they are representatives of all distinct conjugacy classes of the group $(G, *)$. Consider the subgroup $C_G(a_j)$ of $(G, *)$. As $a_j \notin Z(G)$, we have $C_G(a_j)$ is a proper subgroup of $(G, *)$. According to our assumption in this case $p^n \nmid |C_G(a_j)$. Therefore we have for each $1 \leq j \leq k$,

$$p \mid \frac{|G|}{|C_G(a_j)|}.$$

This implies that $p \mid |Z(G)|$. So by Cauchy's theorem $Z(G)$ has an element of order $p$. We call it $a$. Consider $H := \langle a \rangle$. Clearly $H$ is normal as $H \subseteq Z(G)$. Therefore we can consider the quotient group $(G/H, *)$. Now note that the order of $G/H$ is strictly less than $|G|$. Also $p^{n-1}$ divides $|G/H|$ as $p^n$ divides $|G|$ and $|H| = p$. Thus by induction hypothesis, $G/H$ has a subgroup say $T$ of order $p^{n-1}$. Now $T = K/H$ for some subgroup $K$ of $(G, *)$ containing $H$. Note that $|K| = |T||H| = p^n$. This proves that $(G, *)$ has a subgroup of order $p^i$.

**Definition 2.15.2 (Sylow subgroup)** Let $(G, *)$ be a group such that $|G| = p^n m$, where $p$ is a prime and $m, n \in \mathbb{N}$ be such that $\gcd(p, m) = 1$. Then the subgroups of $(G, *)$ of order $p^n$ are called the $p$ - **Sylow subgroups**.

Before we state Sylow's second theorem, we define the notion of double coset. Let $H, K$ be two subgroups of a group $(G, *)$. Let $x, y \in G$. We define a relation $\sim$ on $(G, *)$ as follows: define $x \sim y$ if and only if there exist an $h \in H$ and $k \in K$ such that $y = h * x * k$. One can check that $\sim$ is an equivalence relation. Note that the equivalence class of $x \in G$ is given by the set $HxK := \{h * x * k : h \in H, k \in K\}$. We call the set $HxK$, a **double coset** of $H, K$ in $(G, *)$. Now $G$ is the union of the distinct double cosets of $H, K$ in $(G, *)$ as we know that distinct equivalence classes form a partition on

$G$.

**Lemma 2.15.3** Let $H, K$ be two finite subgroups of a group $(G, *)$. Then for any $x \in G$,

$$|HxK| = \frac{|H||K|}{|H \cap xKx^{-1}|}.$$

*Proof.* Let $x \in G$. We define a map $f : HxK \to HxKx^{-1}$ by $f(h*x*k) := h*x*k*x^{-1}$ for all $h \in H, k \in K$.

First we note that $f$ is well-defined. Let $h * x * k = h' * x * k'$ for $h, h' \in H$ and $k, k' \in K$. Therefore, $h * x * k * x^{-1} = h' * x * k' * x^{-1}$ i.e. $f(h * x * k) = f(h' * x * k')$. Also $f(h * x * k) = f(h' * x * k')$, implies that $h * x * k = h' * x * k'$. Thus $f$ is injective. For $y \in HxKx^{-1}$, write $y = h * x * k * x^{-1}$ for some $h \in H, k \in K$. Then consider $h * x * k \in HxK$, so we get a pre-image of $y$. Thus $f$ is bijective. This proves that,

$$\begin{aligned} |HxK| &= |HxKx^{-1}| \\ &= \frac{|H||xKx^{-1}|}{|H \cap xKx^{-1}|} \\ &= \frac{|H||K|}{|H \cap xKx^{-1}|}. \end{aligned}$$

**Remark 2.15.4** From the first theorem of Sylow we know that $p$ - Sylow subgroups exist in a group whose order is divisible by $p$. Note that if $H$ is a $p$ - Sylow subgroup then $gHg^{-1}$ is also a $p$ - Sylow subgroup for each $g \in G$. So the collection of subgroups $\{gHg^{-1} : g \in G\}$ consists of $p$ - Sylow subgroups of $(G, *)$. The second theorem of Sylow tells us that they are all the $p$ - Sylow subgroups.

**Theorem 2.15.5 (Sylow's second theorem)** Let $H, K$ be two $p$ - Sylow subgroups of a group $(G, *)$ with $|G| = p^n m$ and $\gcd(p, m) = 1$. Then there exists $g \in G$ such that $K = gHg^{-1}$.

*Proof.* Consider the distinct double cosets of $H, K$ in $(G, *)$. Since $|G|$ is finite, the number of distinct double cosets are also finite. Therefore we can write,

$$|G| = \sum_{1 \le i \le r} |Hx_iK|, \text{ where } Hx_iK\text{'s are distinct double cosets.}$$

We want to show that there exists $g \in G$ such that $K = gHg^{-1}$. Suppose no such $g$

exists. Then $H \neq x_i K x_i^{-1}$ for all $1 \leq i \leq r$. Therefore $H \cap x_i K x_i^{-1} \subset H$ and hence $|H \cap x_i K x_i^{-1}| = p^{m_i}$ where $m_i < n$. So we have,

$$\begin{aligned} |G| = \sum_{1 \leq i \leq r} |H x_i K| &= \sum_{1 \leq i \leq r} \frac{|H||K|}{|H \cap x_i K x_i^{-1}|} \\ &= \sum_{1 \leq i \leq r} \frac{p^{2n}}{p^{m_i}} \\ &= \sum_{1 \leq i \leq r} p^{n+(n-m_i)}. \end{aligned}$$

This implies that $p^{n+1}$ divides $|G|$, a contradiction. So there exists $g \in G$ such that $K = gHg^{-1}$.

**Remark 2.15.6** For two subgroups $H, K$ of $(G, *)$ if there exists $g \in G$ such that $K = gHg^{-1}$, then $H, K$ are called **conjugates** of each other.

**Remark 2.15.7** The third theorem of Sylow determines the number of distinct $p$ - Sylow subgroups. Before we state and prove the theorem, we define the **normaliser** of a subgroup $H$ of $(G, *)$. It is denoted by $N(H)$ and defined as the set $N(H) := \{g \in G : gHg^{-1} = H\}$. One can check that $N(H)$ is a subgroup of $(G, *)$ and $H$ is a normal subgroup $(N(H), *)$.

**Theorem 2.15.8 (Sylow's third theorem)** Let $(G, *)$ be a group of order $p^n m$, where $p$ is a prime and $n, m \in \mathbb{N}$ with $\gcd(p, m) = 1$. Then the number of $p$ - Sylow subgroups of $(G, *)$, $n_p$ is of the form $1 + kp$ where $k \in \mathbb{N} \cup \{0\}$ and $n_p \mid |G|$.

*Proof.* Let $H$ be a $p$ - Sylow subgroup of $(G, *)$. From the second theorem of Sylow we know that all the $p$ - Sylow subgroups are of the form $gHg^{-1}$ for $g \in G$. First we claim that,

$$n_p = [G : N(H)].$$

Let $\mathcal{S}$ be the set of all distinct $p$ - Sylow subgroups of $(G, *)$ and $\mathcal{N}$ be the set of all distinct left cosets of $N(H)$ in $(G, *)$. Define, $f : \mathcal{S} \to \mathcal{N}$ by $f(gHg^{-1}) := gN(H)$. Note that,

$$\begin{aligned} & aHa^{-1} = bHb^{-1} \\ \iff & (a^{-1} * b)H(b^{-1} * a) = H \\ \iff & (a^{-1} * b)H(a^{-1} * b)^{-1} = H \end{aligned}$$

$$\iff a^{-1} * b \in N(H)$$
$$\iff aN(H) = bN(H).$$

Therefore the map $f$ is well-defined and injective. Also $f$ is onto as for any $gN(H) \in \mathcal{N}$, we consider $gHg^{-1} \in \mathcal{S}$ so that $f(gHg^{-1}) = gN(H)$. This proves that $f$ is bijective and hence $n_p = [G : N(H)]$. From here we get that $n_p \mid |G|$ as $n_p|N(H)| = |G|$.

Next we consider the double cosets of $H, H$ in $(G, *)$. Let $\mathcal{I}$ be the set of all $g \in G$ such that $HgH$ are the distinct double cosets of $H, H$. Therefore,

$$G = \bigcup_{g\in\mathcal{I}} HgH.$$

We can write,

$$|G| = \sum_{\substack{g\in\mathcal{I},\\ g\in N(H)}} |HgH| + \sum_{\substack{g\in\mathcal{I},\\ g\notin N(H)}} |HgH|.$$

Now for $g \notin N(H)$, we have $H \cap gHg^{-1} \subset H$, so $|H \cap gHg^{-1}| = p^r$ where $r < n$. Thus,

$$|HgH| = \frac{|H||H|}{|H \cap gHg^{-1}|} = p^{2n-r} = p^{n+(n-r)}.$$

Therefore we have,

$$\sum_{\substack{g\in\mathcal{I},\\ g\notin N(H)}} |HgH| = p^{n+1}u, \text{ where } u \in \mathbb{N} \cup \{0\}.$$

If the second sum is empty sum we get $u = 0$. Note that for $g \in N(H)$, $gH = Hg$, so $HgH = gH$. Therefore,

$$|G| = \sum_{\substack{g\in\mathcal{I}\\ g\in N(H)}} |gH| + p^{n+1}u.$$

Denote $\mathcal{L} := \{gH : g \in \mathcal{I}, g \in N(H)\}$. Let $g_1, g_2 \in N(H)$. Note that, $g_1H = g_2H$ if and only if $Hg_1H = Hg_2H$. This proves that $|\mathcal{L}| = [N(H) : H]$. So we have,

$$|G| = [N(H) : H]|H| + p^{n+1}u, \quad \text{as} \quad |gH| = |H|.$$

Therefore,

$$\frac{|G|}{|N(H)|} = 1 + \frac{p^{n+1}u}{|N(H)|} \in \mathbb{N}, \text{ as } |N(H)| \text{ divides } |G|.$$

So we have,

$$\frac{p^{n+1}u}{|N(H)|} \in \mathbb{N} \cup \{0\}.$$

Now $p^{n+1} \nmid |N(H)|$ as $N(H) \subseteq G$ and $p^{n+1} \nmid |G|$. This proves that,

$$p \mid \frac{p^{n+1}u}{|N(H)|}.$$

Therefore finally we have,

$$n_p = \frac{|G|}{|N(H)|} = 1 + kp \text{ for some } k \in \mathbb{N} \cup \{0\}.$$

### 2.15.1 Applications of Sylow's theorems

1) Let $H$ be a $p$ - Sylow subgroup of the group $(G, *)$. Then $n_p = 1$ if and only if $H$ is a normal subgroup.

   First suppose that $H$ is a normal $p$ - Sylow subgroup of $(G, *)$. Then $gHg^{-1} = H$ for all $g \in G$. By the second theorem of Sylow we know that all the other Sylow subgroups are conjugates of $H$, therefore we have only one $p$ - Sylow subgroup $H$, i.e. $n_p = 1$. Next suppose that $n_p = 1$. We are given $H$ is a $p$ - Sylow subgroup. As $n_p = 1$, we have $gHg^{-1} = H$ for all $g \in G$ by the aid of the second theorem of Sylow. This proves that $H$ is a normal subgroup of $(G, *)$.

2) Let $n_p = 1$ for all primes $p \mid |G|$. Then $(G, *)$ is the direct product of its Sylow subgroups.

   Let $|G| = n$. We write $n = p_1^{a_1} \cdots p_k^{a_k}$ where $p_i$'s are distinct primes and $a_i \in \mathbb{N}$. Given that $n_{p_i} = 1$ for all $1 \leq i \leq k$. We denote each $p_i$ - Sylow subgroup by $H_i$. As $n_{p_i} = 1$ for all $1 \leq i \leq k$, we have $H_i$'s are normal. We show that $G \simeq H_1 \times \cdots \times H_k$. Note that $H_1 \cdots H_k$ is a subgroup of $(G, *)$. Also $H_i \cap H_1 \cdots H_{i-1} H_{i+1} \cdots H_k = \{e\}$ as orders of $H_i$'s are mutually co-prime. Therefore $|H_1 \cdots H_k| = |G|$ and hence $H_1 \cdots H_k = G$. This prove us that $(G, *)$ is an internal direct product of $H_1, \ldots, H_k$. Thus $G \simeq H_1 \times \cdots \times H_k$.

3) Let $(G, *)$ be a group of order $pq$ where $p, q$ are primes with $p > q$. Also let $q \nmid (p - 1)$. Then $(G, *)$ is a cyclic group.

*Proof.* We know that $n_p$ can be $1, p, q$ or $pq$ as $n_p$ divides$|G|$. Also we know that $n_p = 1 + kp$ for some integer $k \geq 0$. Therefore $n_p$ is either 1 or $q$. Now $q \neq (1 + kp)$ for $k \geq 0$ as $q < p$. Therefore $n_p = 1$ and this implies that the $p$ - Sylow subgroup of $(G, *)$ say $H$ is normal.

Similarly $n_q$ can be $1, p, q$ or $pq$ as $n_q$ divides $|G|$. Also $n_q = 1 + k'q$ for some integer $k' \geq 0$. Therefore $n_q$ is either 1 or $p$. Now $p \neq (1 + k'q)$ for $k' \geq 1$ as $q \nmid (p - 1)$. Therefore $n_q = 1$ and this implies that the $q$ - Sylow subgroup of $(G, *)$ say $K$ is normal.

Since $H, K$ are of co-prime orders we also get that $H \cap K = \{e\}$. Thus $(G, *)$ is an internal direct product of $H, K$ and therefore $G \simeq H \times K$. Note that $H \simeq \mathbb{Z}/p\mathbb{Z}$ and $K \simeq \mathbb{Z}/q\mathbb{Z}$ as every group of prime order is cyclic and up to isomorphism there is only one group of a fixed finite order. As $p, q$ are co-prime, we have,

$$G \simeq H \times K \simeq \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z} \simeq \mathbb{Z}/pq\mathbb{Z}.$$

This completes the proof.

**Remark 2.15.9** The last application plays a very important in the classification of finite abelian groups. For example, it tells us that every group of order 15 is cyclic.

**Remark 2.15.10** If in the last application we take $q \mid (p - 1)$, then $n_q$ can be 1 or $p$. If $n_q = 1$ then again we get the group to be cyclic. But if $n_q = p$, then there exists a unique non-abelian group of order $pq$. To find such non-abelian example we need to study semi-direct product.

**Theorem 2.15.11** Every group of order 12 has either a normal 3 - Sylow subgroup or it is isomorphic to the alternating group of 4 symbols $A_4$.

*Proof.* Let $(G, *)$ be a group of order 12. We have $|G| = 2^2 \cdot 3$. So the possible choices for $n_3$ are $1, 2, 3, 4, 6, 12$. If $n_3 = 1$, then $(G, *)$ has a normal 3 - Sylow subgroup. So let $n_3 > 1$. As $n_3 = 1 + 3k$ for some $k \in \mathbb{N}$, then $n_3$ can only be 4. Let $X$ be the set of all distinct 3 - Sylow subgroups of $(G, *)$. We denote the distinct 3 - Sylow subgroups by $P_1, P_2, P_3, P_4$. First we note that the normaliser of $P_i$, $N(P_i) = P_i$ for all $1 \leq i \leq 4$. We have seen that $[G : N(P_i)] = n_3 = 4$. So $|N(P_i)| = 3$. As $N(P_i) \subseteq P_i$ and $|N(P_i)| = 3 = |P_i|$, we obtain that $N(P_i) = P_i$. We define $\sigma_g : X \to X$ by $\sigma_g(P_i) := gP_ig^{-1}$ for all $1 \leq i \leq 4$. Clearly $\sigma_g$ is a bijection as $gP_ig^{-1} = gP_jg^{-1}$ implies

that $P_i = P_j$, which is not true. Now we define a map $\phi : G \to S_4$ by $\phi(g) = \sigma_g$. We can check that $\phi(g * g') = \sigma_{g*g'} = \sigma_g \circ \sigma'_g$. Thus $\phi$ is a group homomorphism. Now,

$$\begin{aligned}
\text{Ker}\phi &= \{g \in G : gP_ig^{-1} = P_i \text{ for all } 1 \leq i \leq 4\} \\
&= \{g \in G : g \in N(P_i) \text{ for all } 1 \leq i \leq 4\} \\
&= \{g \in G : g \in P_i \text{ for all } 1 \leq i \leq 4\}, \text{ as } N(P_i) = P_i \text{ for all } 1 \leq i \leq 4. \\
&= \{e\}, \text{ as } P_i\text{'s are all distinct.}
\end{aligned}$$

Because $|P_i \cap P_j|$ can be 1 or 3. If $|P_i \cap P_j| = 3$, then it turns out that $P_i = P_j$. This proves that $\phi$ is injective. Now we show that $\phi$ is onto. Each $P_i$ contains 2 elements of order 3, therefore $(G, *)$ has exactly 8 elements of order 3. Since $\phi$ is injective, there are exactly 8 elements of order 3. In $S_4$, an element has order 3 if and only if it is a 3 - cycle. Now in $S_4$, there are total 8 many 3 - cycles and 3 - cycles are all even permutation. So $|\text{Im}\phi \cap A_4| \geq 8$. Also $|A_4| = 12$ and $|\text{Im}\phi \cap A_4| \mid |A_4|$. Therefore $\text{Im}\phi \cap A_4 = A_4$. This implies that $A_4 \subseteq \text{Im}\phi$. Again,

$$|\text{Im}\phi| \leq |G| = 12 = |A_4|.$$

Hence $\text{Im}\phi = A_4$. This proves that $G \simeq A_4$.

**Theorem 2.15.12** Let $(G, *)$ be a group of order $p^2q$ where $p, q$ are distinct primes. Then $(G, *)$ has a normal Sylow subgroup.

*Proof.* The possible choices for $n_p$ are $1, p, p^2, q, pq, p^2q$. Also as $n_p = 1+kp$ for some integer $k \geq 0$, $n_p$ can be either 1 or $q$. Similarly the possible choices for $n_q$ are $1, p, p^2, q, pq, p^2q$ and as $n_q = 1 + k'q$ for some integer $k' \geq 0$, $n_q$ can be $1, p$ or $p^2$.
Now since $p, q$ are distinct primes, either $p < q$ or $q < p$.
**Case I:** Let $p > q$. Then $n_p = 1 + kp \neq q$. So $n_p = 1$ and hence $(G, *)$ has a normal $p$ - Sylow subgroup.
**Case II:** Let $p < q$. In this case if $n_q = 1$, then $(G, *)$ has a normal $q$ - Sylow subgroup. Let $n_q > 1$. Now $n_q = 1 + k'q \neq p$ as $p < q$. So let $n_q = p^2$. Therefore, $k'q = p^2 - 1 = (p - 1)(p + 1)$. As $q$ is a prime number, we have $q \mid (p - 1)$ or $q \mid (p + 1)$. Note that $q \nmid (p - 1)$ as $q > p$. So $q \mid (p + 1)$. Therefore we have,

$$p < q \leq (p + 1).$$

This is possible only if $p = 2, q = 3$. This implies that $(G, *)$ is a group of order 12.

Now as $(G, *)$ is a group of order 12 either $(G, *)$ has a normal 3 - Sylow subgroup or $G \simeq A_4$. If $G \simeq A_4$, then note that $\{e, (1\ 2)(3\ 4), (1\ 3)(2\ 4), (1\ 4)(2\ 3)\}$ is a normal 2-Sylow subgroup of $A_4$.

## 2.16 Structure theorem of finite abelian groups

For a given $n \in \mathbb{N}$, we know there are only finitely many groups of order $n$ up to isomorphism as every group of order $n$ is isomorphic to some subgroup of $S_n$ and there are only finitely many subgroups of $S_n$ of order $n$. The structure theorem of finite abelian groups helps us to determine the number of abelian groups of order $n$ up to isomorphism.

**Theorem 2.16.1 (Structure theorem of finite abelian groups)** Any finite abelian group can be written as a direct product of finite cyclic groups.

For a given $n \in \mathbb{N}$, by writing down the prime factorisation of $n$, we can list down all the finite abelian groups of order $n$ up to isomorphism. Let

$$n = p_1^{a_1} \cdots p_r^{a_r},$$

where $p_i$'s are distinct prime numbers and $a_i \in \mathbb{N}$ for all $1 \leq i \leq r$. If we denote the partition of the exponent $a_i$ by $P(a_i)$ for all $1 \leq i \leq r$, then up to isomorphism there are $P(a_1) \cdots P(a_r)$ many groups of order $n$. We illustrate a few examples in the following table:

| $n$ | All abelian groups of order $n$ (up to isomorphism) | Number of groups |
|---|---|---|
| $2 = 2$ | $\mathbb{Z}/2\mathbb{Z}$ | $P(1) = 1$ |
| $3 = 3$ | $\mathbb{Z}/3\mathbb{Z}$ | $P(1) = 1$ |
| $4 = 2^2$ | $\mathbb{Z}/4\mathbb{Z}$, $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ | $P(2) = 2$ |
| $5 = 5$ | $\mathbb{Z}/5\mathbb{Z}$ | $P(1) = 1$ |
| $6 = 2 \times 3$ | $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \simeq \mathbb{Z}/6\mathbb{Z}$ | $P(1)P(1) = 1$ |
| $7 = 7$ | $\mathbb{Z}/7\mathbb{Z}$ | $P(1) = 1$ |
| $8 = 2^3$ | $\mathbb{Z}/8\mathbb{Z}$, $\mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ | $P(3) = 3$ |
| $9 = 3^2$ | $\mathbb{Z}/9\mathbb{Z}$, $\mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ | $P(2) = 2$ |
| $10 = 2 \times 5$ | $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z} \simeq \mathbb{Z}/10\mathbb{Z}$ | $P(1)P(1) = 1$ |
| $11$ | $\mathbb{Z}/11\mathbb{Z}$ | $P(1) = 1$ |
| $12 = 2^2 \times 3$ | $\mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \simeq \mathbb{Z}/12\mathbb{Z}$, $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \simeq \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/6\mathbb{Z}$ | $P(2)P(1) = 2$ |
| $13$ | $\mathbb{Z}/13\mathbb{Z}$ | $P(1) = 1$ |
| $14 = 2 \times 7$ | $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/7\mathbb{Z} \simeq \mathbb{Z}/14\mathbb{Z}$ | $P(1)P(1) = 1$ |
| $15 = 3 \times 5$ | $\mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z} \simeq \mathbb{Z}/15\mathbb{Z}$ | $P(1) = 1$ |
| $16 = 2^4$ | $\mathbb{Z}/16\mathbb{Z}$, $\mathbb{Z}/8\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, $\mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$, $\mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ | $P(4) = 5$ |

In this context, let us define what do we mean by a finitely generated group.

**Definition 2.16.2** A group $(G, *)$ is called **finitely generated** if there exists a finite set $A \subseteq G$ such that every element of $G$ can be written as a composition of the elements of the set $A$.

We know that all cyclic groups are generated by a single element, hence they are finitely generated. Other than cyclic groups we can consider the group $(\mathbb{Z} \times \mathbb{Z}, +)$. This group is finitely generated as this group is generated by the set $\{(1, 0), (0, 1)\}$. We also have structure theorem for finitely generated abelian groups, which states that **a finitely generated abelian group is isomorphic to $\mathbb{Z}^r \times K$ where $r \geq 0$ and $(K, *)$ is a finite abelian group**. So if $(G, *)$ is a finitely generated abelian group, then there exists an $r \in \mathbb{N} \cup \{0\}$ and $n_i \in \mathbb{N}$ for all $1 \leq i \leq t$ such that

$$G \simeq \mathbb{Z}^r \times \mathbb{Z}/n_1\mathbb{Z} \times \cdots \mathbb{Z}/n_t\mathbb{Z}.$$

### 2.16.1 Applications of the structure theorem of finite abelian groups

1) Up to isomorphism there are only two abelian groups of order 12. By structure theorem of finite abelian groups we have up to isomorphism there are only two groups of order 12 namely

$$\mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z} \simeq \mathbb{Z}/12\mathbb{Z} \quad \text{and} \quad \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \simeq \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/6\mathbb{Z}.$$

2) Up to isomorphism there is only one abelian group of order 6 namely

$$\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \simeq \mathbb{Z}/6\mathbb{Z} \text{ as } \gcd(2,3) = 1.$$

3) We now know that every group of order $2^2 = 4$ is abelian. From the structure theorem of finite abelian groups we know that up to isomorphism there are only two abelian groups of order 4. Therefore up to isomorphism there are only two groups of order 4.

### 2.16.2 Proof of the structure theorem of finite abelian groups

In an abelian group, every subgroup is normal. So all the Sylow subgroups are normal. Let $(G, *)$ be an abelian group of order $n = p_1^{a_1} \cdots p_k^{a_k}$ where $p_i$'s are distinct primes and $a_i \in \mathbb{N}$. Let $H_i$ denote the $p_i$ - Sylow subgroup of $(G, *)$ for all $1 \leq i \leq k$. Therefore we can write

$$G \simeq H_1 \times \cdots \times H_k.$$

So now if we can show that each $H_i$ can be written as a product of finite cyclic groups, then we are done. Therefore suffices to prove the structure theorem for a finite abelian group $(G, *)$ of order $p^n$ for some prime $p$ and $n \in \mathbb{N}$. We choose $a_1 \in G$ such that $o(a_1)$ is maximal. Say $o(a_1) = p^{n_1}$ where $1 \leq n_1 \leq n$. Consider $A_1 := \langle a_1 \rangle$. As $(G, *)$ is abelian, we have $A_1$ is a normal subgroup. So we can consider the quotient group $(G/A_1, *)$. We choose $b_2 \in G$ such that $o(b_2A_1)$ is maximal in $(G/A_1, *)$. Say $o(b_2A_1) = p^{n_2}$. We know that $f : G \to G/A_1$ defined by $f(g) := gA_1$ for all $g \in G$ is a group homomorphism. Therefore, $o(b_2A_1) \mid o(b_2)$. So $o(b_2A_1) \leq o(b_2)$. Since $p^{n_1}$ is the maximal order in $(G, *)$, we have $p^{n_2} \leq p^{n_1}$ i.e. $n_2 \leq n_1$. Note that since $o(b_2A_1) = p^{n_2}$, so $p^{n_2}$ is the first power

of $b_2$ that falls into $A_1$. Now we can write,

$$b_2^{p^{n_2}} = a_1^i \text{ for some } i \in \mathbb{Z}.$$

Thus,

$$b_2^{p^{n_1}} = a_1^{i\,p^{n_1-n_2}}.$$

Also $b_2^{p^{n_1}} = e$ as $p^{n_1}$ is the maximum of orders of all elements of the group $(G, *)$. Therefore we get that $a_1^{i\,p^{n_1-n_2}} = e$. Hence $p^{n_1} \mid ip^{n_1-n_2}$ i.e. $p^{n_2} \mid i$. Write $i = jp^{n_2}$ for some $j \in \mathbb{Z}$. So $b_2^{p^{n_1}} = a_1^i = a_1^{jp^{n_2}}$ i.e. $(b_2 a_1^{-j})^{p^{n_2}} = e$. We set $a_2 := b_2 a_1^{-j} \in G$. Note that $o(a_2) = p^{n_2}$. Then consider $A_2 := \langle a_2 \rangle$. We show that $A_1 \cap A_2 = \{e\}$. Let $a_2^t \in A_1$. Then we have $b_2^t a_1^{-jt} \in A_1$. This implies that $b_2^t \in A_1$. So $b_2^t A_1 = A_1$. Since $o(b_2 A_1) = p^{n_2}$, we get that $p^{n_2} \mid t$. Therefore, $a_2^t = a_2^{rp^{n_2}} = e$ where $r \in \mathbb{Z}$. This proves that $A_1 \cap A_2 = \{e\}$.

Now note that $A_1 A_2$ is a normal subgroup of $(G, *)$. We then choose $b_3 \in G$ so that the order of $b_3 A_1 A_2$ is maximal in $G/A_1 A_2$. Say $o(b_3 A_1 A_2) = p^{n_3}$. We first claim that $n_3 \leq n_2 \leq n_1$. Since we already have $n_2 \leq n_1$, it is enough to show that $n_3 \leq n_2$. Since $b_2 \in G$ is such that $o(b_2 A_1) = p^{n_2}$ is maximal in $(G/A_1, *)$, we get $(b_3 A_1)^{p^{n_2}} = A_1$. So,

$$b_3^{p^{n_2}} \in A_1 \subset A_1 A_2.$$

This gives us that,

$$o(b_3 A_1 A_2) \leq p^{n_2}, \text{ i.e. } p^{n_3} \leq p^{n_2}.$$

Hence $n_3 \leq n_2$. Now $b_3^{p^{n_3}} \in A_1 A_2$, so we can write,

$$b_3^{p^{n_3}} = a_1^{i_1} a_2^{i_2} \text{ for some integers } i_1, i_2.$$

As $(b_3)^{p^{n_2}} \in A_1$, we get

$$(a_1^{i_1} a_2^{i_2})^{p^{n_2-n_3}} = ((b_3)^{p^{n_2}})^{p^{n_2-n_3}} = (b_3)^{p^{n_2}} \in A_1.$$

Thus $(a_2^{i_2})^{p^{n_2-n_3}} \in A_1$. Since $o(a_2 A) = p^{n_2}$, we have $p^{n_2} \mid i_2 p^{n_2-n_3}$, i.e. $p^{n_3} \mid i_2$. Also $b_3^{p^{n_1}} = e$. Hence,

$$(a_1^{i_1} a_2^{i_2})^{p^{n_1-n_3}} = b_3^{p^{n_1}} = e.$$

Hence $a_1^{i_1 p^{n_1-n_3}} \in A_1 \cap A_2$ and we have $A_1 \cap A_2 = \{e\}$. Therefore,

$$p^{n_1} \mid i_1 p^{n_1-n_3} \text{ i.e. } p^{n_3} \mid i_1.$$

Write $i_1 = j_1 p^{n_3}, i_2 = j_2 p^{n_3}$. Set $a_3 := b_3 a_1^{-j_1} a_2^{-j_2}$, so that $a_3^{p^{n_3}} = e$, moreover $o(a_3) = p^{n_3}$. Set $A_3 := \langle a_3 \rangle$. We now show that $A_3 \cap A_1 A_2 = \{e\}$. Let $a_3^t \in A_1 A_2$, then

$$(b_3 a_1^{-j_1} a_2^{-j_2})^t \in A_1 A_2.$$

Hence $b_3^t \in A_1 A_2$. Therefore $p^{n_3} | t$. Since $a_3^{p^{n_3}} = e$, we get $a_3^t = e$. This proves that $A_3 \cap A_1 A_2 = \{e\}$.

We repeat the process with the normal subgroup $A_1 A_2 A_3$. Since $(G, *)$ is a finite group, we finally end up getting cyclic subgroups $A_i = \langle a_i \rangle$ of order $p^{n_i}$ for $1 \leq i \leq s$, with $n_1 \geq n_2 \geq \cdots \geq n_s$ such that $G = A_1 A_2 \cdots A_s$ and for every $2 \leq i \leq s$, $A_i \cap A_1 \cdots A_{i-1} = \{e\}$. This proves that $(G, *)$ is an internal direct product of the cyclic subgroups $A_1, \ldots, A_s$. Therefore $G \simeq A_1 \times \cdots \times A_s$.

CHAPTER 3

# Ring theory

## 3.1 Definition and examples

**Definition 3.1.1** A **Ring** is an ordered triple $(R, +, \cdot)$ where $R$ is a non-empty set and '$+, \cdot$' are two binary operations on $R$ such that the following hold:

- $(R, +)$ is an abelian group.
- $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ for all $a, b, c \in R$.
- Distributive property: $a \cdot (b + c) = a \cdot b + a \cdot c$ and $(a + b) \cdot c = a \cdot c + b \cdot c$ for all $a, b, c \in R$.

**Remark 3.1.2** The notations '$+, \cdot$' are standardly used to denote the first and the second operations of a ring respectively. The identity element with respect to the binary operation '$+$' is denoted by 0. The first operation is called 'addition' and the second operation is called 'multiplication'.

**Definition 3.1.3** A ring $(R, +, \cdot)$ is called a **commutative ring** if $a \cdot b = b \cdot a$ for all $a, b \in R$.

**Definition 3.1.4** A ring $(R, +, \cdot)$ is said to have **unity** if there exists an element $1 \in R$ such that $1 \cdot a = a = a \cdot 1$ for all $a \in R$.

**Remark 3.1.5** In the above definition the notation 1 is standardly used to denote the identity element with respect to the operation '$\cdot$'.

**Exercise 3.1.6** Let $(R, +, \cdot)$ be a ring. Show that the following properties hold:

1. $0 \cdot a = 0 = a \cdot 0$ for all $a \in R$.

2. $-(a \cdot b) = (-a) \cdot b = a \cdot (-b)$ for all $a, b \in R$.

3. $n(a \cdot b) = (na) \cdot b = a \cdot (nb)$, for all $a, b \in R$ and $n \in \mathbb{Z}$.

4. $(mn)a = m(na) = n(ma)$, for all $n, m \in \mathbb{Z}$ and $a \in R$, where

$$na = \begin{cases} \underbrace{(a + \cdots + a)}_{n \text{ times}} & \text{if } n \geq 0, \\ -\underbrace{(a + \cdots + a)}_{-n \text{ times}} & \text{otherwise.} \end{cases}$$

**Remark 3.1.7** Let $(R, +, \cdot)$ be a ring with unity. Then $R = \{0\}$ if and only if $1 = 0$.

**Examples 3.1.8** We give some examples of rings:

- With respect to usual addition and multiplication of numbers, $(\mathbb{Z}, +, \cdot)$, $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$ and $(\mathbb{C}, +, \cdot)$ are commutative rings with unity.

- For $n \in \mathbb{N}$, consider the set $\mathbb{Z}/n\mathbb{Z}$, i.e. the set of congruence classes modulo $n$. We have seen that under the operation '$+$' defined as $[a] + [b] := [a + b]$, $\mathbb{Z}/n\mathbb{Z}$ is an abelian group. Define the operation '$\cdot$' on the congruence classes modulo $n$ as follows: $[a] \cdot [b] := [a \cdot b]$, where the operation $\cdot$ in the right hand side denotes the usual multiplication of integers. Then $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ is a commutative ring with unity.

- Let us denote the set of all continuous function from $[0, 1]$ to $\mathbb{R}$ by $\mathcal{C}([0, 1], \mathbb{R})$. We define the following operations on $\mathcal{C}([0, 1], \mathbb{R})$: Let $f, g \in \mathcal{C}([0, 1], \mathbb{R})$. Then

$$(f + g)(x) := f(x) + g(x), \text{ and } (f \cdot g)(x) := f(x) \cdot g(x),$$

  where the operations on the right hand side of both the equations are usual addition and multiplication in $\mathbb{R}$ respectively. One can check that $\mathcal{C}([0, 1], \mathbb{R})$ with respect to the operations defined above is a commutative ring with unity.

- Consider the set of all $2 \times 2$ matrices with entries from $\mathbb{Z}$ i.e. the set $M_2(\mathbb{Z})$. With respect to matrix addition and multiplication $(M_2(\mathbb{Z}), +, \cdot)$ is a non-commutative ring with unity.

- The set of even numbers with respect to usual addition and multiplication of numbers i.e. $(2\mathbb{Z}, +, \cdot)$ is a commutative ring without unity. In fact for each integer $n > 1$, $(n\mathbb{Z}, +, \cdot)$ is a commutative ring without unity.
- Let $(G, +)$ be an abelian group. We define the operation '$\cdot$' on $G$ as folllows: $a \cdot b = 0$ for all $a, b \in G$. Then $(G, +, \cdot)$ is a commutative ring without unity. This operation '$\cdot$' is called the **trivial multiplication** on $G$.
- The set $(M_2(2\mathbb{Z}), +, \cdot)$ is a non-commutative ring without unity.
- Let us consider the set
$$R = \left\{ \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix} : a, b \in \mathbb{Z} \right\}.$$
The set $R$ with respect to matrix addition and multiplication forms a non-commutative ring without unity.

**Definition 3.1.9** Let $(R, +, \cdot)$ be a ring with unity. An element $a \in R$ is called a **unit** if there exists an element $b \in R$ such that $a \cdot b = 1 = b \cdot a$.

**Example 3.1.10** The elements $1, -1$ are units in the ring $(\mathbb{Z}, +, \cdot)$.

**Exercise 3.1.11** Show that in the ring $(\mathbb{Z}, +, \cdot)$, an element $a$ is a unit if and only if $a = 1$ or $a = -1$.

**Exercise 3.1.12** Let $n \in \mathbb{N}$. Show that in the ring $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$, an element $[a]$ is a unit if and only if $\gcd(a, n) = 1$.

**Remarks 3.1.13**

- Addition of two units need not be a unit. For example $1, -1$ are units in $(\mathbb{Z}, +, \cdot)$. But their sum is 0, which is not a unit in $(\mathbb{Z}, +, \cdot)$.
- Multiplication of two unit is always a unit. Let $a, b$ be two units in a ring $(R, +, \cdot)$. Then there exist $a', b' \in R$ such that
$$a \cdot a' = 1 = a' \cdot a \quad \text{and} \quad b \cdot b' = 1 = b' \cdot b.$$
Note that,
$$(a \cdot b) \cdot (b' \cdot a') = 1 = (b' \cdot a') \cdot (a \cdot b).$$
This proves that $(a \cdot b)$ is a unit.

**Exercise 3.1.14** Let $(R, +, \cdot)$ be a ring with unity. Denote the set of units of $(R, +, \cdot)$ by $U(R)$. Check that $(U(R), \cdot)$ is a group.

## 3.2 Zero divisors

**Definition 3.2.1** Let $(R, +, \cdot)$ be a ring. An element $z \in R$ is said to be a **left zero divisor** if there exists an element $v \neq 0$ such that $z \cdot v = 0$ and a **right zero divisor** if $v \cdot z = 0$. We call an element $z \in R$, a **zero divisor** if $z$ is a left zero divisor or a right zero divisor.

**Examples 3.2.2**

- The element 0 is always a zero divisor in every non-zero ring.
- In $(\mathbb{Z}/4\mathbb{Z}, +, \cdot)$, the element $[2]$ is a zero divisor.
- In $(\mathbb{Z}/6\mathbb{Z}, +, \cdot)$, the elements $[2], [3]$ are zero divisors.
- In $\mathcal{C}([0,1], \mathbb{R})$, consider the elements

$$f(x) := \begin{cases} (1-2x) & \text{if } x \in [0, \frac{1}{2}], \\ 0 & \text{if } x \in [\frac{1}{2}, 1], \end{cases}$$

and

$$g(x) := \begin{cases} 0 & \text{if } x \in [0, \frac{1}{2}], \\ (2x-1) & \text{if } x \in [\frac{1}{2}, 1]. \end{cases}$$

Note that $f, g$ are zero divisors.

- In $(M_2(\mathbb{Z}), +, \cdot)$, consider the elements

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Note that both are non-zero matrices and

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Therefore they are zero divisors.

**Remarks 3.2.3**

- A unit can never be a zero divisor and vice-versa. Let $u$ be a unit and also a zero divisor. Then there exists $v \neq 0$ such that $u \cdot v = 0$. Note that,

$$0 = u^{-1} \cdot 0 = u^{-1} \cdot (u \cdot v) = (u^{-1} \cdot u) \cdot v = 1 \cdot v = v.$$

  A contradiction.

- Addition of two zero divisors need not be a zero divisor. For example, in $(\mathbb{Z}/6\mathbb{Z}, +, \cdot)$, we know $[2], [3]$ are zero divisors but $[5]$ is not a zero divisor. In fact $[5]$ is a unit in $(\mathbb{Z}/6\mathbb{Z}, +, \cdot)$.

- Let $a, b$ be two zero divisors. Let $b' \in R \setminus \{0\}$ be such that $b \cdot b' = 0$. Then,

$$(a \cdot b) \cdot (b') = a \cdot (b \cdot b') = a \cdot 0 = 0,$$

  i.e. $a \cdot b$ is a zero divisor. If $b' \cdot b = 0$, then $b \cdot a$ is a zero divisor. If there exists $a' \neq 0$ such that $a \cdot a' = 0$, then $b \cdot a$ is a zero divisor. If $a' \cdot a = 0$, then $a \cdot b$ is a zero divisor. So if we have that the ring $(R, +, \cdot)$ is commutative then $a \cdot b, b \cdot a$ both are always zero divisors.

## 3.3 Idempotent elements in a ring

**Definition 3.3.1** Let $(R, +, \cdot)$ be a ring. An element $a \in R$ is called **idempotent** if $a^2 := a \cdot a = a$.

**Remark 3.3.2** The additive identity 0 is always an idempotent in every ring.

**Examples 3.3.3**

- In the ring $(\mathbb{Z}, +, \cdot)$, the elements $0, 1$ are the only idempotent elements.
- In $(\mathbb{Z}/4\mathbb{Z}, +, \cdot)$, the elements $[0], [1]$ are only idempotent.
- In the ring $(\mathbb{Z}/6\mathbb{Z}, +, \cdot)$, the elements $[0], [1], [3], [4]$ are all the idempotent elements.

**Remarks 3.3.4**

- Sum of two idempotent elements need not be an idempotent element. For example, in $(\mathbb{Z}/6\mathbb{Z}, +, \cdot)$, the elements $[1], [4]$ are idempotent elements. But $[1] + [4] = [5]$ is not an idempotent.

- In a commutative ring, product of two idempotent elements is always an idempotent. Let $a, b$ be two idempotent elements in a commutative ring $(R, +, \cdot)$. Then,

$$(a \cdot b)^2 = a^2 \cdot b^2 = a \cdot b.$$

- If the ring is not commutative, then product of two idempotent elements need not be an idempotent. For example, in $(M_2(\mathbb{Z}), +, \cdot)$, consider the elements

$$\begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

  Note that both are idempotent but

$$\begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

  is not idempotent.

- In a ring $(R, +, \cdot)$ with unity, for an idempotent $a$, the element $1 - a$ is always an idempotent. Note that,

$$(1-a)^2 = (1-a) \cdot (1-a) = 1 - a \cdot 1 - 1 \cdot a + a^2 = 1 - a - a + a = (1-a).$$

## 3.4 Nilpotent elements in a ring

**Definition 3.4.1** Let $(R, +, \cdot)$ be a ring. An element $a \in R$ is called **nilpotent** if there exists $n \in \mathbb{N}$ such that $a^n := \underbrace{a \cdot \cdots \cdot a}_{n \text{ times}} = 0$. The least positive integer $n$ so that $a^n = 0$, is called the **degree of nilpotency** of $a$.

**Remark 3.4.2** The additive identity 0 is always a nilpotent in every ring.

**Example 3.4.3** In $(\mathbb{Z}/4\mathbb{Z}, +, \cdot)$, the element $[2]$ is nilpotent as $[2]^2 = [4] = [0]$ and the degree of nilpotency of $[2]$ is 2.

**Exercise 3.4.4** Let $(R, +, \cdot)$ be a ring and $a \in R$ be an idempotent as well as nilpotent. Show that $a = 0$.

So the above exercise tells us that in a ring the intersection of the set of idempotent elements and the set of nilpotent elements is $\{0\}$.

**Remarks 3.4.5**

- Sum of two nilpotent elements need not be a nilpotent element. For example, in $(M_2(\mathbb{Z}), +, \cdot)$, consider the elements

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

  Note that both are nilpotent but

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

  is not nilpotent.

- In a commutative ring, product of two nilpotent elements is always a nilpotent. Let $a, b \in R$ be such that $a^m = 0$ and $b^n = 0$. Let $\text{lcm}(m, n) = d$. Then,

$$(a \cdot b)^d = a^d \cdot b^d = 0.$$

- If the ring is not commutative, then product of two nilpotent elements need not be a nilpotent. For example, in $(M_2(\mathbb{Z}), +, \cdot)$, consider the elements

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

  Note that both are nilpotent but

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

  is not nilpotent.

- In a ring $(R, +, \cdot)$ with unity, for a nilpotent $a$, the element $1 - a$ is always a unit. Let $n$ be the least positive integer so that $a^n = 0$. Note that,

$$(1 - a) \cdot (1 + a + \cdots + a^{n-1}) = 1 - a^n = (1 + a + \cdots + a^{n-1}) \cdot (1 - a).$$

  We have $1 - a^n = 1$. So the element $(1 + a + \cdots + a^{n-1}) \in R$ is the inverse of $(1 - a) \in R$.

  Note that $1 + a$ is also a unit as we can write $1 + a = 1 - (-a)$ and $(-a)$ is also nilpotent if $a$ is nilpotent.

- In a commutative ring with unity, sum of a unit and a nilpotent is always a unit.

  Let $u \in R$ be a unit and $a^n = 0$ for some $n \in \mathbb{N}$. Then $u + a$ is a unit. We can write $u + a = u \cdot (1 + u^{-1} \cdot a)$. Now,

  $$(u^{-1} \cdot a)^n = (u^{-1})^n \cdot a^n = (u^{-1})^n \cdot 0 = 0, \text{ as the ring is commutative.}$$

  Therefore $(1 + u^{-1} \cdot a)$ is a unit and hence $u \cdot (1 + u^{-1} \cdot a) = u + a$ is a unit.

## 3.5 Subrings

**Definition 3.5.1** Let $(R, +, \cdot)$ be a ring. A non-empty subset $S$ of $R$ is called a **subring** of $(R, +, \cdot)$ if $(S, +, \cdot)$ itself is a ring.

**Remark 3.5.2** Note that to check a subset $S$ of a ring $(R, +, \cdot)$ is a subring or not we need to check only the following:

1. $S \neq \varnothing$,
2. For $a, b \in S$, $a - b \in S$,
3. For $a, b \in S$, $a \cdot b \in S$.

The other properties will be inherited to $S$ from the ring $(R, +, \cdot)$.

**Examples 3.5.3**

- In a ring $(R, +, \cdot)$, we always have the subrings $(\{0\}, +, \cdot)$ and $(R, +, \cdot)$. These two subrings are called the **trivial subrings**.
- Under standard addition '+' and multiplication '$\cdot$' ,
  1. The set of integers $\mathbb{Z}$ is a subring of the set of rational numbers $\mathbb{Q}$.
  2. The set of rational numbers $\mathbb{Q}$ is a subring of the set of real numbers $\mathbb{R}$.
- Note that if $S$ is a subring of $R$ and $R$ is a subring of $R'$ under the same operations, then $S$ is a subring of $R'$. So $\mathbb{Z}$ is a subring of $\mathbb{R}$ too.

**Definition 3.5.4** Let $(R, +, \cdot)$ be a ring and define

$$Z(R) := \{a \in R : a \cdot x = x \cdot a \text{ for all } x \in R\}.$$

The set $Z(R)$ is called the **center** of the ring $(R, +, \cdot)$.

**Remark 3.5.5** If $(R, +, \cdot)$ is commutative then $Z(R) = R$.

Note that $Z(R) \neq \varnothing$ as $0 \in Z(R)$.

**Exercise 3.5.6** Show that $Z(R)$ is a subring of $(R, +, \cdot)$.

**Remarks 3.5.7**

- If $(R, +, \cdot)$ is commutative, then a subring $S$ of $(R, +, \cdot)$ is always commutative. But a subring $S$ of $(R, +, \cdot)$ may be commutative but $(R, +, \cdot)$ need not be commutative. For example, consider $(R, +, \cdot)$ as the ring $(M_2(\mathbb{Z}), +, \cdot)$, which is not commutative. Consider the subring $(S, +, \cdot)$ as the ring of all $2 \times 2$ diagonal matrices with entries in $\mathbb{Z}$. Note that $(S, +, \cdot)$ is commutative.

- A ring $(R, +, \cdot)$ may have unity, but a subring of it may not have. For example $(\mathbb{Z}, +, \cdot)$ has unity which is 1 but $(2\mathbb{Z}, +, \cdot)$ is a subring of $(\mathbb{Z}, +, \cdot)$ and $(2\mathbb{Z}, +, \cdot)$ does not have unity.

- A subring $S$ of $(R, +, \cdot)$ may have unity, but the ring may not have. For example, consider $R = \mathbb{Z} \times 2\mathbb{Z}$ with component wise standard addition and multiplication. Note that this ring has no unity. Consider the subring $S = \mathbb{Z} \times \{0\}$ of $R = \mathbb{Z} \times 2\mathbb{Z}$. Note that this subring has unity which is $(1, 0)$.

- A ring and its subring both may have unity but different. For example, the ring $R = M_2(\mathbb{R})$ has unity which is
$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$
Consider,
$$S = \left\{ \begin{bmatrix} a & a \\ a & a \end{bmatrix} : a \in \mathbb{R} \right\}.$$
One can check that the set $S$ is a subring of the ring $M_2(\mathbb{R})$. Note that the subring $S$ also has unity and the unity is
$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

- Suppose a ring $R$ and its subring $S$ have same unity. Then $a \in S$ is a unit implies that $a \in R$ is a unit. But the converse is not true. For example let $S = \mathbb{Z}$ and $R = \mathbb{R}$. Then 2 is a unit in $\mathbb{R}$ but not in $\mathbb{Z}$.

- If an element is a zero-divisor in a subring, then it is also a zero-divisor in the ring. But the converse is not true. For example, let $R = \mathbb{Z}$ and $S = \{0\}$. Note that 0 is a zero-divisor in $R$ but not in $S$.

## 3.6 Boolean ring

A ring is called a **Boolean** ring if every element of the ring is idempotent. For example, $\{0\}$ with addition and multiplication is a Boolean ring. The ring $(\mathbb{Z}/2\mathbb{Z}, +, \cdot)$ is a Boolean ring. We give another example: Let $X$ be a non-empty set. Consider the power set of $X$ which is $\mathcal{P}(X)$, i.e. the set of all subsets of $X$. We define for any $A, B \in \mathcal{P}(X)$,

$$A + B := (A \setminus B) \cup (B \setminus A), \quad \text{and} \quad A \cdot B = A \cap B.$$

Check that with this two operations $\mathcal{P}(X)$, is a commutative ring with unity. In fact, the unity is the set $X$ itself. Below we list down some properties of Boolean rings.

1. Subring of a Boolean ring is Boolean.

2. A Boolean ring need not have unity. For example let $X$ be an infinite set. Then consider $R = \mathcal{P}(X)$ and $S$ as the set of all finite subsets of $X$. Note that $S$ is a subring of $\mathcal{P}(X)$ but $S$ does not have unity. Consider $A_x := \{x\}$ for all $x \in X$. Now $A_x \in S$ for all $x \in X$. Let $T \in S$ be such that $A_x \cap T = A_x$ for all $x \in X$. Then $X \subseteq T$. A contradiction as $X$ is infinite.

3. Every Boolean ring $(R, +, \cdot)$ is commutative. Let $a, b \in R$. Then $(a+b)^2 = (a+b)$. Also,

   $$(a + b)^2 = (a + b) \cdot (a + b) = a \cdot a + a \cdot b + b \cdot a + b \cdot b = a + a \cdot b + b \cdot a + b.$$

   Therefore we get, $a \cdot b = -b \cdot a$. Now in $(R, +, \cdot)$, for every $a \in R$ we have $a = -a$. Because,

   $$a + a = (a + a)^2 = (a + a) \cdot (a + a) = a \cdot a + a \cdot a + a \cdot a + a \cdot a. = a + a + a + a.$$

   Hence $a + a = 0$ and thus $a = -a$ for all $a \in R$. Therefore ,

   $$a \cdot b = b \cdot a.$$

## 3.7 Polynomial ring

Let $(R, +, \cdot)$ be a ring and $X$ be a symbol. Consider the following collection of symbols

$$R[X] := \{a_0 + a_1X + \cdots + a_nX^n : n \in \mathbb{N} \text{ and } a_i \in R \text{ for all } 0 \leq i \leq n\}.$$

Note that $R \subset R[X]$. Let $P(X) = p_0+p_1X+\cdots+p_rX^r$ and $Q(X) = q_0+q_1X+\cdots+q_sX^s$. We say $P(X) = Q(X)$ if and only if $r = s$ and $p_i = q_i$ for all $0 \leq i \leq r$.
Now we define two operations '$+$' and '$\cdot$' on $R[X]$ as follows:
For $P(X) = p_0 + p_1X + \cdots + p_rX^r$ and $Q(X) = q_0 + q_1X + \cdots + q_sX^s$, set $p_i = q_j = 0$ for any $i > r$ and $j > s$, and define

$$P(X) + Q(X) := \sum_{i=0}^{\max(r,s)} (p_i + q_i)X^i, \text{ and}$$

$$P(X) \cdot Q(X) := p_0q_0 + (p_0q_1 + p_1q_0)X + \cdots + (p_0q_i + \cdots + p_iq_0)X^i + \cdots + p_rq_sX^{r+s},$$

i.e. $P(X) \cdot Q(X) = a_0 + a_1X + \cdots + a_nX^n$, where $n = r + s$ and $a_i = \sum_{k=0}^{i} p_kq_{i-k}$ for $0 \leq i \leq n$. Note that the symbol $X^n = \underbrace{X \cdot X \cdots X}_{n \text{ times}}$. Check that $(R[X], +, \cdot)$ is a ring. This ring is called the **ring of polynomials in $X$ over the ring $R$**. For any element $P(X) = p_0+p_1X+\cdots+p_rX^r \in R[X]$, the largest $r \in \mathbb{N}\cup\{0\}$ such that $p_r \neq 0$ is called the degree of $P(X)$. If $P(X) = 0$, then by convention we take the degree of $P(X)$ to be $-\infty$.

**Exercise 3.7.1** If $(R, +, \cdot)$ is commutative, then $(R[X], +, \cdot)$ is commutative.

**Exercise 3.7.2** If $(R, +, \cdot)$ has unity, then $(R[X], +, \cdot)$ has unity.

We can also define polynomial ring in multiple symbols. Let $X, Y$ be two different symbols. Let $R[X, Y]$ be the collection of all finite linear combinations of the symbols $X^mY^n$ where $m$ and $n$ non-negative integers, with coefficients in $R$, i.e.

$$R[X, Y] := \left\{ \sum_{m,n\geq 0} a_{m,n}X^mY^n : a_{m,n} \in R \text{ with all but finitely many } a_{m,n} = 0 \right\}.$$

We can define two operations '$+$' and '$\cdot$' on $R[X, Y]$ in a similar manner as follows: let $P(X, Y) = \sum_{m,n\geq 0} a_{m,n}X^mY^n$ and $Q(X, Y) = \sum_{m,n\geq 0} b_{m,n}X^mY^n$ such that all but

finitely many $a_{m,n}$ and $b_{m,n}$ are zero. Define

$$P(X,Y)+Q(X,Y) := \sum_{m,n\geq 0}(a_{m,n}+b_{m,n})X^mY^n, \text{ and}$$

$$P(X,Y)\cdot Q(X,Y) := \sum_{m,n\geq 0} c_{m,n}X^mY^n$$

where $c_{m,n} = \sum_{i=0}^{m}\sum_{j=0}^{n} a_{i,j}b_{m-i,n-j}$. As before $(R[X,Y],+,\cdot)$ is a ring which is called the **ring of polynomials in** $X,Y$ **over** $R$. For $P(X,Y) = \sum_{m,n\geq 0} a_{m,n}X^mY^n$, we can rewrite it as

$$a_0(X) + a_1(X)Y + \cdots + a_r(X)Y^r$$

where $a_i(X) \in R[X]$ for each $0 \leq i \leq r$, by adding all the symbols with a fixed power of $Y$ together. Thus we see that $R[X,Y]$ is nothing but $R[X][Y]$, the **ring of polynomials in** $Y$ **over** $R[X]$. Equivalently, we can also write it as $R[Y][X]$, the **ring of polynomials in** $X$ **over** $R[Y]$. Therefore, $R[X]$ and $R[Y]$ are subrings of $R[X,Y]$.

Similarly, for $n$ distinct symbols $X_1,\ldots,X_n$, we can analogously define the **ring of polynomials in** $n$ **symbols over** $R$, denoted by $R[X_1,\ldots,X_n]$, which can also be viewed as $R[X_1,\ldots,X_{i-1},X_{i+1},\ldots,X_n][X_i]$, the ring of polynomials in $X_i$ over the ring $R[X_1,\ldots,X_{i-1},X_{i+1},\ldots,X_n]$, for any $1 \leq i \leq n$.

More interestingly, for an arbitrary collection of symbols $\{X_\alpha\}_{\alpha\in I}$, where $I$ is a set of (possibly infinitely many) indices, we can define the ring of polynomials in these symbols over $R$. For this we consider the symbols $\prod X_\alpha^{n_\alpha}$ where $n_\alpha$'s are non-negative integers and all but finitely many of them are zero. Then we consider the collection of all finite linear combinations of these symbols with coefficients in $R$, denoted by $R[X_\alpha : \alpha \in I]$ and as before, for any chosen symbol $X_{\alpha_0}$, we can also view it as $R[X_\alpha : \alpha \in I \setminus \{\alpha_0\}][X_{\alpha_0}]$.

## 3.8 Power series ring

Let $(R,+,\cdot)$ be a ring and $X$ be a symbol. Consider the following collection of symbols

$$R[[X]] := \{a_0 + a_1X + \cdots + a_nX^n + \cdots : a_i \in R \text{ for all } i \geq 0\}.$$

Note that $R \subset R[X] \subset R[[x]]$. Let $P(X) = p_0 + p_1X + \cdots + p_rX^r + \cdots$ and $Q(X) = q_0 + q_1X + \cdots + q_sX^s + \cdots$. We say $P(X) = Q(X)$ if and only if $p_i = q_i$ for all $i \geq 0$. Now we define two operations '+' and '$\cdot$' on $R[[X]]$ as follows:

For $P(X) = p_0 + p_1X + \cdots + p_rX^r + \cdots$ and $Q(X) = q_0 + q_1X + \cdots + q_sX^s + \cdots$, define

$$P(X) + Q(X) := \sum_{i\geq 0}(p_i + q_i)X^i, \text{ and}$$

$$P(X) \cdot Q(X) := p_0q_0 + (p_0q_1 + p_1q_0)X + \cdots + (p_0q_i + \cdots + p_iq_0)X^i + \cdots,$$

i.e. $P(X)\cdot Q(X) = a_0 + a_1X + \cdots + a_nX^n + \cdots$, and $a_i = \sum_{k=0}^{i} p_kq_{i-k}$ for all $i \geq 0$. Check that $(R[[X]], +, \cdot)$ is a ring. This ring is called the **ring of formal power series in $X$ over the ring $R$.**

**Exercise 3.8.1** If $(R, +, \cdot)$ is commutative, then $(R[[X]], +, \cdot)$ is commutative.

**Exercise 3.8.2** If $(R, +, \cdot)$ has unity, then $(R[[X]], +, \cdot)$ has unity.

Let $X, Y$ be two different symbols. We can analogously define the ring of formal power series in two variables over $R$. Let $R[[X, Y]]$ be the collection of symbols

$$R[[X, Y]] := \left\{ \sum_{m,n\geq 0} a_{m,n}X^mY^n : a_{m,n} \in R \right\}.$$

Imitating the operations as defined in the ring of polynomials in $X, Y$ over $R$, we will get a ring structure on $R[[X, Y]]$, which is called the **ring of formal power series in $X, Y$ over $R$.** And likewise one can define the ring of formal power series in arbitrarily many symbols $\{X_\alpha\}_{\alpha\in I}$ over $R$, by allowing any (possibly infinite) linear combinations of the the symbols $\prod X_\alpha^{n_\alpha}$ where $n_\alpha$'s are non-negative integers and all but finitely many of them are zero.

## 3.9 Ring of Gaussian integers

A **Gaussian integer** is a complex number of the form $a + \iota b$ where $a, b \in \mathbb{Z}$. Consider the set,

$$\mathbb{Z}[\iota] := \{a + \iota b : a, b \in \mathbb{Z}\}.$$

We can observe that $\mathbb{Z} \subset \mathbb{Z}[\iota]$. The addition and multiplication of two Gaussian integers are defined as follows:

$$(a + \iota b) + (c + \iota d) := (a + c) + \iota(b + d),$$

$$(a + \iota b) \cdot (c + \iota d) := (ac - bd) + \iota(bc + ad).$$

Note that $\iota^2 = -1$.

**Exercise 3.9.1** Show that the set $\mathbb{Z}[\iota]$ with the addition and multiplication defined above forms a commutative ring with unity.

The ring $\mathbb{Z}[\iota]$ is called the **ring of Gaussian integers**.

## 3.10 Integral domain

**Definition 3.10.1** A non-zero commutative ring is called an **integral domain**, if there is no non-zero zero divisor in it.

**Remark 3.10.2** By convention the zero ring is not an integral domain.

**Remark 3.10.3** For any two integral domains $(R_1, +, \cdot)$, $(R_2, +, \cdot)$, the ring $(R_1 \times R_2, +, \cdot)$ with component wise addition and multiplication, is never an integral domain. Note that for any non-zero $a \in R_1$ and a non-zero $b \in R_2$,

$$(a, 0) \cdot (0, b) = (0, 0).$$

**Exercise 3.10.4** Show that the ring $\mathbb{Z}[\iota]$ is an integral domain.

**Exercise 3.10.5** If $(R, +, \cdot)$ is an integral domain, then $(R[X], +, \cdot)$ is an integral domain.

**Exercise 3.10.6** Subring of an integral domain is an integral domain.

**Exercise 3.10.7** Let $(R, +, \cdot)$ be an integral domain with unity. A polynomial $P(X) = p_0 + p_1 X + \cdots + p_n X^n \in R[X]$ is a unit if and only if $p_0$ is a unit in $(R, +, \cdot)$ and $n = 0$.

**Exercise 3.10.8** Let $R$ be an integral domain with unity. A power series $\sum_{i \geq 0} a_i X^i$ is a unit in $R[[X]]$ if and only if $a_0$ is a unit in $(R, +, \cdot)$

**Examples 3.10.9**

- The ring $(\mathbb{Z}, +, \cdot)$ is an integral domain.
- The ring $(n\mathbb{Z}, +, \cdot)$ for $n \in \mathbb{Z}$ is an integral domain.

- The ring $(\mathbb{Z}/2\mathbb{Z}, +, \cdot)$ is an integral domain.
- The ring $(\mathbb{Z}/3\mathbb{Z}, +, \cdot)$ is an integral domain.
- The ring $(\mathbb{Z}/4\mathbb{Z}, +, \cdot)$ is not an integral domain.
- The ring $(\mathbb{Z}/5\mathbb{Z}, +, \cdot)$ is an integral domain.
- The ring $(\mathbb{Z}/6\mathbb{Z}, +, \cdot)$ is not an integral domain.

In the above examples, note that whenever $n$ is prime, we are getting $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ is an integral domain. We have the following theorem:

**Theorem 3.10.10** Let $n \in \mathbb{N}$. Then $\mathbb{Z}/n\mathbb{Z}$ is an integral domain if and only if $n$ is prime.

*Proof.* First assume that $\mathbb{Z}/n\mathbb{Z}$ is an integral domain. We want to show that $n$ is prime. If not, then we can write $n = ab$ where $a, b \in \mathbb{N}$ and $1 < a, b < n$. Consider $[a], [b]$ in $\mathbb{Z}/n\mathbb{Z}$. Note that both are non-zero elements in $\mathbb{Z}/n\mathbb{Z}$ and $[a][b] = [ab] = [n] = [0]$. A contradiction to the fact that $\mathbb{Z}/n\mathbb{Z}$ is an integral domain. Therefore $n$ must be prime.

Next assume that $n$ is prime. We shall show that $\mathbb{Z}/n\mathbb{Z}$ is an integral domain. If not, let $[a], [b] \in \mathbb{Z}/n\mathbb{Z}$ be two non-zero elements such that $[a][b] = [0]$. Therefore we have $[ab] = [0]$ i.e. $n \mid ab$. Since $n$ is a prime by prime factorization of $ab$ in $\mathbb{Z}$, we get $n \mid a$ or $n \mid b$. This implies that $[a] = [0]$ or $[b] = [0]$, a contradiction to our assumption that both are non-zero. Therefore $\mathbb{Z}/n\mathbb{Z}$ is an integral domain.

The next theorem tells us that the cancellation laws hold in an integral domain. From now on we will use only the notation $R$ to refer to a ring instead of $(R, +, \cdot)$ unless otherwise stated.

**Theorem 3.10.11** Let $R$ be a non-zero commutative ring. Then $R$ is an integral domain if and only if $a \cdot b = a \cdot c$ with $a \neq 0$ implies that $b = c$.

*Proof.* Suppose that $R$ is an integral domain and $a \cdot b = a \cdot c$ with $a \neq 0$. Then $a \cdot (b - c) = 0$. Since $a \neq 0$, we have $b - c = 0$, i.e. we get that $b = c$.

Next suppose that $a \cdot b = a \cdot c$ with $a \neq 0$ implies that $b = c$ in $R$, where $R$ is a non-zero commutative ring. We want to show that $R$ is an integral domain. Let $a, b \in R$ be such that $a \cdot b = 0$. Without loss of generality let $a \neq 0$. Then we can write, $a \cdot b = 0 = a \cdot 0$. Therefore $b = 0$. This proves that there is no non-zero zero divisor in $R$. Hence $R$ is an integral domain.

### 3.10.1 Fields

**Definition 3.10.12** A **field** $F$ is a non-zero commutative ring with unity, where each non-zero element has a multiplicative inverse.

**Example 3.10.13** The set of rational numbers $\mathbb{Q}$, the set of real numbers $\mathbb{R}$ are examples of field.

**Theorem 3.10.14** Every field is an integral domain.

*Proof.* Let $F$ be a field. We show that $F$ has no non-zero zero divisors. Let $a, b \in F$ be such that $a \cdot b = 0$. Without loss of generality let, $a \neq 0$, then we get $b = a^{-1} \cdot (a \cdot b) = a^{-1} \cdot 0 = 0$. So $F$ is an integral domain.

**Remark 3.10.15** We know that the converse of the above theorem is not true, i.e. every integral domain need not be a field. For example, the set of integers $\mathbb{Z}$ is an integral domain which is not a field as we have already seen that the only units in $\mathbb{Z}$ are $\{1, -1\}$. Note that $\mathbb{Z}$ is an infinite set. The below theorem gives us a sufficient condition for an integral domain to be a field.

**Theorem 3.10.16** Every finite integral domain is a field.

*Proof.* Let $R$ be a finite integral domain. We want to show that $R$ is a field i.e. $R$ contains unity and every non-zero element of $R$ has a multiplicative inverse.

We write $R = \{a_1, \ldots, a_n\}$. Let $a \in R \setminus \{0\}$. Now $aR = \{a \cdot a_1, \ldots, a \cdot a_n\}$. Note that $|aR| = n$ because if $a \cdot a_i = a \cdot a_j$ for some $i \neq j$, then since $R$ is an integral domain we get that $a_i = a_j$. Therefore all the elements of the set $aR$ are distinct. Now as $aR \subseteq R$ and $|R| = |aR| = n$, we obtain $R = aR$. Therefore for $a \in R$, there exists $a_i \in R$ so that $a = a \cdot a_i = a_i \cdot a$ as $R$ is commutative. We show that $a_i$ is the unity of $R$. Let $b \in R$. Then we have,

$$b = a \cdot a_j = a_j \cdot a \text{ for some } 1 \leq j \leq n.$$

Thus,

$$b = (a \cdot a_i) \cdot a_j = a_i \cdot (a \cdot a_j) = a_i \cdot b.$$

As $R$ is commutative, we get $b = b \cdot a_i = a_i \cdot b$. This proves that $a_i \in R$ is the unity. We denote $a_i$ by 1. Now there exists $a_k \in R$ such that $1 = a \cdot a_k = a_k \cdot a$. This proves that $a$ has a multiplicative inverse in $R$ which is $a_k$. This completes the proof.

**Remark 3.10.17** We can now conclude that $\mathbb{Z}/n\mathbb{Z}$ is a field if and only if $n$ is a prime.

## 3.11 Characteristic of a non-zero commutative ring with unity

Let $R$ be a non-zero commutative ring with unity. Then the least positive integer $n$ (if it exists) such that $na = 0$ for all $a \in R$ is called the characteristic of the ring $R$. If no such positive integer exists, then we define it to be 0. Note that if the characteristic of $R$ is non-zero then it is same as the additive order of the unity in $R$. We denote it by $\mathrm{Char}(R)$.

**Examples 3.11.1**

- $\mathrm{Char}(\mathbb{Z}/2\mathbb{Z}) = 2$.
- $\mathrm{Char}(\mathbb{Z}/4\mathbb{Z}) = 4$.
- $\mathrm{Char}(\mathbb{Z}) = 0$.
- $\mathrm{Char}(\mathbb{Q}) = 0$.

**Exercise 3.11.2** Prove that the characteristic of an integral domain is either 0 or a prime number.

**Remark 3.11.3** The characteristic of a non-zero finite ring is always non-zero. But for infinite rings both cases can arise. For example, for $n \geq 2$, $(\mathbb{Z}/n\mathbb{Z})[X]$ is an infinite ring with characteristic $n$ and $\mathbb{Z}$ is an infinite ring having characteristic 0.

**Exercise 3.11.4** For a prime number $p$, consider

$$(\mathbb{Z}/p\mathbb{Z})(X) := \left\{ \frac{f}{g} : f, g \in (\mathbb{Z}/p\mathbb{Z})[X] \text{ with } g \neq 0 \right\}.$$

Define $+, \cdot$ in $(\mathbb{Z}/p\mathbb{Z})(X)$ as follows:

$$\frac{f_1}{g_1} + \frac{f_2}{g_2} := \frac{f_1 g_2 + g_1 f_2}{g_1 g_2},$$

$$\frac{f_1}{g_1} \cdot \frac{f_2}{g_2} := \frac{f_1 f_2}{g_1 g_2},$$

where $f + g, fg$ denote the usual addition and multiplication of two polynomials $f, g$ in $(\mathbb{Z}/p\mathbb{Z})[X]$. Prove that $(\mathbb{Z}/p\mathbb{Z})(X)$ with $+, \cdot$ is an infinite field with characteristic $p$.

**Remark 3.11.5** We have $\mathbb{Q}$ is an infinite field with characteristic 0 and $(\mathbb{Z}/p\mathbb{Z})(X)$ is an infinite field with characteristic $p$ for a prime number $p$. Also we know, every field is an integral domain, the characteristic of a field is either 0 or a prime number.

## 3.12 Ideals

**Definition 3.12.1** Let $R$ be a ring. A subring $I$ of $R$ is called

- a **left ideal** if $r \cdot x \in I$ for all $r \in R$ and $x \in I$.
- a **right ideal** if $x \cdot r \in I$ for all $r \in R$ and $x \in I$.
- an **ideal** if $x \cdot r, r \cdot x \in I$ for all $r \in R$ and $x \in I$.

**Examples 3.12.2**

1. In any ring $R$, the subrings $\{0\}$ and $R$ are always ideals. They are called **trivial ideals**.
2. In the ring of integers $\mathbb{Z}$, the subrings $n\mathbb{Z}$ for all $n \in \mathbb{Z}$ are ideals.

**Exercise 3.12.3** Consider the ring of continuous functions $\mathcal{C}([0,1],\mathbb{R})$. Let $a, b \in [0,1]$. Show that $M_{a,b} := \{f \in \mathcal{C}([0,1],\mathbb{R}) : f(a) = 0 = f(b)\}$ is an ideal of $\mathcal{C}([0,1],\mathbb{R})$.

**Exercise 3.12.4** Let $R$ be a ring and $I$ be an ideal in $R$. Show that $M_2(I)$ is an ideal in the matrix ring $M_2(R)$.

**Remark 3.12.5** We know that an ideal is a subring with some additional properties. But not every subring is an ideal. For example the set of integers $\mathbb{Z}$ is a subring of the ring of rational numbers $\mathbb{Q}$. But the subring $\mathbb{Z}$ is not an ideal of the ring $\mathbb{Q}$.

**Exercise 3.12.6** Consider the matrix ring $M_2(\mathbb{Z})$. Show that

$$\mathcal{I} := \left\{ \begin{bmatrix} a & a \\ a & a \end{bmatrix} : a \in \mathbb{Z} \right\}$$

is not an ideal of the matrix ring $M_2(\mathbb{Z})$.

**Remark 3.12.7** Let $R$ be a ring and $I, J$ be two ideals of $R$. Then the union $I \cup J$ need not be an ideal of $R$. For example $2\mathbb{Z}, 3\mathbb{Z}$ are ideals of $\mathbb{Z}$ but $2\mathbb{Z} \cup 3\mathbb{Z}$ is not an ideal of $\mathbb{Z}$. In fact $I \cup J$ is a subgroup of $(R, +)$ if and only if $I \subseteq J$ or $J \subseteq I$. Suppose

that $I \cup J$ a subgroup of $(R, +)$. Let $I$ be not a subset of $J$ and $J$ be not a subset of $I$. Then there exist $a \in I \setminus J$ and $b \in J \setminus I$. Now $a + b \in I \cup J$ as $I \cup J$ is a subgroup. So $a + b \in I$ or $a + b \in J$. If $a + b \in I$, then we get $b \in I$, a contradiction. If $a + b \in J$, then we get $a \in J$, a contradiction.

**Exercise 3.12.8** Show that intersection of any arbitrary family of ideals in a ring is again an ideal.

### 3.12.1 Sum and product of ideals

Let $R$ be a ring and $I, J$ be two ideals of $R$. We define the sum of two ideals $I, J$ as follows:

$$I + J := \{a + b : a \in I, b \in J\}.$$

Note that $0 \in I + J$. For any two elements $\alpha, \beta \in I + J$, we can write $\alpha = a_1 + b_1$ and $\beta = a_2 + b_2$ for $a_1, a_2 \in I$ and $b_1, b_2 \in J$. Now as $I, J$ are ideals we have $a_1 - a_2 \in I$ and $b_1 - b_2 \in J$. Therefore $\alpha - \beta = (a_1 - a_2) + (b_1 - b_2) \in I + J$. Also using the distributive law and the facts that $I, J$ are ideals we get that

$$\alpha \cdot \beta = (a_1 + b_1) \cdot (a_2 + b_2) = a_1 \cdot (a_2 + b_2) + b_1 \cdot (a_2 + b_2) \in I + J.$$

Similarly for any $r \in R$, $r \cdot \alpha = r \cdot a_1 + r \cdot b_1 \in I + J$. This proves that $I + J$ is an ideal of $R$.

We define the product of two ideals $I, J$ of $R$ as follows:

$$IJ := \left\{ \sum_{i=1}^{n} a_i b_i : n \in \mathbb{N}, a_i \in I, b_i \in J \right\}.$$

Note that $0 \in IJ$. One can check that $IJ$ is an ideal of $R$.

**Exercise 3.12.9** Let $I, J$ be two ideals in a ring $R$. Show that $IJ \subseteq I + J$.

**Exercise 3.12.10** Let $I, J$ be two ideals in a ring $R$. Show that $I \cup J \subseteq I + J$.

**Remarks 3.12.11**

1. The ideal $I + J$ is the smallest ideal containing $I \cup J$. Let $K$ be another ideal of $R$ such that $I \cup J \subseteq K$. Therefore $I \subseteq K$ and $J \subseteq K$. This implies that $I + J \subseteq K$.

2. Since $I, J$ are ideals, clearly we see that $IJ \subseteq I \cap J$. Now as $0 \in I, J$, we have $I \subseteq I + J$ and $J \subseteq I + J$. Therefore $I \cap J \subseteq I + J$. Thus we have,

$$IJ \subseteq I \cap J \subseteq I + J.$$

3. We can find examples so that

$$IJ \subsetneq I \cap J \text{ and } I \cap J \subsetneq I + J.$$

In the ring $\mathbb{Z}$, consider $I = 2\mathbb{Z}$ and $J = 4\mathbb{Z}$. Here $IJ = 8\mathbb{Z}$ and $I \cap J = 4\mathbb{Z}$. In the ring $\mathbb{Z}$, consider $I = 2\mathbb{Z}$ and $J = 3\mathbb{Z}$. Here $I \cap J = 6\mathbb{Z}$ and $I + J = \mathbb{Z}$.

4. Also we can find an ideal $K$ in $R$ such that

$$IJ \subsetneq K \subsetneq I \cap J.$$

Consider $I = 4\mathbb{Z}$ and $J = 8\mathbb{Z}$ in the ring $\mathbb{Z}$. Then $IJ = 32\mathbb{Z}$ and $I \cap J = 8\mathbb{Z}$. Consider $K = 16\mathbb{Z}$.

**Theorem 3.12.12** Let $R$ be a ring with unity and $I$ be an ideal in $R$. Then the following are equivalent:

1. $I = R$.
2. $1 \in I$.
3. $u \in I$ for some unit $u \in R$.

*Proof.* Clearly we can see that 1) $\Rightarrow$ 2) $\Rightarrow$ 3). Now note that 3) $\Rightarrow$ 2) as $1 = u^{-1} \cdot u \in I$. Now we show that 2) $\Rightarrow$ 1). Let $r \in R$. Then $r = r \cdot 1 \in I$ as $1 \in I$. Therefore $R = I$.

**Remark 3.12.13** From the above theorem we can see that any non-zero proper ideal of any ring can never contain a unit.

### 3.12.2 Principal ideals

Let $R$ be a ring and $A$ be a subset of $R$. Let $\{I_\alpha\}_{\alpha \in \Lambda}$ be the collection of all ideals of $R$ containing $A$. Clearly the collection is non-empty as $R$ is an ideal containing $A$. By the ideal generated by the set $A$ denoted by $(A)$ we mean, $\cap_{\alpha \in \Lambda} I_\alpha$. We know that intersection of any arbitrary collection of ideals in any ring is again an ideal, so $\cap_{\alpha \in \Lambda} I_\alpha$

is an ideal of $R$ containing $A$. So $(A)$ is the smallest ideal of $R$ containing the set $A$ in the sense that if any ideal of $R$ contains $A$, then it contains $(A)$.

**Definition 3.12.14** Let $I$ be an ideal of $R$. If there exists a set $A$ of $R$ such that $I = (A)$, then we say that the ideal $I$ is generated by the set $A$. If there exists a finite set $A \subseteq R$ such that $I = (A)$, then $I$ is called **finitely generated**. If $A$ has only one element then $I$ is called a **principal ideal**.

**Exercise 3.12.15** Let $x \in R$. Consider the following subset of $R$,

$$I := \{r \cdot x + x \cdot s + \sum_{i=1}^{m} r_i \cdot x \cdot s_i + nx : r, s, r_i, s_i \in R \text{ for all } 1 \leq i \leq m, m \in \mathbb{N}, n \in \mathbb{Z}\}.$$

Here,

$$nx = \begin{cases} \underbrace{(x + \cdots + x)}_{n \text{ times}} & \text{if } n \geq 0, \\ -\underbrace{(x + \cdots + x)}_{-n \text{ times}} & \text{otherwise.} \end{cases}$$

1. Show that $I$ is an ideal of $R$ containing $x$.
2. Prove that $I = (x)$.

**Remarks 3.12.16** We can observe the following:

- If $R$ is commutative then $(x) = \{r \cdot x + nx : r \in R, n \in \mathbb{Z}\}$.
- If $R$ is commutative with unity then $(x) = \{r \cdot x : r \in R\} = Rx$.

**Remark 3.12.17** Let $R$ be a commutative ring with unity and $a, b \in R$. Then the smallest ideal containing $\{a, b\}$ is $Ra + Rb$.

Since $Ra+Rb$ is an ideal containing $\{a, b\}$ and $(\{a, b\})$ is the smallest ideal containing $\{a, b\}$, we get $(\{a, b\}) \subseteq Ra+Rb$. Now let $ra+sb \in Ra+Rb$. We know that $ra$ belongs to every ideal containing $a$. Also $sb$ belongs to every ideal containing $b$. So $ra, sb \in (\{a, b\})$. Since $(\{a, b\})$ is an ideal we have, $ra+sb \in (\{a, b\})$. This proves that $Ra+Rb \subseteq (\{a, b\})$. Altogether we get that $(\{a, b\}) = Ra + Rb$.

**Example 3.12.18** In the ring of integers $\mathbb{Z}$, the ideals $n\mathbb{Z}$ for $n \in \mathbb{Z}$ are all principal ideals. We know that $\mathbb{Z}$ is a commutative ring with unity. So we can easily see that $n\mathbb{Z} = (n)$.

**Example 3.12.19** In the polynomial ring over $\mathbb{Z}$ i.e. in $\mathbb{Z}[X]$ consider the set

$$\{a_1X + \cdots + a_nX^n : a_i \in \mathbb{Z} \text{ for all } 1 \leq i \leq n, n \in \mathbb{N}\}.$$

Note that the above set is an ideal of $\mathbb{Z}[X]$. Also we can see that

$$\{a_1X + \cdots + a_nX^n : a_i \in \mathbb{Z} \text{ for all } 1 \leq i \leq n, n \in \mathbb{N}\} = \{P(X)X : P(X) \in \mathbb{Z}[X]\} = (X).$$

**Example 3.12.20** In the ring $\mathbb{Z}[X]$, the smallest ideal containing $\{2, X\}$ i.e. $(2, X)$ is not a principal ideal. If there exists $P(X) \in \mathbb{Z}[X]$ be such that $(2, X) = (P(X))$, then we can write,

$$2 = P(X)Q(X), \quad \text{where } Q(X) \in \mathbb{Z}[X].$$

Comparing both sides we get that $P(X), Q(X)$ both have to be constant polynomials i.e. elements of $\mathbb{Z}$. Say $P(X) = p_0$ and $Q(X) = q_0$. Now $p_0q_0 = 2$ in $\mathbb{Z}$ implies that $p_0$ is one of the values $1, -1, 2$ or $-2$. Now if $p_0$ is 1 or $-1$, then $P(X)$ is a unit. Therefore $(2, X) = \mathbb{Z}[X]$. This is not possible as $1 \notin (2, X)$. Because $1 \neq 2A(X) + XB(X)$ for all $A(X), B(X) \in \mathbb{Z}(X)$. Therefore $p_0$ is either 2 or $-2$. Now note that $X$ can never be equal to $2A(X)$ or $-2A(X)$ for any $A(X) \in \mathbb{Z}[X]$. This completes the proof that $(2, X)$ is not generated by any single polynomial over $\mathbb{Z}$.

**Example 3.12.21** In the ring of polynomials over $\mathbb{Q}$ i.e. in $\mathbb{Q}[X]$, the ideal $(2, X)$ is a principal ideal. Note that 2 is a unit in $\mathbb{Q}[X]$, therefore $(2, X) = \mathbb{Q}[X]$. Now $\mathbb{Q}[X] = (u)$ for any unit $u \in \mathbb{Q}[X]$. We know that all non-zero rational numbers are the only units in $\mathbb{Q}[X]$ as $\mathbb{Q}$ is an integral domain . Therefore we can write $(2, X) = (u)$ for any non-zero rational number $u$.

### 3.12.3 Maximal ideals

**Definition 3.12.22** Let $R$ be a non-zero ring. An ideal $\mathcal{M}$ of $R$ is called a **maximal ideal** if the following hold:

- $\mathcal{M} \neq R$,
- if there exists any ideal $I$ of $R$ such that $\mathcal{M} \subseteq I$, then either $I = \mathcal{M}$ or $I = R$.

In other words, a proper ideal of $R$ is called a maximal ideal if that is not contained in any other proper ideal of $R$.

**Examples 3.12.23**

1. In the ring $\mathbb{Z}$, the trivial ideal $\{0\}$ is not a maximal ideal as it is contained in all ideals of $\mathbb{Z}$.

2. In the ring $\mathbb{Z}$, the ideal $4\mathbb{Z}$ is not maximal as it is contained in the proper ideal $2\mathbb{Z}$.

3. In the ring $\mathbb{Z}$, the ideal $2\mathbb{Z}$ is maximal. Suppose there exists an ideal $I$ of $\mathbb{Z}$ such that $2\mathbb{Z} \subsetneq I$. Since $2\mathbb{Z} \subsetneq I$, there exists an integer $a \in I$ such that $a \notin 2\mathbb{Z}$ i.e. $a = 2m + 1$ for some $m \in \mathbb{Z}$. Note that $2m \in 2\mathbb{Z}$. So $2m \in I$ and therefore $2m + 1 - 2m = 1 \in I$. This proves that $I = \mathbb{Z}$.

4. Consider the ring of continuous functions from $[0, 1]$ to $\mathbb{R}$, $\mathcal{C}([0, 1], \mathbb{R})$. Let $a \in [0, 1]$. Then consider the set $M_a := \{f \in \mathcal{C}([0, 1], \mathbb{R}) : f(a) = 0\}$. Now $M_a \neq \varnothing$ as the function $f \equiv 0 \in M_a$. Now for any $f_1, f_2 \in M_a$, we have

$$(f_1 - f_2)(a) = f_1(a) - f_2(a) = 0,$$

Therefore $f_1 - f_2 \in M_a$. Also for any $g \in \mathcal{C}([0, 1], \mathbb{R})$, and any $h \in M_a$, we have

$$(gh)(a) = g(a)h(a) = 0 = h(a)g(a) = (hg)(a),$$

and hence $gh = hg \in M_a$. This shows that $M_a$ is an ideal of $\mathcal{C}([0, 1], \mathbb{R})$.

We shall show that $M_a$ is a maximal ideal of $\mathcal{C}([0, 1], \mathbb{R})$. Let $M$ be an ideal of $\mathcal{C}([0, 1], \mathbb{R})$ such that $M_a \subsetneq M$. So there exists $f \in M$ such that $f(a) \neq 0$. Consider $g : [0, 1] \to \mathbb{R}$ defined as follows:

$$g(x) = f(x) - f(a).$$

Clearly $g$ is continuous as $f$ is continuous and $g(a) = 0$. So $g \in M_a$. Therefore $g \in M$. So the function $h := f - g \in M$. Now note that $h(x) = f(a)$ for all $x \in [0, 1]$. So $h$ is never 0 and hence it is a unit. This proves that $M = \mathcal{C}([0, 1], \mathbb{R})$, and hence $M_a$ is a maximal ideal.

**Remark 3.12.24** In fact, all the maximal ideals of $\mathcal{C}([0, 1], \mathbb{R})$ are of the form $M_a$ for $a \in [0, 1]$. Let $M$ be a maximal ideal of $\mathcal{C}([0, 1], \mathbb{R})$ such that $M \neq M_a$ for all $a \in [0, 1]$. Then for each $a \in [0, 1]$, there exists a function $f_a \in M$ such that $f_a(a) \neq 0$. Now as $f_a$ is continuous, there exists a $\delta_a > 0$ such that $f_a(x) \neq 0$ for all $x \in (a - \delta_a, a + \delta_a) \cap [0, 1]$. Now,

$$[0, 1] = \bigcup_{a \in [0,1]} (a - \delta_a, a + \delta_a) \cap [0, 1].$$

We know that $[0,1]$ is compact i.e. every open cover of $[0,1]$ has finitely many sub-covers. Therefore we obtain,

$$[0,1] = \bigcup_{i=1}^{n} (a_i - \delta_{a_i}, a_i + \delta_{a_i}) \cap [0,1],$$

for finitely many points $a_1, \ldots, a_n \in [0,1]$. Consider the function $g : [0,1] \to \mathbb{R}$ by

$$g := \sum_{i=1}^{n} f_{a_i}^2.$$

The function $g \in M$ as $M$ is an ideal. Also $g(x) > 0$ for all $x \in [0,1]$ as each $f_{a_i}^2(x) > 0$ for all $x \in (a_i - \delta_{a_i}, a_i + \delta_{a_i}) \cap [0,1]$. Therefore $g$ is a unit and hence $M = \mathcal{C}([0,1], \mathbb{R})$. Thus $M$ is not a maximal ideal. This proves that $\{M_a : a \in [0,1]\}$ is the set of all the maximal ideals of $\mathcal{C}([0,1], \mathbb{R})$.

**Exercise 3.12.25** Show that maximal ideals of $\mathbb{Z}$ are of the form $p\mathbb{Z}$, where $p$ is a prime number.

**Theorem 3.12.26** Let $R$ be a ring with unity and $I$ be a proper ideal of $R$. Then there exists a maximal ideal $\mathcal{M}$ in $R$ such that $I \subseteq \mathcal{M}$.

*Proof.* We consider the set,

$$\mathcal{F} := \{J : J \text{ is an ideal of } R, I \subseteq J, J \neq R\}.$$

First we note that $\mathcal{F}$ is a non-empty set as $I \in \mathcal{F}$. We define the relation '$\leq$' on $\mathcal{F}$ as $A \leq B$ if and only $A \subseteq B$. As indicated in Section 1.1.2, this relation is a partial order relation on the non-empty set $\mathcal{F}$. Let $T$ be a non-empty chain in $\mathcal{F}$. We show that $T$ is bounded above. Consider

$$T' := \bigcup_{I_\alpha \in T} I_\alpha.$$

Clearly $I \subseteq T'$ as $I \subseteq I_\alpha$ for all $I_\alpha \in T$. Now we show that $T'$ is an ideal of $R$. Let $x, y \in T$. So $x \in I_{\alpha_1}$ and $y \in I_{\alpha_2}$ for some $I_{\alpha_1}, I_{\alpha_1} \in T$. Since $T$ is a chain, we have,

$$I_{\alpha_1} \subseteq I_{\alpha_2} \text{ or } I_{\alpha_2} \subseteq I_{\alpha_1}.$$

Without loss of generality let $I_{\alpha_1} \subseteq I_{\alpha_2}$. Therefore $x, y \in I_{\alpha_2}$. Now as $I_{\alpha_2}$ is an ideal of $R$, we get $x - y \in I_{\alpha_2}$. Hence $x - y \in T'$. Again for any $x \in T'$ and $r \in R$, we have $x \in I_\alpha$ for some $I_\alpha \in T$. So $r \cdot x \in I_\alpha$ and hence $r \cdot x \in T'$. This proves that $T'$ is an ideal of $R$.

Now we show that $T'$ is a proper ideal of $R$. If $T' = R$, then as $R$ is a ring with 1, we get $1 \in T'$. Therefore $1 \in I_\alpha$ for some $I_\alpha \in T$. Thus $I_\alpha = R$, a contradiction as $I_\alpha \in \mathcal{F}$. Therefore $T' \neq R$. Hence $T' \in \mathcal{F}$. This proves that the chain $T$ in $\mathcal{F}$ is bounded above by $T' \in \mathcal{F}$. So we have any non-empty set in the poset $\mathcal{F}$ is bounded above. Therefore by **Zorn's lemma**, there exists a maximal element say $\mathcal{M}$ in $\mathcal{F}$. We show that $\mathcal{M}$ is a maximal ideal of $R$ containing the proper ideal $I$. Since $\mathcal{M} \in \mathcal{F}$, we have that $\mathcal{M}$ is a proper ideal of $R$ and $I \subseteq \mathcal{M}$. Suppose that, $\mathcal{M} \subseteq J \subsetneq R$. Now,

$$I \subseteq \mathcal{M} \subseteq J \subsetneq R.$$

So as $J \in \mathcal{F}$ and $\mathcal{M}$ is a maximal element in $\mathcal{F}$, we get that $J = \mathcal{M}$. This proves that $\mathcal{M}$ is a maximal ideal of $R$.

**Corollary 3.12.27** Let $R$ be a non-zero ring with unity. Then there exists at least one maximal ideal in $R$. Because $\{0\}$ is always a proper ideal in a ring with unity, therefore by the aid of the above theorem there exists a maximal ideal containing the ideal $\{0\}$ in $R$.

**Remark 3.12.28** If a non-zero ring $R$ is not having unity, then even if $R$ is commutative, there many not exist a maximal ideal in $R$.

Let us consider the set of all rational numbers $\mathbb{Q}$. Consider the standard addition of rational numbers as the additive operation on $\mathbb{Q}$ and the trivial multiplication i.e. $a \cdot b = 0$ for all $a, b \in \mathbb{Q}$ as the multiplicative operation on $\mathbb{Q}$. Note that since for any non-zero $a \in \mathbb{Q}$, $a \cdot b = 0 \neq a$ for all $b \in \mathbb{Q}$, this ring does not have any unity. Therefore $(\mathbb{Q}, +, \cdot)$ is a non-zero commutative ring without unity. We show that this ring does not contain any maximal ideal. Suppose $\mathcal{M}$ is a maximal ideal in this ring $\mathbb{Q}$. Therefore we have the following:

1. $\mathcal{M} \neq \mathbb{Q}$,
2. $\mathcal{M} \neq \{0\}$ as $\{0\} \subsetneq \mathbb{Z}$ and $\mathbb{Z}$ is a proper ideal of $\mathbb{Q}$.

Since $\mathcal{M} \neq \mathbb{Q}$, we can find a rational number $\frac{r}{s} \notin \mathcal{M}$ with $s > 0$ (it is standard to represent a rational number as $a/b$ where $a \in \mathbb{Z}$ and $b \in \mathbb{N}$). Then we consider the ideal

$$J := \mathcal{M} + \left(\frac{r}{s}\right) \subseteq \mathbb{Q},$$

where $\left(\frac{r}{s}\right)$ is the principal ideal generated by the rational number $\frac{r}{s}$. Recall that in a commutative ring $R$ without unity, the principal ideal generated by $x \in R$ is $(x) =$

$\{r \cdot x + nx : r \in R, n \in \mathbb{Z}\}$. Now note that $\mathcal{M} \subsetneq J$ as $\frac{r}{s} \notin \mathcal{M}$. Otherwise if $\mathcal{M} = J$, then

$$\frac{r}{s} = 0 + \frac{r}{s} \in J = \mathcal{M},$$

a contradiction. We assumed that $\mathcal{M}$ is a maximal ideal, so $J = \mathbb{Q}$. Therefore any rational number $\frac{x}{y}$ can be written as

$$\frac{x}{y} = m + r' \cdot \frac{r}{s} + n\frac{r}{s} = m + n\frac{r}{s}, \text{ where } m \in \mathcal{M}, r' \in \mathbb{Q}, n \in \mathbb{Z}.$$

We observed that $\mathcal{M} \neq 0$, therefore we can choose a non-zero rational number $\frac{a}{b} \in \mathcal{M}$ with $b > 0$. Since $\mathcal{M}$ is an ideal,

$$a = \underbrace{\frac{a}{b} + \cdots + \frac{a}{b}}_{b \text{ times}} \in \mathcal{M}.$$

Now we consider the number $\frac{r}{sas} \in \mathbb{Q}$. Therefore there exist $m' \in \mathcal{M}$ and $n' \in \mathbb{Z}$ such that

$$\frac{r}{sas} = m' + n'\frac{r}{s}.$$

So we obtain,

$$\frac{r}{s} = \underbrace{\frac{r}{sas} + \cdots + \frac{r}{sas}}_{as \text{ times}} = asm' + asn'r \in \mathcal{M},$$

as $a, m' \in \mathcal{M}$ and $sn'r \in \mathbb{Z}$. This is a contradiction as $r/s \notin \mathcal{M}$ and hence $\mathcal{M}$ can not be a maximal ideal.

### 3.12.4 Prime ideals

**Definition 3.12.29** Let $R$ be a non-zero ring and $\mathcal{P}$ be an ideal of $R$. We say that $\mathcal{P}$ is a **prime ideal** of $R$ if the following hold:

1. $\mathcal{P} \neq R$,
2. if there exist ideals $\mathcal{I}, \mathcal{J}$ of $R$ such that $\mathcal{I}\mathcal{J} \subseteq \mathcal{P}$, then $\mathcal{I} \subseteq \mathcal{P}$ or $\mathcal{J} \subseteq \mathcal{P}$.

**Theorem 3.12.30** Let $R$ be a commutative ring with unity. Then a proper ideal $\mathcal{P}$ is prime ideal if and only if for any $a, b \in R$ whenever $a \cdot b \in \mathcal{P}$ then $a \in \mathcal{P}$ or $b \in \mathcal{P}$.

*Proof.* First suppose that $\mathcal{P}$ be a proper ideal of $R$ such that for $a, b \in R$ with $a \cdot b \in \mathcal{P}$ implies that $a \in \mathcal{P}$ or $b \in \mathcal{P}$. We show that $\mathcal{P}$ is a prime ideal. Let $\mathcal{I}, \mathcal{J}$ be two ideals

of $R$ such that $\mathcal{I}\mathcal{J} \subseteq \mathcal{P}$. Suppose that $\mathcal{I} \nsubseteq \mathcal{P}$. Therefore there exists an element $a \in \mathcal{I}$ such that $a \notin \mathcal{P}$. Now for any $b \in \mathcal{J}$, we have $a \cdot b \in \mathcal{I}\mathcal{J} \subseteq \mathcal{P}$. Thus $a \in \mathcal{P}$ or $b \in \mathcal{P}$. Since $a \notin \mathcal{P}$, we get $b \in \mathcal{P}$. This is true for all $b \in \mathcal{J}$. This proves that $\mathcal{J} \subseteq \mathcal{P}$ and hence $\mathcal{P}$ is a prime ideal.

Next suppose that $\mathcal{P}$ is a prime ideal of $R$. Let $a, b \in R$ be such that $a \cdot b \in \mathcal{P}$. We want to show that $a \in \mathcal{P}$ or $b \in \mathcal{P}$. We have $a \cdot b \in \mathcal{P}$, therefore $(a \cdot b) := R(a \cdot b) \subseteq \mathcal{P}$. Since $R$ is commutative, we get $(a \cdot b) = (a)(b)$. Thus we obtain,

$$(a)(b) \subseteq \mathcal{P}.$$

Since $\mathcal{P}$ is a prime ideal we can conclude that $(a) \subseteq \mathcal{P}$ or $(b) \subseteq \mathcal{P}$. Therefore $a \in \mathcal{P}$ or $b \in \mathcal{P}$.

**Examples 3.12.31**

- In $\mathbb{Z}$, the trivial ideal $\{0\}$ is a prime ideal. Let $a \cdot b \in \{0\}$. Now $a \cdot b = 0$ implies that $a = 0$ or $b = 0$ as $\mathbb{Z}$ is an integral domain. So $a \in \{0\}$ or $b \in \{0\}$.
- In $\mathbb{Z}$, for any prime number $p$, $p\mathbb{Z}$ is a prime ideal. Let $a \cdot b \in p\mathbb{Z}$. Therefore $a \cdot b = p \cdot m$ for some $m \in \mathbb{Z}$ i.e. $p \mid (a \cdot b)$. Since $p$ is a prime, we get $p \mid a$ or $p \mid b$. Therefore $a \in p\mathbb{Z}$ or $b \in p\mathbb{Z}$.

**Theorem 3.12.32** Let $R$ be a commutative ring with unity. Then every maximal ideal of $R$ is also a prime ideal.

*Proof.* Let $M$ be a maximal ideal in $R$. Let $a, b \in R$ be such that $a \cdot b \in M$. We want to show that $a \in M$ or $b \in M$. Suppose that $a \notin M$. So $M \subsetneq (a) + M$. Since $M$ is a maximal ideal, we get $(a) + M = R$. Now $R$ has unity so we can write $1 = r \cdot a + m$ for some $r \in R$ and $m \in M$. Therefore,

$$b = 1 \cdot b = r \cdot (a \cdot b) + m \cdot b \in M, \text{ as } a \cdot b, m \in M.$$

This proves that $M$ is a prime ideal.

**Remark 3.12.33** The converse of the above theorem is not true. For example in $\mathbb{Z}$, the trivial ideal $\{0\}$ is a prime ideal but not a maximal ideal.

**Theorem 3.12.34** In a Boolean ring with unity, every prime ideal is maximal ideal.

*Proof.* Let $R$ be a Boolean ring with unity and $P$ be a prime ideal in $R$. We show that $P$ is a maximal ideal. Let $J$ be an ideal of $R$ such that

$$P \subsetneq J \subseteq R.$$

We show that $J = R$. Since $P \subsetneq J$, we get there exists $a \in J$ such that $a \notin P$. Now $a^2 = a$ implies that $a \cdot (a-1) = 0$. Therefore $a \cdot (a-1) \in P$ as $0 \in P$. Recall that every Boolean ring is commutative. Now since $P$ is a prime ideal and $a \notin P$, we get that $(a-1) \in P$. Therefore $(a-1) \in J$. Thus $a - (a-1) = 1 \in J$. This proves that $J = R$ and hence $P$ is a maximal ideal.

## 3.13 Quotient rings

Let $R$ be a ring and $I$ be an ideal of the ring $R$. Then consider the set,

$$R/I := \{a + I : a \in R\}.$$

We define the operations $+, \cdot$ on $R/I$ using the operations of $R$: for $a + I, b + I \in R/I$,

$$(a + I) + (b + I) := (a + b) + I, \quad (a + I) \cdot (b + I) := a \cdot b + I.$$

With these two operations, $(R/I, +, \cdot)$ is a ring called the **quotient ring** of $R$ modulo the ideal $I$.
Since $I$ is an additive subgroup of $(R, +)$, we know that the additive operation on $R/I$ is well-defined and $(R/I, +)$ is a group. Here the element $I$ is the additive identity. Since $(R, +)$ is commutative, we have $(R/I, +)$ is a commutative group. Now we show that the operation '$\cdot$' on $R/I$ is well-defined. Let $a + I = a' + I$ and $b + I = b' + I$. Therefore,

$$a - a',\ b - b' \in I.$$

Now,

$$a \cdot b - a' \cdot b' = a \cdot b - a' \cdot b + a' \cdot b - a' \cdot b' = (a - a') \cdot b + a' \cdot (b - b').$$

Since $I$ is an ideal we get, $a \cdot b - a' \cdot b' \in I$ i.e. $a \cdot b + I = a' \cdot b' + I$. This proves that the operation '$\cdot$' is well-defined on $R/I$. It is easy to note that $R/I$ is closed under the operation '$\cdot$'. Also as $\cdot$ is associative on $R$ we get the associativity of '$\cdot$' on $R/I$. This

completes the proof that $(R/I, +, \cdot)$ is a ring.

**Exercise 3.13.1** If $R$ is commutative, show that $R/I$ is commutative.

**Remark 3.13.2** If $R/I$ is commutative, then $R$ need not be commutative. For example, consider a non-commutative ring $R$ and the trivial ideal $I = R$. In this case $R/I = \{0\}$, which is commutative whereas $R$ is not commutative.

**Exercise 3.13.3** If $R$ has unity, show that $R/I$ has unity.

**Remark 3.13.4** If $R/I$ has unity, then also $R$ may not have unity. For example, consider a ring $R$ without unity and the trivial ideal $I = R$. In this case $R/I = \{0\}$. Therefore $R/I$ has unity which is $0 = 1$ but $R$ does not have unity.

## 3.14 Characterisation of prime and maximal ideals with quotient rings

The theorems illustrated here tells us the connection of a prime ideal and a maximal ideal of a commutative ring $R$ with unity with the associated quotient rings respectively.

**Theorem 3.14.1** Let $R$ be a commutative ring with unity. Then an ideal $\mathcal{P}$ of $R$ is a prime ideal if and only if the quotient ring $R/\mathcal{P}$ is an integral domain.

*Proof.* First we assume that $\mathcal{P}$ is a prime ideal of $R$. We show that $R/\mathcal{P}$ is an integral domain. We already have $R/\mathcal{P}$ is a commutative ring as $R$ is a commutative ring. Also since $\mathcal{P}$ is a prime ideal, we have $\mathcal{P}$ is a proper ideal of $R$ and hence $R/\mathcal{P}$ is a non-zero ring. We only need to show now that $R/\mathcal{P}$ has no non-zero zero divisor. Let $a+\mathcal{P}, b+\mathcal{P} \in R/\mathcal{P}$ be such that $(a+\mathcal{P})\cdot(b+\mathcal{P}) = \mathcal{P}$. This implies that $a\cdot b+\mathcal{P} = \mathcal{P}$ i.e. $a\cdot b \in \mathcal{P}$. Since $\mathcal{P}$ is a prime ideal in a commutative ring with unity, we can conclude that $a \in \mathcal{P}$ or $b \in \mathcal{P}$. Therefore we obtain that $a+\mathcal{P} = \mathcal{P}$ or $b+\mathcal{P} = \mathcal{P}$. This proves that $R/\mathcal{P}$ is an integral domain.

Next we assume that $R/\mathcal{P}$ is an integral domain. We show that $\mathcal{P}$ is a prime ideal of $R$. We have $R/\mathcal{P}$ is a non-zero ring, so $\mathcal{P}$ is a proper ideal of $R$. Let $a, b \in R$ be such that $a\cdot b \in \mathcal{P}$. Therefore $a\cdot b+\mathcal{P} = \mathcal{P}$ in $R/\mathcal{P}$. Therefore,

$$(a+\mathcal{P})\cdot(b+\mathcal{P}) = \mathcal{P}.$$

Now since $R/\mathcal{P}$ is an integral domain, $(a+\mathcal{P}) = \mathcal{P}$ or $(b+\mathcal{P}) = \mathcal{P}$. This implies that $a \in \mathcal{P}$ or $b \in \mathcal{P}$. Hence $\mathcal{P}$ is a prime ideal of $R$.

**Corollary 3.14.2** We have already seen that in the ring of integers $\mathbb{Z}$, the ideals $\{0\}$, $p\mathbb{Z}$, where $p$ is prime number, are the prime ideals. We know that all the ideals of $\mathbb{Z}$ are of the form $n\mathbb{Z}$ where $n \in \mathbb{Z}$. So since $\mathbb{Z}/\{0\}$ is an integral domain, we immediately conclude that $\{0\}$ is a prime ideal of $\mathbb{Z}$. Now for any $n \in \mathbb{N}$, $\mathbb{Z}/n\mathbb{Z}$ is an integral domain if and only if $n$ is a prime number. Therefore we get that $p\mathbb{Z}$ for $p$ a prime number, are the only non-zero prime ideals of $\mathbb{Z}$. This establishes that the set of all non-zero prime numbers and the set of all non-zero prime ideals of $\mathbb{Z}$ are in bijective correspondence.

**Theorem 3.14.3** Let $R$ be a commutative ring with unity. Then an ideal $\mathcal{M}$ of $R$ is a maximal ideal if and only if the quotient ring $R/\mathcal{M}$ is a field.

*Proof.* First we assume that $\mathcal{M}$ is a maximal ideal of $R$. We show that $R/\mathcal{M}$ is a field. As $R$ is a commutative ring with unity and $\mathcal{M} \subsetneq R$, we have $R/\mathcal{M}$ is a non-zero commutative ring with unity. We only need to show that every non-zero element of $R/\mathcal{M}$ has an inverse in the ring $R/\mathcal{M}$. Let $a+\mathcal{M} \in R/\mathcal{M}$ be such that $a \notin \mathcal{M}$. We consider the ideal $J := (a) + \mathcal{M}$ in $R$. Now note that $\mathcal{M} \subsetneq J$ as $a \notin \mathcal{M}$ because if $\mathcal{M} = J$, then $a = a + 0 \in (a) + \mathcal{M} = J = \mathcal{M}$, a contradiction. Therefore $J = R$ as $\mathcal{M}$ is a maximal ideal of $R$. Thus we have,

$$1 = r \cdot a + m, \text{ for some } r \in R, m \in \mathcal{M}.$$

So,

$$1 + \mathcal{M} = r \cdot a + m + \mathcal{M} = (r + \mathcal{M}) \cdot (a + \mathcal{M}).$$

This proves that $r + \mathcal{M} \in R/\mathcal{M}$ is the inverse of $a + \mathcal{M}$ in $R/\mathcal{M}$. Hence $R/\mathcal{M}$ is a field.

Next we assume that $R/\mathcal{M}$ is a field. We show that $\mathcal{M}$ is a maximal ideal of $R$. We have $R/\mathcal{M}$ is a field, therefore $R/\mathcal{M}$ is a non-zero ring. Therefore $\mathcal{M} \subsetneq R$. Let $J$ be an ideal of $R$ such that $\mathcal{M} \subseteq J \subseteq R$. Suppose that $\mathcal{M} \subsetneq J$. We need to show that $J = R$. Since $\mathcal{M} \subsetneq J$, there exists $a \in J$ such that $a \notin \mathcal{M}$. Consider $a + \mathcal{M} \in R/\mathcal{M}$. Clearly $a + \mathcal{M} \neq \mathcal{M}$ as $a \notin \mathcal{M}$. Since $R/\mathcal{M}$ is a field, there exists $b + \mathcal{M} \neq \mathcal{M}$ such that

$$(a + \mathcal{M}) \cdot (b + \mathcal{M}) = 1 + \mathcal{M}, \text{ where } 1 + \mathcal{M} \text{ is the unity of } R/\mathcal{M}.$$

This implies that $a \cdot b - 1 \in \mathcal{M} \subsetneq J$. Therefore $a \cdot b - (a \cdot b - 1) = 1 \in J$ as $a \cdot b, a \cdot b - 1 \in J$. Thus $J = R$. This proves that $\mathcal{M}$ is a maximal ideal of $R$.

**Corollary 3.14.4** In the ring of integers $\mathbb{Z}$, the ideals $p\mathbb{Z}$, where $p$ is prime number, are the maximal ideals. Since $\mathbb{Z}/\{0\}$ is not a field, we immediately conclude that $\{0\}$ is not a maximal ideal of $\mathbb{Z}$. For any $n \in \mathbb{N}$, $\mathbb{Z}/n\mathbb{Z}$ is a field if and only if $n$ is a prime number. Thus we get that $p\mathbb{Z}$, for $p$ a prime number, are the maximal of $\mathbb{Z}$. This establishes that the set of all maximal ideals of $\mathbb{Z}$ and the set of all prime numbers are in bijective correspondence.

**Corollary 3.14.5** We have seen that in a non-zero commutative ring with unity, every maximal ideal is also a prime ideal. We can derive this as an application of the above two theorems also. Let $\mathcal{M}$ be a maximal ideal of a non-zero commutative ring $R$ with unity. Therefore $R/\mathcal{M}$ is a field and hence an integral domain. Therefore $\mathcal{M}$ is a prime ideal of $R$.

## 3.15 Homomorphism of rings

**Definition 3.15.1** Let $R$ and $S$ be two rings. A map $f : R \to S$ is called a **ring homomorphism** or a **homomorphism of rings from** $R$ **to** $S$ if for all $a, b \in R$ we have,

1. $f(a + b) = f(a) + f(b)$,
2. $f(a \cdot b) = f(a) \cdot f(b)$.

If a homomorphism is injective, then it is called a **monomorphism**. If a homomorphism is surjective, then it is called an **epimorphism**. If a homomorphism is bijective, then it is called an **isomorphism**.

**Remark 3.15.2** Let $R, S$ be two rings, then $f : R \to S$ defined by $f(r) = 0_S$ for all $r \in R$ is a homomorphism of rings. This ring homomorphism is called the **trivial ring homomorphism** from $R$ to $S$.

**Examples 3.15.3**

1. Consider the rings $\mathbb{Z}$ and $\mathbb{Z}/n\mathbb{Z}$ for some integer $n \in \mathbb{N}$. Define $f : \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ by $f(a) := [a]$, where $[a]$ denotes the congruence class of the integer $a$ modulo $n$. As we know that $[a + b] = [a] + [b]$ and $[ab] = [a][b]$ for all $a, b \in \mathbb{Z}$, we can conclude that $f$ is a ring homomorphism from $\mathbb{Z}$ to $\mathbb{Z}/n\mathbb{Z}$.

2. Consider the rings $\mathcal{C}([0,1],\mathbb{R}),\mathbb{R}$. Define $\Phi : \mathcal{C}([0,1],\mathbb{R}) \to \mathbb{R}$ by $\Phi(f) := f(0)$ for all $f \in \mathcal{C}([0,1],\mathbb{R})$. Note that for any $f, g \in \mathcal{C}([0,1],\mathbb{R})$,

$$\Phi(f+g) = (f+g)(0) = f(0) + g(0) = \Phi(f) + \Phi(g),$$

and

$$\Phi(fg) = (fg)(0) = f(0)g(0) = \Phi(f)\Phi(g).$$

Therefore $\Phi$ is a ring homomorphism from $\mathcal{C}([0,1]\mathbb{R})$ to $\mathbb{R}$.

**Remark 3.15.4** The map $\Phi$ defined above from $\mathcal{C}([0,1],\mathbb{R})$ to $\mathbb{R}$ is surjective. Let $a \in \mathbb{R}$. Define $f_a : [0,1] \to \mathbb{R}$ by $f_a(x) := a$ for all $x \in [0,1]$. Note that $f_a \in \mathcal{C}([0,1],\mathbb{R})$ as constant functions are continuous. Now $\Phi(f_a) = f_a(0) = a$. This proves that $\Phi$ is surjective.

**Exercise 3.15.5** Let $R, S$ be two rings and $f$ be a ring homomorphism from $R$ to $S$. Show that,

1. $f(0_R) = 0_S$ where $0_R, 0_S$ denote the additive identities of $R, S$ respectively.
2. $f(-a) = -f(a)$ for all $a \in R$.
3. $f(a-b) = f(a) - f(b)$ for all $a, b \in R$.

**Exercise 3.15.6** Let $R, S$ be two rings and $f$ be a ring homomorphism from $R$ to $S$. Show that $\mathrm{Im} f := f(R) := \{f(a) : a \in R\}$ is a subring of $S$.

**Remarks 3.15.7** Let $R, S$ be two rings and $f$ be a ring homomorphism from $R$ to $S$.

- Suppose that $R$ has unity and denote the unity of $R$ by $1_R$, then $f(1_R)$ is the unity of the subring $f(R)$ of the ring $S$. Let $x \in f(R)$. Then there exists $y \in R$ such that $f(y) = x$. Now $f(1_R) \in f(R)$. Now,

$$x \cdot f(1_R) = f(y) \cdot f(1_R) = f(y \cdot 1_R) = f(y) = x.$$

Similarly we have,

$$f(1_R) \cdot x = f(1_R) \cdot f(y) = f(1_R \cdot y) = f(y) = x.$$

Therefore $f(1_R)$ is the unity of the subring $f(R)$ of the ring $S$.

- Even if $R$ and $S$ both have unity, the unity of $f(R)$ need not be same as the unity of the ring $S$. Consider the ring homomorphism $f : \mathbb{Z} \to \mathbb{Z} \times \mathbb{Z}$ defined by $f(n) := (n, 0)$ for all $n \in \mathbb{Z}$. Here the unity of the subring $f(\mathbb{Z})$ is $(1, 0)$ whereas the unity of $\mathbb{Z} \times \mathbb{Z}$ is $(1, 1)$.

- If $R$ and $S$ both have unity and $f$ is surjective, then $f(1_R) = 1_S$. Since $f$ is surjective, for $1_S \in S$, we have $r \in R$ such that $f(r) = 1_S$. Now $r = 1_R \cdot r$. So,

$$1_S = f(r) = f(1_R \cdot r) = f(1_R) \cdot f(r) = f(1_R) \cdot 1_S = f(1_R).$$

**Exercise 3.15.8** Let $R, S$ be two rings and $f$ be a ring homomorphism from $R$ to $S$. Show that if $R$ is commutative then $f(R)$ is also commutative.

**Definition 3.15.9** Let $R, S$ be two rings and $f$ be a ring homomorphism from $R$ to $S$. Then the set $\{a \in R : f(a) = 0_S\}$ is called the **kernel** of the ring homomorphism $f$ and is denoted by $\mathrm{Ker} f$.

**Exercise 3.15.10** Let $R, S$ be two rings and $f$ be a ring homomorphism from $R$ to $S$. Show that $\mathrm{Ker} f$ is an ideal of the ring $R$.

**Exercise 3.15.11** Let $R, S$ be two rings and $f$ be a ring homomorphism from $R$ to $S$. Show that $f$ is injective if and only if $\mathrm{Ker} f = \{0_R\}$.

### 3.15.1 Isomorphism theorem of rings

Analogous to the isomorphism theorems in group theory we have isomorphism theorems in ring theory as well. In ring theory the role of normal subgroups of a group is replaced by the ideals of a ring.

**Theorem 3.15.12 (First isomorphism theorem)** Let $R, S$ be two rings. Let $f : R \to S$ be a ring homomorphism. Then we have,

$$R/\mathrm{Ker} f \simeq \mathrm{Im} f.$$

*Proof.* We know that $\mathrm{Ker} f$ is an ideal of $R$. So we can consider the quotient ring $R/\mathrm{Ker} f$. Define, $\psi : R/\mathrm{Ker} f \to \mathrm{Im} f$ as follows:

$$\psi(a + \mathrm{Ker} f) := f(a) \text{ for all } a \in R.$$

First we show that the map $\psi$ is well-defined. Let $a + \mathrm{Ker}f = b + \mathrm{Ker}f$. Therefore $a - b \in \mathrm{Ker}f$. This implies that $f(a - b) = 0_R$. Now as $f$ is a ring homomorphism, we get $f(a) = f(b)$, i.e. $\psi(a + \mathrm{Ker}f) = \psi(b + \mathrm{Ker}f)$ and hence $\psi$ is well-defined.

Note that for $a + \mathrm{Ker}f, b + \mathrm{Ker}f \in R/\mathrm{Ker}f$ we have,

$$\psi\left((a + \mathrm{Ker}f) + (b + \mathrm{Ker}f)\right) = \psi(a+b+\mathrm{Ker}f) = f(a+b) = f(a)+f(b) = \psi(a+\mathrm{Ker}f)+\psi(b+\mathrm{Ker}f),$$

and

$$\psi\left((a + \mathrm{Ker}f) \cdot (b + \mathrm{Ker}f)\right) = \psi(a\cdot b+\mathrm{Ker}f) = f(a\cdot b) = f(a)\cdot f(b) = \psi(a+\mathrm{Ker}f)\cdot\psi(b+\mathrm{Ker}f).$$

Therefore $\psi$ is a ring homomorphism.

For $s \in \mathrm{Im}f$, we have an element $r \in R$ such that $f(r) = s$. Now we consider the element $r + \mathrm{Ker}f \in R/\mathrm{Ker}f$. We then have,

$$\psi(r + \mathrm{Ker}f) = f(r) = s.$$

This proves that $\psi$ is surjective. Now,

$$\begin{aligned}\mathrm{Ker}\psi &:= \{a + \mathrm{Ker}f : \psi(a + \mathrm{Ker}f) = 0_S\}\\ &= \{a + \mathrm{Ker}f : f(a) = 0_S\}\\ &= \{a + \mathrm{Ker}f : a \in \mathrm{Ker}f\}\\ &= \{\mathrm{Ker}f\}.\end{aligned}$$

This implies that $\psi$ is injective. Thus we have an isomorphism map $\psi$ between the rings $R/\mathrm{Ker}f$ to the ring $\mathrm{Im}f$. Hence,

$$R/\mathrm{Ker}f \simeq \mathrm{Im}f.$$

**Corollary 3.15.13** Let $f : R \to S$ be a surjective ring homomorphism with $R$ being a commutative ring with unity. Then $S$ is an integral domain if and only if $\mathrm{Ker}f$ is a prime ideal in $R$. We have $R/\mathrm{Ker}f \simeq \mathrm{Im}f$ and then we use the fact that in a commutative ring $R$ with unity, an ideal $I$ is a prime ideal if and only if $R/I$ is an integral domain.

**Corollary 3.15.14** Let $f : R \to S$ be a surjective ring homomorphism with $R$ being a commutative ring with unity. Then $S$ is a field if and only if $\mathrm{Ker}f$ is a maximal ideal in

$R$.

**Theorem 3.15.15 (Second isomorphism theorem)** Let $A$ be a subring and $J$ an ideal of a ring $R$. Then,

$$A\Big/(A \cap J) \simeq (A + J)\Big/J,$$

where $A + J := \{a + j : a \in A, j \in J\}$.

*Proof.* Note that $J$ is an ideal of $A+J$ as $J \subseteq A+J$. We define a map $f : A \to (A+J)/J$ by

$$f(a) := a + J \text{ for all } a \in A.$$

For $a, b \in A$, we have,

$$f(a + b) = a + b + J = (a + J) + (b + J) = f(a) + f(b),$$

and

$$f(a \cdot b) = a \cdot b + J = (a + J) \cdot (b + J) = f(a) \cdot f(b).$$

Therefore $f$ is a ring homomorphism. Now let $x + J \in (A + J)/J$. Therefore we can write $x = a + b$ where $a \in A$ and $b \in J$. So $x + J = a + b + J = a + J$. Then $f(a) = a + J = x + J$. This proves that $f$ is surjective. Note that,

$$\begin{aligned} \mathrm{Ker} f &:= \{a \in A : f(a) = J\} \\ &= \{a \in A : a + J = J\} \\ &= \{a \in A : a \in J\} \\ &= A \cap J. \end{aligned}$$

Therefore by first isomorphism theorem we obtain,

$$A\Big/(A \cap J) \simeq (A + J)\Big/J.$$

**Theorem 3.15.16 (Third isomorphism theorem)** Let $I, J$ be two ideals of a ring $R$ so that $I \subseteq J$. Then,

$$(R/I)\Big/(J/I) \simeq R/J.$$

*Proof.* Define a map $f : R/I \to R/J$ as follows:

$$f(a+I) := a+J \text{ for all } a \in R.$$

Let $a+I = b+I$. Then $a-b \in I$. Since $I \subseteq J$, we get $a-b \in J$. Therefore $a+J = b+J$ and hence the map $f$ is well-defined. Note that for $a+I, b+I \in R/I$,

$$f\left((a+I)+(b+I)\right) = f(a+b+I) = a+b+J = (a+J)+(b+J) = f(a+I)+f(b+I),$$

and

$$f\left((a+I)\cdot(b+I)\right) = f(a\cdot b+I) = a\cdot b+J = (a+J)\cdot(b+J) = f(a+I)\cdot f(b+I).$$

So $f$ is a ring homomorphism. For any $r+J \in R/J$, we consider $r+I \in R/I$, then $f(r+I) = r+J$. Therefore $f$ is surjective. Also,

$$\begin{aligned}\mathrm{Ker} f &:= \{a+I \in R/I : f(a+I) = J\}\\ &= \{a+I \in R/I : a+J = J\}\\ &= \{a+I \in R/I : a \in J\}\\ &= J \cap I.\end{aligned}$$

Therefore by first isomorphism theorem we obtain,

$$(R/I)\Big/(J/I) \simeq R/J.$$

### 3.15.2 Correspondence theorem

**Theorem 3.15.17 (Correspondence theorem)** Let $R, S$ be two rings and $f : R \to S$ be a surjective ring homomorphism. Then,

1. For any ideal $I$ of $R$, $f(I)$ is an ideal of $S$.
2. For any ideal $J$ of $S$, $f^{-1}(J) := \{a \in R : f(a) \in J\}$ is an ideal of $R$ containing $\mathrm{Ker} f$.

*Proof.*

1. Note that $f(I) \neq \varnothing$ as $0_S = f(0_R) \in f(I)$. Let $a, b \in f(I)$. Then there exist $a', b' \in I$ such that $f(a') = a$ and $f(b') = b$. Now, $a - b = f(a') - f(b') = f(a' - b')$. As $I$ is an ideal, $a' - b' \in I$. Therefore $a - b \in f(I)$. Now let $s \in S$, since $f$ is surjective, there exists $r \in R$ such that $f(r) = s$. So for $a = f(a') \in f(I)$, we have $a \cdot s = f(a') \cdot f(r) = f(a' \cdot r) \in f(I)$. Similarly $s \cdot a \in f(I)$. This shows that $f(I)$ is an ideal of $S$.

2. Since $0_S \in J$ and $\mathrm{Ker} f := \{x \in R : f(x) = 0_S\}$, we have,

$$\mathrm{Ker} f \subseteq \{a \in R : f(a) \in J\} := f^{-1}(J).$$

   Let $a, b \in f^{-1}(J)$. So $f(a), f(b) \in J$. As $J$ is an ideal of $S$, we have $f(a) - f(b) \in J$, i.e. $f(a-b) \in J$. This implies that $a-b \in f^{-1}(J)$. Let $r \in R$. Then for $a \in f^{-1}(J)$, $f(r \cdot a) = f(r) \cdot f(a) \in J$ as $f(a) \in J$ and $f(r) \in S$. Therefore $r \cdot a \in f^{-1}(J)$. Similarly $a \cdot r \in f^{-1}(J)$. This shows that $f^{-1}(J)$ is an ideal of $R$.

**Corollary 3.15.18** Let $R$ be a ring and $I$ an ideal of $R$. Any ideal of $R/I$ is of the form $J/I$ where $J$ is an ideal of $R$ containing $I$.

*Proof.* Let $R$ be a ring and $I$ be an ideal of $R$. Then we can define the map $f : R \to R/I$ by $f(a) := a + I$ for all $a \in I$. This map $f$ is a surjective ring homomorphism with $\mathrm{Ker} f = I$. Let $K$ be an ideal of $R/I$. Then $f^{-1}(K)$ is an ideal of $R$ containing $I$. Call $f^{-1}(K)$ as $J$. Now we note that $K = f(J) = J/I$.

**Corollary 3.15.19** Let $R$ be a ring and $I$ an ideal of $R$. Then the ideals of $R/I$ and ideals of $R$ containing $I$ are in bijective correspondence.

*Proof.* Let $A$ be the set of all ideals of $R$ containing $I$ and $B$ be the set of all ideals of $R/I$. We define $\phi : A \to B$ by $\phi(J) = J/I$. From the previous corollary we have, $\phi$ is surjective. Now let $\phi(J_1) = \phi(J_2)$ i.e. $J_1/I = J_2/I$. Let $a \in J_1$. Then $a + I \in J_1/I = J_2/I$. So we can write $a + I = b + I$ for some $b \in J_2$. Therefore $a - b \in I \subseteq J_2$. Since $b \in J_2$, we get $a \in J_2$. Similarly $J_2 \subseteq J_1$ and hence $J_1 = J_2$. Hence $\phi$ is injective.

## 3.16 Quotient field/Field of fractions

**Definition 3.16.1** Let $R, S$ be two rings. We say that a ring $R$ is **embedded** in a ring $S$ if there exists an injective ring homomorphism $f : R \to S$. In other words we also

say, $S$ contains $R$ as an **embedding**, if there exists an injective ring homomorphism $f : R \to S$.

**Remark 3.16.2** If $f$ is an injective ring homomorphism from the ring $R$ to the ring $S$, then by the first isomorphism theorem,

$$R \simeq R/\mathrm{Ker} f \simeq f(R).$$

We know that $f(R)$ is a subring of $S$, so we can view $R$ as a subring of $S$ by identifying $R$ with $f(R)$.

### 3.16.1 Construction of the field of fraction of an integral domain

Let $R$ be an integral domain. We know that every field is an integral domain but the converse is not true. But we can always construct a field in which the integral domain $R$ can be embedded.

Denote $X := R \times R \setminus \{0\}$. We define a relation $\sim$ on $X$ as follows:

$$(a, x) \sim (b, y) \text{ if and only if } a \cdot y = b \cdot x.$$

We show that the relation $\sim$ on $X$ is an equivalence relation.

- We have, $(a, x) \sim (a, x)$ as $a \cdot x = a \cdot x$.
- If $(a, x) \sim (b, y)$ then $(b, y) \sim (a, x)$ as $a \cdot y = b \cdot x$.
- Let $(a, x) \sim (b, y)$ and $(b, y) \sim (c, z)$ i.e. we have, $a \cdot y = b \cdot x$ and $b \cdot z = c \cdot y$. We have to show that $(a, x) \sim (c, z)$ i.e. $a \cdot z = c \cdot x$. Since $b \cdot z = c \cdot y$, then $a \cdot (b \cdot z) = a \cdot (c \cdot y)$. Therefore, $b \cdot (a \cdot z) = c \cdot (a \cdot y) = c \cdot (b \cdot x) = b \cdot (c \cdot x)$, by using commutativity and associativity of $R$. Now as $R$ is an integral domain, by the cancellation law we obtain $a \cdot z = c \cdot x$. Therefore, $(a, x) \sim (c, z)$.

Therefore $\sim$ is an equivalence relation on $X$. We denote the set of all the equivalence classes on $X$ by $Q(R)$. Let us denote the equivalence class of $(a, x)$ by $a/x$. We define two binary operations $\oplus$ and $\odot$ in $Q(R)$ as follows:

$$a/x \oplus b/y := (a \cdot y + b \cdot x)/(x \cdot y),$$

$$a/x \odot b/y := (a \cdot b)/(x \cdot y).$$

First we show the operations $\oplus$ and $\odot$ are well-defined on $Q(R)$. Let,

$$a/x = a'/x' \quad \text{and} \quad b/y = b'/y'.$$

We have to show that

$$(a \cdot y + b \cdot x)/(x \cdot y) = (a' \cdot y' + b' \cdot x')(x' \cdot y') \quad \text{and} \quad (a \cdot b)(x \cdot y) = (a' \cdot b')(x' \cdot y').$$

We have $a/x = a'/x'$, i.e. $(a, x) \sim (a', x')$ and hence $a \cdot x' = a' \cdot x$. Similarly we have, $b \cdot y' = b' \cdot y$. Note that,

$$\begin{aligned}(a \cdot y + b \cdot x) \cdot (x' \cdot y') &= (a \cdot y) \cdot (x' \cdot y') + (b \cdot x) \cdot (x' \cdot y') \\ &= (a \cdot x') \cdot (y \cdot y') + (b \cdot y') \cdot (x \cdot x') \\ &= (a' \cdot x) \cdot (y \cdot y') + (b' \cdot y) \cdot (x \cdot x') \\ &= (a' \cdot y) \cdot (x \cdot y) + (b' \cdot x') \cdot (x \cdot y) \\ &= (a' \cdot y' + b' \cdot x') \cdot (x \cdot y).\end{aligned}$$

Also,

$$\begin{aligned}(a \cdot b) \cdot (x' \cdot y') &= (a \cdot x') \cdot (b \cdot y') \\ &= (a' \cdot x) \cdot (b' \cdot y) \\ &= (a' \cdot b') \cdot (x \cdot y).\end{aligned}$$

This proves that,

$$(a \cdot y + b \cdot x)(x \cdot y) = (a' \cdot y' + b' \cdot x')(x' \cdot y') \quad \text{and} \quad (a \cdot b)(x \cdot y) = (a' \cdot b')(x' \cdot y').$$

**Theorem 3.16.3** We have, $(Q(R), \oplus, \odot)$ is a field.

*Proof.* For $a/x, b/y, c/z \in Q(R)$, we have,

$$\begin{aligned}(a/x \oplus b/y) \oplus c/z &= \big((a \cdot y + b \cdot x)/(x \cdot y)\big) \oplus c/z \\ &= \big((a \cdot y + b \cdot x) \cdot z + c \cdot (x \cdot y)\big)/\big((x \cdot y) \cdot z\big) \\ &= \big((a \cdot y) \cdot z + (b \cdot z) \cdot x + (c \cdot y) \cdot x\big)/\big(x \cdot (y \cdot z)\big) \\ &= \big(a \cdot (y \cdot z) + (b \cdot z + c \cdot y) \cdot x\big)/\big(x \cdot (y \cdot z)\big) \\ &= a/x \oplus \big((b \cdot z + c \cdot y)/(y \cdot z)\big)\end{aligned}$$

$$= a/x \oplus (b/y \oplus c/z)\,.$$

This shows that $\oplus$ satisfies associativity law in $Q(R)$. Define,

$$0_{Q(R)} := 0/x \text{ for all } x \in R \setminus \{0\}.$$

If $R$ has unity then we can take $0_{Q(R)}$ as $0/1$. Now for any $a/x \in Q(R)$,

$$a/x \oplus 0/x = (a \cdot x + x \cdot 0)/(x \cdot x) = (a \cdot x)/(x \cdot x) = a/x = 0/x \oplus a/x.$$

This shows that $0_{Q(R)}$ is the identity with respect to $\oplus$. For $a/x \in Q(R)$, the element $(-a)/x = a/(-x)$ serves as the inverse of $a/x$ as

$$a/x \oplus (-a)/x = (a \cdot x - a \cdot x)/(x \cdot x) = 0/x^2 = 0_{Q(R)}.$$

Now note that for $a/x, b/y \in Q(R)$,

$$a/x \oplus b/y = b/y \oplus a/x, \text{ as } R \text{ is a commutative ring.}$$

Therefore $(Q(R), \oplus)$ is an abelian group.

For $a/x, b/y, c/z \in Q(R)$,

$$\begin{aligned}(a/x \odot b/y) \odot c/z &= (a \cdot b)/(x \cdot y) \odot c/z \\ &= \big((a \cdot b) \cdot c\big)\big/\big((x \cdot y) \cdot z\big) \\ &= (a \cdot (b \cdot c))\,\big/\big(x \cdot (y \cdot z)right\big) \\ &= a/x \odot (b \cdot c)/(y \cdot z) \\ &= a/x \odot (b/y \odot c/z)\,.\end{aligned}$$

This shows that $\odot$ satisfies associativity law. Also $\odot$ satisfies commutativity,

$$\begin{aligned}a/x \odot b/y &= (a \cdot b)/(x \cdot y) \\ &= (b \cdot a)/(y \cdot x) \\ &= b/y \odot a/x.\end{aligned}$$

We define,

$$1_{Q(R)} := x/x \text{ for all } x \in R \setminus \{0\}.$$

If $R$ has unity then we can take $1_{Q(R)}$ as $1/1$. Now,

$$a/x \odot x/x = (a \cdot x)/(x \cdot x) = a/x.$$

Therefore $(Q(R), \oplus, \odot)$ is a commutative ring with unity. Let $\alpha \neq 0_{Q(R)}$ be an element in $Q(R)$. Therefore,

$$\alpha = a/x \quad \text{for some } a, x \in R \setminus \{0\}.$$

We consider $x/a \in Q(R)$. Note that

$$a/x \odot x/a = (a \cdot x)/(x \cdot a) = 1_{Q(R)}.$$

This proves that $(Q(R), \oplus, \odot)$ is a field.

**Definition 3.16.4** Let $R$ be an integral domain. The field $Q(R)$ is called the **quotient field** or the **field of fractions** of $R$.

**Theorem 3.16.5** Let $R$ be an integral domain. Then $R$ can be embedded in the field $Q(R)$. In other words, $Q(R)$ contains an isomorphic copy of $R$ as a subring.

*Proof.* Let $a \in R\setminus\{0\}$. We define a map $f : R \to Q(R)$ by $f(r) := (a \cdot r)/a$ for all $r \in R$. Note that,

$$\begin{aligned} f(r_1 + r_2) = (a \cdot (r_1 + r_2))/a &= (a \cdot r_1 + a \cdot r_2)/a \\ &= (a \cdot (a \cdot r_1 + a \cdot r_2))/(a \cdot a) \\ &= (a \cdot r_1)/a \oplus (a \cdot r_2)/a \\ &= f(r_1) \oplus f(r_2), \end{aligned}$$

$$\begin{aligned} f(r_1 \cdot r_2) &= (a \cdot (r_1 \cdot r_2))/a \\ &= ((a \cdot r_1) \cdot (a \cdot r_2))/(a \cdot a) \\ &= (a \cdot r_1)/a \odot (a \cdot r_2)/a \\ &= f(r_1) \odot f(r_2). \end{aligned}$$

Hence $f$ is a ring homomorphism. Now,

$$\begin{aligned} \mathrm{Ker} f &:= \left\{r \in R : f(r) = 0_{Q(R)}\right\} \\ &= \{r \in R : (a \cdot r)/a = 0/a\} \end{aligned}$$

$$\begin{aligned}&= \{r \in R : (a \cdot r) \cdot a = 0 \cdot a = 0\}\\ &= \{r \in R : a \cdot r = 0\}\\ &= \{0\}.\end{aligned}$$

Hence we have, $Q(R)$ contains $R$ as an embedding.

**Theorem 3.16.6** Let $K$ be a field containing the integral domain $R$. Then $K$ contains $Q(R)$ as an embedding.

*Proof.* Let $K$ be a field such that $R \subseteq K$. We know $R \subseteq K$ and $K$ is a field, so $x^{-1} \in K$ for all $x \neq 0$ in $R$. We define a map $f : Q(R) \to K$ by

$$f(a/x) := a \cdot x^{-1} \quad \text{for all} \quad a/x \in Q(R).$$

First we show that the map $f$ is well-defined. Let $a/x = b/y \in Q(R)$. Therefore, $a \cdot y = b \cdot x$. Thus $a \cdot x^{-1} = b \cdot y^{-1}$ and hence $f$ is well-defined. Now for $a/x, b/y \in Q(R)$,

$$\begin{aligned}f(a/x \oplus b/y) &= f\left((a \cdot y + b \cdot x)/(x \cdot y)\right)\\ &= (a \cdot y + b \cdot x) \cdot (x \cdot y)^{-1}\\ &= a \cdot x^{-1} + b \cdot y^{-1}\\ &= f(a/x) + f(b/y).\end{aligned}$$

Also,

$$\begin{aligned}f(a/x \odot b/y) &= f\big((a \cdot b)/(x \cdot y)\big)\\ &= (a \cdot b) \cdot (x \cdot y)^{-1}\\ &= (a \cdot x^{-1}) \cdot (b \cdot y^{-1})\\ &= f(a/x) \cdot f(b/y).\end{aligned}$$

Therefore $f$ is a ring homomorphism. Now,

$$\begin{aligned}\mathrm{Ker} f &:= \{a/x \in Q(R) : f(a/x) = 0\}\\ &= \{a/x \in Q(R) : a \cdot x^{-1} = 0\}\\ &= \{a/x \in Q(R) : a = 0\} \text{ as } K \text{ is a field and hence an integral domain}\\ &= \{0_{Q(R)}\}.\end{aligned}$$

This proves that $K$ contains the field $Q(R)$ as an embedding.

**Exercise 3.16.7** Let $R, S$ be two integral domains such that $R \simeq S$. Show that $Q(R) \simeq Q(S)$.

**Remark 3.16.8** The field $Q(R)$ is the smallest field containing an embedding of the integral domain $R$ i.e. if any other field contains an embedding of $R$, then that contains $Q(R)$ as an embedding. If $K$is a field that contains an embedding of $R$ say $R'$, then $Q(R')$ is embedded in $K$. Since $Q(R) \simeq Q(R')$, we get that $Q(R)$ is embedded in $K$.

**Corollary 3.16.9** For any field $K$, $Q(K) = K$.

*Proof.* We know that $Q(K)$ is the smallest field containing $K$ as an embedding. Now $K$ itself is a field containing $K$, therefore $Q(K) = K$.

**Exercise 3.16.10** Show that $Q(\mathbb{Z}) = Q(2\mathbb{Z}) = \mathbb{Q}$.

**Exercise 3.16.11** Show that $Q(\mathbb{Z}[\iota]) = \mathbb{Q}(\iota)$, where $\mathbb{Q}(\iota) := \{a + \iota b : a, b \in \mathbb{Q}\}$.

## 3.17 Division in a commutative ring

**Definition 3.17.1** Let $R$ be a commutative ring and $a, b \in R$ be such that $a \neq 0$. Then we say that '$a$ **divides** $b$' if there exists an element $c \in R$ such that $b = a \cdot c$. We denote $a \mid b$. The element $a$ is called a **divisor** of $b$. If $a \neq 0$ is not a divisor of $b$ then we denote $a \nmid b$.

**Remarks 3.17.2**

- Note that $c$ is also a divisor of $b$ if $c \neq 0$.
- For all $a \neq 0$, $a \mid 0$ as $a \cdot 0 = 0$.

**Examples 3.17.3**

- In the ring $\mathbb{Z}$, $2 \mid 4$ but $4 \nmid 2$. Also $n \mid -n$ and $-n \mid n$ for all $n \in \mathbb{Z}$.
- In the ring $\mathbb{Z}/6\mathbb{Z}$, $[2] \mid [4]$ as $[4] = [2] \cdot [2]$. Note that in the ring $\mathbb{Z}/6\mathbb{Z}$, $[4] \mid [2]$ as $[2] = [8] = [2] \cdot [4]$.

**Remark 3.17.4** If $R$ is a non-zero ring with unity, then $1 \mid a$ as $a \cdot 1 = a$. If $u \in R$ is a unit then $u \mid a$ for all $a \in R$ as $a = a \cdot 1 = a \cdot (u^{-1} \cdot u) = (a \cdot u^{-1}) \cdot u$.

**Theorem 3.17.5** Let $R$ be an integral domain. Let $a \in R \setminus \{0\}$ and $b, c \in R$ be such that $b = a \cdot c$. The element $c$ is unique.

*Proof.* Let $c_1, c_2 \in R$ be such that $b = a \cdot c_1 = a \cdot c_2$. Then $a \cdot (c_1 - c_2) = 0$. As $a \neq 0$ and $R$ is an integral domain, we get $c_1 = c_2$.

### 3.17.1 Associates

**Definition 3.17.6** Let $R$ be a commutative ring and $a, b \in R \setminus \{0\}$. We say that two elements $a$ and $b$ are **associates** of each other if $a \mid b$ and $b \mid a$. We denote $a \sim b$.

**Examples 3.17.7**

- In the ring $\mathbb{Z}$, $n \sim -n$ for all $n \in \mathbb{Z}$.
- In the ring $\mathbb{Z}/6\mathbb{Z}$, $[2] \sim [4]$.

**Theorem 3.17.8** Let $R$ be a commutative ring with unity and $a, b \in R \setminus \{0\}$. Then the following are equivalent:

1. $a \mid b$.
2. $b \in (a)$.
3. $(b) \subseteq (a)$.

*Proof.* Note that $a \mid b$ implies that $b = a \cdot c$ for some $c \in R$. So $b \in (a)$. Since $b \in (a)$, we get $b \cdot r \in (a)$ for all $r \in R$. Therefore $(b) \subseteq (a)$. Now $(b) \subseteq (a)$ implies that $b \in (a)$ and thus $b = a \cdot r$ for some $r \in R$ and thus $a \mid b$.

**Corollary 3.17.9** Note that in a commutative ring with unity $R$, for $a, b \in R \setminus \{0\}$, we have $a \sim b$ if and only if $(a) = (b)$.

**Exercise 3.17.10** Let $R$ be a commutative ring with unity. We have $a \sim b$ if and only if $(a) = (b)$ for all $a, b \in R \setminus \{0\}$. Show that $\sim$ is an equivalence relation on $R \setminus \{0\}$.

**Theorem 3.17.11** Let $R$ be an integral domain with unity and $a, b \in R \setminus \{0\}$. Then $a \sim b$ if and only if there exists a unit $u \in R$ such that $b = a \cdot u$.

*Proof.* Suppose, $b = a \cdot u$ for some unit $u \in R$. Therefore $a \mid b$. Also $a = b \cdot u^{-1}$. This implies that $b \mid a$. Therefore $a \sim b$.

Next suppose that $a \sim b$. Therefore we have $a \mid b$ and $b \mid a$. Let $a = b \cdot u_1$ and $b = a \cdot u_2$ for some $u_1, u_2 \in R$. We then get,

$$b = b \cdot (u_1 \cdot u_2).$$

Since $R$ is an integral domain with unity, we obtain $u_1 \cdot u_2 = 1$. This proves that $u_1, u_2$ are units. Therefore $b = a \cdot u_2$ for some unit $u_2 \in R$.

### 3.17.2 Greatest common divisor and least common multiplier

Now following the notions of greatest common divisor and least common multiplier of two integers, we define them in a commutative ring analogously. We will see that in general, in an arbitrary commutative ring greatest common divisor and least common multiplier of two elements may not exist.

**Definition 3.17.12** Let $R$ be a commutative ring and $a, b \in R \setminus \{0\}$. An element $d \in R$ is called **a greatest common divisor** of $a$ and $b$ if the following hold:

- $d \mid a$ and $d \mid b$,
- if there exists $e \in R$ such that $e \mid a$ and $e \mid b$, then $e \mid d$.

Note that if $d$ is a greatest common divisor of $a$ and $b$, then $u \cdot d$ for any unit $u \in R$, is also a greatest common divisor of $a$ and $b$. In fact if $d_1, d_2$ are two greatest common divisors of $a$ and $b$ then from the definition it follows that $d_1 \mid d_2$ and $d_2 \mid d_1$, i.e. $d_1, d_2$ are associates of each other. So we can say that **the** greatest common divisor of any two elements is unique up to associates. We denote $d = \gcd(a, b)$.

**Remark 3.17.13** In the ring $\mathbb{Z}[\iota\sqrt{5}]$, the greatest common divisor of 6 and $2 \cdot (1 + \iota\sqrt{5})$ does not exist.

Let $(a + \iota b\sqrt{5}) \neq 0$ and $(a + \iota b\sqrt{5}) \mid 6$ where $a, b \in \mathbb{Z}$. So there exist $c, d \in \mathbb{Z}$ such that

$$(a + b\iota\sqrt{5}) \cdot (c + d\iota\sqrt{5}) = 6. \tag{3.17.1}$$

Therefore,

$$(ac - 5bd) + (ad + bc)\iota\sqrt{5} = 6.$$

This implies that,

$$(ac - 5bd) = 6 \text{ and } (ad + bc) = 0.$$

**Case I:** If $b = 0$, then $a \neq 0$ as $(a + \iota b\sqrt{5}) \neq 0$. From $(ad + bc) = 0$, we obtain that $d = 0$. Hence $ac = 6$. So when $b = 0$, the possible choices for $a$ are $\pm 1, \pm 2, \pm 3, \pm 6$.

**Case II:** If $b \neq 0$. In this case $d \neq 0$, because if $d = 0$, then $ad = 0$. Therefore $bc = 0$ as $(ad + bc) = 0$. Since $b \neq 0$, we get $c = 0$. Now $d = 0, c = 0$ implies $(c + \iota d\sqrt{5}) = 0$, which is not possible. Therefore $d \neq 0$. We write,

$$\frac{a}{b} = -\frac{c}{d} = \lambda.$$

Then substituting $a = b\lambda$ and $c = -d\lambda$ in (3.17.1), we obtain

$$6 = -bd(5 + \lambda^2). \tag{3.17.2}$$

Now note that,

$$(a - b\iota\sqrt{5}) \cdot (c - d\iota\sqrt{5}) = 6. \tag{3.17.3}$$

From (3.17.1) and (3.17.3) we have, $(a + b\iota\sqrt{5}) \mid 6$ and $(a - b\iota\sqrt{5}) \mid 6$. Thus we get, $(a^2 + 5b^2) \mid 36$. Since $(a^2 + 5b^2)$ is a positive integer, the possible choices of $(a^2 + 5b^2)$ are $1, 2, 3, 4, 6, 9, 12, 18, 36$. So for $b = \pm 1$, possible choices for $a$ are $\pm 1, \pm 2$. For $b = \pm 2$, possible choices for $a$ are $\pm 4$. For $|b| \geq 3$, we do not have any possible choice of $a$. Also from (3.17.2) we know that $\lambda^2$ can be at most 1. Therefore only possible choices are $b = \pm 1, a = \pm 1$.

Also assume, $(a + \iota b\sqrt{5}) \mid 2(1 + \iota\sqrt{5})$. Then there exist $e, f \in \mathbb{Z}$ such that

$$(a + b\iota\sqrt{5}) \cdot (e + f\iota\sqrt{5}) = 2 \cdot (1 + \iota\sqrt{5}). \tag{3.17.4}$$

Therefore,

$$(ae - 5bf) + (af + be)\iota\sqrt{5} = 2 \cdot (1 + \iota\sqrt{5}).$$

This implies that,

$$(ae - 5bf) = 2 \text{ and } (af + be) = 2. \tag{3.17.5}$$

**Case I:** If $b = 0$, then possible choices of $a$ are $\pm 1, \pm 2$.

**Case II:** If $b = 1$, then $a$ can be only 1. We can note that $a$ cannot be $-1$ as then putting $b = 1, a = -1$ in (3.17.5) we get

$$e - f = 2 \text{ and } -e - 5f = 2,$$

i.e.

$$-6f = 4, \text{ which is not possible.}$$

**Case III:** If $b = -1$, then $a$ can be $-1$ but $a$ can not be 1 as then again from (3.17.5) we will get $6f = 4$, which does not have any solution in $\mathbb{Z}$.

Therefore the possible common divisors of 6 and $2 \cdot (1+\iota\sqrt{5})$ are $\pm 1, \pm 2, \pm(1+\iota\sqrt{5})$. Now 2 and $(1+\iota\sqrt{5})$ both are common divisors but neither $2 \mid (1+\iota\sqrt{5})$ nor $(1+\iota\sqrt{5}) \mid 2$. So none of $(1+\iota\sqrt{5})$ and 2 can be the greatest common divisor of 6 and $2 \cdot (1+\iota\sqrt{5})$. Also $2, (1+\iota\sqrt{5})$ do not divide 1 in $\mathbb{Z}[\iota\sqrt{5}]$, so 1 can not be also the greatest common divisor of 6 and $2 \cdot (1+\iota\sqrt{5})$. This proves that the greatest common divisor of 6 and $2 \cdot (1+\iota\sqrt{5})$ does not exist in $\mathbb{Z}[\iota\sqrt{5}]$.

**Definition 3.17.14** Let $R$ be a commutative ring and $a, b \in R \setminus \{0\}$. An element $l \in R$ is called **a least common multiplier** of $a$ and $b$ if the following hold:

- $a \mid l$ and $b \mid l$,
- if there exists $m \in R$ such that $a \mid m$ and $b \mid m$, then $l \mid m$.

Again note that if $l$ is a least common multiplier of $a$ and $b$, then $u \cdot l$ for any unit $u \in R$ is also a least common multiplier of $a$ and $b$. In fact if $l_1, l_2$ are two least common multipliers of $a$ and $b$ then from the definition it follows that $l_1 \mid l_2$ and $l_2 \mid l_1$, i.e. $l_1, l_2$ are associates of each other. So we can say that **the** least common multiplier of any two elements is unique up to associates. We denote $l = \operatorname{lcm}(a, b)$.

**Exercise 3.17.15** Show that in the ring $\mathbb{Z}[\iota\sqrt{5}]$, the least common multiplier of 6 and $2 \cdot (1+\iota\sqrt{5})$ does not exist.

## 3.18 Irreducible and prime elements in a commutative ring with unity

Keeping the analogy of the prime numbers in the ring of integers $\mathbb{Z}$ in mind, we have the following notions of irreducible and prime elements in a commutative ring with unity.

**Definition 3.18.1** Let $R$ be a commutative ring with unity. A non-zero non-unit element $a \in R$ is called an **irreducible element** if $a = b \cdot c$ for some $b, c \in R$, then either $b$ or $c$ is a unit.

**Examples 3.18.2**

- Let $p$ be a prime number in $\mathbb{Z}$. Then $p = p \cdot 1$ and $p = (-p) \cdot (-1)$. As $1, -1$ are units in $\mathbb{Z}$, we conclude that $p$ is an irreducible element in $\mathbb{Z}$. Similarly $-p$ is also an irreducible element.

- In $\mathbb{Z}$, the element 4 is non-zero non-unit. But 4 is not irreducible as $4 = 2 \cdot 2$ and 2 is not a unit.

- In $\mathbb{Z}/6\mathbb{Z}$, consider the element $[3]$. Note that $[3]$ is non-zero and also non-unit as $\gcd(3, 6) \neq 1$. In $\mathbb{Z}/6\mathbb{Z}$, $[3] = [1] \cdot [3]$ and $[3] = [3] \cdot [5]$. Since $[1], [5]$ are units in $\mathbb{Z}/6\mathbb{Z}$, we can conclude that $[3]$ is an irreducible element in $\mathbb{Z}/6\mathbb{Z}$.

- In $\mathbb{Z}/6\mathbb{Z}$, consider the element $[2]$. Now $[2]$ is non-zero and also non-unit as $\gcd(2, 6) \neq 1$. In $\mathbb{Z}/6\mathbb{Z}$ we can write $[2] = [8] = [2] \cdot [4]$. Both the elements $[2], [4]$ are non-units in $\mathbb{Z}/6\mathbb{Z}$ and hence $[2]$ is not an irreducible element in $\mathbb{Z}/6\mathbb{Z}$.

**Definition 3.18.3** Let $R$ be a commutative ring with unity. A non-zero non-unit element $\mathfrak{p} \in R$ is called a **prime element** if $\mathfrak{p} \mid b \cdot c$ where $b, c \in R$, then $\mathfrak{p} \mid b$ or $\mathfrak{p} \mid c$.

**Examples 3.18.4**

- Let $p$ be a prime number in $\mathbb{Z}$. Note that $p, -p$ both are prime elements in $\mathbb{Z}$.

- In the ring $\mathbb{Z}/6\mathbb{Z}$ consider the element $[2]$. Note that this element is non-zero and non-unit in $\mathbb{Z}/6\mathbb{Z}$. Let $[2] \mid ([a] \cdot [b])$. Therefore $[ab] = [2k]$ for some integer $k \in \mathbb{Z}$. Thus we have $6 \mid (ab - 2k)$ in $\mathbb{Z}$. This implies that $ab - 2k = 6c$ for some integer $c$. So $ab = 2(k + 3c)$ in $\mathbb{Z}$. This implies that $2 \mid ab$ in $\mathbb{Z}$. Now since 2 is a prime in $\mathbb{Z}$, we obtain that $2 \mid a$ or $2 \mid b$. Therefore $[2] \mid [a]$ or $[2] \mid [b]$. This proves that $[2]$ is a prime element in $\mathbb{Z}/6\mathbb{Z}$.

**Remark 3.18.5** In a field every non-zero element in unit. So there is no irreducible, no prime element in a field.

**Exercise 3.18.6** Let $R$ be an integral domain with unity. Show that an associate of a prime element in $R$ is a prime element and an associate of an irreducible element is an irreducible element.

**Theorem 3.18.7** Let $R$ be a commutative ring with unity and $\mathfrak{p} \in R$. Then $\mathfrak{p}$ is a prime element in $R$ if and only if $(\mathfrak{p}) := \mathfrak{p}R$ is a non-zero prime ideal in $R$.

*Proof.* Suppose $\mathfrak{p} \in R$ is a prime element. Therefore $\mathfrak{p}$ is non-zero and non-unit and hence $(\mathfrak{p}) \neq 0$ and $(\mathfrak{p}) \subsetneq R$. Let $a \cdot b \in (\mathfrak{p})$. Then we can write,

$$a \cdot b = \mathfrak{p} \cdot c, \quad \text{for some} \quad c \in R.$$

So we get, $\mathfrak{p} \mid a \cdot b$ in $R$. Now since $\mathfrak{p}$ is a prime element, we can conclude that $\mathfrak{p} \mid a$ or $\mathfrak{p} \mid b$. Thus $a = \mathfrak{p} \cdot a'$ or $b = \mathfrak{p} \cdot b'$ for some $a', b' \in R$. This proves that $a \in (\mathfrak{p})$ or $b \in (\mathfrak{p})$. Therefore $(\mathfrak{p})$ is a non-zero prime ideal in $R$.

Next suppose that $(\mathfrak{p})$ is a non-zero prime ideal in $R$. We want to show that $\mathfrak{p}$ is a prime element in $R$. Since $(\mathfrak{p})$ is a non-zero prime ideal, we can immediately conclude that $\mathfrak{p}$ is non-zero and non-unit in $R$. Let $\mathfrak{p} \mid a \cdot b$ in $R$. So,

$$a \cdot b = \mathfrak{p} \cdot c, \quad \text{for some} \quad c \in R.$$

This implies that $a \cdot b \in (\mathfrak{p})$ and since $(\mathfrak{p})$ is a prime ideal we get $a \in (\mathfrak{p})$ or $b \in (\mathfrak{p})$. Therefore we obtain, $\mathfrak{p} \mid a$ or $\mathfrak{p} \mid b$. This completes the proof that $\mathfrak{p}$ is a prime element in $R$.

**Theorem 3.18.8** Let $R$ be an integral domain with unity and $\mathfrak{p} \in R$ be a prime element. Then $\mathfrak{p}$ is an irreducible element in $R$.

*Proof.* Since $\mathfrak{p}$ is prime, we already have $\mathfrak{p}$ is non-zero and non-unit. Let $\mathfrak{p} = a \cdot b$. Then $\mathfrak{p} \mid (a \cdot b)$. So we get, $\mathfrak{p} \mid a$ or $\mathfrak{p} \mid b$. Therefore we have,

$$a = \mathfrak{p} \cdot a' \quad \text{or} \quad b = \mathfrak{p} \cdot b', \text{ for some } \quad a', b' \in R.$$

Without loss of generality let $a = \mathfrak{p} \cdot a'$. Then $\mathfrak{p} = \mathfrak{p} \cdot (a' \cdot b)$. This implies that $a' \cdot b = 1$ as $R$ is an integral domain. Hence $b$ is a unit. If $b = \mathfrak{p} \cdot b'$, then similarly we can conclude that $a$ is a unit. Therefore $\mathfrak{p}$ is an irreducible element.

**Remarks 3.18.9**

- In general the statement is not true unless $R$ is an integral domain. For example $\mathbb{Z}/6\mathbb{Z}$ is not an integral domain and we have seen that the element $[2] \in \mathbb{Z}/6\mathbb{Z}$ is prime but not irreducible.
- In $\mathbb{Z}$, the prime and irreducible elements are same.
- In general, the converse of the above theorem is not true i.e. there are integral domains in which every irreducible need not be prime. For example consider the

element 3 in the integral domain $\mathbb{Z}[\iota\sqrt{5}]$. We can show that 3 is irreducible but not prime in $\mathbb{Z}[\iota\sqrt{5}]$. First let us find out all the units in this integral domain. Let $(a + b\iota\sqrt{5})$ be a unit. So there exist $c, d \in \mathbb{Z}$ such that

$$(a + b\iota\sqrt{5}) \cdot (c + d\iota\sqrt{5}) = 1.$$

Then,

$$(a - b\iota\sqrt{5}) \cdot (c - d\iota\sqrt{5}) = 1.$$

Thus,

$$(a^2 + 5b^2)(c^2 + 5d^2) = 1.$$

Since $a, b, c, d \in \mathbb{Z}$, we obtain from the above equation that $a = \pm 1$ and $b = 0$. Therefore $1, -1$ are the only units in $\mathbb{Z}[\iota\sqrt{5}]$. Now we show that 3 is irreducible. Note that 3 is non-zero and 3 is also not a unit in $\mathbb{Z}[\iota\sqrt{5}]$. Suppose,

$$3 = (a + b\iota\sqrt{5}) \cdot (c + d\iota\sqrt{5}) \text{ for some } a, b, c, d \in \mathbb{Z}.$$

Then,

$$3 = (a - b\iota\sqrt{5}) \cdot (c - d\iota\sqrt{5}).$$

Therefore we get,

$$9 = (a^2 + 5b^2)(c^2 + 5d^2).$$

So $(a^2 + 5b^2)$ can be 1 or 9 as $(a^2 + 5b^2) > 0$ and $a, b \in \mathbb{Z}$. If $(a^2 + 5b^2) = 1$, then $a = \pm 1$, $b = 0$, and hence $(a + b\iota\sqrt{5})$ is a unit. If $(a^2 + 5b^2) = 9$, then $(c^2 + 5d^2) = 1$. Therefore $c = \pm 1$ and $d = 0$ and hence $(c + d\iota\sqrt{5})$ is a unit. This proves that 3 is an irreducible element. Now we show that 3 is not a prime element in $\mathbb{Z}[\iota\sqrt{5}]$. As $6 = 3 \cdot 2$, we can conclude that $3 \mid 6$. Now in $\mathbb{Z}[\iota\sqrt{5}]$, we can write,

$$6 = (1 + \iota\sqrt{5}) \cdot (1 - \iota\sqrt{5}).$$

Note that $3 \nmid (1 + \iota\sqrt{5})$ in $\mathbb{Z}[\iota\sqrt{5}]$. Because if $3 \mid (1 + \iota\sqrt{5})$, then there exist $u, v \in \mathbb{Z}$ such that

$$3 \cdot (u + v\iota\sqrt{5}) = (1 + \iota\sqrt{5}).$$

This implies that $3u = 1$ and $3v = 1$, which is not possible as $u, v \in \mathbb{Z}$. Similarly $3 \nmid (1 - \iota\sqrt{5})$ in $\mathbb{Z}[\iota\sqrt{5}]$. Therefore 3 is not a prime element in $\mathbb{Z}[\iota\sqrt{5}]$.

**Exercise 3.18.10** Show that the elements $2, (1 + \iota\sqrt{5}), (1 - \iota\sqrt{5})$ are irreducible but

not prime in the integral domain $\mathbb{Z}[\iota\sqrt{5}]$.

## 3.19 Principal ideal domains

**Definition 3.19.1** An integral domain $R$ is called a **principal ideal domain** (in short PID) if every ideal in $R$ is a principal ideal.

**Examples 3.19.2**

- The ring $\mathbb{Z}$ is a PID as every ideal in $\mathbb{Z}$ is of the form $n\mathbb{Z}$ where $n \in \mathbb{Z}$.
- All fields are principal ideal domains. Let $F$ be a field. Then every non-zero element of $F$ is a unit. So $F$ has only trivial ideals which are principal ideals $(0)$ and $F = (1)$.
- Subring of a PID need not be a PID. For example consider $Z[\iota\sqrt{5}]$ as a subring of $\mathbb{Q}(\iota\sqrt{5})$. Here $Z[\iota\sqrt{5}]$ is not a PID but $\mathbb{Q}(\iota\sqrt{5})$ is PID.

**Theorem 3.19.3** Let $R$ be a PID with unity and $\mathfrak{p} \in R$ be a non-zero non-unit. Then $\mathfrak{p}$ is irreducible if and only if $\mathfrak{p}$ is prime.

*Proof.* We have already seen that in every integral domain with unity every prime is irreducible, so therefore in a PID with unity it holds true.

Now let us assume that $\mathfrak{p} \in R$ is irreducible. We show that $\mathfrak{p}$ is prime. Let $\mathfrak{p} \mid a \cdot b$. Therefore,

$$a \cdot b = \mathfrak{p} \cdot c \quad \text{for some} \quad c \in R.$$

Let $\mathfrak{I} =:= (a) + (\mathfrak{p}) \subseteq R$. Since $R$ is a PID, there exists $d \in R$ such that $\mathfrak{I} = (d)$. Therefore $\mathfrak{p} \in (d)$ and so $\mathfrak{p} = d \cdot r$ for some $r \in R$. Since $\mathfrak{p}$ is irreducible, we get either $d$ or $r$ is a unit.

If $d$ is a unit then $\mathfrak{I} = R$ and so $(a) + (\mathfrak{p}) = R$. We know $R$ has unity, thus there exists $r_1, r_2 \in R$ so that $1 = r_1 \cdot a + r_2 \cdot \mathfrak{p}$. So,

$$\begin{aligned} b &= r_1 \cdot (a \cdot b) + (r_2 \cdot b) \cdot \mathfrak{p} \\ &= \mathfrak{p} \cdot (r_1 \cdot c) + (r_2 \cdot b) \cdot \mathfrak{p}, \quad \text{as} \quad a \cdot b = \mathfrak{p} \cdot c \\ &= \mathfrak{p} \cdot (r_1 \cdot c + r_2 \cdot b). \end{aligned}$$

Note that $(r_1 \cdot c + r_2 \cdot b) \in R$, therefore $\mathfrak{p} \mid b$.

If $r$ is a unit then, $\mathfrak{p} \cdot r^{-1} = d$. This implies that $d \in (\mathfrak{p}) \subseteq \mathfrak{I}$. Therefore,

$$\mathfrak{I} = (d) \subseteq (\mathfrak{p}) \subseteq \mathfrak{I}.$$

Hence we get, $\mathfrak{I} = (\mathfrak{p})$. Therefore $a \in (\mathfrak{p})$, and hence $\mathfrak{p} \mid a$.

This completes the proof that $\mathfrak{p}$ is a prime element in $R$.

**Corollary 3.19.4** The integral domain $\mathbb{Z}[\iota\sqrt{5}]$ is not a PID.

**Theorem 3.19.5** Let $R$ be a PID with unity and $a \in R$. The following are equivalent:

1. The element $a$ is irreducible.
2. The ideal $(a)$ is a non-zero maximal ideal.
3. The ideal $(a)$ is a non-zero prime ideal.
4. The element $a$ is prime.

*Proof.*

1) $\Rightarrow$ 2) First we assume that $a$ is irreducible. We want to show that the ideal $(a)$ is a non-zero maximal ideal. Since $a \neq 0$, we have $(a) \neq 0$. As $R$ has unity and $a$ is a non-unit, $(a) \subsetneq R$. Let $J$ be an ideal of $R$ such that $(a) \subseteq J \subsetneq R$. Since $R$ is a PID, we can find $b \in R$ such that $J = (b)$. Now $(a) \subseteq (b)$ implies that $a \in (b)$ and therefore $a = b \cdot c$ where $c \in R$. Since $a$ is irreducible either $b$ or $c$ is a unit. Since $J \subsetneq R$, the element $b$ is not a unit. Thus we get, $c$ is a unit. Hence $b = a \cdot c^{-1} \in (a)$ and so $(b) \subseteq (a)$. This proves that $J = (a)$, and hence $(a)$ is a non-zero maximal ideal.

2) $\Rightarrow$ 3) Since $R$ is a PID with unity, it is a non-zero commutative ring with unity. We have already seen that in a non-zero commutative ring with unity, every maximal ideal is a prime ideal. Therefore as $(a)$ is a non-zero maximal ideal, it is also a non-zero prime ideal.

3) $\Rightarrow$ 4) We have seen that in a commutative ring with unity a principal ideal $(a)$ is a non-zero prime ideal if and only if $a$ is a prime element. So therefore $(a)$ is a non-zero prime ideal implies that the element $a$ is a prime.

4) $\Rightarrow$ 1) We know in an integral domain with unity every prime element is irreducible. So $a$ is prime in $R$ implies that $a$ is an irreducible element.

**Theorem 3.19.6** Let $R$ be a PID with unity and $a, b \in R \setminus \{0\}$. Then $\gcd(a, b)$ exists and if $d = \gcd(a, b)$ then there exist $x, y \in R$ such that $d = a \cdot x + b \cdot y$.

*Proof.* Let $\mathfrak{I} := (a) + (b)$. Since $R$ is a PID, there exists an element $d \in R$ such that $\mathfrak{I} = (d)$. We show that $d$ is the gcd of $a$ and $b$ up to associates. Note that,

$$a = a \cdot 1 + b \cdot 0 \in (d) \quad \text{and} \quad b = a \cdot 0 + b \cdot 1 \in (d).$$

This proves that $d \mid a$ and $d \mid b$. Let $c \in R$ be such that $c \mid a$ and $c \mid b$. Therefore $(a) \subseteq (c)$ and $(b) \subseteq (c)$. Hence,

$$(d) = (a) + (b) \subseteq (c).$$

So $d \in (c)$ and hence $c \mid d$. This proves that $d = \gcd(a, b)$. This completes the proof that $\gcd(a, b)$ exists in $R$. Also note that $(d) = (a) + (b)$ implies that $d \in (a) + (b)$, and hence we can find $x, y \in R$ such that $d = a \cdot x + b \cdot y$.

**Theorem 3.19.7** Let $R$ be a PID with unity and $a, b \in R \setminus \{0\}$. Then $\operatorname{lcm}(a, b)$ exists.

*Proof.* In this case let $\mathfrak{I} := (a) \cap (b)$. Since $R$ is a PID, there exists an element $l \in R$ such that $\mathfrak{I} = (l)$. We show that $l$ is the lcm of $a$ and $b$ up to associates. Note that, $l \in (a) \cap (b)$ and thus $a \mid l$ and $b \mid l$. Let $m \in R$ be such that $a \mid m$ and $b \mid m$. Therefore $(m) \subseteq (a)$ and $(m) \subseteq (b)$. Hence,

$$(m) \subseteq (a) \cap (b) = (l).$$

So $m \in (l)$ and hence $l \mid m$. This proves that $l = \operatorname{lcm}(a, b)$.

## 3.20 Factorization domains

In $\mathbb{Z}$, we know that every positive integer bigger than 1 can be written uniquely as a product of distinct prime powers. Also we have seen that in $\mathbb{Z}$, prime elements and irreducible elements are same. We will now define a domain in which every non-zero non-unit element can be written as a product of irreducible factors.

**Definition 3.20.1** An integral domain $R$ with unity is called a **factorization domain** if every non-zero non-unit element can be written as a product of irreducible elements. In short we call it FD.

**Examples 3.20.2**

- The ring $\mathbb{Z}$ is an FD as every integer which is not in the set $\{0, -1, 1\}$ can be written as product of irreducible factors. The set $\{\pm p : \text{ where p is a prime number }\}$ is the set of all irreducible elements in $\mathbb{Z}$.
- All fields are factorization domains. Let $F$ be a field. Then every non-zero element of $F$ is a unit. So in $F$ there is no non-zero non-unit element which can not be written as a product of irreducible elements. In fact, in a field there is no irreducible elements.

**Theorem 3.20.3** Every PID with unity is an FD.

*Proof.* Let $R$ be a PID. Consider,

$$\Omega := \{x \in R : x \neq 0, x \text{ is non-unit and not a product of irreducibles}\}.$$

To show that $R$ is an FD, it is sufficient to show that $\Omega = \varnothing$. If $\Omega \neq \varnothing$. Let us consider,

$$\mathfrak{F} := \{(x) : x \in \Omega\}.$$

Note that the collection of principal ideals $\mathfrak{F}$ is non-empty as we assumed $\Omega$ is non-empty. Under the operation set inclusion, $\mathfrak{F}$ is a partially ordered set. Let $T$ be a non-empty chain in $\mathfrak{F}$. Define,

$$T_0 := \bigcup_{\mathfrak{I} \in T} \mathfrak{I}.$$

We show that $T_0$ is an ideal of $R$. We have $T_0 \neq \varnothing$ as $T \neq \varnothing$. Let $x, y \in T_0$. Therefore $x \in \mathfrak{I}_\alpha$ and $y \in \mathfrak{I}_\beta$ for some $\mathfrak{I}_\alpha, \mathfrak{I}_\beta \in T$. Without loss of generality let $\mathfrak{I}_\alpha \subseteq \mathfrak{I}_\beta$. So $x, y \in \mathfrak{I}_\beta$ and hence $x - y \in \mathfrak{I}_\beta$ and so $x - y \in T_0$. Also for any $r \in R$, the element $x \cdot r \in \mathfrak{I}_\alpha$, so it is in $T_0$. This proves that $T_0$ is an ideal of $R$. Since $R$ is a PID, there exists an element $d \in R$ such that $T_0 = (d)$. Now $d \in T_0$ implies that $d \in \mathfrak{I}$ for some $\mathfrak{I} \in T$. Hence,

$$(d) \subseteq \mathfrak{I} \subseteq T_0 = (d).$$

Thus, $T_0 = \mathfrak{I}$. Therefore $T$ is bounded above by $T_0$ in $\mathfrak{F}$. Now using Zorn's lemma we can conclude that $\mathfrak{F}$ has a maximal element say $\mathfrak{M}$. Again as $R$ is a PID we can find $m \in R$ so that $\mathfrak{M} = (m)$. As $\mathfrak{M} \in \mathfrak{F}$, we have in fact $m \in \Omega$. In particular $m$ is not an irreducible. Therefore there exist two non-zero non-units $b, c \in R$ such that $m = b \cdot c$. Note that $(m) \subsetneq (b)$ and $(m) \subsetneq (c)$. Because if $(m) = (b)$, then $b = m \cdot r$ for some

$r \in R$. Then $m = (m \cdot r) \cdot c$ and hence $r \cdot c = 1$, which contradicts the fact that $c$ is a non-unit. Similarly if $(m) = (c)$, then we will get a contradiction to the fact that $b$ is a non-unit. Now as $(m) \subsetneq (b)$ and $(m) \subsetneq (c)$ and $(m)$ is a maximal element of $\mathfrak{F}$, we get that $b, c \notin \Omega$. Therefore each of $b, c$ can be written as product of irreducible elements and so $m$ also can be written as product of irreducible elements as $m = b \cdot c$. This is a contradiction as $m \in \Omega$.

Therefore the set $\Omega$ has to be empty. This proves that $R$ is an FD.

**Exercise 3.20.4** Let $d \in \mathbb{N}$. Show that the ring $\mathbb{Z}[\iota\sqrt{d}]$ is an FD.

**Hint :** First note that for $d = 1$, the set of units is $\{\pm 1, \pm\iota\}$ and for $d > 1$ the set of units is $\{\pm 1\}$. Let $x = (a + b\iota\sqrt{d}) \in \mathbb{Z}[\iota\sqrt{d}]$. Define,

$$N(x) := (a + b\iota\sqrt{d}) \cdot (a - b\iota\sqrt{d}) = a^2 + b^2 d \in \mathbb{N}.$$

Observe that $N(xy) = N(x)N(y)$ for all $x, y \in \mathbb{Z}[\iota\sqrt{d}]$. Use induction on $N(x)$ to complete the proof.

## 3.21 Unique factorization domains

Now we define below a class of integral domains in which every non-zero non-unit there can be written as a product of irreducible elements uniquely up to associates.

**Definition 3.21.1** An integral domain $R$ with unity is called a **unique factorization domain** if

- $R$ is an FD,
- If $a \in R \setminus \{0\}$ be a non-unit such that $a = a_1 \cdots a_r = b_1 \cdots b_s$ where for each $1 \leq i \leq r$ and $1 \leq j \leq s$, $a_i, b_j$ are irreducible elements of $R$, then $r = s$ and each $a_i$ is an associate of a $b_j$ and vice-versa.

In short we call it a UFD.

**Theorem 3.21.2** In a UFD, every irreducible element is prime.

*Proof.* Let $R$ be a UFD and $a$ be an irreducible element of $R$. We show that $a$ is a prime element. Let $x, y \in R \setminus \{0\}$ be such that $a \mid x \cdot y$. We take $x, y$ both to be non-zero as we know that every non-zero element divides 0. As $a \mid x \cdot y$, there exists $b \in R$ such that

$a \cdot b = x \cdot y$. First we note that $x, y$ both can not be simultaneously units. Because if $x, y$ both are units then $x \cdot y$ is a unit. We know that a unit divides all the elements, so $x \cdot y \mid a$. This gives us, $a \sim x \cdot y$, which is not possible as $a$ is a non-unit. So at least one of $x, y$ is not a unit.
**Case I:** Suppose that $x$ is a unit and $y$ is a non-unit. Since $R$ is a UFD, we have $y = y_1 \cdots y_s$ where $y_1, \ldots, y_s$ are irreducible elements. Therefore,

$$a \cdot b = x \cdot y = x \cdot y_1 \cdots y_s.$$

Now $x \cdot y_1$ is also an irreducible element and we rename it as $y_1$ for simplicity. By the uniqueness of factorization we obtain,

$$a \sim y_i, \text{ for some } 1 \leq i \leq s.$$

This proves that $a \mid y$ as $y_i \mid y$.
**Case II:** If $x$ is a non-unit and $y$ is a unit. Then in this case we get $a \mid x$.
**Case III:** Suppose that $x, y$ both are non-units. In this case we can write, $x = x_1 \cdots x_r$ and $y = y_1 \cdots y_s$ where $x_1, \ldots, x_r$ and $y_1, \ldots, y_s$ are irreducible elements. Therefore,

$$a \cdot b = x \cdot y = x_1 \cdots x_r \cdot y_1 \cdots y_s.$$

Then by the uniqueness of the facorization, $a \sim x_i$ for some $1 \leq i \leq r$ or $a \sim y_j$ for some $1 \leq j \leq s$. This proves that $a \mid x$ or $a \mid y$.

**Corollary 3.21.3** We know that the integral domain $\mathbb{Z}[\iota\sqrt{5}]$ is an FD. But this is not a UFD, as we have seen that $3 \in \mathbb{Z}[\iota\sqrt{5}]$ is an irreducible element which is not prime. This also shows that subring of a UFD need not be a UFD. Consider $\mathbb{Z}[\iota\sqrt{5}]$ inside its field of fractions which is clearly a UFD, as every field is a UFD.

**Theorem 3.21.4** An integral domain is a UFD if and only if it is an FD where all irreducible elements are prime elements.

*Proof.* We know that a UFD is by definition an FD and by the previous theorem all irreducible elements are prime there. So we only need to prove the other part i.e. assuming $R$ is an FD where every irreducible element is prime, we shall show that $R$ is a UFD. Let $x \neq 0$ be a non-unit. Suppose,

$$x = x_1 \cdots x_r = x'_1 \cdots x'_s,$$

where $x_i, x'_j$ are irreducible elements for all $1 \leq i \leq r, 1 \leq j \leq s$. We prove that $r = s$ and each $x_i$ is an associate of some $x'_j$ and vice-versa. We have that $x_i, x'_j$ are all prime elements. We have, $r, s \geq 1$ as $x$ is a non-zero non-unit element. We prove the result by induction on $r$.

If $r = 1$, then $x = x_1$, so $x_1 \mid x'_j$ for some $1 \leq j \leq s$. Also $x'_j \mid x_1$ as $x'_j \mid x'_1 \cdots x'_s$. Therefore we have $x_1 \sim x'_j$. Now if $s > 1$, then we get $x'_1 \cdots x'_{j-1} x'_{j+1} \cdots x'_s$ is a unit, that implies $x'_1$ is a unit, which is not true. Hence $s = 1$.

Assume the induction hypothesis for $r - 1$. Now $x_1 \mid x_1 \cdots x_r$. Therefore $x_1 \mid x$ and so $x_1 \mid x'_1 \cdots x'_s$. Since $x_1$ is a prime, $x_1 \mid x'_j$ for some $1 \leq j \leq s$. Write,

$$x'_j = x_1 \cdot a.$$

Also $x'_j$ is a prime so $x'_j \mid x_1$ or $x'_j \mid a$. Now $x'_j \nmid a$ as if $x'_j \mid a$ then $a = b \cdot x'_j$ for some $b \in R$, so $x'_j = x_1 \cdot (b \cdot x'_j)$ and it gives us $x_1 \cdot b = 1$ i.e. $x_1$ is a unit, which is not true as $x_1$ is a prime. So $x'_j \mid x_1$. Therefore $x_1 \sim x'_j$. Now we have (after renaming),

$$x' = x_2 \cdots x_r = x'_1 \cdots x'_{j-1} \cdot x'_{j+1} \cdots x'_s.$$

By induction hypothesis we get $r - 1 = s - 1$ and each $x_i$ for $2 \leq i \leq r$ is an associate of $x'_t$ for some $1 \leq t \leq s$ and $t \neq j$. This completes the proof.

**Theorem 3.21.5** Every PID is a UFD.

*Proof.* We have already proved that every PID is an FD and in a PID every irreducible is prime. So therefore by the aid of the above theorem we are done showing that every PID is a UFD.

**Theorem 3.21.6** Let $R$ be a UFD with unity and $a, b \in R \setminus \{0\}$. Then $\gcd(a, b)$ and $\operatorname{lcm}(a, b)$ exist in $R$.

**Remark 3.21.7** We know that this is true in the case of a PID. But the proof does not extend in this case. We saw that in case of a PID, $\gcd(a, b)$ is a generator of the ideal $(a) + (b)$ and $\operatorname{lcm}(a, b)$ is a generator of the ideal $(a) \cap (b)$. But in general whenever $\gcd(a, b)$ and $\operatorname{lcm}(a, b)$ exist, we can only have,

$$(a) + (b) \subseteq (\gcd(a, b)) \quad \text{and} \quad (\operatorname{lcm}(a, b)) \subseteq (a) \cap (b) \quad \text{respectively.}$$

For example, consider $R = \mathbb{Q}[X, Y] = \mathbb{Q}[X][Y] = \mathbb{Q}[Y][X]$. Clearly $\gcd(X, Y) = 1$. But $(X) + (Y) \subsetneq (1) = R$ as $(X) + (Y)$ does not have any of the non-zero constant

polynomials. The fact that $R$ is a UFD follows from a theorem of Gauss, which is discussed in the following subsection.

**Proof of Theorem 3.21.6.** Let $a, b \in R \setminus \{0\}$. We can take them to be non-units, as otherwise if one of them, say $a$, is a unit then we get $\gcd(a, b) = 1$ and $\operatorname{lcm}(a, b) = b$. Now if both $a, b$ are non-units then we can write them as product of irreducible elements. Say

$$a = p_1^{e_1} \cdots p_r^{e_r} \text{ and } b = p_1^{f_1} \cdots p_r^{f_r},$$

where $e_i, f_i$ are non-negative integers for each $1 \leq i \leq r$. We set $d_i = \min(e_i, f_i)$ and $l_i = \max(e_i, f_i)$ for each $1 \leq i \leq r$. Consider the elements

$$d = p_1^{d_1} \cdots p_r^{d_r} \text{ and } l = p_1^{l_1} \cdots p_r^{l_r}.$$

We now show that $d = \gcd(a, b)$ and $l = \operatorname{lcm}(a, b)$.

Clearly, $d \mid a$ and $d \mid b$. Suppose $e$ is such that $e \mid a$ and $e \mid b$. If $e$ is a unit, we have nothing to show as $e$ then clearly divides $d$ as $d = e \cdot e^{-1} \cdot d$. So let $e$ be a non-unit. We can therefore write $e$ as product of irreducible elements. Since $e \mid a$ and $e \mid b$, the factorization of $e$ will be of the form

$$e = p_1^{g_1} \cdots p_r^{g_r}$$

with $g_i \leq e_i, f_i$ for each $1 \leq i \leq r$. Thus $g_i \leq \min(e_i, f_i) = d_i$ for each $1 \leq i \leq r$, equivalently $e \mid d$. Hence $d = \gcd(a, b)$.

Similarly, for $l$, we see that $a \mid l$ and $b \mid l$. Now suppose $m$ is such that $a \mid m$ and $b \mid m$. Clearly $m$ is a non-unit and we can therefore write $m$ as product of irreducible elements. Since $a \mid m$ and $b \mid m$, the factorization of $m$ will be of the form

$$m = p_1^{h_1} \cdots p_r^{h_r}$$

with $h_i \geq e_i, f_i$ for each $1 \leq i \leq r$. Thus $h_i \geq \max(e_i, f_i) = l_i$ for each $1 \leq i \leq r$, equivalently $l \mid m$. Hence $l = \operatorname{lcm}(a, b)$.

### 3.21.1 Primitive polynomials over a UFD

**Definition 3.21.8** Let $R$ be a UFD and $f$ be a non-zero polynomial in $R[X]$. By the **content** of $f$ we mean the greatest common divisor of all the non-zero co-efficients of $f$. We denote it by $c(f)$.

**Definition 3.21.9** Let $R$ be a UFD and $f$ be a non-zero polynomial in $R[X]$. If the content of $f$ is a unit, we call it a **primitive polynomial**.

**Example 3.21.10** Non-constant irreducible polynomials in $R[X]$ are primitive.

Next we show that product of two primitive polynomial is also a primitive polynomial.

**Theorem 3.21.11** Let $R$ be a UFD and $f(X), g(X) \in R[X]$ are two primitive polynomials. Then $f \cdot g$ is also a primitive polynomial.

*Proof.* Let us write,

$$f(X) = \sum_{i=0}^{m} a_i X^i \quad \text{and} \quad g(X) = \sum_{j=0}^{n} b_j X^j.$$

Hence,

$$(f \cdot g)(X) = \sum_{k=0}^{m+n} c_k X^k, \quad \text{where} \quad c_k = \sum_{i=0}^{k} (a_i \cdot b_{k-i}) \quad \text{for all } 0 \leq k \leq m+n.$$

Suppose $c(f \cdot g)$ is not a unit. Hence there exists a prime element $p \in R$ such that $p \mid c(f \cdot g)$, i.e. $p \mid c_k$ for all $0 \leq k \leq m+n$.

Since $c(f)$ and $c(g)$ are units, $p \nmid c(f)$ and $p \nmid c(g)$. Hence there are $a_i$ and $b_j$ such that $p \nmid a_i$ and $p \nmid b_j$. Let $r, s$ be the least non-negative integers such that $p \nmid a_r$ and $p \nmid b_s$. Hence $p \mid a_i$ for all $0 \leq i < r$ and $p \mid b_j$ for all $0 \leq j < s$. Consider

$$c_{r+s} = \sum_{i=0}^{r+s} (a_i \cdot b_{r+s-i}).$$

We rewrite it as

$$c_{r+s} = \sum_{0 \leq i < r} (a_i \cdot b_{r+s-i}) + a_r \cdot b_s + \sum_{r < i \leq r+s} (a_i \cdot b_{r+s-i}).$$

Note that when $i > r$, then $r + s - i < s$. Therefore we have,

$$p \mid \sum_{0 \leq i < r} (a_i \cdot b_{r+s-i}) \quad \text{and} \quad p \mid \sum_{r < i \leq r+s} (a_i \cdot b_{r+s-i}).$$

We also know $p \mid c_{r+s}$. Thus $p \mid a_r \cdot b_s$, which is a contradiction as $p$ is a prime and $p \nmid a_r$, $p \nmid b_s$. Hence $c(f \cdot g)$ is a unit, i.e. $f \cdot g$ is a primitive polynomial.

As a corollary we get that the content is multiplicative up to units.

**Corollary 3.21.12** Let $R$ be a UFD and $f(X), g(X) \in R[X]$ be two non-zero polynomials. Then $c(f \cdot g) = c(f) \cdot c(g)$ up to a unit.

*Proof.* Let us write $f(X) = c(f)f_1(X)$ and $g(X) = c(g)g_1(X)$. Then $f_1$ and $g_1$ are primitive polynomials. Now $(f \cdot g)(X) = (c(f) \cdot c(g))\,(f_1(X) \cdot g_1(X))$. From the above lemma we know that $f_1(X) \cdot g_1(X)$ is also a primitive polynomial. Hence by taking content of $(f \cdot g)(X)$ we get that $c(f \cdot g) = (c(f) \cdot c(g)) \cdot c(f_1 \cdot g_1)$ where we have $c(f_1 \cdot g_1)$ is a unit, this completes the proof.

### 3.21.2 Gauss's lemma

**Lemma 3.21.13 (Gauss's lemma)** Let $R$ be a UFD and $K$ be its field of fractions. A primitive polynomial $f \in R[X]$ is irreducible in $R[X]$ if and only if $f$ is irreducible in $K[X]$.

*Proof.* We first assume that $f$ is irreducible in $K[X]$, then we show that $f$ is irreducible in $R[X]$. Since $f$ is irreducible in $K[X]$, it is a non-zero non-unit in $K[X]$. As $R[X] \subseteq K[X]$, $f$ is also a non-zero non-unit in $R[X]$. Suppose $f = g \cdot h$ in $R[X]$. Now this factorization holds in $K[X]$ as well. Since $f$ is irreducible in $K[X]$, therefore exactly one of $g$ and $h$ is a unit in $K[X]$. Without loss of generality let us assume $g$ is a unit in $K[X]$. We know units in $K[X]$ are nothing but a non-zero constant polynomial. So $g$ is a constant in $K$. Since $g \in R[X]$, we get that $g \in R$. As we have, $f = g \cdot h$ we therefore can write

$$c(f) = c(g)c(h)u, \text{ where } u \text{ is a unit in } R.$$

Now since $g$ is a constant, we obtain, $c(g) = g$. This implies, $g \mid c(f)$. As $f$ is a primitive polynomial, we then conclude that $c(f)$ is a unit in $R$. Thus $g$ must be a unit in $R$. This means $g$ is a unit in $R[X]$. Hence $f$ must be irreducible in $R[X]$.

Next let $f$ be a primitive irreducible polynomial in $R[X]$. Hence $f$ is non-zero and non-unit in $R[X]$. We shall show that $f$ is an irreducible polynomial in $K[X]$. If possible suppose $f$ is not irreducible in $K[X]$. Then we have two possibilities, either $f$ is a unit in $K[X]$, or $f$ is a product of two non-units in $K[X]$.

**Case I:** Let $f$ be a unit in $K[X]$. Since $K$ is an integral domain, the only units in $K[X]$ are non-zero constant polynomials. Hence $f \in R[X]$ is a constant polynomial. Since $f$ is a primitive polynomial in $R[X]$, this constant must be a unit in $R$ and therefore $f$ can not be irreducible in $R[X]$. So we get a contradiction in this case.

**Case II:** Let $f$ be a product of two non-units $g, h$ in $K[X]$. Note that if

$$g = \frac{a_0}{b_0} + \frac{a_1}{b_1}X + \cdots + \cdots + \frac{a_n}{b_n}X^n, \quad \text{with} \quad a_i, b_i \in R \quad \text{for all} \quad 0 \leq i \leq n,$$

then by clearing denominator we can write,

$$g = \frac{a'}{b'} \cdot \hat{g} \quad \text{with } \hat{g} \in R[X] \quad \text{and} \quad a', b' \in R \setminus \{0\}.$$

Taking out the content of $\hat{g}$ we write,

$$g = \frac{a}{b} \cdot p,$$

where $p$ is a primitive polynomial in $R[X]$ and $a, b \in R \setminus \{0\}$. Since $g$ is a non-unit in $K[X]$, we have $p$ is a non-unit in $R[X]$. Similarly we write,

$$h = \frac{c}{d} \cdot q,$$

where $q$ is a non-unit and a primitive polynomial in $R[X]$ and $c, d \in R \setminus \{0\}$. Hence we have,

$$f = \frac{ac}{bd} \cdot (p \cdot q)\,.$$

i.e.

$$bdf = ac\,(p \cdot q)\,. \tag{3.21.1}$$

Since $p, q$ are primitive polynomials in $R[X]$, so is $p \cdot q$. Now taking content on both sides of (3.21.1) we get that

$$bd \cdot c(f) = ac \cdot c(p \cdot q),$$

where $c(f), c(p \cdot q)$ are all units in $R$. Therefore by multiplying by $(bd)^{-1} \in K$ in both sides we get,

$$\frac{ac}{bd} = c(f)c(p \cdot q)^{-1}.$$

Thus the ratio $ac/bd$ is a unit in $R$ as $c(f)c(p \cdot q)^{-1} \in R$ is unit. We call the ratio by $v$. So we have

$$f = v \cdot (p \cdot q)\,,$$

where $v \in R$ and $p, q$ are two non-units in $R[X]$. It contradicts the fact that $f$ is an irreducible polynomial in $R[X]$. Therefore we get that $f$ is a non-zero non-unit polynomial which can not be written as a product of two non-units in $K[X]$. So $f$ is

irreducible in $K[X]$.

### 3.21.3 Gauss's theorem on UFD

We prove here the Gauss's theorem on UFD. This enables us to construct new UFDs from existing UFDs by considering polynomial rings over them. We need the following theorems to prove Gauss's theorem on UFD.

**Theorem 3.21.14** Let $K$ be a field and $f \in K[X]$. For any non-zero polynomial $p \in K[X]$, there exist $q, r \in K[X]$ such that $f = p \cdot q + r$ where $r = 0$ or $\deg r < \deg p$.

*Proof.* Let $\deg p$ be $m \in \mathbb{N}$. If $f = 0$ or $\deg f < m$, then we take $q = 0$ and $r = f$ so that $f = p \cdot q + r$. So we prove the result for all polynomials of degree $\geq m$. We shall prove the theorem by induction on $\deg f$. We write,

$$f = \sum_{i=0}^{n} a_i X_i \quad \text{and} \quad p = \sum_{i=0}^{m} b_i X_i, \quad \text{with} \quad a_n, b_m \neq 0.$$

For $n = m$, we take,

$$q = a_m b_m^{-1}$$

and set

$$r = f - p \cdot q = \sum_{i=0}^{m} (a_i - b_i a_m b_m^{-1}) X_i.$$

Then we get that

$$f = p \cdot q + r \quad \text{with} \quad r = 0 \quad \text{or} \quad \deg r < m = \deg p.$$

Assume induction hypothesis for all polynomials of degree $\leq k - 1$. Let $n = k$, we take

$$q_1 = a_k b_m^{-1} X^{k-m}.$$

Consider,

$$g = f - p \cdot q_1.$$

Note that $\deg g < k$. If $g = 0$ or $\deg g < m = \deg p$, then we are done by setting $q = q_1$ and $r = g$. If we get

$$\deg g \geq m = \deg p,$$

then we use the induction hypothesis to write $g = p \cdot q_2 + r$ with $r = 0$ or $\deg r < \deg p$.

Hence we obtain,

$$f = p \cdot q + r$$

where

$$q = q_1 + q_2.$$

This completes the proof.

**Theorem 3.21.15** Let $K$ be a field. Then $K[X]$ is a PID and hence a UFD.

*Proof.* Let $I$ be an ideal in $K[X]$. If $I$ is the zero ideal then it is generated by the zero polynomial. Let $I \neq 0$. Consider $p(X)$ to be a non-zero polynomial of the least degree in $I$. We claim that $I$ is generated by $p(X)$. Let $f(X) \in I$, then by Theorem 3.21.14, there exist $q(X), r(X) \in K[X]$ such that $f(X) = p(X) \cdot q(X) + r(X)$ where $r(X) = 0$ or $\deg(r(X)) < \deg(p(X))$. Suppose $r(X) \neq 0$. Then we get,

$$\deg(r(X)) < \deg(p(X)).$$

But $r(X) = f(X) - p(X) \cdot q(X) \in I$. This contradicts the fact that $p(X)$ is a non-zero polynomial of the least degree in $I$. Hence $r(X) = 0$ and $f(X) = p(X) \cdot q(X)$. Therefore $f(X) \in (p(X))$ and so $I$ is a principal ideal generated by $p(X)$.

**Theorem 3.21.16 (Gauss)** Let $R$ be an integral domain. Then $R$ is a UFD if and only if $R[X]$ is a UFD.

*Proof.* We first assume that $R[X]$ is a UFD. We shall show that $R$ is an FD. Let $a \in R$ be a non-zero non-unit. Since $R \subset R[X]$ and $R$ is an ID, the units of $R[X]$ are precisely the units of $R$. Hence we get that $a$ is a non-zero non-unit in $R[X]$. Since $R[X]$ is a UFD, we can therefore write,

$$a = a_1(X) \cdots a_r(X)$$

where $a_i(X)$'s are irreducible elements in $R[X]$. By comparing the degrees on both the sides, we see that $a_i(X)$'s are constant polynomial for each $1 \leq i \leq r$. We write

$$a_i(X) = a_i \in R \quad \text{for } 1 \leq i \leq r$$

and so

$$a = a_1 \cdots a_r.$$

We now show that $a_i$'s are irreducible in $R$. Clearly $a_i$'s are non-units in $R$, as the units of $R[X]$ are only the units of $R$. Suppose $a_i$ is not irreducible in $R$ for some $i$. We write,

$$a_i = b_i \cdot c_i, \quad \text{where} \quad b_i, c_i \quad \text{are two non-units in} \quad R.$$

Since $b_i, c_i$ are non-units in $R[X]$ as well, we get that $a_i = a_i(X)$ is not irreducible in $R[X]$. This is a contradiction to our assumption in the factorization of $a$ in $R[X]$. Hence $a_i$'s are all irreducible elements of $R$ for each $1 \leq i \leq r$ and we have the factorization $a = a_1 \cdots a_r$ in $R$. This proves that $R$ is an FD. Hence by Theorem 3.21.4 it is now enough to prove that in $R$ every irreducible element is prime. Let $a \in R$ be an irreducible element. Hence $a$ is a non-unit in $R$ and therefore in $R[X]$. Note that $a$ is also irreducible as an element of $R[X]$. Now in the hypothesis we have $R[X]$ is a UFD and as $a$ is irreducible in $R[X]$, we therefore conclude that $a$ is a prime in $R[X]$. Now since $R \subset R[X]$, $a$ must be prime in $R$. This shows that $R$ is a UFD.

Next we assume that $R$ is a UFD. We shall show that $R[X]$ is a UFD. We first prove that $R[X]$ is an FD. Let $f \in R[X]$ be a non-zero non-unit element. Without loss of generality, we may assume that $f$ is a primitive polynomial. If $f$ is not a primitive polynomial, by taking out the content we can write $f = c(f)g$, where $g$ is a primitive polynomial. Now $c(f)$ is a non-zero non-unit in $R$ and therefore can be written as a product of irreducible elements in $R$. We have seen already that an irreducible element of $R$ is also irreducible in $R[X]$. Therefore $c(f)$, which is also a non-zero non-unit in $R[X]$, can be written as a product of irreducible elements in $R[X]$. Hence we are left to factorize primitive polynomials into irreducible elements of $R[X]$. We do this by induction on the degree of $f$. Note that since $f$ is a non-unit and a primitive polynomial, it can not be constant. Hence $\deg f \geq 1$. Let $\deg f = 1$. In this case we show that $f$ is an irreducible polynomial. If possible let $f = g \cdot h$ in $R[X]$. Since $\deg f = 1$, one among $g, h$ is a constant polynomial. Without loss of generality let $g$ be a constant polynomial. So $c(g) = g$. Now by taking content of $f$, we get

$$c(f) = c(g)c(h)u, \quad \text{where} \quad u \quad \text{is a unit .}$$

Since $f$ is a primitive polynomial, we know $c(f)$ is a unit. As

$$c(g)c(h)uc(f)^{-1} = 1,$$

we get that $g = c(g)$ is also a unit. Therefore, $f$ is irreducible in $R[X]$. Now let us assume the induction hypothesis for every polynomial of degree strictly less than $\deg f$.

If $f$ is irreducible, we are done. Let us assume that $f$ is not irreducible. Hence we can write $f = g \cdot h$ in $R[X]$, where $g, h$ are non-units. Note that $g, h$ are non-constant polynomials. For instance, if $g$ is a constant polynomial, then by taking content of $f$, we will get that $c(g)c(h)$ is a unit and hence $g = c(g)$ is a unit, which is a contradiction as we have assumed $g$ to be a non-unit. Since both $g, h$ are non-constant polynomials, $\deg g, \deg h < \deg f$. By induction hypothesis then each of them can be written as a product of irreducible polynomials and therefore $f$ can be written as a product of irreducible polynomials. This shows that $R[X]$ is an FD.

By Theorem 3.21.4 it is now enough to show that in $R[X]$, every irreducible element is prime. Let $p \in R[X]$ be an irreducible polynomial. If $\deg p = 0$, then $p \in R$ and since $R$ is a UFD, we get that $p$ is a prime element in $R$. Now if $p \mid fg$, then by taking content we see that $p = c(p) \mid c(f)c(g)$. Since $p$ is a prime element in $R$, we get $p \mid c(f)$ or $p \mid c(g)$. Hence $p \mid f$ or $p \mid g(X)$, as every polynomial is product of its content and some primitive polynomial. Hence $p$ is a prime element in $R[X]$. So let $\deg p \geq 1$ and $p \mid fg$ in $R[X]$. Note that since $p$ is an irreducible polynomial, it must be primitive, otherwise we can write it as product of its content and some primitive polynomial of degree $\geq 1$, which would contradict that $p$ is irreducible.

Now that $p$ is a primitive irreducible polynomial in $R[X]$, it is also irreducible in $K[X]$ by Gauss lemma, where $K$ is the field of fractions of $R$. By Theorem 3.21.15, $K[X]$ is a PID and hence a UFD. Therefore $p$ is in fact a prime element in $K[X]$. Hence we get $p \mid f$ or $p \mid g$ in $K[X]$. Without loss of generality, let us assume $p \mid f$. Hence we can write $f = pq$ for some polynomial $q \in K[X]$. Clearing the denominator we get a polynomial over $R$ and then taking out the content, we can write

$$q = \frac{a}{b} r$$

such that $r$ is a primitive polynomial in $R[X]$ and $a, b \in R$. Hence

$$f = \frac{a}{b} pr.$$

Taking content we get,

$$c(f) = \frac{a}{b} c(pr).$$

Since $p, r$ are primitive polynomials in $R[X]$, we get that $c(pr)$ is a unit in $R$. Hence

$$\frac{a}{b} = c(f)v,$$

where $v$ is a unit in $R$. Thus

$$f = c(f)vpr$$

in $R[X]$. Hence $p \mid f$ in $R[X]$. Hence $p$ is a prime element in $R[X]$. This shows that every irreducible element in $R[X]$ is prime and hence by Theorem 3.21.4 $R[X]$ is a UFD.

**Remark 3.21.17** We have already seen that for a field $K$, $K[X]$ is a PID and hence a UFD. Therefore, by Gauss's theorem $K[X,Y] = K[X][Y]$ is also a UFD. In particular $\mathbb{Q}[X,Y]$ is a UFD, as mentioned in 3.21.7.

**Remark 3.21.18** As a corollary we also get that $\mathbb{Z}[X]$ is a UFD. We have seen that $\mathbb{Z}[\iota\sqrt{5}]$ is not a UFD as 3 is an irreducible element in $\mathbb{Z}[\iota\sqrt{5}]$, which is not prime. Note that $\mathbb{Z}[\iota\sqrt{5}] \simeq \mathbb{Z}[X]/I$, where $I$ is the ideal generated by the polynomial $x^2+5$ is $\mathbb{Z}[X]$. Hence we get that quotient of a UFD need not be a UFD.

**Corollary 3.21.19** More generally for a UFD $R$ and an arbitrary collection of symbols $\{X_\alpha\}_{\alpha\in I}$, where $I$ is a set of (possibly infinitely many) indices, $R[X_\alpha : \alpha \in I]$ is also a UFD. Note that for a polynomial $f$ in $R[X_\alpha : \alpha \in I]$, we have only finitely many $X_\alpha$'s, say $X_{\alpha_1}, \ldots, X_{\alpha_n}$, such that $f$ belongs to $R[X_{\alpha_1}, \ldots, X_{\alpha_n}]$. Now it is easy to see that by recurrent use of Gauss's theorem, we get $R[X_{\alpha_1}, \ldots, X_{\alpha_n}]$ is a UFD. Hence any non-zero non-units can be written as product of irreducible elements. Hence $R[X_\alpha : \alpha \in I]$ is an FD. Now since for any polynomial in $R[X_{\alpha_1}, \ldots, X_{\alpha_n}]$, the irreducible factors of it also lie in $R[X_{\alpha_1}, \ldots, X_{\alpha_n}]$ and since $R[X_{\alpha_1}, \ldots, X_{\alpha_n}]$ is a UFD, the factorization of $f$ in $R[X_\alpha : \alpha \in I]$ is unique up to associates. Therefore $R[X_\alpha : \alpha \in I]$ is a UFD.

### 3.21.4 Eisenstein's criterion

We have seen that any non-constant irreducible polynomial is primitive. Eisenstein provided a criterion for a non-constant primitive polynomial over a UFD to be irreducible.

**Theorem 3.21.20 (Eisenstein's criterion)** Let $R$ be a UFD and $f \in R[X]$ be a non-constant primitive polynomial. Let $f = \sum_{i=0}^{n} a_i X^i$ with $a_n \neq 0$. Suppose there exists a prime $p \in R$ such that $p \mid a_i$ for each $0 \leq i \leq n-1$ and $p \nmid a_n$, $p^2 \nmid a_0$. Then $f$ is irreducible in $R[X]$.

*Proof.* Given $f \in R[X]$ is a non-constant polynomial. So $f$ is non-zero non-unit in $R[X]$. If possible let us assume that $f$ is not irreducible in $R[X]$. We then write,

$$f = g \cdot h, \quad \text{where} \quad g, h \quad \text{are non-units in} \quad R[X].$$

Now $c(f) = c(g)c(h)u$ where $u$ is a unit. We have $f$ is primitive, so $c(f)$ is unit. Hence $c(g), c(h)$ both are units in $R$. Hence $g, h$ are non-constant polynomials. Because if say $g$ is constant then we get $g = c(g)$. This is not possible as we assumed $g$ is non-unit. We write,

$$g = \sum_{i=0}^{m} b_i X^i \quad \text{and} \quad h = \sum_{i=0}^{r} c_i X^i \quad \text{with } b_m, c_r \neq 0.$$

Since $f = g \cdot h$, we have $a_0 = b_0 \cdot c_0$. Now as $p \mid a_0$, we get $p$ divides at least one of $b_0$ and $c_0$. But $p$ can not divide both of them, as otherwise $p^2$ would divide $a_0$. Hence either $p$ divides $b_0$ or $p$ divides $c_0$. Without loss of generality let us assume $p \mid b_0$ and $p \nmid c_0$. We have $p \nmid a_n$ and $a_n = b_m \cdot c_r$. So $p \nmid b_m$. Now let $1 \leq l \leq m$ be the least positive integer such that $p \nmid b_l$. Since $g, h$ both are non-constant polynomials, we have $l \leq m < n$. This implies, $p \mid a_l$. Note that

$$a_l = \sum_{i=0}^{l} b_i \cdot c_{l-i}.$$

Since $l$ is the least positive integer such that $p \nmid b_l$, we get $p \mid \sum_{i=0}^{l-1} b_i \cdot c_{l-i}$. We have,

$$a_l = \sum_{i=0}^{l-1} (b_i \cdot c_{l-i}) \; + \; b_l \cdot c_0.$$

Hence $p \mid b_l \cdot c_0$. But we know that $p \nmid b_l$ and $p \nmid c_0$. This is a contradiction as $p$ is a prime. Hence $f$ must be irreducible in $R[X]$. This completes the proof.

**Corollary 3.21.21** As an immediate corollary, we get that there are infinitely many irreducible polynomials of any given degree $\geq 1$. For a given degree $n \geq 1$, consider the polynomials $X^n + p \in \mathbb{Z}[X]$, where $p$ is a prime number. As there are infinitely many prime numbers in $\mathbb{Z}$, we get infinitely many irreducible polynomials of degree $n$.

**Remark 3.21.22** Eisenstein's criterion is only a sufficient condition to ensure irreducibility of a polynomial. It is not necessary. For example, the element $1 + x + x^2$ in $\mathbb{Z}[X]$ is irreducible and primitive. But there does not exist any prime satisfying the conditions of Eisenstein's criterion. However, the irreducibility of $1 + x + x^2$ in $\mathbb{Z}[X]$ can be proved using Eisenstein's criterion.

**An application of Eisenstein's criterion:** For any prime number $p$, the polynomial $f_p(X) = 1 + X + \cdots + X^{p-1}$ is irreducible in $\mathbb{Z}[X]$.
Clearly $f_p(X)$ is a non-constant primitive polynomial in $\mathbb{Z}[X]$. We first show that the

polynomial $f_p(X+1)$ is irreducible. Note that,

$$f_p(X) = \frac{X^p - 1}{X - 1}.$$

Therefore,

$$f_p(X+1) = \frac{(X+1)^p - 1}{X} = \sum_{i=0}^{p-1} \binom{p}{p-1-i} X^i.$$

This is a monic polynomial and hence non-constant and primitive. Further $p$ is a prime satisfying Eisenstein's criterion. Hence $f_p(X+1)$ is irreducible. Now if we write $f_p(X) = g(X) \cdot h(X)$, where $g(X), h(X) \in \mathbb{Z}[X]$, then by a change of variable we get

$$f_p(X+1) = g(X+1) \cdot h(X+1).$$

Since $f_p(X+1)$ is irreducible, exactly one of $g(X+1), h(X+1)$ is a unit. Without loss of generality we assume $g(X+1)$ is a unit. Since in $\mathbb{Z}[X]$, only possible units are $\pm 1$, we get $g(X+1) = \pm 1$. Therefore $g(X) = \pm 1$ and hence $f_p(X)$ is irreducible.

## 3.22 Euclidean domains

**Definition 3.22.1** An integral domain $R$ is said to be a Euclidean domain (in short ED) if there exists a map $d : R \setminus \{0\} \to \mathbb{N} \cup \{0\}$ such that

(i) for all $a, b \in R \setminus \{0\}$, $a \mid b$ implies $d(a) \leq d(b)$,

(ii) for all $a \in R \setminus \{0\}$ and $b \in R$, there exist $q, r \in R$ such that $b = aq + r$ with either $r = 0$ or $d(r) < d(a)$.

The map $d$ is called a Euclidean algorithm map. The elements $q, r$ in (ii) above are called quotient and remainder (or residue), respectively.

**Examples 3.22.2**

- $\mathbb{Z}$ is an ED. Consider the map $d : \mathbb{Z} \setminus \{0\} \to \mathbb{N} \cup \{0\}$ defined by

$$d(n) := |n| = \begin{cases} n & \text{if } n > 0, \\ -n & \text{if } n < 0. \end{cases}$$

Clearly if $a, b \in \mathbb{Z} \setminus \{0\}$ be such that $a \mid b$, then $b = ac$ for some $c \in \mathbb{Z} \setminus \{0\}$. Hence $|b| = |ac| = |a||c|$. Since $|c| \geq 1$, we have $|a| \leq |b|$. Now for any given integer $a > 0$, we can write $\mathbb{Z}$ as a disjoint union of blocks of $a$ many consecutive integers. To be precise,

$$\mathbb{Z} = \bigcup_{k \in \mathbb{Z}} \{ka + 1, ka + 2, \ldots, (k + 1)a\}.$$

So for any given $n \in \mathbb{Z}$, there exists $k_0 \in \mathbb{Z}$ such that $k_0 a + 1 \leq n \leq (k_0 + 1)a$. If $n = (k_0 + 1)a$, we take $q = (k_0 + 1)$ and $r = 0$, so that $n = aq + r$. If $n \neq (k_0 + 1)a$, then $n = k_0 a + r$ with $1 \leq r < a$. For $a < 0$, we repeat the above process with $-a$. Since $d(a) = d(-a)$, this shows that $d$ is a Euclidean algorithm map.

- Every field is an ED. Let $K$ be a field and consider the map $d : K \setminus \{0\} \to \mathbb{N} \cup \{0\}$ defined by $d(a) := 1$ for all $a \in K \setminus \{0\}$. Since in $K$, every non-zero element is a unit, every non-zero element divides every other element in $K$ and hence clearly $d$ is a Euclidean algorithm map.

- For a field $K$, the ring of polynomial $K[X]$ is an ED. Consider the map $d : K[X] \setminus \{0\} \to \mathbb{N} \cup \{0\}$ defined by $d(f) := \deg f$. Then for two non-zero polynomials $f, g$ if $f \mid g$, then we can write $g = f \cdot h$ for some non-zero polynomial $h$. Since $K$ is an integral domain, we get

$$\deg g = \deg f + \deg h.$$

  Hence $\deg g \geq \deg f$. Now for given $f \in K[X]$ and any non-zero polynomial $p \in K[X]$, by Theorem 3.21.14 there exist $q, r \in K[X]$ such that $f = p \cdot q + r$ where $r = 0$ or $\deg r < \deg p$. Hence $d$ is a Euclidean algorithm map.

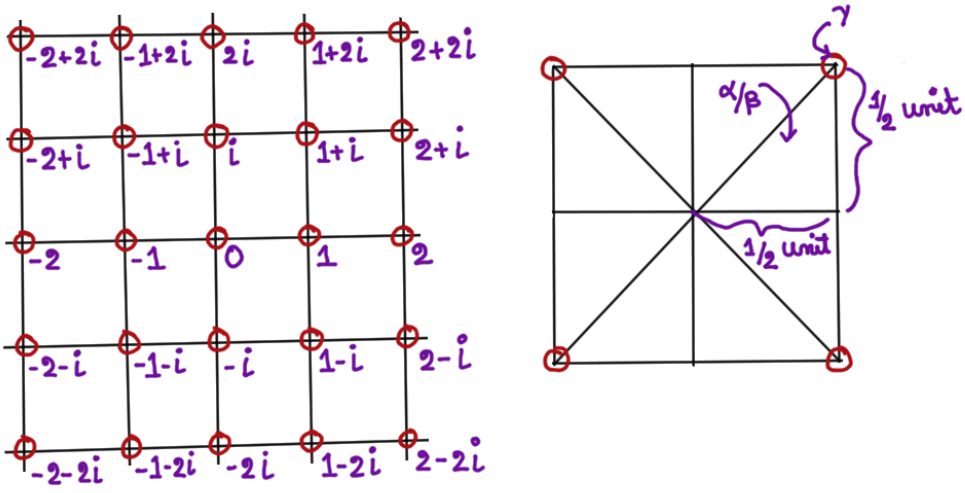
- The ring $\mathbb{Z}[\iota] := \{a + \iota b : a, b \in \mathbb{Z}\}$ is an ED. Consider the map $d : \mathbb{Z}[\iota] \setminus \{0\} \to \mathbb{N} \cup \{0\}$ defined by $d(a + \iota b) := a^2 + b^2 = |a + \iota b|^2$, where the modulus is the modulus of a complex number. Now let $a_1 + \iota b_1, a_2 + \iota b_2$ be two non-zero elements of $\mathbb{Z}[\iota]$ and $(a_1 + \iota b_1) \mid (a_2 + \iota b_2)$. Then we can write $a_2 + \iota b_2 = (a_1 + \iota b_1) \cdot (a_3 + \iota b_3)$ for some non-zero $a_3 + \iota b_3 \in \mathbb{Z}[\iota]$. Hence $|a_2 + \iota b_2| = |a_1 + \iota b_1||a_3 + \iota b_3|$ and hence $d(a_2 + \iota b_2) = d(a_1 + \iota b_1)d(a_3 + \iota b_3)$. Now $d(a_3 + \iota b_3) = a_3^2 + b_3^2 \geq 1$. Hence $d(a_2 + \iota b_2) \geq d(a_1 + \iota b_1)$.

  Now let $\alpha, \beta \in \mathbb{Z}[\iota]$ with $\beta \neq 0$. We have to find $\gamma, \delta \in \mathbb{Z}[\iota]$ such that $\alpha = \beta\gamma + \delta$ with either $\delta = 0$ or $d(\delta) < d(\beta)$. Let us note that it is enough to find $\gamma \in \mathbb{Z}[\iota]$ such that the distance of the complex number $\alpha/\beta$ from $\gamma$ is less than 1, i.e.

$$|\alpha/\beta - \gamma| < 1.$$

For this we set $\delta := \alpha - \beta\gamma$. So if $\alpha = \beta\gamma$, then $\delta = 0$ and if $\alpha \neq \beta\gamma$, then $\delta \neq 0$. Further $|\alpha/\beta - \gamma| < 1$ implies $|\alpha - \beta\gamma| < |\beta|$, i.e. $d(\delta) < d(\beta)$.

Now to find $\gamma \in \mathbb{Z}[\iota]$ such that $|\alpha/\beta - \gamma| < 1$, we divide the complex plane by an infinite grid made of the lines $x = a$ and $y = b$ as $a, b$ varies in $\mathbb{Z}$ (see the picture below). Note that the intersection point of two such lines is an element of $\mathbb{Z}[\iota]$. Now the complex number $\alpha/\beta$ lies in one of the unit squares of this infinite grid. That unit square can further be subdivided into eight equal sized right angle triangles (see the picture below).




This is done by drawing the chords and then bisecting each of the four resulting triangles. Hence $\alpha/\beta$ lies in one of these right angle triangles which are of height and base of $1/2$ unit each. Note that since the vertices of the concerned square are elements of $\mathbb{Z}[\iota]$, each triangle contains an element $\mathbb{Z}[\iota]$. We would like to calculate an upper bound of the distance of $\alpha/\beta$ from this point, say $\gamma$. We know that in a right angle triangle the maximum distance between any two points is same as the length of the hypotenuse. Since the right angle triangles are of height and base of $1/2$ unit, the hypotenuse is of $1/\sqrt{2}$ unit. Hence $|\alpha/\beta - \gamma| \leq 1/\sqrt{2} < 1$. Hence $d$ is a Euclidean algorithm map.

**Theorem 3.22.3** Every ED is a PID.

*Proof.* Let $R$ be an ED and $d : R \setminus \{0\} \to \mathbb{N}$ be an Euclidean algorithm map. Let $I$ be an ideal of $R$. If $I$ is the zero ideal then $I = (0)$. Now if $I \neq \{0\}$, then the set $d(I \setminus \{0\})$ is a non-empty subset of $\mathbb{N}$. So by well-ordering principle of $\mathbb{N}$, the set $d(I \setminus \{0\})$ has

a least element, say $m$. We choose $a \in I \setminus \{0\}$ be such that $d(a) = m$. We claim that $I = (a)$. For this let $b \in I$, then we can find $q, r \in R$ such that $b = aq + r$ with either $r = 0$ or $d(r) < d(a)$. If $r \neq 0$, then $d(r) < d(a)$. As $r = b - aq \in I$, it contradicts the minimality of $d(a)$ in $d(I \setminus \{0\})$. Hence $r$ has to be zero and thus $b = aq \in (a)$.

Recall that in our definition of an integral domain, we only considered it to be a non-zero commutative ring which does not have any non-zero zero divisor. But an ED always have a unity.

**Proposition 3.22.4** Every ED contains unity.

*Proof.* Let $R$ be an ED and $d : R \setminus \{0\} \to \mathbb{N}$ be a Euclidean algorithm map. The set $d(R \setminus \{0\})$ is a non-empty subset of $\mathbb{N}$. Hence it has a least element, say $m$. We choose $a \in R \setminus \{0\}$ be such that $d(a) = m$. Hence for all $b \neq 0$ in $R$, $d(a) \leq d(b)$. We claim that for all $b \in R$, $a \mid b$. We know that there exist $q, r \in R$ such that $b = aq + r$ with either $r = 0$ or $d(r) < d(a)$. If $r \neq 0$, then $d(r) < d(a)$. This contradicts the minimality of $d(a)$ in $d(R \setminus \{0\})$. Hence $r$ must be zero and thus $a \mid b$. In particular, $a \mid a$. So there exists $e \in R \setminus \{0\}$ such that $a = ae$. Hence for any $b \in R$, $ab = aeb$, i.e. $a(b - eb) = 0$. Since $a \neq 0$ and $R$ is an integral domain we have $b - eb = 0$, i.e. $b = eb$ for all $b \in R$. This proves that $e$ is the unity in $R$.

**Remark 3.22.5** This enables us to give a simple example of the fact that every PID need not be an ED. Note that $2\mathbb{Z}$ is a PID, but it has no unity and hence is not an ED. This also shows that subring of an ED need not be an ED. It is possible to find example of a PID with unity which is not an ED. Perhaps the most popular one is $\mathbb{Z}\left[\frac{1+\iota\sqrt{19}}{2}\right]$. The proof that $\mathbb{Z}\left[\frac{1+\iota\sqrt{19}}{2}\right]$ is a PID requires tools from algebraic number theory and hence out of the scope of this course.

**Exercise 3.22.6** Consider the map $d : 2\mathbb{Z} \setminus \{0\} \to \mathbb{N} \cup \{0\}$ defined by $d(n) := |n|$. Show that $d$ is not a Euclidean algorithm map.

We have already seen that for a field $K$, the polynomial ring $K[X]$ is a PID, more precisely an ED. The following theorem says that all these statements are in fact equivalent.

**Theorem 3.22.7** Let $R$ be an integral domain with unity. Then the following are equivalent:

1. $R$ is a field.

2. $R[X]$ is an ED.

3. $R[X]$ is a PID.

*Proof.* We know that if $R$ is a field then $R[X]$ is an ED and hence a PID. Now let us assume $R[X]$ is a PID. We define a map $\phi : R[X] \to R$ by

$$\phi(f(X)) := \phi(f(0)),$$

i.e. if

$$f(X) = \sum_{i=0}^{n} a_i X^i, \quad \text{then} \quad \phi(f(X)) = a_0.$$

It is easy to check that $\phi$ is a ring homomorphism. Further, considering the constant polynomials we can see that the map $\phi$ is surjective. Hence

$$R[X]/\ker\phi \simeq R.$$

Let us know find out $\ker\phi$. Let $f(X) = \sum_{i=0}^{n} a_i X^i \in \ker\phi$. Hence $\phi(f(X)) = a_0 = 0$. Therefore,

$$f(X) = \sum_{i=1}^{n} a_i X^i = X \sum_{i=0}^{n-1} a_{i+1} X^i.$$

Hence $f(X)$ belongs to the ideal generated by $X$. Note that $\phi(X) = 0$. Hence $X \in \ker\phi$. So $\ker\phi = (X)$, the ideal generated by $X$. Hence

$$R[X]/(X) \simeq R.$$

Since $R$ is an integral domain, by Theorem 3.14.1 we get that $(X)$ is a prime ideal. As it is a non-zero ideal, by Theorem 3.19.5, $(X)$ is also a maximal ideal. Hence $R$ is a field by Theorem 3.14.3.

## 3.23 Some counter examples

We end this chapter with the following examples. We have the following implications

$$\text{ED} \implies \text{PID with unity} \implies \text{UFD} \implies \text{FD} \implies \text{ID}.$$

However the reverse inclusions do not hold.

**PID which is not an ED** : $2\mathbb{Z}$ is a PID but not an ED as it has no unity, $\mathbb{Z}\left[\frac{1+\iota\sqrt{19}}{2}\right]$ is a PID but not ED (this requires tools from algebraic number theory).

**UFD which is not a PID** : $\mathbb{Z}[X]$, by Gauss's theorem we get $\mathbb{Z}[X]$ is a UFD, but since $\mathbb{Z}$ is not a field, by Theorem 3.22.7 it is not a PID.

**FD which is not a UFD** : $\mathbb{Z}[\iota\sqrt{5}]$ is a FD but not a UFD since 3 is an irreducible element which is not prime.

**ID which is not an FD** : Consider the set

$$\overline{\mathbb{Z}} := \{a \in \mathbb{C} : a \text{ is a root of a monic polynomial in } \mathbb{Z}[X]\}.$$

It is called the integral closure of $\mathbb{Z}$ in $\mathbb{C}$ (or the ring of algebraic integers). The proof that this set is indeed a subring of $\mathbb{C}$, requires tools from module theory. Clearly $\mathbb{Z} \subset \bar{\mathbb{Z}}$ as for any $n \in \mathbb{Z}$, it is root of the polynomial $X - n$. Note that, $\iota \in \overline{\mathbb{Z}}$ as it is a root of the polynomial $X^2 + 1$. So $\mathbb{Z} \subsetneq \overline{\mathbb{Z}}$. Since it is a subring of $\mathbb{C}$, it is an integral domain. But it is not a field as $1/2 \notin \overline{\mathbb{Z}}$. Hence $\overline{\mathbb{Z}}$ has non-zero non-units. Now we show that any non-unit that we choose from $\overline{\mathbb{Z}}$, is not irreducible. Let $r \neq 0$ be a non-unit in $\overline{\mathbb{Z}}$. As it is a zero of a monic polynomial in $\mathbb{Z}[X]$, let

$$r^n + a_1 r^{n-1} + \cdots + a_{n-1} r + a_n = 0.$$

We can rewrite this as

$$(\sqrt{r})^{2n} + a_1(\sqrt{r})^{2(n-1)} + \cdots + a_{n-1}(\sqrt{r})^2 + a_n = 0.$$

Hence $\sqrt{r}$ is in $\overline{\mathbb{Z}}$. Since $r$ is a non-unit, so is $\sqrt{r}$, as $r = (\sqrt{r})^2$. Hence $r$ can be written as a product of two non-units. So there are non-units in $\overline{\mathbb{Z}}$ but there is no irreducible element in $\overline{\mathbb{Z}}$ and hence $\overline{\mathbb{Z}}$ is not an FD.

# Bibliography

[1] Abstract Algebra by Dummit, Foote.

[2] Topics in Algebra by Herstein.

[3] Introduction to Rings and Modules by Musili.

[4] Topics in Abstract Algebra by Sen, Ghosh, Mukhopadhyay.